

FIRST-PRINCIPLES QUANTUM
SIMULATIONS OF MANY-MODE
OPEN INTERACTING BOSE GASES
USING STOCHASTIC GAUGE
METHODS

A THESIS SUBMITTED FOR
THE DEGREE OF DOCTOR OF PHILOSOPHY
AT THE UNIVERSITY OF QUEENSLAND IN
JUNE 2004



Piotr Paweł Deuar, BSc. (Hons)

School of Physical Sciences

Statement of Originality

Except where acknowledged in the customary manner, the material presented in this thesis is, to the best of my knowledge and belief, original and has not been submitted in whole or in part for a degree in any university.

Piotr P. Deuar

Peter D. Drummond

Statement of Contribution by Others

The original concepts of the gauge P representation and of (drift) stochastic gauges were due to Peter Drummond, as was the suggestion to investigate boundary term removal and thermodynamics of uniform 1D gases. The concepts in Section 5.7 and Subsection 6.1.4, are also due to Peter Drummond, but are included because they are important for a background understanding.

The XMDS program [<http://www.xmds.org/>] (at the time, in 2001, authored by Greg Collecutt and Peter Drummond) was used for the calculations of Chapter 6, while the calculations in the rest of the thesis were made with programs written by me but loosely based on that 2001 version of XMDS.

The exact Yang & Yang solutions in Figure 11.10 were calculated with a program written by Karen Kheruntsyan.

Piotr P. Deuar

Peter D. Drummond

Acknowledgements

Above all I thank Prof. Peter Drummond, my principal Ph.D. supervisor. You led by example, and showed me what true scientific work is all about — along with the sheer glee of it. Our many physics discussions were always extremely illuminating and productive. Thank you for your patience (which I hope not to have stretched too thin) and positive attitude.

I also owe much to Drs. Bill Munro and Karen Kheruntsyan, my secondary supervisors in the earlier and later parts of my candidature, respectively. Your assistance through this whole time has been invaluable. I would also like to thank Dr. Margaret Reid, who first introduced me to real scientific work in my Honours year.

My thanks go also to my fellow Ph.D. students (mostly now Drs.) at the physics department, for the innumerable discussions and intellectually stimulating atmosphere. Especially Damian Pope, Timothy Vaughan, Joel Corney, and Greg Collett by virtue of the greater similarity of our research. It is much easier to write a thesis with a good example — thank you Joel. I am also conscious that there are many others at the department who have assisted me in innumerable ways.

From outside of U.Q., I thank especially Prof. Ryszard Horodecki from Gdańsk University both for your outstanding hospitality and the opportunity of scientific collaboration. I also thank Dr. Marek Trippenbach and Jan Chwedeńczuk from Warsaw University for stimulating discussions during the final stages of my thesis.

More personally, I thank my wife Maria, for her support and above all her extreme tolerance of the husband engrossed in physics, particularly while I was writing up.

Publications by the Candidate Relevant to the Thesis but not forming part of it

Some of the research reported in this thesis has been published in the following refereed publications:

1. P. Deuar and P. D. Drummond. Stochastic gauges in quantum dynamics for many-body simulations. *Computer Physics Communications* **142**, 442–445 (Dec. 2001).
2. P. Deuar and P. D. Drummond. Gauge P-representations for quantum-dynamical problems: Removal of boundary terms. *Physical Review A* **66**, 033812 (Sep. 2002).
3. P. D. Drummond and P. Deuar. Quantum dynamics with stochastic gauge simulations. *Journal of Optics B — Quantum and Semiclassical Optics* **5**, S281–S289 (June 2003).
4. P. D. Drummond, P. Deuar, and Kheruntsyan K. V. Canonical Bose gas simulations with stochastic gauges. *Physical Review Letters* **92**, 040405 (Jan. 2004).

This is noted in the text where appropriate.

Abstract

The quantum dynamics and grand canonical thermodynamics of many-mode (one-, two-, and three-dimensional) interacting Bose gases are simulated from first principles. The model uses a lattice Hamiltonian based on a continuum second-quantized model with two-particle interactions, external potential, and interactions with an environment, with no further approximations. The interparticle potential can be either an (effective) delta function as in Bose-Hubbard models, or extended with a shape resolved by the lattice.

Simulations are of a set of stochastic equations that in the limit of many realizations correspond exactly to the full quantum evolution of the many-body systems. These equations describe the evolution of samples of the gauge P distribution of the quantum state, details of which are developed.

Conditions under which general quantum phase-space representations can be used to derive stochastic simulation methods are investigated in detail, given the criteria: 1) The simulation corresponds exactly to quantum mechanics in the limit of many trajectories. 2) The number of equations scales linearly with system size, to allow the possibility of efficient first-principles quantum mesoscopic simulations. 3) All observables can be calculated from one simulation. 4) Each stochastic realization is independent to allow straightforward use of parallel algorithms. Special emphasis is placed on allowing for simulation of open systems. In contrast to typical Monte Carlo techniques based on path integrals, the phase-space representation approach can also be used for dynamical calculations.

Two major (and related) known technical stumbling blocks with such stochastic simulations are instabilities in the stochastic equations, and pathological trajectory distributions as the boundaries of phase space are approached. These can (and

often do) lead to systematic biases in the calculated observables. The nature of these problems are investigated in detail.

Many phase-space distributions have, however, more phase-space freedoms than the minimum required for exact correspondence to quantum mechanics, and these freedoms can in many cases be exploited to overcome the instability and boundary term problems, recovering an unbiased simulation. The stochastic gauge technique, which achieves this in a systematic way, is derived and heuristic guidelines for its use are developed.

The gauge P representation is an extension of the positive P distribution, which uses coherent basis states, but allows a variety of useful stochastic gauges that are used to overcome the stability problems. Its properties are investigated, and the resulting equations to be simulated for the open interacting Bose gas system are derived.

The dynamics of the following many-mode systems are simulated as examples: 1) Uniform one-dimensional and two-dimensional Bose gases after the rapid appearance of significant two-body collisions (e.g. after entering a Feshbach resonance). 2) Trapped bosons, where the size of the trap is of the same order as the range of the interparticle potential. 3) Stimulated Bose enhancement of scattered atom modes during the collision of two Bose-Einstein condensates. The grand canonical thermodynamics of uniform one-dimensional Bose gases is also calculated for a variety of temperatures and collision strengths. Observables calculated include first to third order spatial correlation functions (including at finite interparticle separation) and momentum distributions. The predicted phenomena are discussed.

Improvements over the positive P distribution and other methods are discussed, and simulation times are analyzed for Bose-Hubbard lattice models from a general perspective. To understand the behavior of the equations, and subsequently optimize the gauges for the interacting Bose gas, single- and coupled two-mode dynamical and thermodynamical models of interacting Bose gases are investigated in detail. Directions in which future progress can be expected are considered.

Lastly, safeguards are necessary to avoid biased averages when exponentials of Gaussian-like trajectory distributions are used (as here), and these are investigated.

Contents

List of Figures	xv
List of Tables	xvi
Thesis Rationale and Structure	1
Rationale	1
Structure	6
1 Introduction: Quantum many-body simulations	8
1.1 Hilbert space complexity	9
1.2 Bulk properties are the key	10
1.3 Sampling system configurations	11
1.4 Path integral Monte Carlo and static thermal calculations	13
1.5 phase-space evolution methods	14
1.6 Open systems and the mode formulation	15
2 System of choice: Open interacting Bose gases	19
2.1 Motivation	20
2.2 Field model with binary interactions	23
2.3 Reduction to a lattice model	23
2.4 Locally interacting lattice model	26
2.5 Open dynamic equations of motion	28
2.6 Thermodynamic equations of motion	30

Part A Phase-space distribution methods for many-mode quantum mechanics 32

3	Generalized quantum phase-space representations	33
3.1	Introduction	33
3.2	Representation of a density matrix	34
3.2.1	Properties of the distribution	35
3.2.2	Splitting into subsystems	36
3.2.3	Dual configuration space of off-diagonal kernels	36
3.2.4	An example: the positive P distribution	38
3.3	Stochastic interpretation of the distribution	39
3.3.1	Calculating observables	39
3.3.2	Assessing estimate accuracy	40
3.3.3	Calculating non-static observables	41
3.3.4	Overcomplete vs. orthogonal bases	43
3.3.5	Positive P distribution example continued	46
3.3.6	Non-uniqueness of distributions	47
3.4	Equations of motion	49
3.4.1	Master equation to Fokker-Planck equation	50
3.4.2	Fokker-Planck equation to stochastic equations	53
3.4.3	Ensuring positivity of diffusion for analytic kernels	55
3.4.4	Positive P example: dilute lattice Bose gas equations	57
3.5	Convenience for parallel computation	58
3.6	Summary of representation requirements	59
4	Stochastic gauges	62
4.1	Generalized kernel stochastic gauges	63
4.2	Weighting stochastic gauges	65
4.3	Drift stochastic gauges	67
4.3.1	Mechanism	67
4.3.2	Real drift gauges, and their conceptual basis	70
4.3.3	Complex gauges	73

4.3.4	Weight spread	74
4.4	Diffusion stochastic gauges	80
4.4.1	Noise matrix freedoms with general and complex analytic kernels	81
4.4.2	Standard form of diffusion gauges for analytic kernels	83
4.4.3	Real standard diffusion gauges and noise mixing	85
4.4.4	Imaginary standard diffusion gauges and statistics	86
4.4.5	Non-standard diffusion gauges and further freedoms	87
4.4.6	Distribution broadening gauges	88
4.4.7	Diffusion from different physical processes	91
4.5	Summary of standard gauges	91
5	The gauge P representation	94
5.1	Properties of the representation	95
5.2	Observables	96
5.3	Dynamics of locally-interacting Bose gas	98
5.4	Comparison to Gross-Pitaevskii semiclassical equations	101
5.5	Extended interparticle interactions	102
5.6	Thermodynamics of interacting Bose gas	107
5.7	Comparison with historical distributions	110
6	Boundary term errors and their removal	113
6.1	Boundary term errors of the first kind	114
6.1.1	Origin and general form	114
6.1.2	Symptoms indicating boundary term errors	116
6.1.3	Moving singularities in a complex phase-space	118
6.1.4	Generic presence in many-mode simulations	119
6.1.5	Boundary terms in the gauge P representation	120
6.2	Boundary term errors of the second kind	122
6.2.1	Mechanism and overview	122
6.2.2	Example: un-normalized positive P representation	123
6.2.3	In the gauge P representation	126
6.2.4	Locally-interacting Bose gas calculations	129

6.2.5	Unresolved issues	130
6.3	Removal of the errors using kernel gauges	131
6.3.1	Conceptual justification	131
6.3.2	Practical implementation using drift gauges	133
6.3.3	Ito vs. Stratonovich moving singularities	135
6.4	Removal example 1: Two-particle absorber	137
6.4.1	The single-mode model	137
6.4.2	Two-boson absorber	139
6.4.3	One- and two-boson absorber	142
6.4.4	Driven two-boson absorber	144
6.4.5	Relevance to interacting Bose gas	145
6.5	Removal example 2: Single-mode laser	145
6.5.1	The laser model	146
6.5.2	Initial conditions	148
6.5.3	Drift gauged equations	149
6.5.4	Correcting for the moving singularity	150
6.5.5	Non-optimal initial conditions	152
6.6	Summary of main boundary term results	153

Part B Development of gauges for interacting Bose gases **155**

7	Gauges for single-mode interacting Bose gas dynamics	156
7.1	Motivation	156
7.2	The single-mode interacting Bose gas	157
7.2.1	Physical model	157
7.2.2	Stochastic gauge P equations	158
7.2.3	Anharmonic oscillator	159
7.2.4	Behaviour of the positive P simulation	161
7.2.5	Coupling to other modes and moving singularities	164
7.3	Drift gauges: Removal of instability	165

7.4	Exponentials of Gaussian random variables	168
7.5	Optimization of diffusion gauges	171
7.5.1	Aims	171
7.5.2	Variables to be optimized	172
7.5.3	Optimization of g''	175
7.5.4	Important special cases	178
7.5.5	Suggested approximate form of diffusion gauge	180
7.5.6	Relationship between target time and mode occupation	180
7.5.7	Boundary term issues	182
7.5.8	Particle gain	183
7.6	Estimates of simulation times	184
7.7	Numerical investigation of improvement	185
7.7.1	Procedure	186
7.7.2	Features seen	191
7.8	Diffusion gauges on their own	194
7.8.1	Motivation	194
7.8.2	Optimization of g'' and comments	195
7.8.3	Numerical investigation of performance	197
7.9	Comparison to recent related work	199
7.9.1	The work [1] of Carusotto, Castin, and Dalibard	199
7.9.2	The work [2] of Plimak, Olsen, and Collett	202
7.9.3	The work [3] of Deuar and Drummond	203
7.10	Gauge recommendation	204
7.11	Summary	205
8	Dynamics of two coupled Bose gas modes	206
8.1	The model	206
8.2	Case 1: Coupling to a vacuum mode	208
8.2.1	Description	208
8.2.2	Procedure	209
8.2.3	Features seen	210

8.2.4	Comparison to Fock state stochastic wavefunctions	217
8.3	Case 2: Coherent mixing of two identical modes	218
8.3.1	Description	218
8.3.2	Features seen	218
8.4	Analysis and Conclusions	222
9	Single-mode interacting Bose gas thermodynamics	227
9.1	The single-mode model	228
9.1.1	Quantum description	228
9.1.2	Gauge P stochastic equations	228
9.1.3	Exact solution	229
9.2	Moving singularities and removal with gauges	230
9.2.1	Moving singularity	230
9.2.2	Minimal drift gauge	231
9.2.3	Radial drift gauge	233
9.3	Numerical simulations	234
9.4	Chemical potential as a free gauge parameter	237
9.5	Summary	240
Part C	Examples of many-mode simulations	241
10	Many-mode dynamics	242
10.1	Simulation procedure	243
10.2	Lattice size and simulation performance	246
10.2.1	Scaling due to interparticle scattering	246
10.2.2	Effect of kinetic interactions	247
10.2.3	Lattice size and diffusion gauge usefulness	250
10.2.4	Drift gauges and many-mode dynamics	251
10.3	Correlation functions	253
10.4	Example 1: Correlation waves in a uniform gas	254
10.4.1	The system	254

10.4.2	Correlation waves in a one-dimensional gas	255
10.4.3	Lattice dependence of simulation time	261
10.4.4	Diffusion gauge dependence of simulation time	261
10.4.5	Two-dimensional gas	263
10.5	Example 2: Extended interactions in a uniform gas	264
10.6	Example 3: Correlations in a trap	266
10.7	Example 4: Bosonic enhancement of scattered atoms	270
10.7.1	The Vogels <i>et al</i> four wave-mixing experiment	270
10.7.2	Simulating Bose-enhanced scattering into initially empty modes	271
10.7.3	Comparison to imaginary scattering length estimate	274
10.8	Summary	276
11	Thermodynamics of a one-dimensional Bose gas	278
11.1	One-dimensional uniform Bose gas at finite temperature	278
11.1.1	Exact solutions	278
11.1.2	Correspondence with trapped gases	279
11.1.3	Uniform gas regimes	280
11.2	Simulation details	282
11.2.1	Overview	282
11.2.2	Drift Gauge	283
11.2.3	Importance sampling	283
11.2.4	Momentum cutoff	287
11.2.5	Chemical potential at intermediate temperatures	288
11.2.6	Scaling of weight variance	289
11.3	Physical regimes simulated	289
11.3.1	Parameter targeting	289
11.3.2	Regimes attained	290
11.4	Observable predictions	291
11.4.1	Spatial correlation functions	291
11.4.2	Momentum distributions	295

11.4.3 Comparison to exact Yang & Yang solution	297
12 Conclusions	299
12.1 Overview of main results	299
12.2 Future directions of improvement	303
Bibliography	314
A Exponentials of Gaussian random variables	315
B Some details of Stochastic Calculus	318
C Quotients of means or random variables	321

List of Figures

2.1	Articles on BECs per year	22
6.1	Unbounded integrands in an un-normalized P representation	127
6.2	Comparison of two-boson absorption simulations	142
6.3	Steady state expectation values of boson number	143
6.4	Simulation with both one- and two-particle absorption	145
6.5	Driven two-boson absorber simulations	146
6.6	Single-mode laser simulation	152
7.1	Sample size growth for exponentials of a Gaussian random variable	170
7.2	Discrepancy between exact g''_{opt} and its approximation	181
7.3	Maximum useful simulation time	188
7.4	Comparison of target time and actual simulation time	190
7.5	Uncertainty in $G^{(1)}(0, t)$ for various gauges	190
7.6	Phase correlations using various gauges	191
7.7	Discrepancy between g''_{opt} and its approximation for $\mathcal{G} = 0$ scheme	196
7.8	Comparison of target time and actual simulation time for $\mathcal{G} = 0$ schemes	198
7.9	Uncertainty in $G^{(1)}(0, t)$ for various gauges in $\mathcal{G} = 0$ schemes	198
7.10	Comparison of drift gauge to that in Carusotto <i>et al</i> [1]	201
8.1	Simulation times for coupling to vacuum mode	211
8.2	Coupling to vacuum mode: $n_0 = 1$	212
8.3	Coupling to vacuum mode: $n_0 = 17$	213
8.4	Coupling to vacuum mode: $n_0 = 200$	214
8.5	Coupling to vacuum mode: $n_0 = 10^4$	215

8.6	Coupling to vacuum mode t_{sim} vs. t_{opt}	216
8.7	Simulation times for mixing of two identical modes	219
8.8	Mixing of two identical modes t_{sim} vs. t_{opt}	220
8.9	Mixing of two identical modes $\omega_{12} = 0.0005$	221
8.10	Mixing of two identical modes $\omega_{12} = 5$	222
8.11	Mixing of two identical modes $\omega_{12} = 500$	223
9.1	Deterministic phase space for single-mode thermodynamics	232
9.2	Efficiency of two gauges	234
9.3	Mean occupation and two-body correlations	235
9.4	Mode occupations at low temperature	235
9.5	Performance of un-gauged and mean field calculations	236
9.6	Dependence of efficiency on starting occupation \bar{n}_0	239
9.7	Dependence of efficiency on form of μ_e	239
10.1	Evolution of second order correlations $g^{(2)}$ in a uniform 1D gas	256
10.2	Second order correlations $g^{(2)}$ at short times in a uniform 1D gas	257
10.3	Evolution of velocity distribution in a uniform 1D gas	257
10.4	Correlation functions in a uniform 1D gas	258
10.5	Correlation functions in a dense uniform 1D gas	261
10.6	Simulation time vs. lattice spacing / occupation in uniform 1D gas	262
10.7	Improvement of uniform 1D gas simulation time with diffusion gauges	262
10.8	Correlation functions in 2D uniform gases	263
10.9	$\bar{g}^{(2)}(x)$ in a uniform 1D gas with extended interparticle interactions	265
10.10	Uniform 1D gas with extended interparticle interactions at $t = 35t_\xi$	265
10.11	Correlation functions in a breathing trapped condensate	268
10.12	Local two-particle correlations in the breathing trapped condensate	269
10.13	Scattered atoms in a condensate collision	272
10.14	Velocity distribution of atoms during condensate collision	275
11.1	Physical regimes attained for the uniform interacting 1D Bose gas	291
11.2	Correlations in a classical gas	292

11.3	Correlations in a decoherent degenerate quantum gas	293
11.4	Correlations in a fermionized gas (1)	294
11.5	Correlations in a fermionized gas (2)	295
11.6	Correlations in a transition gas	296
11.7	Three-particle correlations in a fermionized gas (2)	296
11.8	Momentum distribution in a transition gas	297
11.9	Momentum distribution in the far tails	298
11.10	Comparison of density and energy to exact solution	298
A.1	Finite sample estimates with multiplicative noise	316
C.1	Error estimates with several subensemble sizes	322

List of Tables

1.1	Comparison of quantum dynamics simulation methods	16
1.2	Comparison of quantum thermodynamics calculation methods	17
4.1	Tally of drift gauge freedoms	74
4.2	Tally of standard gauge freedoms	92
4.3	Tally of diffusion gauge freedoms	93
5.1	Properties of several representations for the interacting Bose gas	112
7.1	Maximum useful simulation time	189
7.2	Empirical fitting parameters for simulation time	189

Thesis Rationale and Structure

Rationale

It is a common view that first-principles quantum simulations of mesoscopic dynamics are intractable because of the complexity and astronomical size of the relevant Hilbert space. The following quotes illustrate the significance of the problem: include[4, 5]:

“Can a quantum system be probabilistically simulated by a classical universal computer? . . . the answer is certainly, No!” (Richard P. Feynman, 1982).

“One is forced to either simulate very small systems (i.e. less than five particles) or to make serious approximations” (David M. Ceperley, 1999).

This is certainly true if one wishes to follow all the intricate details of a wavefunction that completely specifies the state of the system. Hilbert space size grows exponentially as more subsystems (e.g. particles or modes) are added, and methods that calculate state vectors or density matrix elements bog down very quickly. Path integral Monte Carlo methods also fail because of the well-known destructive interference between paths that occurs when one attempts dynamics calculations.

Such a situation appears very unfortunate because for many complex physical systems a reliable simulation method is often the only way to obtain accurate quantitative predictions or perhaps even a well-grounded understanding. This is particularly so in situations where several length/time/energy scales or processes are of comparable size/strength, or when non-equilibrium phenomena are important.

The need for reliable quantum dynamics simulations can be expected to become ever more urgent as more mesoscopic systems displaying quantum behavior are accessed experimentally. A pioneering system in this respect are the Bose-Einstein condensates of alkali-atom gases realized in recent years[6, 7, 8, 9].

There is, however, a very promising simulation method using phase-space representations that works around this complexity problem. In brief, a correspondence is made between the full quantum state and a distribution over separable operator kernels, each of which can be specified by a number of variables *linear* in the number of subsystems (e.g. modes). If one then samples the operator kernels according to their distribution, then as the number of samples grows, observable averages of these operators approach the exact quantum values, i.e. “emerge from the noise”. In principle one could reach arbitrary accuracy, but in practice computer power often severely limits the number of samples. Nevertheless, if one concentrates only on bulk properties, and is prepared to sacrifice precision beyond (typically) two to four significant digits, many first-principles quantum mesoscopic dynamics results can be obtained. The mesoscopic region can be reached because simulation time scales only log-linearly¹ ($N_{\text{var}} \log N_{\text{var}}$) with variable number N_{var} , and so still scales log-linearly with system size. Some initial examples of calculations with this method, particularly with the positive P representation[10, 11] based on a separable coherent state basis, are many-mode quantum optics calculations in Kerr dispersive media[12, 13] evaporative cooling of Bose gases with repulsive delta-function interactions[14, 15], and breathing of a trapped one-dimensional Bose gas[1].

In summary, first-principles mesoscopic quantum dynamics is hard, but some progress can (and has) been made in recent years. In broad terms, *the aim of the work reported in this thesis is to advance the phase-space simulation methods some more* . The investigation here is carried out in several directions:

¹This is because discrete Fourier transforms, usually required for kinetic energy evaluation, can be calculated on timescales proportional to $N_{\text{var}} \log N_{\text{var}}$. For some particularly demanding models, the scaling may be log-polynomial in N_{var} due to an increased number of terms in the equation for each variable if there is complicated coupling between all subsystem pairs, triplets, etc. In any case, simulation time never scales exponentially, as it would for brute force methods based on density matrix or state vector elements.

-
1. **Stochastic gauges** Many phase-space distributions have more phase-space degrees of freedom than the minimum required for exact correspondence to quantum mechanics. Each such degree of freedom leads to possible modifications of the stochastic equations of motion by the insertion of arbitrary functions or *gauges* in an appropriate way. While the choice of these gauge functions does not influence the correspondence to quantum mechanics in the limit of infinitely many samples, it can have an enormous effect on the efficiency and/or statistical bias of a simulation with a finite number of samples. (Hence the use of the word “gauge”, in analogy with the way electromagnetic gauges do not change the physical observables but can have an important effect on the ease with which a calculation proceeds). In particular, non-standard choices of these gauge functions can lead to enormous improvements in simulation efficiency, or can be essential to allow any unbiased simulation at all.

Here, a systematic way to include these freedoms is derived and ways of making an advantageous gauge choice are considered in some detail. As a corollary, some results present in the literature[2, 1, 3] are found to be examples of non-standard gauge choices.

The gauge P representation, which is a generalization of the successful positive P representation based on a coherent state basis to allow a variety of useful gauges, is explained. Application of it to interacting Bose gases is developed.

2. **Removal of systematic biases using stochastic gauges.** Two major (and related) stumbling blocks for phase-space distribution methods have been instabilities in the stochastic equations, and pathological trajectory distributions as the boundaries of phase space are approached. These occur for nonlinear systems and can (in fact, apart from special cases, *do*) lead to overwhelming noise or systematic biases in the calculated observables, preventing dependable simulations. To date this has been “problem number one” for these methods. In this thesis it is shown how appropriate stochastic gauges can be used to overcome the instability and boundary term problems in a wide range of models, recovering an unbiased simulation, and opening the way for reliable simula-

tions. Heuristic ways of achieving this in general cases are considered in detail, and examples are given for known cases in the literature.

- 3. Improvement of efficiency using stochastic gauges** The stochastic gauge method also has the potential to significantly improve the efficiency of simulations when appropriate gauges are chosen. In particular, the time for which useful precision is obtained can be extended in many cases by retarding the growth of noise. This is investigated for the case of gauge P simulations of interacting Bose gases and some useful gauges obtained. The regimes in which improvements can be seen with the gauges developed here are characterized. Heuristic guidelines for gauge choice in more general representations and models are also given. Example simulations are made.
- 4. General requirements for usable distributions in open systems** Another aim here is to determine what are the bare necessities for a phase-space distribution approach to be successful, so that other details of the representation used (e.g. choice of basis) can be tailored to the model in question. This can be essential to get any meaningful results, as first-principles mesoscopic simulations are often near the limit of what can be tractably calculated. To this end, general features of the correspondence between quantum mechanics and stochastic equations for the variables specifying the operator kernels are considered in some detail.

Particular concern is given to simulating open systems, as most experimentally realizable systems do not exist in isolation and will have significant thermal and also particle-exchange interactions with external environments. These must usually be taken into account for an accurate description. (The alternative is to make wholesale approximations to the model whose precise effect is often difficult to ascertain).

- 5. Application to thermodynamic calculations** A separate issue are static calculations of thermodynamic equilibrium ensembles. These have traditionally been the domain of quantum Monte Carlo methods, the most versatile

of which have been those based on the path integral approach. While path integral methods are generally not useful for dynamical calculations because destructive interference between the paths occurs extremely rapidly, masking any dynamics, good results can be obtained for static thermodynamic ensemble calculations. If the model is kind, even $\mathcal{O}(10^4)$ particles or modes can be successfully simulated, if one again concentrates on bulk properties and only several significant digits.

Phase-space representations that include dynamically changing trajectory weights can also be used for thermodynamics calculations. The thermal density matrix is evolved with respect to inverse temperature after starting with the known high temperature state. (Such simulations are sometimes said to be “in imaginary time” because of a similarity between the resulting equations and the Schrödinger equation after multiplying time by i). This approach appears competitive with path integral Monte Carlo methods in terms of efficiency, but offers two distinct advantages: Firstly, all observables can in principle be calculated in a single simulation run, which is not the case in path integral methods. These latter require separate algorithms for e.g. observables in position space, momentum space, or observables not diagonal in either. Secondly, a single simulation run gives results for a range of temperatures, while in path integral methods a new simulation is needed for each temperature value.

6. **Demonstration with non-trivial examples** Lastly, but perhaps most importantly, one wishes to demonstrate that the results obtained actually are useful in non-trivial cases. Simulations of dynamics and thermodynamics in a variety of mesoscopic interacting systems are carried out, and their physical implications considered. The emphasis will be foremost on simulations of interacting Bose gases. It is chosen to concentrate on these because they are arguably the systems where quantum effects due to collective motion of atoms are most clearly seen and most commonly investigated by contemporary experiments. This refers, of course, to the celebrated Bose-Einstein condensate experiments on cold, trapped, rarefied, alkali-metal gases. (For a recent

overview see, for example, the collection of articles in Nature *Insight on ultracold matter*[16, 17, 18, 19, 20].) As these gases are an extremely dynamic field of research the need for first-principle simulation methods appears urgent here.

The “holy grail” of quantum simulations, which would be a universally-applicable black-box tractable simulation method², probably does not exist. However, simulations are of such fundamental importance to reliable predictions in mesoscopic physics that any significant progress has the potential to be a catalyst for far-ranging discoveries (and in the past often has).

Structure

This thesis begins with two introductory chapters, which explain in more detail the background issues. That is, Chapter 1 discusses the fundamentals of why many-body quantum simulations are difficult, compares several approaches, and motivates the choice to work on the mode-based phase-space representation simulation methods considered in this thesis. Chapter 2 summarizes the interacting Bose gas model that will be considered in all the simulation examples, and discusses under what circumstances a first-principles rather than a semiclassical calculation is needed to arrive at reliable predictions.

The body of the thesis is then divided into three parts of a rather different nature. Part A investigates what general properties of a phase-space representation are needed for a successful simulation, and explains the gauge P representation, which will be used in later parts. Part B calibrates this method on some toy problems relevant to the interacting Bose gas case, while Part C applies them to non-trivial mesoscopic systems (dynamics and thermodynamics of interacting Bose gases).

Accordingly, Chapter 3 presents a generalized formalism for phase-space representations of mixed quantum states and characterizes the necessary properties for an exact correspondence between quantum mechanics and the stochastic equations.

²In Newtonian dynamics this is just the usual “start with initial conditions and integrate the differential equations”.

The stochastic gauge technique, which forms the basis of the rest of the developments in this thesis, is developed in Chapter 4. In Chapter 5, properties of the gauge P representation are investigated, and its application to interacting Bose gases developed. In Chapter 6 it is explained how stochastic gauges can be used to overcome “technical difficulty number one”: the systematic “boundary term” biases that otherwise prevent or hinder many attempted phase-space distribution simulations. Some relevant technical issues regarding stochastic simulations have been relegated to Appendices.

One- and two-mode toy models are used in Part B to check correctness of the method, and to optimize the gauge functions in preparation for the target aim of simulating multi-mode systems. These models are the single-mode interacting Bose gas dynamics in Chapter 7, the dynamics of two such Rabi-coupled modes in Chapter 8, and grand canonical single-mode thermodynamics in Chapter 9.

Finally, Chapters 10 and 11 give examples of nontrivial many-mode simulations of many-mode interacting Bose gas dynamics and grand canonical thermodynamics, respectively.

Chapter 12 summarizes and concludes the work along with some speculation on fruitful directions of future research.

Chapter 1

Introduction:

Why are quantum many-body simulations difficult and what can be done about it?

Here it is considered how stochastic methods can allow first-principles quantum simulations despite the seemingly intractably large Hilbert space required for such a task. The main approaches: path integral Monte Carlo, quantum trajectories, and phase-space distributions are briefly compared (without going into mathematical details of the methods themselves). This thesis develops and applies the stochastic gauge technique to make improvements to the phase-space distribution methods, and the choice to concentrate on these methods is motivated by their greater versatility. The emphasis throughout this thesis on methods applicable to open systems (leading to mostly mode-based formulations) is also discussed.

1.1 Hilbert space complexity

In classical mechanics, an exhaustive simulation of an interacting N -particle system is (in principle) computable¹ on a classical computer. For the trivial case of a non-interacting three-dimensional (3D) gas in a potential, one would start from initial conditions in $6N$ real variables², and then generically simulate by solving $6N$ differential equations. For particles with ranged binary interactions³ there are $6N$ real differential equations, each with $\mathcal{O}(N)$ terms, so simulation times scale as N^2 — grudgingly computable.

Let us now look at the quantum case. Consider a generic mesoscopic system, divided up (arbitrarily) into N subsystems. Most typically these subsystems would be N particles, N spatial or momentum modes, or N orbitals. If we allow each mode or orbital to be occupied by up to P_{\max} particles ($P_{\max} = 1$ for fermions, and more for bosons), then the full state vector of a pure system in such a model has $(P_{\max} + 1)^N$ complex number components. In the other popular formulation with a set number of particles N , one would discretize \mathcal{D} -dimensional space into M_{latt} lattice points in each dimension, leading to $M_{\text{latt}}^{\mathcal{D}}$ lattice points in all. The state vector of such a system would then have $M_{\text{latt}}^{N\mathcal{D}}$ complex number components. Both of these cases are *exponential* in the system size N , and the amount of memory needed to hold the state (let alone the time needed to evolve the corresponding number of coupled differential equations) becomes intractable very rapidly. This is the Hilbert space complexity problem, and is generic to all mesoscopic quantum calculations.

For example, if one had 8000 Gigabytes of memory just to hold the state vector⁴, one could have 40 fermion modes, or 12 boson modes with each one occupied by at most 9 atoms. In the set particle number formulation, one could get away with 4 particles in a (very sparse!) $10 \times 10 \times 10$ three-dimensional lattice, or 12 particles in such a one-dimensional lattice. This indicates where the estimate of about “five

¹“Computable” here having the usual meaning that simulation time scales linearly, or (grudgingly) polynomially with N . Naturally if N is truly macroscopic such a system still cannot be simulated in practice, but in such large classical systems the primary hindrance to accurate predictions is usually deterministic chaos rather than simulation speed.

²Position and momentum in each degree of freedom.

³As in the majority of first-principles classical models.

⁴Assuming 4 byte floating point data.

particles” quoted in the Thesis Rationale on page 1 came from.

For open systems, one needs even more data space to hold the state, as the density matrix $\hat{\rho}$ will have approximately half the square of the number of components in a state vector — exponential in system size again, but with twice the exponent. The general conclusion is that for mesoscopic (and even many microscopic) systems a brute force approach is doomed to failure, except for perhaps a few special cases that are separable⁵.

1.2 Bulk properties are the key

The exponential growth of Hilbert space mentioned above appears prohibitive, but let us first stand back from the quantum problem for a moment. In classical mechanics there are enormous amounts of variables to keep track of once a system has macroscopic numbers of particles — Avogadro’s number is very big! The reason that statistical mechanics is so successful at dealing with this issue is that for a large system, one is only really interested in *bulk* properties (temperature, pressure, two-particle correlations, ...). The exact position or momentum of particle number 9501292851748 is not of interest to anyone. Even more importantly, knowledge of such a “raw” observable calculated from a model has no physical predictive power, because in such a large system it is not feasible to measure the initial conditions exhaustively. Only a statistical description (based on bulk properties) can be input as initial conditions into a model and because of this, only bulk statistical properties of the model will correspond accurately to the behavior of the physical system under study.

In quantum systems, we have (in a general sense) an analogous problem, but it is much more acute because it turns up when we have $\mathcal{O}(5)$ subsystems, not $\mathcal{O}(10^5)$. This is simply because in quantum mechanics there is much more to keep track of. The full classical description contains the positions and momenta of all the particles (e.g. system variables record things like “the position and momentum of particle 753024”), whereas a full quantum description contains the amplitude and phase of

⁵e.g. an ideal gas

each one of all the possible configurations between the entire set of particles⁶. (e.g. a quantum mechanics variable would be expressing something like “the probability and phase (relative to a general reference) of the particular configuration of particles for which: number 1 is at x_1 , \dots , number 753024 is at x_{753024} , \dots , and number N is at x_N ”). In classical mechanics the entire description is just one of these possible correlations.

Conceptually, what lets one vanquish the enormous numbers of variables is similar in both cases: We’re only interested in *bulk properties*, be they themselves classical or quantum in nature⁷.

1.3 Sampling system configurations

There are two kinds of conceptually distinct kinds of correlations to sample (rather than characterize exhaustively) to estimate the position of the quantum system in Hilbert space.

Firstly, in models that include interactions with an external environment, the system is described fully by a density matrix rather than a pure state vector. Here, the mixture of pure states that combine to the reduced density matrix description of the system must be sampled. While this kind of sampling reduces the number of variables substantially (We go from some huge number to its square root), it is not enough to make a qualitative difference. The pure state vectors still reside in a space whose dimension scales exponentially with system size N .

An example of an approach that carries out only this first kind of sampling is the quantum trajectories method for open systems. (Some reviews of this approach are e.g. the articles of Zoller and Gardiner[21], and Plenio and Knight[22], or Wiseman and Milburn[23] for continuous variable measurements.). This method can be successful in simulations of several particles, especially when microscopic details of their entanglement are important — for example, simulations with up to 24 qubits have been reported[24].

⁶Or, more generally, “subsystems”, particularly when particles are not conserved.

⁷An experimentally accessible example of a macroscopic property of a “quantum” nature is the condensate phase at a given position in a Bose-Einstein Condensate.

What is really needed to attack at least mesoscopic systems, is to change the scaling with number of subsystems from exponential down to at least polynomial. This is achieved by going from the pure system states (which are in general superpositions of various separable system configurations) to a *sample of the separable system configurations*. (examples of such separable system configurations are: the set of occupations of all modes, or the set of positions of all the particles.) The essence of the two sampling steps can be schematically expressed as

$$\begin{aligned} \hat{\rho} &\longrightarrow \sum_{j'=1}^{\mathcal{S}'} |\psi_{j'}\rangle \langle \tilde{\psi}_{j'}|, \\ |\psi_{j'}\rangle &\longrightarrow \sum_{j=1}^{\mathcal{S}} \otimes_{k=1}^N |C_k^{(j)}\rangle_k, \end{aligned} \tag{1.1}$$

where the sums are over \mathcal{S} and \mathcal{S}' samples, and the tensor products over N subsystems, having local subsystem parameters C_k (C stands for separable Configuration). The pure states $|\psi_{j'}\rangle$ and $|\tilde{\psi}_{j'}\rangle$ are not necessarily separable product states, but the samples $\otimes_k |C_k^{(j)}\rangle_k$ are. In the limit of many samples ($\mathcal{S}, \mathcal{S}' \rightarrow \infty$), the correspondence becomes exact.

This second sampling is really the crucial step to allow any kind of first-principles quantum simulation of realistic mesoscopic models⁸, as the number of variables to keep track of now finally scales proportionally to N .

The unavoidable price paid is loss of accuracy. Fortunately most of this price gets transferred onto the detailed non-bulk properties, which one is not interested in anyway. This occurs almost automatically, because there are just so many more of these detailed properties. In many cases bulk properties of quite large systems can be calculated with useful precision, where no simulation at all was possible with brute force methods. The calculation time required for a given uncertainty Δ now scales (by the central limit theorem) roughly proportionally to N/Δ^2 , whereas for a brute force non-stochastic method it would be⁹ proportional to $e^N(\log \Delta)^2$. Clearly, one won't be getting more than two- or three-digit accuracy with the stochastic methods in most cases, but this is often sufficient for a good comparison between theory and experiment.

⁸Obviously an enormous amount of physical prediction can be made with mean field, or other semiclassical methods, but by first-principles models I refer to those where no semiclassical approximations are made.

⁹The time required for multiplication operations on floating point numbers with $d_{10} \approx -\log_{10} \Delta$ digits scales as d_{10}^2

1.4 Path integral Monte Carlo and static thermal calculations

The most widely used approach for many-mode/body calculations is path integral Monte Carlo. The basis for these methods were Feynman's imaginary time path integrals[25]. This method, however, is only successful for static calculations of thermal equilibrium states (reasons explained below).

In the static thermal case, the un-normalized density matrix at temperature T is $\hat{\rho}_u(T) = \exp\left[-\hat{H}/k_B T\right]$, and path integrals make use of the simple (and exact) property that $\hat{\rho}_u(T) = \hat{\rho}_u(\mathcal{M}T)^{\mathcal{M}}$. If \mathcal{M} is sufficiently large that $\mathcal{M}k_B T \gg \langle \hat{H} \rangle$, then one can use a high-temperature approximation¹⁰ to evaluate $\hat{\rho}_u(\mathcal{M}T)$. The density matrix $\hat{\rho}_u(T)$ can be written as a convolution of the high temperature density matrix elements (in complete bases of separable system configurations C labeled by $C_{(0)}, \dots, C_{(\mathcal{M})}$):

$$\hat{\rho}_u(T) = \int \dots \int dC_{(0)} dC_{(1)} \dots dC_{(\mathcal{M})} \quad (1.2)$$

$$\langle C_{(0)} | \hat{\rho}_u(\mathcal{M}T) | C_{(1)} \rangle \langle C_{(1)} | \hat{\rho}_u(\mathcal{M}T) | C_{(2)} \rangle \dots \langle C_{(\mathcal{M}-1)} | \hat{\rho}_u(\mathcal{M}T) | C_{(\mathcal{M})} \rangle.$$

To obtain properties diagonal in the configuration basis chosen, one takes samples while setting $C_{(0)} = C_{(\mathcal{M})}$, and ends up sampling a ‘‘ring polymer’’ of configurations, the number of variables in a full sample scaling as $\mathcal{M}N$. The distribution of the samples is given by the product of the expectation values:

$$P(C_{(0)}, C_{(1)}, \dots, C_{(\mathcal{M})}) = \otimes_{j=1}^{\mathcal{M}} \langle C_{(j-1)} | \hat{\rho}_u(\mathcal{M}T) | C_{(k)} \rangle, \quad (1.3)$$

which is real positive, or largely so. Off-diagonal properties require separate calculations with slightly different sample properties. An extensive discussion of the details of the path integral approach can be found e.g. in Ceperley[27, 28] and references therein.

While, formally, this approach might appear to be also applicable for dynamical evolution via $\hat{\rho}(t) = \hat{\rho}(0) \exp\left[-i\hat{H}t/\hbar\right]$, in practice this does not work. The reason

¹⁰Typically one uses the Feynman-Kacs or Trotter formula[26] $\exp\left[-(\hat{H}_V + \hat{H}_K)/k_B T\right] = \lim_{\mathcal{M} \rightarrow \infty} \left(\exp\left[-\hat{H}_V/\mathcal{M}k_B T\right] \exp\left[-\hat{H}_K/\mathcal{M}k_B T\right]\right)^{\mathcal{M}}$ to evaluate the effect of the potential \hat{H}_V and kinetic \hat{H}_K terms of the Hamiltonian separately.

is that now $P(C_{(j)})$ is primarily a product of complex phase factors, rather than real positive values. One obtains a very wide range of configurations, all with similar weighting, but mutually canceling phase factors. In the limit of infinite trajectories, almost all of these configurations would destructively interfere leaving just the observed dynamics, but for finite numbers of samples in a complex system there is no automatic way to tell which configurations are the important ones. The end result is that no sensible observable averages emerge from the noise even for very short times.

1.5 phase-space evolution methods

The density matrix of a system can often be written in a form

$$\hat{\rho} = \int P(C) \hat{\Lambda}(C) dC, \quad (1.4)$$

where C is a separable system configuration¹¹, containing some number of variables linear in N (the number of subsystems). $P(C)$ is a positive function, and $\hat{\Lambda}(C)$, the *kernel*, is a projector or similar tensor product operator parameterized by C .

In a generic case where pure quantum states of each subsystem (labeled k) can be fully described by the set of parameters C_k , a viable kernel would be $\hat{\Lambda}(C) = e^{i\theta} \otimes_k |C_k\rangle_k \langle \tilde{C}_k|_k$. Here $|C_k\rangle_k$ is the pure state of subsystem k given by parameters C_k , and θ is a phase. The full configuration then would be $C = \{\theta, C_1, \dots, C_N, \tilde{C}_1, \dots, \tilde{C}_N\}$.

The basic idea is that $P(C)$ can be interpreted as a probability distribution of the configuration C . Then, the evolution equation for $\hat{\rho}$ (with time or with temperature) produces a corresponding evolution equation for P , and finally evolution equations for the configuration samples C . Depending on the system and the kernel $\hat{\Lambda}$ chosen, this procedure can often be carried out, and evolution equations for the configuration samples of the system derived. In particular, *dynamical evolution can be simulated* while the number of variables remains linear or polynomial in N .

In fact phase-space distribution methods are much more versatile than path integral approaches on a number of counts:

¹¹The variables of C can be discrete or continuous. In the former case the integral over C is replaced by a sum.

1. Dynamics can be simulated.
2. Results are given for a range of times (or temperatures) by one simulation
3. All observables can be calculated in one simulation

In both phase-space distribution and path integral approaches, the amount of computer effort scales linearly with system size N . More details of how this is achieved are given in Chapter 3. See also a tabular comparison of the various quantum many-body/mode simulation approaches in Tables 1.1 and 1.2.

Unfortunately, a major stumbling block for phase-space distribution methods is that the equations for the configuration variables in C often have stability problems, which can lead to large noise and/or statistical bias. This is one of the main reasons that these methods have not been as widely investigated nor used as path integral Monte Carlo, despite their other advantages.

This thesis concerns itself with methods of overcoming this stability problem using the stochastic gauge technique, and some demonstration of subsequent applications. It is also intended to show that when stability issues are taken care of, these phase-space distribution methods can be competitive with path integrals for static thermal ensemble calculations as well.

1.6 Open systems and the mode formulation

Actual physical systems do not exist in isolation, and there is always some degree of coupling to the external environment. This coupling becomes ever more pronounced as one moves from microscopic to mesoscopic systems, and is considered responsible for the reduction of quantum phenomena to classical behavior as the macroscopic limit is reached.

Traditionally in most many-body studies the system has been taken to be composed of a set number of N particles. A different formalism is required if one is to investigate open systems that can exchange particles with a reservoir having chemical potential μ , or for dynamical calculations if particles are lost or injected into the system. A natural choice is to consider lattice models where rather than dividing

Table 1.1: A comparison of different methods to exactly simulate quantum dynamics from first principles. As applicable to many-body/mode simulations with N subsystems (e.g. particles, modes).

Variety:	EXPLICIT DENSITY MATRIX	QUANTUM TRAJECTORIES	PATH INTEGRAL	PHASE-SPACE DISTRIBUTIONS
Stochastic	No	Yes	Yes	Yes
Sample size	$\propto e^{2N}$	$\propto e^N$	$\propto \mathcal{M}N$	$\propto N$
Phase factor catastrophe ^a	No	No	Yes	No ^b
Unstable equations or sampling bias	No	No	No	Often ^c
Observables from one calculation	All	All	Limited subset	All
Simulation scope	Range of t	Range of t	Single t	Range of t

^aWhere samples of similar weight but different phase factors interfere, masking any physically meaningful results.

^bAt long times some gauged schemes may develop interfering phase factors. This occurs only for some gauges in some systems.

^cHowever, in a wide range of cases the instability can be removed using the stochastic gauge technique, as shown in this thesis.

the system into N subsystems consisting of a particle, we divide the system up into N spatial (or momentum) lattice points that can be occupied by a variable number of particles. The kernel $\hat{\Lambda}$ then becomes a separable tensor product of lattice-point states, described by some local parameters. Their total number for a system configuration sample is again proportional to N . Because we want the methods developed in this thesis to be applicable to systems that allow particle exchange with an environment, this is the kind of model that we will mostly concentrate on here. The general methods developed are, however, by no means restricted to mode-based models.

A feature to be aware of in such a model is that in general the state will be in a superposition or mixture of pure states with different numbers of total particles. This is not expected to hinder comparison with experiment, however. A subsystem

Table 1.2: A comparison of different methods to exactly calculate static quantum properties of equilibrium thermodynamic ensembles from first principles. As applicable to many-body/mode simulations with N subsystems (e.g. particles, modes).

Variety:	EXPLICIT DENSITY MATRIX	QUANTUM TRAJECTORIES	PATH INTEGRAL	PHASE-SPACE DISTRIBUTIONS
Stochastic	No	Yes	Yes	Yes
Sample size	$\propto e^{2N}$	$\propto e^N$	$\propto MN$	$\propto N$
Phase factor catastrophe ^b	No	No	For fermions ^a $k_B T \lesssim \Delta E_{FB}$	Can occur for low T
Unstable equations or sampling bias	No	No	No	Often ^c
Observables from one calculation	All	All	Limited subset	All
Simulation scope	Range of T	Range of T	Single T	Range of T

^aFor fermion calculations the “fermion sign problem” appears after $k_B T \lesssim \Delta E_{FB}$, where ΔE_{FB} is the free energy difference between the fermion system, and an analogous system of bosons. For a review see Schmidt and Kalos[29]

^bWhere samples of similar weight but different phase factors interfere, masking any physically meaningful results.

^cHowever, in a wide range of cases the instability can be removed using the stochastic gauge technique, as shown in this thesis.

such as e.g. a trapped gas is perfectly entitled to be in a superposition of number states, and e.g. evaporative cooling of Bose atoms will produce BECs of different sizes in each individual experimental run.

For open systems with particle transfer to the environment, the boundary where the system ends and the environment starts will be at best hazy, and particles will take some time to travel from one to the other. There will always be a gray area, and in it some particles, where the labeling of particles as belonging to the system or environment is dubious. Superpositions and/or mixtures of different particle number states are actually trying to reflect this labeling difficulty which is encountered in experiments.

Also, in many situations the effect of such superpositions will be small: A rea-

sonably accurate description of a realistic mesoscopic system in most cases requires a fairly large number of lattice points to resolve spatial and momentum variations in the system. In such a case, for statistical reasons, the variation in particle number of the whole system will be small in comparison with the total number of particles.

Lastly, it should also be pointed out that if desired, one can always simply turn off the external couplings, and obtain results for closed number-conserving Hamiltonian models, which are just special cases with external interaction strength zero. (Although a simulation method that is “hard-wired” to a conservative system is likely to be somewhat more efficient than an open systems method.)

Chapter 2

System of choice: Open interacting Bose gases

Simulations presented in this thesis mostly¹ concern themselves with simulations of interacting Bose gases in a lattice model with two-particle interactions². Here those basic details of the model that are repeatedly referred to in the body of the thesis are introduced. Also, this choice of mesoscopic system from the wide variety whose simulations could have been attempted is motivated.

This chapter does not attempt to present a review of the state of knowledge on interacting Bose gases, as this is currently very extensive and doing so would unnecessarily lengthen the thesis without introducing any additional novelty. Instead, I recommend to the reader the excellent reviews on both theoretical and experimental aspects of Bose Einstein condensation in cold alkali gases by Leggett[30], Dalfovo *et al*[31], Parkins and Walls[32]. Topical also are the 1998 Nobel lectures of Chu[33], Cohen-Tannoudji[34], and Phillips[35] on laser cooling and magneto-optical trapping, as well as the reviews of evaporative cooling by Ketterle and van Druten[36] and Walraven[37]. An overview of recent developments can be found in e.g. the collection of articles in Nature *Insight on ultracold matter*[16, 17, 18, 19, 20]. Physical properties of interacting Bose gases pertinent to the analysis or interpretation of the

¹Although not exclusively, see for example Chapter 6.

²In the most commonly used cold alkali-metal gas models these interactions are local as in a Bose-Hubbard model, and are then often referred to as “delta-function” interactions.

simulations in Parts B and C of this thesis are given in the relevant chapters.

2.1 Motivation

The choice to concentrate on cold dilute interacting Bose gas models when trying out the simulation methods developed in this thesis is motivated by the following considerations:

- The prime motivation is that mesoscopic systems that are very well described by these models are experimentally accessible. These are the cold dilute alkali-metal gases at temperatures below or near the Bose-Einstein condensation (BEC) temperature. In the last several years production of cold Bose gases of alkali-metal atoms has become almost routine, occurring in many tens of labs around the world. The largest interest (see Figure 2.1) has been in gases below the condensation temperature where quantum effects come to dominate the system even though it is mesoscopic or even approaching macroscopic in size. These were first realized in 1995 by Anderson *et al*[6] (^{87}Rb), Davis *et al*[7] (^{23}Na), and Bradley *et al*[8] (^7Li), by laser cooling the gas in a trap, and subsequently lowering the temperature even further by evaporative cooling. Three years later the first signatures of BEC were seen by Fried *et al*[9] in the somewhat different case of hydrogen. Presently, many variants of these systems are being investigated experimentally. For example, quasi-one-dimensional gases have been achieved[38, 39, 40], and coherent four-wave mixing between atom clouds has been observed[41], just to name a few of the experimental directions relevant to the models considered later in this thesis.
- A related argument is that there are phenomena in the experimental systems that are not precisely understood and appear quite resistant to conclusive theoretical analysis by approximate methods. This invites a first-principles investigation.

Most of these occur in situations where there is a transition or interplay between disordered high temperature behavior well above condensation temper-

ature T_c , and the pure many-body quantum states approached as $T \rightarrow 0$. The first can be treated with kinetic theories, while the latter is very well described by the Gross-Pitaevskii (GP)[42, 43, 44] mean-field equations, or Bogoliubov theory[45]³. As usual in physical systems, the intermediate regime is the most complicated and is the place where first-principles simulations are the most needed.

Examples of such issues where first-principles simulations can make a step forward to full understanding are, for example:

1. The behavior of the system during initial condensation, when the $T \approx T_c$ barrier is crossed. This includes issues such as “does the always condensate form in the ground state, or is (at least metastable) condensation into excited states possible?” There are some indications that condensates may occur e.g. in states of non-zero centre-of-mass motion[15].
2. Is the exact ground state of a BEC the same as that obtained by semi-classical GP methods?
3. Dynamics of condensates during atom interferometry.
4. The decoupling of atoms from the trapped condensate in an atom laser arrangement. How is the emitting condensate disturbed, and in what state are the pulses emitted?
5. Scattering of atoms into empty momentum modes when condensates collide[41].
6. Heating and other processes that eventually destroy the condensate in experiments.

It is worth noting that most of the above processes usually involve coupling of the condensate to an external environment. This is further indication that a simulation method capable of handling open systems is desirable.

- Not only is the field of Bose-Einstein condensation experimentally accessible, but the sheer amount of research on the topic has been growing (and continues

³At finite temperatures, but significantly below T_c

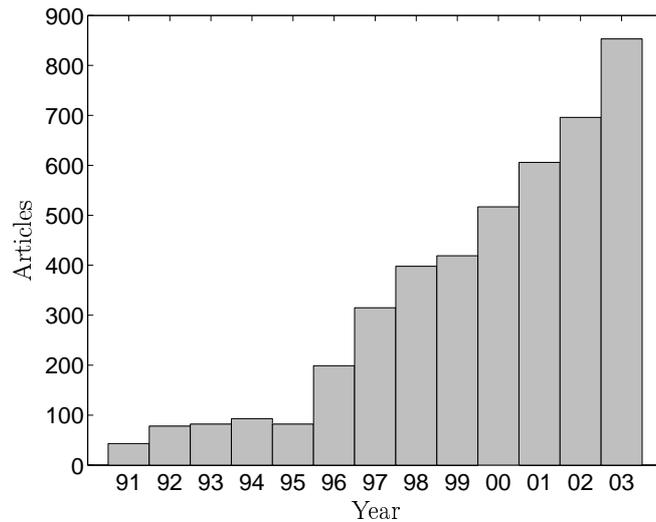


Figure 2.1: **Papers published in refereed journals** (by year) on the topic of Bose-Einstein condensation. Based on a search for the keywords “Bose Einstein condens*” or “Bose condens*” through titles and abstracts in the ISI “Web of Science” citation database (<http://wos.isiglobalnet.com/>) on 11 Feb 2004. Note the rapid increase after the first successful BEC experiments in 1995.

to grow) at a great rate (see Figure 2.1). This would indicate that some of the predictions made may be verified rapidly.

- Lastly, but importantly, many-mode simulations of similar systems have been successful with a phase-space method (the positive P distribution[10, 11]). Some examples of this include simulations of evaporative cooling of a Bose gas to the beginning of condensation[14, 15, 46], as well as simulations of optical soliton propagation in nonlinear Kerr media[12, 13], which have a similar form of nonlinearity to (2.17) — quadratic in local particle density. The gauge distributions developed in this thesis include the positive P distribution as a special case, and improve on it. This previous success is an indication that the newer methods are likely to lead to significant results with these systems.

2.2 Field model with binary interactions

Consider a \mathcal{D} -dimensional boson field $\widehat{\Psi}(\mathbf{x})$ undergoing two-particle interactions. After second quantization the Hamiltonian of the system can be written as

$$\begin{aligned} \widehat{H} = & \int d^{\mathcal{D}}\mathbf{x} \left\{ \frac{\hbar^2}{2m} \frac{\partial \widehat{\Psi}^\dagger(\mathbf{x})}{\partial \mathbf{x}} \frac{\partial \widehat{\Psi}(\mathbf{x})}{\partial \mathbf{x}} + V^{\text{ext}}(\mathbf{x}) \widehat{\Psi}^\dagger(\mathbf{x}) \widehat{\Psi}(\mathbf{x}) \right\} \\ & + \frac{1}{2} \int d^{\mathcal{D}}\mathbf{x} d^{\mathcal{D}}\mathbf{y} U(\mathbf{x} - \mathbf{y}) \widehat{\Psi}^\dagger(\mathbf{x}) \widehat{\Psi}^\dagger(\mathbf{y}) \widehat{\Psi}(\mathbf{x}) \widehat{\Psi}(\mathbf{y}). \end{aligned} \quad (2.1)$$

The bosons have mass m , move in \mathcal{D} -dimensional space whose coordinates are \mathbf{x} , experience an external conservative potential $V_{\text{ext}}(\mathbf{x})$, and an interparticle potential $U(\mathbf{r}) = U(-\mathbf{r})$ where \mathbf{r} is the spacing between two particles. The field operators $\widehat{\Psi}^\dagger(\mathbf{x})$ and $\widehat{\Psi}(\mathbf{x})$ are creation and destruction operators on bosons at \mathbf{x} , and obey the usual commutation relations

$$\left[\widehat{\Psi}(\mathbf{x}), \widehat{\Psi}^\dagger(\mathbf{y}) \right] = \delta^{\mathcal{D}}(\mathbf{x} - \mathbf{y}). \quad (2.2)$$

The total number of bosons is

$$\widehat{N} = \int d^{\mathcal{D}}\mathbf{x} \widehat{\Psi}^\dagger(\mathbf{x}) \widehat{\Psi}(\mathbf{x}). \quad (2.3)$$

In a typical Bose-Einstein condensation experiment with cold trapped alkali-metal gases the trapping potential V_{ext} is parabolic, in general non-spherical.

Three-body⁴ processes are not included in the model. While these can be neglected for most contemporary experiments with cold rarefied alkali-metal gases or BECs, they can become significant when density is high enough, or when one operates near a Feshbach resonance, requiring modifications to the model. (For example, three-body recombination into molecules[47, 48] can play an important role in loss of atoms from a BEC).

2.3 Reduction to a lattice model

To allow computer simulations, (2.1) must be reduced to a lattice form. Provided the lattice is sufficiently fine to resolve all features, no change in physical interpretation occurs.

⁴Strictly speaking, all $N > 2$ body.

One issue to keep in mind, however, is that some care must be taken when going from a continuum to a lattice model to make sure that the behavior of the latter at length scales longer than the lattice spacing is the same — i.e. that the lattice model has been renormalized.

One way to do this is to impose:

1. A spatial box size in each dimension L_d ,
2. A momentum cutoff in Fourier space k_d^{\max} ,

where spatial dimensions have been labeled by $d = 1, \dots, \mathcal{D}$. In this case one proceeds by expressing all spatially-varying functions (generically called $f(\mathbf{x})$) in Fourier modes as

$$f(\mathbf{x}) = \left(\prod_d \frac{\sqrt{2\pi}}{L_d} \right) \sum_{\tilde{\mathbf{n}}} \tilde{f}_{\tilde{\mathbf{n}}} e^{i\mathbf{k}_{\tilde{\mathbf{n}}}\cdot\mathbf{x}}, \quad (2.4)$$

where the index $\tilde{\mathbf{n}} = \{\tilde{n}_1, \dots, \tilde{n}_{\mathcal{D}}\}$ has an integer component $\tilde{n}_d = 0, \pm 1, \pm 2, \dots$ for each dimension. The wave vector $\mathbf{k}_{\tilde{\mathbf{n}}} = \{k_1(\tilde{n}_1), \dots, k_{\mathcal{D}}(\tilde{n}_{\mathcal{D}})\}$ also contains \mathcal{D} components

$$k_d(\tilde{n}_d) = 2\pi\tilde{n}_d/L_d. \quad (2.5)$$

This has imposed the spatial box size, and now we also impose a momentum cutoff k_d^{\max} by truncating all terms in (2.4) for which any $|k_d(\tilde{n}_d)| > k_d^{\max}$. That is,

$$f(\mathbf{x}) \rightarrow f_{\text{cut}}(\mathbf{x}) = \left(\prod_d \frac{\sqrt{2\pi}}{L_d} \right) \sum_{|\tilde{n}_1| \leq L_1 k_1^{\max}/2\pi} \dots \sum_{|\tilde{n}_{\mathcal{D}}| \leq L_{\mathcal{D}} k_{\mathcal{D}}^{\max}/2\pi} \tilde{f}_{\tilde{\mathbf{n}}} e^{i\mathbf{k}_{\tilde{\mathbf{n}}}\cdot\mathbf{x}}. \quad (2.6)$$

This leaves an M_d point lattice in each dimension, with $M_d = 1 + 2 \text{int}[L_d k_d^{\max}/2\pi]$, when $\text{int}[\cdot]$ gives the integer value (rounded down).

The lattice points in Fourier space k_d have been defined by (2.5), the function values there are $\tilde{f}_{\tilde{\mathbf{n}}}$, while the function values in normal space $f_{\mathbf{n}}$ are given by the discrete inverse Fourier transform of the $\tilde{f}_{\tilde{\mathbf{n}}}$:

$$f_{\mathbf{n}} = f_{\text{cut}}(\mathbf{x}_{\mathbf{n}}). \quad (2.7)$$

Here the spatial lattice is $\mathbf{x}_{\mathbf{n}} = \{x_1(n_1), \dots, x_{\mathcal{D}}(n_{\mathcal{D}})\}$ given by

$$x_d(n_d) = n_d \Delta x_d, \quad (2.8)$$

with spacing $\Delta x_d = L_d/M_d$. The index \mathbf{n} is composed of non-negative integers $\mathbf{n} = \{n_1, \dots, n_{\mathcal{D}}\} : n_d = 0, 1, \dots, M_d - 1$.

The lattice model is physically equivalent to the continuum field model provided that

1. No features occur on length scales $\approx L_d$ or greater.
2. The truncated Fourier components ($\tilde{f}_{\mathbf{n}}$ corresponding to any $k_d > k_d^{\max}$) are negligible so that for all practical purposes $f(\mathbf{x}) = f_{\text{cut}}(\mathbf{x})$. In other words, no features occur on length scales $\approx \Delta x_d$ or smaller.

Conversely, in cases when some processes occur entirely on length scales smaller than the lattice spacings Δx_d , the lattice model may be renormalized in non-trivial ways (so that $f_{\mathbf{n}} \neq f(\mathbf{x}_{\mathbf{n}})$) to still be physically equivalent to the continuum model on lattice spacing length scales. An example is outlined in Section 2.4.

Proceeding in this manner for the case of extended interparticle interactions (2.1), and using the lattice notation of (2.4), (2.6), and (2.7), let us define the lattice annihilation operators

$$\hat{a}_{\mathbf{n}} = \hat{\Psi}_{\mathbf{n}} \sqrt{\prod_d \Delta x_d}, \quad (2.9)$$

which obey the boson commutation relations

$$[\hat{a}_{\mathbf{n}}, \hat{a}_{\mathbf{m}}^\dagger] = \delta_{\mathbf{nm}}. \quad (2.10)$$

The particle number operator at lattice point \mathbf{n} is

$$\hat{n}_{\mathbf{n}} = \hat{a}_{\mathbf{n}}^\dagger \hat{a}_{\mathbf{n}}. \quad (2.11)$$

With these, one obtains the expression:

$$\hat{H} \rightarrow \sum_{\mathbf{nm}} \left[\hbar \omega_{\mathbf{nm}} \hat{a}_{\mathbf{n}}^\dagger \hat{a}_{\mathbf{m}} + \frac{1}{2} u_{\mathbf{nm}} \hat{n}_{\mathbf{n}} (\hat{n}_{\mathbf{m}} - \delta_{\mathbf{nm}}) \right]. \quad (2.12)$$

In this normally ordered expression, the frequencies $\omega_{\mathbf{nm}} = \omega_{\mathbf{mn}}^*$ come from the kinetic energy and external potential. They produce a local particle number dependent energy, and linear coupling to other sites, the latter arising only from kinetic

processes. The frequencies $\omega_{\mathbf{nm}}$ are explicitly

$$\omega_{\mathbf{nm}} = \omega_{\mathbf{mn}}^* = \frac{\delta_{\mathbf{nm}}}{\hbar} V_{\mathbf{n}}^{\text{ext}} + \frac{\hbar}{2m} \left(\prod_d \frac{\Delta x_d}{\sqrt{2\pi}} \right) K_{\mathbf{n}-\mathbf{m}}^{(2)}, \quad (2.13)$$

where the function $K^{(2)}$ is the \mathcal{D} -dimensional discrete inverse Fourier transform of $|\mathbf{k}|^2 = \sum_d k_d^2$. That is,

$$K_{\mathbf{n}}^{(2)} = \left(\prod_d \frac{\sqrt{2\pi}}{L_d} \right) \sum_{\tilde{\mathbf{n}}} |\mathbf{k}_{\tilde{\mathbf{n}}}|^2 e^{i\mathbf{k}_{\tilde{\mathbf{n}}}\cdot\mathbf{x}_{\mathbf{n}}}. \quad (2.14)$$

The interparticle potential can also be discretized on the lattice, however there are a few subtleties caused by the finite box lengths L_d . Let us define the lattice mutual interaction strength $U_{\mathbf{n}} = U_{\text{cut}}(\mathbf{x}_{\mathbf{n}})$ in the notation of (2.7) and (2.6). At first glance, the lattice scattering terms might be $u_{\mathbf{nm}} = U_{\mathbf{n}-\mathbf{m}}$ since $x_d(n_d) \propto n_d$, but there are two issues that complicate matters:

1. One has only non-negative lattice labels \mathbf{n} and coordinates $x_d(n_d)$, as seen from (2.8).
2. The Fourier decomposition on a box of side lengths L_d implicitly assumes periodic boundary conditions, so e.g. particles at lattice points $\mathbf{n} = 0$ and $\mathbf{m} = \{M_1 - 1, \dots, M_{\mathcal{D}} - 1\}$ are effectively in very close proximity.

In the continuum field model the interparticle potential is symmetric ($U(\mathbf{r}) = U(-\mathbf{r})$), so to impose periodicity on length scales L_d , the lattice interparticle potential should obey $U_{\mathbf{n}} = U_{\{M_1-n_1, n_2, \dots, n_{\mathcal{D}}\}} = \dots = U_{\{n_1, n_2, \dots, M_{\mathcal{D}}-n_{\mathcal{D}}\}}$. This symmetry then lets us write the scattering terms as

$$u_{\mathbf{nm}} = u_{\mathbf{mn}} = U_{|\mathbf{n}-\mathbf{m}|}, \quad (2.15)$$

where the notation means $|\mathbf{n} - \mathbf{m}| = \{|n_1 - m_1|, \dots, |n_{\mathcal{D}} - m_{\mathcal{D}}|\}$. For the lattice model to be equivalent to an open continuum field model also requires $U_{\mathbf{n}}$ to be negligible when any $n_d \approx \mathcal{O}(M_d/2)$.

2.4 Locally interacting lattice model

The majority of mesoscopic experiments with interacting Bose gases where quantum mechanical features have been seen use alkali-metal atoms (Li, Na, K, Rb, Cs,

and also H) in or near the BEC regime (see also Section 2.1), where a significant simplification of the lattice Hamiltonian (2.12) can be made.

Broadly speaking, what is required is that only binary collisions be relevant, and that these occur on length scales much smaller than all other relevant length scales of the system, *including the lattice spacings* Δx_d . The simple-minded continuum to lattice procedure of the previous Section 2.3 no longer suffices, because the interparticle interactions occur at momenta well above k_d^{\max} . Detailed consideration to this renormalization is given in Section IV of the aforementioned review by Leggett[30], and in even more detail by Dalibard[49] and Wiener *et al*[50]. The reasoning for single-species cold alkali-metal Bose gases is (in brief) as follows:

- The leading term in inter-atomic potentials at long range $\gtrsim 5\text{\AA}$ for the alkali-metal atoms is the van der Waals interaction $\propto 1/|\mathbf{r}|^6$.
- This potential gives a van der Waals length r_{vdw} and energy $E_{\text{vdw}} \sim \hbar^2/mr_{\text{vdw}}^2$ characteristic of the atom species, which are the typical extent and binding energy of the largest bound state. For the alkali-metals E_{vdw} is of order 0.1 – 1mK, except for hydrogen, which has $E_{\text{vdw}} \approx 3\text{K}$.
- The binding energy of these states is much higher than thermal energies at or near the BEC condensation temperature for these atoms (typically, $T_c \approx \mathcal{O}(100\text{nK})$). For atom pairs with relative orbital angular momentum $l_{\text{rel}} \neq 0$, the scattering strength is smaller than s -wave scattering ($l_{\text{rel}} = 0$) by a factor of $(k_B T/E_{\text{vdw}})^{l_{\text{rel}}}$ [30]. Conclusion: for $k_B T \ll E_{\text{vdw}}$ only s -wave scattering interactions need be considered.
- For van der Waals potentials, s -wave scattering at low energies can be characterized by a single parameter, the s -wave scattering length a_s (see e.g. Landau and Lifshitz[51], Sec. 108). This is typically of similar size to r_{vdw} . (e.g. for ^{87}Rb , $a_s = 5.77\text{nm}$ [52]), and can be considered to be roughly the radius scale of an “equivalent” hard sphere (For indistinguishable bosons, the scattering cross-section is $8\pi a_s^2$).
- If the s -wave scattering length is *much smaller than all other relevant length*

scales of the system, then the exact inter-atom interaction can be replaced by a local lattice interaction where $u_{\mathbf{nm}} \propto \delta_{\mathbf{nm}}$ in the formalism of (2.12). Relevant length scales include the de Broglie wavelength of the fastest atoms, the interparticle spacing, the trap size, the transverse thickness (in the case of 2D or 1D systems), and *the lattice spacing* Δx_d .

Following this broad line of argument, it can be shown [49, 50, 30] that for two indistinguishable bosons in 3D, the scale-independent coupling constant g is given by

$$g = \frac{4\pi\hbar^2 a_s}{m}. \quad (2.16)$$

This then leads to a lattice Hamiltonian of the form

$$\hat{H} = \sum_{\mathbf{nm}} \hbar\omega_{\mathbf{nm}} \hat{a}_{\mathbf{n}}^\dagger \hat{a}_{\mathbf{m}} + \hbar\chi \sum_{\mathbf{n}} \hat{n}_{\mathbf{n}}(\hat{n}_{\mathbf{n}} - 1), \quad (2.17)$$

with the (scale-dependent) lattice self-interaction strength being

$$\chi = \frac{g}{2\hbar \prod_d \Delta x_d}. \quad (2.18)$$

In the case of an effectively one- or two-dimensional gas of indistinguishable bosons, the coupling constant is instead given by

$$g = \frac{4\pi\hbar^2 a_s}{\lambda_0 m} \quad (2.19)$$

where λ_0 is, respectively, the effective thickness or cross-section in the transverse (collapsed) dimensions.

2.5 Open dynamic equations of motion

For a wide range of physical situations the interaction of an open system with its environment can be considered Markovian, and in this situation can be modeled with a master equation, where the evolution of the density matrix is dependent only on its present value, with no time lag effects. Broadly speaking, this means that any information about the system received by the environment via interactions is

dissipated much faster than the system dynamics we perceive. Hence, no feedback into the system occurs.

Under these conditions the equation of motion for the density matrix of the system can be written (in Linblad form) as:

$$\frac{\partial \hat{\rho}}{\partial t} = \frac{1}{i\hbar} [\hat{H}, \hat{\rho}] - \frac{1}{2} \sum_j \left[\hat{L}_j^\dagger \hat{L}_j \hat{\rho} + \hat{\rho} \hat{L}_j^\dagger \hat{L}_j \right] + \sum_j \hat{L}_j \hat{\rho} \hat{L}_j^\dagger. \quad (2.20)$$

The Linblad operators \hat{L}_j represent various coupling processes to the environment, and their number depends on the details of the system-environment interactions. Note that $\hat{\rho}$ is the reduced interaction picture density matrix of the (Bose gas + environment) system, traced over the environment. Some examples of common processes are:

Single-particle losses to a standard⁵ zero temperature heat bath with rates $\gamma_{\mathbf{n}}$ at $\mathbf{x}_{\mathbf{n}}$:

$$\hat{L}_{\mathbf{n}} = \hat{a}_{\mathbf{n}} \sqrt{\gamma_{\mathbf{n}}}. \quad (2.21)$$

Exchange with a finite temperature heat bath with coupling strength proportional to $\gamma_{\mathbf{n}}$ at $\mathbf{x}_{\mathbf{n}}$:

$$\begin{aligned} \hat{L}_{\mathbf{n}} &= \hat{a}_{\mathbf{n}} \sqrt{\gamma_{\mathbf{n}}(1 + \bar{n}_{\text{bath}})}, \\ \hat{L}'_{\mathbf{n}} &= \hat{a}_{\mathbf{n}}^\dagger \sqrt{\gamma_{\mathbf{n}} \bar{n}_{\text{bath}}}, \end{aligned} \quad (2.22)$$

Where the sum in (2.20) is taken over both terms in $\hat{L}_{\mathbf{n}}$ and $\hat{L}'_{\mathbf{n}}$. For the standard heat bath with energy $\hbar\omega_{\text{bath}}$ per particle for all modes, \bar{n}_{bath} , the mean number of particles per mode, is given by the Bose-Einstein distribution expression

$$\bar{n}_{\text{bath}} = 1/[\exp(\hbar\omega_{\text{bath}}/k_B T) - 1], \quad (2.23)$$

and hence for a zero temperature bath $\bar{n}_{\text{bath}} = 0$.

Two-particle (at a time) losses to the standard zero temperature heat bath[11] with rate $\gamma_{\mathbf{n}}^{(2)}$ at $\mathbf{x}_{\mathbf{n}}$:

$$\hat{L}_{\mathbf{n}} = \hat{a}_{\mathbf{n}}^2 \sqrt{\gamma_{\mathbf{n}}^{(2)}}. \quad (2.24)$$

⁵The Hamiltonian of the heat bath is taken to be $\hat{H} = \hbar\omega_{\text{bath}} \sum_j (\hat{b}_j^\dagger \hat{b}_j + \frac{1}{2})$ with \hat{b}_j being the annihilation operator for the j th heat bath subsystem.

Coherent gain from a driving field $\varepsilon_{\mathbf{n}}$ at $\mathbf{x}_{\mathbf{n}}$ can be modeled by adding a term of the form

$$\hat{H}_c = i\hbar \sum_{\mathbf{n}} [\varepsilon_{\mathbf{n}} \hat{a}_{\mathbf{n}}^\dagger - \varepsilon_{\mathbf{n}}^* \hat{a}_{\mathbf{n}}] \quad (2.25)$$

to the Hamiltonian.

Details of the derivation of the loss/gain expressions (like those above, or more general) can be found e.g. in Gardiner[53], Chapter 10, or Louisell[54, 55].

2.6 Thermodynamic equations of motion

The (un-normalized) density matrix of a grand canonical ensemble in contact with a bath at temperature T and with chemical potential μ is given by

$$\hat{\rho}_u = \exp \left[-(\hat{H} - \mu \hat{N}) / k_B T \right] = e^{-\hat{K}\tau}. \quad (2.26)$$

In the second expression the “imaginary time”⁶ $\tau = 1/k_B T$ and the “Kamiltonian” \hat{K} have been introduced.

The rate of change of $\hat{\rho}_u$ with τ can be written as⁷

$$\frac{\partial \hat{\rho}_u}{\partial \tau} = -\frac{1}{2} \left[\frac{\partial \hat{K}}{\partial \tau}, \hat{\rho}_u \right]_+ = -\frac{\partial \hat{K}}{\partial \tau} \hat{\rho}_u, \quad (2.27)$$

provided that

$$\left[\frac{\partial \hat{K}}{\partial \tau}, \hat{K} \right] = 0. \quad (2.28)$$

For the Hamiltonian here ((2.12) or (2.17)), the equation of motion can be written

$$\frac{\partial \hat{\rho}_u}{\partial \tau} = \left[\mu_e(\tau) \hat{N} - \hat{H} \right] \hat{\rho}_u, \quad (2.29)$$

where the “effective” chemical potential is

$$\mu_e(\tau) = \frac{\partial [\tau \mu(\tau)]}{\partial \tau}. \quad (2.30)$$

⁶So called because of the similarity of the left equality of (2.27) to (2.20) with t replaced by $i\hbar\tau/2$, and $\hat{L}_j = 0$.

⁷ $[\hat{A}, \hat{B}]_+ = \hat{A}\hat{B} + \hat{B}\hat{A}$.

The condition (2.28) implies, in this case, that only the chemical potential $\mu(T)$ can be temperature dependent⁸

The utility of this equation stems from the fact that the grand canonical ensemble at $\tau = 0$ (i.e. high temperature $T \rightarrow \infty$) is known, and given by the simple expression

$$\hat{\rho}_u(0) = \exp \left[-\lambda_n \hat{N} \right], \quad (2.31)$$

where

$$\lambda_n = -\lim_{\tau \rightarrow 0} [\tau \mu(\tau)], \quad (2.32)$$

and the initial mean number of particles per lattice site is

$$\bar{n}_0 = [e^{\lambda_n} - 1]^{-1}. \quad (2.33)$$

This leads to initial conditions for the stochastic simulation that can be efficiently sampled in most cases. Finite temperature results are then obtained by evolving the system forward in τ .

⁸Strictly speaking, a constant (in space) external potential V^{ext} may also be temperature dependent, but this is physically equivalent to a correction to μ . Quantities such as g and $V^{\text{ext}}(\mathbf{x})$ must be constant with T for (2.29) to hold.

Part A

Phase-space distribution methods for many-mode quantum mechanics

Chapter 3

Generalized quantum phase-space representations

3.1 Introduction

In this chapter a formalism encompassing very general phase-space distributions describing quantum states and their evolution is presented. The correspondence made is between a quantum density operator and a distribution of operators, or “kernels”, as in (1.4). The motivation for this is to investigate what distribution properties are essential for tractable mesoscopic first-principles simulations, and to provide a systematic framework in which comparison between methods is simplified. To this end, the following two conditions on the distribution will be kept in mind throughout:

- **Exact correspondence:** A statistical sample of the kernel operators chosen according to the phase-space distribution must approach the exact quantum density matrix in the limit of infinite samples. This property must be maintained during evolution via stochastic equations.
- **Variable number linear in system size:** The number of variables needed to specify a kernel (and hence the number of stochastic equations needed to model quantum evolution) must scale linearly with N , the number of subsystems.

The formalism is introduced in Section 3.2, and correspondence between density matrix and distribution considered. Subsequently, it is shown how the density matrix corresponds to a statistical sample of kernel operators in Section 3.3, while in 3.4 it is shown how the quantum evolution corresponds to a set of stochastic equations for these. Along the way, examples are given using the positive P representation, commonly used in quantum optics. The resulting requirements on the operator kernels for such a scheme are summarized in Section 3.6.

It is also shown that the phase-space distribution methods have two generic properties that are convenient for practical simulations:

- **Parallel sample evolution:** The individual stochastic realizations of the kernel operator (which are later to be averaged to obtain observables) evolve independently of each other. This allows straightforward and efficient parallel computation (more detail in Section 3.5).
- **All observables in one simulation:** One algorithm is capable of giving estimates for any/all observable averages (see Section 3.3.1).

Some previous generalizations of phase-space representations such as generalized P representations[11], and the (already quite general) discussion of phase-space representations in the article by Drummond and Deuar[56], are contained as special cases of the discussion in this chapter.

3.2 Representation of a density matrix

We expand the density matrix

$$\hat{\rho} = \int P(C)\hat{\Lambda}(C)dC, \quad (3.1)$$

as in (1.4), where C is a set of configuration variables specifying the kernel operator $\hat{\Lambda}(C)$. The idea is that if P is real and positive, it can be considered a probability distribution, and we then can approximate the density matrix using \mathcal{S} samples $\hat{\Lambda}(C)$. Each such sample is fully defined by its set of configuration parameters C .

3.2.1 Properties of the distribution

To be able to interpret the function $P(C)$ as a probability distribution, we require P to be

- REAL.
- NON-NEGATIVE: $P \geq 0$.
- NORMALIZABLE: i.e. $\int P(C)dC$ converges to a finite value.
- NON-SINGULAR. The primary reason why a singular distribution is a problem is that it may be incapable of being sampled in an unbiased way by a finite number of samples. For example, if one initially has a non-singular distribution $P(t)$ but singularities arise through some dynamical process after a time t_{sing} , then a set of samples that were unbiased estimators of $P(t)$ initially ($t \ll t_{\text{sing}}$), will generally be incapable of sampling the singularities at $t \geq t_{\text{sing}}$ correctly.

Some singular behavior can, however, be tolerated in initial distributions (e.g. $P(v) = \sum_{j \ll \mathcal{S}} \delta(v - v_0^{(j)})$ for some real variable v) provided the number of singularities is finite and much less than the number of samples \mathcal{S} . This can even be desirable to achieve a starting set of samples that is compact.

If these conditions are satisfied, then when each sample (labeled by j) is defined by its own configuration parameters $C^{(j)}$,

$$\lim_{\mathcal{S} \rightarrow \infty} \sum_{j=1}^{\mathcal{S}} \hat{\Lambda}(C^{(j)}) = \hat{\rho}. \quad (3.2)$$

Many well-known distributions of the general form (3.1) do not satisfy these conditions for general quantum states. For example the Wigner distribution[57] is commonly negative in some regions of phase space when the system exhibits nonclassical statistics, the complex P distribution[11] is not real, and the Glauber-Sudarshan P distribution[58, 59] can be singular[60] (also when nonclassical statistics are present). These distributions are often very useful, but not for many-body stochastic simulations.

3.2.2 Splitting into subsystems

A mesoscopic system will consist of some number N of subsystems. In this thesis the subsystems are usually spatial or momentum modes, although other common approaches are to label as subsystems individual particles (if their number is conserved), or orbitals.

A crucial aim, as pointed out at the beginning of this chapter and in Section 1.3, is to have system configurations C specifying each sample contain a number of parameters (variables) that is only linear in N . Otherwise any calculations with macro- or mesoscopic N will become intractable due to the sheer number of variables. Hence, the kernel $\widehat{\Lambda}$ should be a separable tensor product of subsystem operators¹:

$$\widehat{\Lambda} = f_{\text{glob}}(C_{\text{glob}}) \otimes_{k=1}^N \widehat{\Lambda}_k(C_k, C_{\text{glob}}). \quad (3.3)$$

Here the *local kernel* for each k th subsystem is described by its own set of local configuration variables C_k . There may also be some additional “global” variables (of number $\ll \mathcal{O}(N)$) in the set C_{glob} affecting global factors $f_{\text{glob}}(C_{\text{glob}})$ or several subsystems. The full configuration variable set for the kernel is $C = \{C_{\text{glob}}, C_1, \dots, C_M\}$.

A separate issue altogether is the choice of basis (i.e. the choice of $\widehat{\Lambda}_k$) for each subsystem. This chapter aims to stay general, and choosing $\widehat{\Lambda}_k$ is deferred to Chapter 5 and later, apart from discussing general features of the $\widehat{\Lambda}$ s necessary for a successful simulation.

3.2.3 Dual configuration space of off-diagonal kernels

Here, arguments will be put forward that for non-trivial simulations the local kernels $\widehat{\Lambda}_k$ are best chosen to include off-diagonal operators.

A density matrix can always be diagonalized in some orthogonal basis, however there are some basic reasons why such a basis is not usually suitable for many-mode simulations of the type considered here.

1. Firstly, a general quantum state will contain entanglement between subsystems, which precludes writing it as a distribution in separable form (3.3) with

¹Strictly speaking, $\widehat{\Lambda}$ could also be a finite sum of ($\ll N$) tensor products of subsystem operators. This may be useful in some situations.

only diagonal kernels.

2. Secondly, while for some situations the state will, after all, be separable, and could be written using only locally diagonal kernel operators, generally the basis that diagonalizes the density matrix will change with time in a nontrivial way. The kernels, on the other hand, do not change. This means that while the initial separable state could, in this case, be sampled with diagonal local kernels, the subsequent exact quantum evolution could not be simulated using those kernels.
3. Thirdly, exceptions to the above arguments occur if one has a system that remains separable while it evolves. Alternatively, if its inseparability has simple time-evolution that could be found exactly by other means, then this time-evolution could be hardwired into a time-dependent kernel $\widehat{\Lambda}(t)$ in such a way that (3.3) continues to hold with diagonal (now-time-dependent) local kernels. In such a case, however, there is no point in carrying out time-consuming stochastic simulations of the whole system, when one could just investigate each subsystem separately.

Still, the above arguments are not a rigorous proof, primarily because non-orthogonal basis sets have not been considered. However, no diagonal distribution that can be used to represent completely general quantum states is presently known. For example the Glauber-Sudarshan P distribution[58, 59], which has $\widehat{\Lambda}_k$ as projectors onto local coherent states (which are non-orthogonal), is defined for all quantum states, but it has been shown that when some nonclassical states are present, the distribution P becomes singular[60], and not amenable to unbiased stochastic simulations.

In summary, it appears that to allow for off-diagonal entangling coherences between subsystems, the local kernels should be of an off-diagonal form

$$\widehat{\Lambda}_k(C_k, C_{\text{glob}}) = |C'_k, C_{\text{glob}}\rangle \langle \widetilde{C}'_k, C_{\text{glob}}|, \quad (3.4)$$

where the C'_k and \widetilde{C}'_k are subsets of *different* independent parameters, and the local parameter set for the kernel is $C_k = \{C'_k, \widetilde{C}'_k\}$.

Finally, an exception to this off-diagonal conjecture might be the dynamics of closed systems — which remain as pure states for the whole duration of a simulation. In such a case, one might try to expand the state vector (rather than the density matrix) $|\psi\rangle$ directly, along the lines of

$$|\psi\rangle = \int P_\psi(C_\psi, \theta) e^{i\theta} |C_\psi\rangle dC_\psi d\theta. \quad (3.5)$$

The global phase factor $e^{i\theta}$ is required to allow for superpositions with a real positive P_ψ .

3.2.4 An example: the positive P distribution

So that the discussion does not become too opaque, let us make a connection to how this looks in a concrete example, the positive P distribution[10, 11]. This distribution has widely used with success in quantum optics, and with Bose atoms as well[15, 61] in mode-based calculations. It also forms the basis of the gauge P distribution, which will be explained and investigated in Chapter 5 and used throughout this thesis. The positive P local kernel at a lattice point (spatial mode — of which there are N) is

$$\widehat{\Lambda}_k = |\alpha_k\rangle_k \langle\beta_k^*|_k \exp\left(-\alpha_k\beta_k + \frac{1}{2}|\alpha_k|^2 + \frac{1}{2}|\beta_k|^2\right), \quad (3.6)$$

where the states $|\alpha_k\rangle_k$ are normalized coherent states at the k th lattice point with amplitude α_k , their form given by

$$|\alpha_k\rangle_k = \exp\left[\alpha_k \widehat{a}_k - \frac{1}{2}|\alpha_k|^2\right] |0\rangle_k. \quad (3.7)$$

Note: the subscript k on state vectors indicates that they are local to the k th subsystem, and $|0\rangle_k$ is the vacuum. The global function is just $f_{\text{glob}}(C_{\text{glob}}) = 1$, and the entire kernel can be written in terms of vectors of parameters $\boldsymbol{\alpha} = \{\alpha_1, \dots, \alpha_N\}$ using

$$|\boldsymbol{\alpha}\rangle = \otimes_{k=1}^N |\alpha_k\rangle_k \quad (3.8)$$

as

$$\widehat{\Lambda} = |\boldsymbol{\alpha}\rangle \langle\boldsymbol{\beta}^*| \exp\left(-\boldsymbol{\alpha} \cdot \boldsymbol{\beta} + \frac{1}{2}|\boldsymbol{\alpha}|^2 + \frac{1}{2}|\boldsymbol{\beta}|^2\right). \quad (3.9)$$

The sets of parameters are $C_{\text{glob}} = \{\}$, $C_k = \{\alpha_k, \beta_k\}$. The kernel is in “dual” off-diagonal operator space, and it has been shown[11] that any density matrix can be written with real, positive $P(C)$ using this kernel. It has also been shown constructively[11] that a non-singular positive P distribution exists, and is given by

$$P_+(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \frac{1}{(4\pi^2)^N} \exp\left(-\frac{1}{4}|\boldsymbol{\alpha} - \boldsymbol{\beta}^*|^2\right) \left\langle \frac{\boldsymbol{\alpha} + \boldsymbol{\beta}^*}{2} \middle| \hat{\rho} \middle| \frac{\boldsymbol{\alpha} + \boldsymbol{\beta}^*}{2} \right\rangle. \quad (3.10)$$

There are four real variables per lattice point.

3.3 Stochastic interpretation of the distribution

To efficiently sample the quantum state we make two correspondences. Firstly, as discussed in Section 3.2, the distribution $P(C)$ corresponds to the density matrix $\hat{\rho}$. The second correspondence, as per (3.2), is between the distribution $P(C)$ that exists in some high-dimensional space, and \mathcal{S} samples $C^{(j)}$ distributed according to $P(C)$.

3.3.1 Calculating observables

Quantum mechanics concerns itself with calculating expectation values of observables, so the equivalence between it and the stochastic equations for the variables in C rests solely on obtaining the same evolution of observable averages.

Suppose one wishes to calculate the expectation value of observable \hat{O} given a set of \mathcal{S} operator samples $\{\hat{\Lambda}(C^{(j)})\}$, with $j = 1, \dots, \mathcal{S}$. (Operationally, one actually has the set of \mathcal{S} operator parameter sets $\{C^{(j)}\}$.) For a normalized density matrix, the expectation value is $\langle \hat{O} \rangle = \text{Tr} [\hat{O}\hat{\rho}]$, however e.g. in the thermodynamic evolution of Section 2.6 one has un-normalized density matrices, in which case

$$\langle \hat{O} \rangle = \frac{\text{Tr} [\hat{O}\hat{\rho}_u]}{\text{Tr} [\hat{\rho}_u]}. \quad (3.11)$$

Using the representation (3.1) this leads to

$$\langle \hat{O} \rangle = \frac{\int P(C) \text{Tr} [\hat{O}\hat{\Lambda}(C)] dC}{\int P(C) \text{Tr} [\hat{\Lambda}(C)] dC}. \quad (3.12)$$

We can make use of the Hermitian nature of both density matrices and observables to get extra use out of non-Hermitian (off-diagonal) kernels, giving

$$\langle \widehat{O} \rangle = \frac{\int P(C) \left(\text{Tr} \left[\widehat{O} \widehat{\Lambda}(C) \right] + \text{Tr} \left[\widehat{O} \widehat{\Lambda}^\dagger(C^*) \right] \right) dC}{\int P(C) \text{Tr} \left[\widehat{\Lambda}(C) + \widehat{\Lambda}^\dagger(C^*) \right] dC}. \quad (3.13)$$

Actually, one could impose hermiticity on the kernel at the representation level via $\widehat{\Lambda} \rightarrow \frac{1}{2}(\widehat{\Lambda} + \widehat{\Lambda}^\dagger)$, but this can complicate the resulting stochastic equations for C — so, let us impose this only at the level of the observable moments.

When samples are taken $P(C)$ is interpreted as a probability distribution, and, lastly, noting that the expectation values of observables are real (since \widehat{O} is Hermitian), one obtains that the expectation value of an arbitrary observable \widehat{O} can be estimated from the samples by the quantity

$$\bar{O} = \frac{\langle \text{Re} \left\{ \text{Tr} \left[\widehat{O} \widehat{\Lambda} \right] \right\} \rangle_{\text{stoch}} + \langle \text{Re} \left\{ \text{Tr} \left[\widehat{O} \widehat{\Lambda}^\dagger \right] \right\} \rangle_{\text{stoch}}}{2 \langle \text{Re} \left\{ \text{Tr} \left[\widehat{\Lambda} \right] \right\} \rangle_{\text{stoch}}}, \quad (3.14)$$

where stochastic averages over samples are indicated by $\langle \cdot \rangle_{\text{stoch}}$. The correspondence² is

$$\lim_{\mathcal{S} \rightarrow \infty} \bar{O} = \langle \widehat{O} \rangle. \quad (3.15)$$

3.3.2 Assessing estimate accuracy

One expects that the accuracy of the averages in the (3.14) expression using \mathcal{S} samples will improve as $\sqrt{\mathcal{S}}$ via the Central Limit Theorem (CLT), since they are just normalized sums over \mathcal{S} terms. The uncertainty in a mean value $\bar{v} = \langle v \rangle_{\text{stoch}}$ calculated with \mathcal{S} samples can then be estimated by

$$\Delta \bar{v} = \sqrt{\frac{\langle v^2 \rangle_{\text{stoch}} - \langle v \rangle_{\text{stoch}}^2}{\mathcal{S}}} \quad (3.16)$$

at the one σ confidence level.

There is, however, a subtlety when several averages over the same variables C are combined as in (3.14), because the quantities averaged may be correlated.

²Provided there are no boundary term errors (see Chapter 6).

Then the accuracy estimate (3.16) may be either too large or too small. A way to overcome this is **subensemble averaging**. While the best estimate \bar{O} is still calculated from the full ensemble \mathcal{S} as in (3.14), the \mathcal{S} samples are also binned into \mathcal{S}_E subensembles with s samples in each ($\mathcal{S} = s\mathcal{S}_E$). The s elements of each subensemble are used as in (3.14) to obtain independent estimates of the observable average $\{\bar{O}^{(1)}, \bar{O}^{(2)}, \dots, \bar{O}^{(\mathcal{S}_E)}\}$, distributed around \bar{O} . One has (approximately)

$$\bar{O} \approx \frac{1}{\mathcal{S}_E} \sum_j \bar{O}^{(j)}. \quad (3.17)$$

Given this, the CLT can be applied to the $\bar{O}^{(j)}$ to estimate the uncertainty in the observable estimate at the one σ level:

$$\Delta\bar{O} = \Delta\langle\hat{O}\rangle = \sqrt{\frac{\frac{1}{\mathcal{S}_E} \sum_j \left(\bar{O}^{(j)}\right)^2 - \frac{1}{\mathcal{S}_E^2} \left(\sum_j \bar{O}^{(j)}\right)^2}{\mathcal{S}_E}}. \quad (3.18)$$

A practical issue to keep in mind is that both the number of samples in a subensemble s and the number of subensembles themselves should be large enough so that both: 1) The subensemble estimates $\bar{O}^{(j)}$ are reasonably close to the accurate value $\bar{O} \approx \langle\hat{O}\rangle$ so that (3.17) is true, and also 2) that there is enough of them that the right hand side of (3.17) has an approximately Gaussian distribution, and the CLT can be applied to obtain (3.18).

An issue that arises when \bar{O} involves a quotient of random variable averages as in (3.14), is that the subensemble size s should be large enough that the denominator is far from zero for all subensembles. Otherwise subensembles for which this denominator is close to zero have an inordinate importance in the final estimate of the mean. More details in Appendix C.

3.3.3 Calculating non-static observables

The observable estimate (3.14) of Section 3.3.1 implicitly assumes that the explicit form of \hat{O} is known *a priori*. When comparing with experiment, these are usually all the observable averages one needs.

Nevertheless, in many theoretical works some other observables that do not fit this mould are considered. Perhaps the most common example of these is the mutual

fidelity between two states $\hat{\rho}_1$ and $\hat{\rho}_2$

$$F(\hat{\rho}_1, \hat{\rho}_2) = \text{Tr} [\hat{\rho}_1 \hat{\rho}_2] = \frac{\text{Tr} [\hat{\rho}_{1u} \hat{\rho}_{2u}]}{\text{Tr} [\hat{\rho}_{1u}] \text{Tr} [\hat{\rho}_{2u}]}, \quad (3.19)$$

which lies between zero (for two orthogonal pure states) and unity for pure $\hat{\rho}_1 = \hat{\rho}_2$. A commonly considered special case is the purity $F_0(\hat{\rho}) = F(\hat{\rho}, \hat{\rho})$.

A difficulty with evaluating such quantities as $F_0(\hat{\rho})$ or $F(\hat{\rho}_1, \hat{\rho}_2)$ on states calculated with stochastic methods is that the density matrices are not known exactly, only their estimates in the form of \mathcal{S} kernel samples. However, one can still proceed by expanding using (3.1) as $\hat{\rho}_j = \int P_j(C_j) \hat{\Lambda}_j(C_j) dC_j$ ($j = 1, 2$) to give

$$F(\hat{\rho}_1, \hat{\rho}_2) = \frac{\int P_1(C_1) P_2(C_2) \text{Tr} [\hat{\Lambda}_1(C_1) \hat{\Lambda}_2(C_2)] dC_1 dC_2}{\int P_1(C_1) \text{Tr} [\hat{\Lambda}_1(C_1)] dC_1 \int P_2(C_2) \text{Tr} [\hat{\Lambda}_2(C_2)] dC_2}. \quad (3.20)$$

(Note that the kernels ($\hat{\Lambda}_1$ and $\hat{\Lambda}_2$, respectively) used for representing the two density matrices do not have to be of the same form. The utility of doing so, however, depends entirely on whether the trace of the product of the two kinds of kernels can be evaluated in closed form for arbitrary sets of parameters C_1 and C_2 .)

As before, off-diagonal kernels can be used twice due to the hermiticity of density matrices, giving

$$F(\hat{\rho}_1, \hat{\rho}_2) = \frac{\int P_1(C_1) P_2(C_2) \text{Tr} [(\hat{\Lambda}_1(C_1) + \hat{\Lambda}_1^\dagger(C_1^*)) (\hat{\Lambda}_2(C_2) + \hat{\Lambda}_2^\dagger(C_2^*))] dC_1 dC_2}{\int P_1(C_1) \text{Tr} [\hat{\Lambda}_1(C_1) + \hat{\Lambda}_1^\dagger(C_1^*)] dC_1 \int P_2(C_2) \text{Tr} [\hat{\Lambda}_2(C_2) + \hat{\Lambda}_2^\dagger(C_2^*)] dC_2}. \quad (3.21)$$

The $P_1(C_1)$ and $P_2(C_2)$ can be interpreted as probability distributions of *independently realized* configuration samples. Denoting the first set of \mathcal{S}_1 samples as $\{C_1^{(j_1)}\}$, and the second set of \mathcal{S}_2 as $\{C_2^{(j_2)}\}$, one obtains an estimate of the mutual fidelity between the two states represented by those two sets of samples as

$$\bar{F} = \frac{\sum_{j_1, j_2} \text{Re} \left\{ \text{Tr} \left[\hat{\Lambda}_1(C_1^{(j_1)}) \left\{ \hat{\Lambda}_2(C_2^{(j_2)}) + \hat{\Lambda}_2^\dagger([C_2^{(j_2)}]^*) \right\} \right] \right\}}{2 \sum_{j_1} \text{Re} \left\{ \text{Tr} \left[\hat{\Lambda}_1(C_1^{(j_1)}) \right] \right\} \sum_{j_2} \text{Re} \left\{ \text{Tr} \left[\hat{\Lambda}_2(C_2^{(j_2)}) \right] \right\}} \quad (3.22)$$

(remembering that observable averages are real), with the large sample limit

$$\lim_{\mathcal{S}_1 \rightarrow \infty, \mathcal{S}_2 \rightarrow \infty} \bar{F} = F(\hat{\rho}_1, \hat{\rho}_2). \quad (3.23)$$

Note that the sums in (3.22) are carried out over *all pairs* of samples $C_1^{(j_1)}$ and $C_2^{(j_2)}$.

If the two states whose mutual fidelity is to be calculated arise in the same simulation, then one can ensure independence of the two sets of samples labeled by j_1 and j_2 by setting aside half of all simulated samples for the set $\{C_1^{(j_1)}\}$, the other for $\{C_2^{(j_2)}\}$. This separation of samples is always necessary if one wants to estimate purity F_0 .

Uncertainty estimates can be obtained in a similar subensemble averaging manner as outlined in Section 3.3.2, by working out (3.22) for each subensemble, then using the CLT to estimate uncertainty in the mean of the subensemble estimates, by the same procedure as in (3.18).

Finally, a major disadvantage of trying to work out fidelity estimates in the above manner from such stochastic simulations is that one must keep all the trajectories in computer memory, so that after all samples have been produced they can be combined in all possible pairs to evaluate the quantities $\text{Tr} \left[\widehat{\Lambda}_1(C_1^{(j_1)}) \widehat{\Lambda}_2(C_2^{(j_2)}) \right]$. Compared to a simulation that only considers static observables as per (3.14), this increases the space required by a factor of \mathcal{S} .

3.3.4 Overcomplete vs. orthogonal bases

Consider observable calculations when using local orthogonal bases for the local kernel operators $\widehat{\Lambda}_k$. In a mode formulation, this would imply that the local parameters C_k consist of discrete quantum numbers for the k th mode (e.g. occupation), while for a particle formulation C_k could, for example, consist of (continuous) positions of the k th particle, since all position eigenstates are orthogonal. Denoting this basis as $|C'_k\rangle_k$, and remembering that off-diagonal kernels should be used (see Section 3.2.3), a typical local kernel (omitting global parameters C_{glob}) will have the form $\widehat{\Lambda}_k = |C'_k\rangle_k \langle \widetilde{C}'_k|_k$ as in (3.4), and the local parameter set is $C_k = \{C'_k, \widetilde{C}'_k\}$.

Suppose one wants to calculate the expectation value of a local observable $\widehat{O} = \widehat{O}_k \otimes_{k' \neq k} \widehat{I}_{k'}$ that is diagonal³ in this $|C'_k\rangle_k$ basis. The observable estimate expression

³Or that contains some components diagonal in the basis $|C'_k\rangle_k$.

(3.14) is proportional to averages of quantities like

$$\mathrm{Tr} \left[\widehat{O} \widehat{\Lambda} \right] = \mathrm{Tr} \left[\widehat{O}_k \widehat{\Lambda}_k \right] f_{\mathrm{glob}}(C_{\mathrm{glob}}) \prod_{k' \neq k} \mathrm{Tr} \left[\widehat{\Lambda}_{k'} \right]. \quad (3.24)$$

If *any one* of the local kernel samples $\widehat{\Lambda}_{k'}$ or $\widehat{\Lambda}_k$ are non-diagonal, then the whole N -subsystem sample contributes nothing to the observable estimate, even if the off-diagonal kernel samples are for a different subsystem than the one in which the local observable is being considered (k). In a large system, *practically all* samples are likely to have at least one subsystem in which the local kernel is non-diagonal⁴. For example, if a proportion $p_{\mathrm{od}} \leq 1$ of samples of each subsystem are off-diagonal then only a proportion $(1 - p_{\mathrm{od}})^N$ will contribute to observable estimates. The result — no reasonable estimate of observables despite lots of calculation. Similar effects occur for off-diagonal local observables (in this case for a sample to contribute all $k' \neq k$ subsystems must have diagonal kernel samples, while the k th subsystem sample must have diagonal $\widehat{O}_k \widehat{\Lambda}_k$), and for observables $\widehat{O} = \otimes_j \widehat{O}_j$ spanning several subsystems (in this case $\widehat{O}_j \widehat{\Lambda}_j$ must be diagonal for all j).

A related issue arises in path integral methods (based on particle positions, for example), and it is the reason that the \mathcal{M} configurations in one “sample” must form a ring polymer structure ($C_{(0)} = C_{(\mathcal{M})}$) if one wants to calculate diagonal observables. Off-diagonal observables (e.g. momentum distributions in position eigenstate bases) require a separate simulation, which contains some open polymer structures.

This inefficient sampling is the main reason why simulations based on particle positions, or occupation numbers at lattice sites, are not usually successful in phase-space distribution based methods. In both cases the position eigenstates or Fock occupation number states used are orthogonal.

To deal with this sampling problem, one can use overcomplete basis sets (in which the basis vectors are not orthogonal) for the description of the subsystems. Then the local off-diagonal kernels $\widehat{\Lambda}_k$ can usually be explicitly normalized so that $\mathrm{Tr} \left[\widehat{\Lambda}_k \right] = 1$. The traces averaged in (3.14) become simply

$$\mathrm{Tr} \left[\widehat{O} \widehat{\Lambda} \right] = \mathrm{Tr} \left[\widehat{O}_k \widehat{\Lambda}_k \right] f_{\mathrm{glob}}(C_{\mathrm{glob}}), \quad (3.25)$$

⁴Apart from some special cases, of course.

and only local kernel parameters C_k are relevant for the calculation of local observables. (Possibly apart from some global factors, which do not lead to sample contributions decaying exponentially with N).

The majority of this thesis considers methods based on coherent state expansions, which are of this kind. A non-orthogonal approach for particle-conserving systems that does use a position wavefunction is to write the density matrix as a distribution over off-diagonal projectors onto orbitals (with arbitrary wavefunctions, not necessarily orthogonal) occupied by all N particles, as recently developed by Carusotto *et al*[1].

Overcomplete expansions go a long way towards working around the problem, but not all non-orthogonal kernels that one might want to use can be normalized. A case in point are kernels that have zero trace for a set of parameters of measure zero. In such a case one should definitely *not* normalize (at least not fully) as boundary term errors tend to result (see Chapter 6). A partial normalization that avoids any singularities in the resulting kernel (potentially caused by normalizing a kernel whose trace tends to zero) may be successful in such cases. However, the inefficient sampling can then recur. The reason is that there is then an effective “overlap” range in configuration space, such that if the “ket” and “bra” parameters C'_k and \tilde{C}'_k differ by more than this range, the resulting kernel will have minimal weight in observable averages in comparison with more diagonal kernels. If the phase-space distributions of C'_k and \tilde{C}'_k have spreads significantly wider than the overlap range, a proportion $p_{\text{od}} \approx \mathcal{O}(1)$ of samples again have largely zero weight in moment calculations, leading again to only $\approx (1 - p_{\text{od}})^N$ significant samples overall. This situation can be additionally insidious because the weight of the samples is never exactly zero. This leads superficially to finite averages in the observable expression (3.14), however if $\mathcal{S}(1 - p_{\text{od}})^N \lesssim 1$ there will be rare “very diagonal” configurations that are not sampled at all, but contribute a great deal to the averages in the limit $\mathcal{S} \rightarrow \infty$. Conclusion: Care must be taken with un-normalized kernels to make sure no sampling bias is introduced.

Summarizing, barring possible developments that may be able to work around the sampling problem discussed in this section, it appears to be necessary to use

overcomplete basis sets for the local kernels $\widehat{\Lambda}_k$ to avoid wasting a lot of computing resources on samples that do not end up contributing to observable estimates. It also appears desirable to make the trace of the local kernels normalized to unity if possible.

3.3.5 Positive P distribution example continued

Returning to the positive P distribution example of Section 3.2.4, all operators of supported states can be written as linear combinations of the moments of the local annihilation and creation operators $\widehat{a}_k, \widehat{a}_k^\dagger$. Thus, to evaluate any observable it suffices to know how to evaluate an expectation value of a general product of the form $\otimes_j \widehat{a}_{L_j}^\dagger \otimes_k \widehat{a}_{L'_k}$ (The L_j and L'_k are subsystem labels (not necessarily unique), while the j and k are “subsystem label counters”).

The coherent states forming the overcomplete kernels are eigenstates of the annihilation operators \widehat{a}_k :

$$\widehat{a}_k |\alpha_k\rangle_k = \alpha_k |\alpha_k\rangle_k, \quad (3.26)$$

and kernels are normalized:

$$\text{Tr} [\widehat{\Lambda}_k] = 1. \quad (3.27)$$

Now the same procedure as in Section 3.3.1 for (3.14) if followed, omitting only taking the real part of the numerator and denominator because $\otimes \widehat{a}^\dagger \otimes \widehat{a}$ is (for convenience) not necessarily a strictly Hermitian observable. This leads to the expectation value estimate

$$\begin{aligned} \langle \otimes_j \widehat{a}_{L_j}^\dagger \otimes_k \widehat{a}_{L'_k} \rangle &= \lim_{S \rightarrow \infty} \overline{\{L_j, L'_k\}} \\ \overline{\{L_j, L'_k\}} &= \frac{1}{2} \left\langle \prod_j \beta_{L_j} \prod_k \alpha_{L'_k} + \prod_j \alpha_{L_j}^* \prod_k \beta_{L'_k}^* \right\rangle_{\text{stoch}}. \end{aligned} \quad (3.28)$$

Hermitian observables are constructed by combining the operator products and stochastic averages of (3.28) with their adjoints and complex conjugates, respectively.

For the fidelity calculations of Section 3.3.3, one finds

$$\mathrm{Tr} \left[\widehat{\Lambda}(\boldsymbol{\alpha}_1, \boldsymbol{\beta}_1) \widehat{\Lambda}(\boldsymbol{\alpha}_2, \boldsymbol{\beta}_2) \right] = \prod_k \mathrm{Tr} \left[\widehat{\Lambda}_k(\alpha_{1k}, \beta_{1k}) \widehat{\Lambda}_k(\alpha_{2k}, \beta_{2k}) \right] \quad (3.29a)$$

$$\mathrm{Tr} \left[\widehat{\Lambda}_k(\alpha_{1k}, \beta_{1k}) \widehat{\Lambda}_k(\alpha_{2k}, \beta_{2k}) \right] = \exp [-(\alpha_{1k} - \alpha_{2k})(\beta_{1k} - \beta_{2k})] \quad (3.29b)$$

$$\mathrm{Tr} \left[\widehat{\Lambda}_k(\alpha_{1k}, \beta_{1k}) \widehat{\Lambda}_k^\dagger(\alpha_{2k}^*, \beta_{2k}^*) \right] = \exp [-(\alpha_{1k} - \beta_{2k}^*)(\beta_{1k} - \alpha_{2k}^*)] \quad (3.29c)$$

for two kernel samples described by variables α_{1k}, β_{1k} and α_{2k}, β_{2k} respectively.

The coherent state basis is non-orthogonal:

$$\langle \beta^* | \alpha \rangle = e^{\alpha\beta - \frac{1}{2}|\alpha|^2 - \frac{1}{2}|\beta|^2}, \quad (3.30)$$

and overcomplete:

$$\int |\alpha\rangle\langle\alpha| d^2\alpha = \widehat{I}\pi. \quad (3.31)$$

(Note: the notation $d^2v = d\mathrm{Re}\{v\} d\mathrm{Im}\{v\}$ is used.)

3.3.6 Non-uniqueness of distributions

When kernels with distinct parameters are non-orthogonal, as is the case when using local overcomplete basis sets, a given quantum state $\widehat{\rho}$ can be described by a whole family of different distributions $P(C)$.

A simple example is the vacuum state with the aforementioned positive P distribution. Any density matrix that satisfies

$$\langle 0 | \widehat{\rho} | 0 \rangle = \mathrm{Tr} [\widehat{\rho}] \quad (3.32)$$

must be the vacuum state $|0\rangle\langle 0|$, since (3.32) specifies that the population of any non-vacuum states must be zero. Expanding as in (3.1) gives

$$\int P(C) \langle 0 | \widehat{\Lambda}(C) | 0 \rangle dC = \int P(C) \mathrm{Tr} [\widehat{\Lambda}(C)] dC, \quad (3.33)$$

which for the positive P distribution is (using $\int P(C) dC = 1$)

$$\int P_+(\boldsymbol{\alpha}, \boldsymbol{\beta}) e^{-\boldsymbol{\alpha}\cdot\boldsymbol{\beta}} d^{2N}\boldsymbol{\alpha} d^{2N}\boldsymbol{\beta} = 1, \quad (3.34)$$

A class of P distributions that satisfy this are Gaussian ensembles of $\alpha_k = \beta_k$ for each subsystem k :

$$\begin{aligned} P_+(\boldsymbol{\alpha}, \boldsymbol{\beta}) &= \prod_k P_k(\alpha_k, \beta_k), \\ P_{+k}(\alpha_k, \beta_k) &= \delta^2(\alpha_k - \beta_k) \left(\frac{1+2\sigma_k^2}{2\pi\sigma_k^2} \right) e^{-|\alpha_k|^2/2\sigma_k^2} \end{aligned}$$

for arbitrary positive real standard deviations $\boldsymbol{\sigma} = \{\sigma_1, \dots, \sigma_N\}$. This can be checked by substitution into (3.34). The most compact such distribution is of course in the limit of $\boldsymbol{\sigma} \rightarrow 0$, when

$$P_+(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \delta^{2N}(\boldsymbol{\alpha})\delta^{2N}(\boldsymbol{\beta}). \quad (3.35)$$

One consequence of this is that the variables α_k and β_k do not necessarily correspond to any physical observable. For example the ‘‘coherent state amplitudes’’ α_k may be highly nonzero, despite representing the vacuum. Only when they appear in the combinations allowed by (3.14) is there a physical meaning. As an example, the estimator of occupation number $\langle \hat{n}_k \rangle$ is given, from (3.14), by

$$\begin{aligned} \bar{n}_k &= \langle \text{Re} \{ \alpha_k \beta_k \} \rangle_{\text{stoch}} \\ &= \int P_+(\boldsymbol{\alpha}, \boldsymbol{\beta}) \text{Re} \{ \alpha_k \beta_k \} d^{2N} \boldsymbol{\alpha} d^{2N} \boldsymbol{\beta} \\ &= \int P_{+k}(\alpha_k, \beta_k) \text{Re} \{ \alpha_k \beta_k \} d^2 \alpha_k d^2 \beta_k \\ &\propto \int e^{-|\alpha_k|^2/2\sigma_k^2} (\text{Re} \{ \alpha_k \}^2 - \text{Im} \{ \alpha_k \}^2) d^2 \alpha_k \\ &= \int e^{-\text{Re} \{ \alpha_k \}^2/2\sigma_k^2} e^{-\text{Im} \{ \alpha_k \}^2/2\sigma_k^2} (\text{Re} \{ \alpha_k \}^2 - \text{Im} \{ \alpha_k \}^2) d\text{Re} \{ \alpha_k \} d\text{Im} \{ \alpha_k \} \\ &= 0, \end{aligned} \quad (3.36)$$

the last line following from $\text{Re} \{ \alpha_k \} \leftrightarrow \text{Im} \{ \alpha_k \}$ symmetry. In practice, however, it is not irrelevant which distribution (3.35) is used, because the low- $\boldsymbol{\sigma}$ distributions are much more compact and thus give much better accuracy when calculating observable estimates from a finite sample.

Generalizing, some consequences of the non-uniqueness of distributions based on overcomplete basis sets are expected to be:

- The variables in the sampled set C will not necessarily correspond to any physical observables. Only when they appear in the observable combinations

allowed by (3.14) is there a physical meaning. Because of this, such distributions are then often said to be in a “non-classical” phase space. (Also because of the number of variables (at least 4) specifying a local off-diagonal kernel (classically only two - e.g. position and momentum)).

- The same state may be represented by many different distributions of the variables. However, some of these will give more precise estimates of observables than others⁵.
- Some distributions representing the same state may be narrower than others in certain directions of phase space, while broader in other directions. This means that a distribution may be better than another for estimating one observable, but worse for a different one.
- Such differences in practical efficiency open the way for various optimizations, which here will be called stochastic gauges. See Chapters 4 and 7.
- It is generally more difficult to obtain an explicit expression for $P(C)$ given an initial state. If the kernel is orthogonal in a basis $|C'\rangle$, thus $\hat{\Lambda}(C = \{C', \tilde{C}'\}) = |C'\rangle \langle \tilde{C}'|$, obtaining the unique $P(\hat{\rho})$ is easy: $P(C) = \langle C' | \hat{\rho} | \tilde{C}' \rangle$. In an over-complete basis, one can obtain expressions such as (3.10), but these are usually only one of many possible.

3.4 Equations of motion

To simulate the evolution of the state one must make correspondences to the quantum master equation⁶, analogous to those made for the density operator in the previous sections. The master equation results in a differential equation⁷ for $P(C)$. This can then often be made to correspond to stochastic equations for the configuration variables C , which specify the evolution of the \mathcal{S} initial samples.

⁵Given the same number of samples.

⁶Or, possibly different non-Markovian kinds of quantum evolution equations. In this thesis, however, only master equations are considered.

⁷A partial differential equation if C is continuous, or a set of coupled (possibly partial) differential equations for the various probabilities $P(C)$ if C is discrete.

Now, while some equation for $P(C)$ can always be obtained, in its raw form it is not very useful, because solving it directly for $P(C, t)$ may be as (or even more) computationally intensive as solving the master equation directly. For example for continuous variables, one would try to evaluate the evolution of $P(C)$ on a lattice of (say) M_{latt} points per variable. Since the number of variables in the set C is linear in N , the number of lattice points on which to evaluate the evolution of $P(C)$ would scale as M_{latt}^N — exponential in N , just like a brute force density matrix element approach.

One needs to make a correspondence between the evolution of $P(C)$ and stochastic evolution of C , its samples. This is not in general a trivial (or perhaps even a feasible) exercise, but when $P(C)$ obeys one of the broad class of Fokker-Planck second order partial differential equations (FPE), there is a well-known “textbook” method[53, 62] to obtain stochastic equations for C . This is the approach that will be henceforth considered here.

It is worth noting that there may also be ways of obtaining stochastic equations for the configuration variables C for different kinds of equations. One direction that has been recently investigated by Olsen *et al*[63] are differential equations for $P(C)$ involving third order partial derivative terms. The resulting equations have stochastic terms with different statistics than the Wiener increments that will be used here. The results are promising but there appear to be serious numerical stability problems.

3.4.1 Master equation to Fokker-Planck equation

Let us investigate in more detail what exactly is required to obtain an FPE for P . These are of the general form

$$\frac{\partial}{\partial t}P(C) = - \sum_j \frac{\partial}{\partial C_j} [A_j(C)P(C)] + \frac{1}{2} \sum_{jk} \frac{\partial^2}{\partial C_j \partial C_k} [D_{jk}(C)P(C)]. \quad (3.37)$$

A point on notation: here, the indices j, k will label all system configuration variables C_j consecutively, *not* implying any relationship between C_j and a j th subsystem.

The first (rather trivial, but significant) comment is that since an FPE has terms

of the differential form $\partial P(C)/\partial C_j$, the system variables C_j should all be continuous so that partial derivatives can be defined.

Additionally, not all combinations of master equations and operator kernels can lead to Fokker-Planck equations. Let us investigate what the exact requirements are. All master equations consist of terms \mathcal{T}_l linear in the density operator, and can be written as

$$\frac{\partial \hat{\rho}}{\partial t} = \sum_l \mathcal{T}_l[\hat{\rho}], \quad (3.38)$$

with the \mathcal{T}_l composed of combinations of system operators \hat{A}_j and linear in⁸ $\hat{\rho}$. When the kernel depends on continuous parameters, such system operators and the kernel can often be found to satisfy mutual second-order differential identities of the forms

$$\hat{A}_j \hat{\Lambda}(C) = \left[A_j^{(0)}(C) + \sum_k A_{jk}^{(1)}(C) \frac{\partial}{\partial C_k} + \sum_{kl} A_{jkl}^{(2)}(C) \frac{\partial}{\partial C_k} \frac{\partial}{\partial C_l} \right] \hat{\Lambda}(C) \quad (3.39a)$$

$$\hat{\Lambda}(C) \hat{A}_j = \left[\tilde{A}_j^{(0)}(C) + \sum_k \tilde{A}_{jk}^{(1)}(C) \frac{\partial}{\partial C_k} + \sum_{kl} \tilde{A}_{jkl}^{(2)}(C) \frac{\partial}{\partial C_k} \frac{\partial}{\partial C_l} \right] \hat{\Lambda}(C) \quad (3.39b)$$

where all differential operators are henceforth defined to act on the right. (a well-known example is in (3.57) for the positive P representation.) If, however, higher than second order differential terms must also be included, no FPE can be obtained.

Assuming (3.39) hold, by substitution of them into the forms of the $\mathcal{T}_l(\{\hat{A}_j\}, \hat{\rho})$, differential expressions for $\mathcal{T}_l[\hat{\Lambda}]$ can be found (note that now \mathcal{T}_l acts on the kernel rather than the density matrix itself). Provided that the master equation does not contain terms of too high order in operators for which the differential terms are nonzero, then the operation of each term on the kernel will be able to (after some algebra) be written in a form:

$$\mathcal{T}_l[\hat{\Lambda}] = \left[T_l^{(0)}(C) + \sum_j T_{lj}^{(1)}(C) \frac{\partial}{\partial C_j} + \frac{1}{2} \sum_{jk} T_{ljk}^{(2)}(C) \frac{\partial}{\partial C_j} \frac{\partial}{\partial C_k} \right] \hat{\Lambda}. \quad (3.40)$$

The factor $\frac{1}{2}$ in second order terms is introduced for later convenience.

It is crucial that this expansion in partial derivatives terminates at second order for all terms \mathcal{T}_l , otherwise the final partial differential equation for P will not be of the

⁸Or, equally well, linear in an un-normalized $\hat{\rho}_u$.

FPE form (3.37). This usually creates a restriction on the complexity of processes in the model that can be considered. For example in the gauge P representation used in later parts of this thesis, one-particle and two-particle interactions are fine, but three-particle collisions cannot be treated. This restriction can often be surmounted if one increases the complexity of the kernel, at the cost of additional complexity in the equations.

Having confirmed that the identities (3.40) hold with no 3rd or higher derivative terms, the FPE for P can be derived as follows. Firstly, note that only kernels that are not explicitly time-dependent are being considered here, although time-dependent kernels can also be treatable by a similar, but more voluminous approach. Having said this, the master equation expanded as (3.38) can be written using (3.40) as

$$\int \hat{\Lambda}(C) \frac{\partial P(C)}{\partial t} dC = \sum_l \int P(C) \times \left[T_l^{(0)}(C) + \sum_j T_{lj}^{(1)}(C) \frac{\partial}{\partial C_j} + \frac{1}{2} \sum_{jk} T_{ljk}^{(2)}(C) \frac{\partial}{\partial C_j} \frac{\partial}{\partial C_k} \right] \hat{\Lambda}(C) dC. \quad (3.41)$$

We now integrate the right hand side by parts, which gives

$$\int \hat{\Lambda}(C) \frac{\partial P(C)}{\partial t} dC = \sum_l \int \hat{\Lambda}(C) \times \left[T_l^{(0)}(C) - \sum_j \frac{\partial}{\partial C_j} T_{lj}^{(1)}(C) + \frac{1}{2} \sum_{jk} \frac{\partial}{\partial C_k} \frac{\partial}{\partial C_j} T_{ljk}^{(2)}(C) \right] P(C) dC \quad (3.42)$$

provided that boundary terms can be discarded. This last is unfortunately not always true, most typically when the distribution $P(C)$ has power-law tails as the $|C_j|$ head to infinity. Chapter 6 is devoted to this issue and ways to ensure that these boundary terms are forced to zero using the stochastic gauge technique. Section 6.1.1 derives the exact form of the boundary terms (6.1), which must be zero for the derivation to succeed. Also, there have been developed numerical tests that allow one to check if nonzero boundary terms may be a problem[64].

Now, there may be many $P(C)$ that satisfy (3.42), but *one* certainly is

$$\frac{\partial P(C)}{\partial t} = \sum_l \left[T_l^{(0)}(C) - \sum_j \frac{\partial}{\partial C_j} T_{lj}^{(1)}(C) + \frac{1}{2} \sum_{jk} \frac{\partial}{\partial C_k} \frac{\partial}{\partial C_j} T_{ljk}^{(2)}(C) \right] P(C). \quad (3.43)$$

This is almost an FPE, except for the possible terms constant in P . These typically appear when the normalization of $\hat{\rho}$ is not conserved. It turns out that in many cases the constant terms are zero, and one can directly obtain a Fokker-Planck equation of the form (3.37), noting the correspondences for the drift vector and diffusion matrices

$$A_j(C) = \sum_l T_{lj}^{(1)}(C) \quad (3.44a)$$

$$D_{jk}(C) = \sum_l T_{ljk}^{(2)}(C). \quad (3.44b)$$

Nonzero constant terms can be easily treated using a gauge method described in Section 4.2.

3.4.2 Fokker-Planck equation to stochastic equations

An appropriate set of diffusive stochastic Langevin equations for the variables C_j is known to be equivalent to the FPE for the distribution $P(\{C_j\})$ in the limit of infinitely many samples of the variables C_j . This textbook topic is described in detail e.g. in Gardiner[53, 62]. An important qualification is that this correspondence only holds if the diffusion matrix $D_{jk}(C)$ when written for real variables C_j is *positive semidefinite*⁹.

If this is the case, then Langevin equations for the (real) C_j , equivalent to the FPE (3.37), are

$$dC_j(t) = A_j(C, t) dt + \sum_k B_{jk}(C, t) dW_k(t) \quad (3.45)$$

in the Ito calculus. Here, the *noise matrices* B_{jk} are any matrices that satisfy the diffusion matrix decomposition

$$D = BB^T, \quad (3.46)$$

i.e. in terms of matrix elements

$$D_{jk} = \sum_l B_{jl} B_{kl}. \quad (3.47)$$

⁹i.e. has no eigenvalues with negative real part.

The $dW_k(t)$ are real stochastic Wiener increments (i.e quantities satisfying the large sample ($\mathcal{S} \rightarrow \infty$) averages)

$$\langle dW_j(t) \rangle_{\text{stoch}} = 0 \quad (3.48a)$$

$$\langle dW_j(t) dW_k(t') \rangle_{\text{stoch}} = \delta_{jk} \delta(t - t') dt^2. \quad (3.48b)$$

For the small but finite time steps $\Delta t \rightarrow dt$ in a calculation, $dW_k(t)$ are usually implemented by real Gaussian noises¹⁰ independent for each k and each time step, having mean zero, and variance Δt . Some more detail of the numerical implementation of stochastic equations with Wiener increments can be found in Appendix B. Lastly, it was chosen to write the equations immediately in difference form, as this is the form in which actual computer calculations are made.

Now, clearly all the coefficients A_j and B_{jk} must be real, since the equations are for real variables C_j . However, the decomposition (3.46) in terms of real B is only possible if the symmetric diffusion matrix $D(C)$ is positive semidefinite. This explains the reason for the positive semidefinite condition on D .

Finally, strictly speaking, a more general set of stochastic equations that remains equivalent to the FPE (3.37) is possible, and will be of use in later parts of this thesis. Consider that only the means and variances (3.48) of the Wiener increments are specified. If one writes (3.45) as

$$dC_j(t) = A_j(C, t) dt + dX_j(C, t) \quad (3.49)$$

then by the properties of Ito stochastic calculus, the only binding relationships for the stochastic terms dX_j are that

$$\langle dX_j(t) \rangle_{\text{stoch}} = 0, \quad (3.50a)$$

$$\langle dX_j(t) dX_k(t') \rangle_{\text{stoch}} = \langle D_{jk} \rangle_{\text{stoch}} \delta(t - t') dt^2. \quad (3.50b)$$

Indeed, $dX_j(t)$ need not even be Markovian (i.e. independent of quantities at times $t' \neq t$), provided the conditions (3.50) are satisfied. The relationship between the

¹⁰Other noise distributions are also possible, simply provided the conditions (3.48) are satisfied. Since by the CLT the sum of several infinitesimal noises will always approach an infinitesimal Gaussian-distributed noise anyway, it is usually more efficient to start with a Gaussian noise right away, rather than make one the hard way by summing several non-Gaussian noises.

Fokker-Planck equation (3.37) and stochastic equations (3.49) is considered in a more rigorous fashion and covered exhaustively in the texts by Gardiner[53, 62], but equations (3.45) through to (3.50) will be sufficient for the considerations of this thesis.

3.4.3 Ensuring positivity of diffusion for analytic kernels

In the important case of a kernel that can be written as analytically dependent on complex parameters z_j (rather than on real C_j)¹¹, then we can use the freedom of interpretation of derivatives with respect to complex arguments to choose such a Fokker-Planck equation that the diffusion matrix D is always positive semidefinite, and Langevin stochastic equations (3.45) can always be derived. The procedure described below is simply an application of the same method used for the special case of a positive P distribution by Drummond and Gardiner[11].

This freedom of derivative choice arises from the property of all analytic functions $f(z_j)$ that

$$\frac{\partial f(z_j)}{\partial z_j} = \frac{\partial f(z_j)}{\partial z'_j} = -i \frac{\partial f(z_j)}{\partial z''_j}. \quad (3.51)$$

(Where real and imaginary parts of the variable have been denoted for brevity by the shorthand $z_j = z'_j + iz''_j$.)

The derivation in Section 3.4.1 follows through formally without change if the real variables C_j are replaced with the complex z_j , the set of parameters now being $C = \{z_j\}$. Using (3.51) the terms in first order derivatives of the kernel in (3.40) can now be equated to

$$T_{lj}^{(1)}(C) \frac{\partial}{\partial z_j} \hat{\Lambda}(C) = \left[\text{Re} \left\{ T_{lj}^{(1)}(C) \right\} \frac{\partial}{\partial z'_j} + \text{Im} \left\{ T_{lj}^{(1)}(C) \right\} \frac{\partial}{\partial z''_j} \right] \hat{\Lambda}(C). \quad (3.52)$$

To deal with the second order terms, one can decompose $\sum_l T_{lj}^{(2)} = D_{jk} = \sum_p B_{jp} B_{kp}$ into the noise matrix forms (which will in general be complex), use the shorthand

¹¹i.e. $\hat{\Lambda}$ depends on z_j but not on any z_j^* .

$B = B' + iB''$, and express them (all together) using (3.51) as

$$\begin{aligned} \sum_l T_{ljk}^{(2)}(C) \frac{\partial}{\partial z_j} \frac{\partial}{\partial z_k} \hat{\Lambda}(C) &= \sum_p B_{jp} B_{kp} \frac{\partial}{\partial z_j} \frac{\partial}{\partial z_k} \hat{\Lambda}(C) \\ &= \sum_p \left[B'_{jp} B'_{kp} \frac{\partial}{\partial z'_j} \frac{\partial}{\partial z'_k} + B'_{jp} B''_{kp} \frac{\partial}{\partial z'_j} \frac{\partial}{\partial z''_k} + B''_{jp} B'_{kp} \frac{\partial}{\partial z''_j} \frac{\partial}{\partial z'_k} + B'_{jp} B''_{kp} \frac{\partial}{\partial z''_j} \frac{\partial}{\partial z''_k} \right] \hat{\Lambda}(C). \end{aligned} \quad (3.53)$$

Let us see what has happened to the diffusion matrix for real variables, as it is *this* that must be positive semidefinite for Langevin equations (3.45) to be obtained. While the indices $j, k, p = 1, \dots, N_z$ have referred to complex variables (of number N_z), let us also define real variables $v_{j'}$ (with primed indices $j', k', p' = 1, \dots, 2N_z$) by $z'_j = v_j$ and $z''_j = v_{(j+N_z)}$. Denoting the diffusion and noise matrices of the real variables $v_{j'}$ as $D^{(v)}$ and $B^{(v)}$, respectively, we have $D^{(v)}_{j',k'} = \sum_{p'=1}^{2N_z} B^{(v)}_{j',p'} B^{(v)}_{k',p'}$, and from (3.53):

$$B^{(v)}_{j,p} = \text{Re} \{ B_{jp} \}, \quad (3.54a)$$

$$B^{(v)}_{j+N_z,p} = \text{Im} \{ B_{jp} \}, \quad (3.54b)$$

$$B^{(v)}_{j,p+N_z} = B^{(v)}_{j+N_z,p+N_z} = 0, \quad (3.54c)$$

which are all explicitly real. Thus, the diffusion matrix is explicitly positive semidefinite, and Langevin equations for all the $v_{j'}$ can be obtained.

From (3.45), (3.52), (3.44a), and (3.54) the resulting stochastic equations for the real variables then are

$$dv_j(t) = \text{Re} \{ A_j(C, t) \} dt + \sum_k B^{(v)}_{j,k}(C, t) dW_k(t) \quad (3.55a)$$

$$dv_{j+N_z}(t) = \text{Im} \{ A_j(C, t) \} dt + \sum_k B^{(v)}_{j+N_z,k}(C, t) dW_k(t), \quad (3.55b)$$

which can be written in shorthand form for the complex variables z_j as

$$dz_j(t) = A_j(C, t) dt + \sum_k B_{jk}(C, t) dW_k(t). \quad (3.56)$$

Note that the Wiener increments dW_k remain real.

The only requirement (apart from obtaining an FPE) to carry out this procedure, which ensures a stochastic equation interpretation, is that the kernel be written as an analytic function of complex variables.

Finally, the procedure described in this section can be carried out on only a part of the total set of variables C , if the kernel can be made analytic in only this part. Provided the resulting combined noise matrix $B^{(v)}$ for all the real variables of the system (real and imaginary parts of the complex variables, and the remaining real variables) is real, the diffusion matrix is positive semidefinite, and Langevin equations for all the variables can still be obtained in the same way as here.

3.4.4 Positive P example: dilute lattice Bose gas equations

Continuing with the positive P representation example, let us obtain the stochastic equations corresponding to the quantum dynamics of the lattice interacting Bose gas Hamiltonian (2.17) with single-particle losses to a zero temperature heat bath. The master equation is (2.20) with Linblad operators (2.21), and can be written in terms of $\hat{\rho}$ and the creation and destruction operators $\hat{a}_{\mathbf{n}}, \hat{a}_{\mathbf{n}}^\dagger$, which can be identified with the \hat{A}_j of Section 3.4.1.

These can be easily verified (via (3.9) and (3.26), and the mutual commutation of local kernels $\hat{\Lambda}_{\mathbf{n}}$) to obey the identities

$$\hat{a}_{\mathbf{n}}\hat{\Lambda} = \alpha_{\mathbf{n}}\hat{\Lambda}, \quad (3.57a)$$

$$\hat{a}_{\mathbf{n}}^\dagger\hat{\Lambda} = \left(\beta_{\mathbf{n}} + \frac{\partial}{\partial\alpha_{\mathbf{n}}}\right)\hat{\Lambda}, \quad (3.57b)$$

$$\hat{\Lambda}\hat{a}_{\mathbf{n}}^\dagger = \beta_{\mathbf{n}}\hat{\Lambda}, \quad (3.57c)$$

$$\hat{\Lambda}\hat{a}_{\mathbf{n}} = \left(\alpha_{\mathbf{n}} + \frac{\partial}{\partial\beta_{\mathbf{n}}}\right)\hat{\Lambda} \quad (3.57d)$$

for all N lattice points \mathbf{n} . It can be seen that to obtain an FPE there can be up to two operations of the kind $\hat{a}^\dagger\hat{\rho}$ or $\hat{\rho}\hat{a}$ per term in the master equation, and any number of $\hat{a}\hat{\rho}$ or $\hat{\rho}\hat{a}^\dagger$. This allows any one- or two-particle processes but rules out three-particle processes, which contain terms of third order in \hat{a} or \hat{a}^\dagger in the Hamiltonian.

The terms in the master equation can be sorted (according to processes) as

$$\mathcal{T}_1[\hat{\rho}] = -i \sum_{\mathbf{nm}} \omega_{\mathbf{nm}} [\hat{a}_{\mathbf{n}}^\dagger \hat{a}_{\mathbf{m}} \hat{\rho} - \hat{\rho} \hat{a}_{\mathbf{n}}^\dagger \hat{a}_{\mathbf{m}}], \quad (3.58a)$$

$$\mathcal{T}_2[\hat{\rho}] = -i\chi \sum_{\mathbf{n}} [\hat{a}_{\mathbf{n}}^\dagger{}^2 \hat{a}_{\mathbf{n}}^2 \hat{\rho} - \hat{\rho} \hat{a}_{\mathbf{n}}^\dagger{}^2 \hat{a}_{\mathbf{n}}^2], \quad (3.58b)$$

$$\mathcal{T}_3[\hat{\rho}] = \sum_{\mathbf{n}} \gamma_{\mathbf{n}} \left[\hat{a}_{\mathbf{n}} \hat{\rho} \hat{a}_{\mathbf{n}}^\dagger - \frac{1}{2} \hat{a}_{\mathbf{n}}^\dagger \hat{a}_{\mathbf{n}} \hat{\rho} - \frac{1}{2} \hat{\rho} \hat{a}_{\mathbf{n}}^\dagger \hat{a}_{\mathbf{n}} \right]. \quad (3.58c)$$

Which, using (3.57), gives

$$\mathcal{T}_1[\hat{\Lambda}] = -i \sum_{\mathbf{nm}} \omega_{\mathbf{nm}} \left(\alpha_{\mathbf{m}} \frac{\partial}{\partial \alpha_{\mathbf{n}}} - \beta_{\mathbf{n}} \frac{\partial}{\partial \beta_{\mathbf{m}}} \right) \hat{\Lambda}, \quad (3.59a)$$

$$\mathcal{T}_2[\hat{\Lambda}] = -i\chi \sum_{\mathbf{n}} \left(2\alpha_{\mathbf{n}}^2 \beta_{\mathbf{n}} \frac{\partial}{\partial \alpha_{\mathbf{n}}} - 2\alpha_{\mathbf{n}} \beta_{\mathbf{n}}^2 \frac{\partial}{\partial \beta_{\mathbf{n}}} + \alpha_{\mathbf{n}}^2 \frac{\partial^2}{\partial \alpha_{\mathbf{n}}^2} - \beta_{\mathbf{n}}^2 \frac{\partial^2}{\partial \beta_{\mathbf{n}}^2} \right) \hat{\Lambda}, \quad (3.59b)$$

$$\mathcal{T}_3[\hat{\Lambda}] = -\frac{1}{2} \sum_{\mathbf{n}} \gamma_{\mathbf{n}} \left(\alpha_{\mathbf{n}} \frac{\partial}{\partial \alpha_{\mathbf{n}}} + \beta_{\mathbf{n}} \frac{\partial}{\partial \beta_{\mathbf{n}}} \right) \hat{\Lambda}. \quad (3.59c)$$

Proceeding in the same manner as in Section 3.4.3 (remember that the $\alpha_{\mathbf{n}}$ and $\beta_{\mathbf{n}}$ variables are complex), we obtain the (Ito) stochastic equations

$$d\alpha_{\mathbf{n}} = -i \sum_{\mathbf{m}} \omega_{\mathbf{nm}} \alpha_{\mathbf{m}} dt - 2i\chi \alpha_{\mathbf{n}}^2 \beta_{\mathbf{n}} dt - \frac{\gamma_{\mathbf{n}}}{2} \alpha_{\mathbf{n}} dt + i\alpha_{\mathbf{n}} \sqrt{2i\chi} dW_{\mathbf{n}}(t), \quad (3.60a)$$

$$d\beta_{\mathbf{n}} = i \sum_{\mathbf{m}} \omega_{\mathbf{mn}} \beta_{\mathbf{m}} dt + 2i\chi \alpha_{\mathbf{n}} \beta_{\mathbf{n}}^2 dt - \frac{\gamma_{\mathbf{n}}}{2} \beta_{\mathbf{n}} dt + \beta_{\mathbf{n}} \sqrt{2i\chi} d\widetilde{W}_{\mathbf{n}}(t). \quad (3.60b)$$

Here the $2N$ real Wiener increments $dW_{\mathbf{n}}(t)$ and $d\widetilde{W}_{\mathbf{n}}(t)$ are all independent at each time step. We have used the “simplest” diagonal noise matrix decomposition $B_{\mathbf{nm}} = B_{\mathbf{nn}} \delta_{\mathbf{nm}}$ here. More on the possibilities with these decompositions in Chapter 4.

3.5 Convenience for parallel computation

There is no cross-talk between separate realizations of the stochastic process in the simulation scheme described in this chapter. This is a highly desirable property for numerical realization because only one trajectory at a time need be stored in memory. In fact, many independent trajectories could be run at the same time on different computers in a cluster, or even on entirely separate computers, and then collected together at the end. The same applies to calculating observable averages

(3.14), where all one needs is to keep track of the running sum of averaged quantities and their number¹².

In combination with the linear scaling of variable number this makes for a potentially powerful and efficient combination for many-body simulations.

An exception to this convenient scaling with trajectory number are non-observable quantities such as fidelity, estimated by methods like in Section 3.3.3. In that kind of calculation we have to keep in memory the variables for all s samples in a subensemble to use all pairs of them in the calculation (3.22). Memory requirements increase by a factor of s to hold the entire subensemble. The actual calculation of estimates like (3.22) also takes a time $\propto s^2$ as opposed to $\propto s$ for static observable estimations (3.14). Since for reasonable subensemble averages one typically needs $\mathcal{O}(100)$ or more samples to a subensemble, this requires an increase of two orders of magnitude in computing resources — highly nontrivial. Fortunately, at least the linear scaling $\propto N$ of the number of variables remains untouched.

3.6 Summary of representation requirements

This chapter has shown or strongly indicated that for a representation to be useful for quantum simulations of many-body systems, it must satisfy the requirements listed below:

1. **Non-negative real distribution:** To interpret $P(C)$ as a probability distribution over kernel operators $\hat{\Lambda}(C)$, it must be real, normalizable to unity, and non-negative.
2. **Non-singular distribution:** Needed for finite probabilities and un-biased dynamic sampling, although delta functions in the initial state can be tolerated.
3. **Complete:** The representation must be able to describe the initial states, and any states reachable by the subsequent evolution, while satisfying the

¹²And possibly the sum of the squares of the averaged properties for assessment of uncertainty in the estimate as per (3.18).

above conditions. In practice this usually implies that the kernel must form a complete or overcomplete basis, and that the kernel be off-diagonal. See Section 3.2.3.

4. **Linear scaling:** The number of system variables in C must scale linearly with the system size N (Typically number of modes, particles, orbitals, ...). In practice this requires the kernel to be a tensor product over local kernels $\widehat{\Lambda}_k$ for each subsystem, or a linear combination of a small ($\ll N$) number of such tensor product terms. See Section 3.2.2.
5. **Most samples significant:** The majority of samples (i.e. $\mathcal{O}(\mathcal{S})$) should contribute significantly to all observable estimates of interest, particularly observables local to the subsystems. This appears to require the use of an overcomplete basis, and kernels normalized locally (i.e. $\text{Tr} [\widehat{\Lambda}_k] = 1$), perhaps apart from some global multiplying functions. These issues are discussed in Section 3.3.4.
6. **Finite kernel trace:** All configurations C must lead to a finite kernel trace, otherwise observable averages will be undefined for $\mathcal{S} \rightarrow \infty$ and biased for finite \mathcal{S} due to divergence of the denominator in (3.14).

If the stochastic interpretation of the quantum evolution is obtained via a Fokker-Planck equation as in Section 3.4 (presently the only approach that has led to nontrivial simulations), then the following requirements must also be satisfied:

7. **Differential equation:** The configuration parameters C should be continuous to allow a differential equation for $P(C)$.
8. **Fokker Planck equation:** The mapping $\frac{\partial \widehat{\rho}}{\partial t} \rightarrow \frac{\partial P(C)}{\partial t}$ must generate only first and second order partial derivatives of $P(C)$. See Section 3.4.1.
9. **Positive propagator:** The short time diffusion in phase space must be non-negative (i.e. the diffusion matrix for real variables $D^{(v)}$ has no eigenvalues with negative real part), to allow a stochastic interpretation. This can be

guaranteed if the kernel is an analytic function of complex variables (see Section 3.4.3), although in some situations non-analytic kernels may also lead to satisfactory diffusion.

Furthermore, there are several additional requirements that appear generically in simulations, and must be satisfied to avoid bias and/or unmanageable noise in the observable estimates.

10. **Stable trajectories:** If trajectories in phase space diverge rapidly, so does the width of the distribution $P(C)$, and all useful precision in the estimates of observables is lost.
11. **Unbiased:** Pathological cases are common (particularly in nonlinear underdamped systems), where the distribution $P(C)$ is too broad to allow unbiased sampling of some or all observable moments — typically when $P(C)$ develops power-law tails. These “boundary term” errors are considered in detail in Chapter 6, and usually go in tandem with unstable trajectories.
12. **UV convergent:** Numerical simulations almost always require one to approximate space by a finite lattice. This does not change the physical predictions of the model provided the lattice spacing is much finer than any other significant length scale in the model. This, in turn, requires that the effect of vacuum fluctuations on the equations of motion and observable estimates abates at large momentum (i.e. at small length scales). A mode-based lattice calculation will typically be non-UV-convergent if there are noise terms in the equations that do not disappear at zero mode occupation.

Chapter 4

Stochastic gauges

The non-uniqueness of representations $P(C)$ leads to many potential sets of stochastic equations representing the same single master equation. In this chapter, the stochastic gauge formalism is introduced, which systematically characterizes the freedoms available in the stochastic equations after the physical model has been completely specified. The term “stochastic gauges” is used here because of an analogy with electromagnetic gauges: In both cases the equations contain arbitrary (in principle) functions, which have no effect on physical observables, but can have a great influence on the ease with which the calculation proceeds. In the case of stochastic gauges, the choice of gauge has no effect¹ on observable quantum expectation values (3.14) calculated in the limit of an infinite number of samples. For a finite number of samples, however, different gauges affect how rapidly this large sample limit is approached. It will be shown that an appropriate gauge choice can improve simulation efficiency by many orders of magnitude, and also correct biases caused by pathological distributions.

During the derivation of stochastic Langevin equations (3.45) for the variables C_j from a master equation for the density matrix, there are two distinct places where stochastic gauges can be introduced. Firstly, null differential identities on the kernel $\hat{\Lambda}$, multiplied by arbitrary functions \mathcal{F} and integrated over $P(C)$ can be introduced into the master equation at the point (3.41), which leads to \mathcal{F} finding its way into the

¹In principle. In practice there is also the issue of boundary term errors, which may be present or not, depending on the gauge chosen. This is discussed in detail in Chapter 6.

FPE, and into the final stochastic equations. These “kernel stochastic gauges” are described in detail in Sections 4.1 to 4.3. Secondly, the diffusion matrix of an FPE does not uniquely specify the noise behavior of the resulting stochastic equations. These degrees of freedom can also be used to introduce arbitrary functions g into the noise matrices B , as described in detail in Section 4.4. A partly restricted set of kernel and diffusion gauges, which will be here dubbed the “standard” gauges, and are sufficient for most purposes, are summarized in Section 4.5.

As in the previous Chapter 3, the derivation is kept as general as possible with the aim of sifting out those elements that are truly necessary for stochastic gauges to be present and to be useful. Several authors have recently (all starting in 2001) proposed phase-space distribution methods that make use of the freedoms that are here dubbed stochastic gauges. This includes the “noise optimization” of Plimak, Olsen, and Collett[2], the “stochastic wavefunction” method of Carusotto, Castin, and Dalibard[1, 65], and the “stochastic gauges” of Deuar, Drummond, and Kheruntsyan[3, 66, 56, 61]. All these are unified within the formalism presented in this chapter, which is itself a generalization of some ideas to be found in the articles by Deuar and Drummond[3] and [56].

4.1 Generalized kernel stochastic gauges

Whenever there are null differential identities on the kernel of the general form

$$\left\{ J^{(0)}(C) + \sum_j J_j^{(1)}(C) \frac{\partial}{\partial C_j} + \frac{1}{2} \sum_{jk} J_{jk}^{(2)}(C) \frac{\partial^2}{\partial C_j \partial C_k} \right\} \hat{\Lambda}(C) = \mathcal{J} [\hat{\Lambda}(C)] = 0, \quad (4.1)$$

a stochastic gauge can be introduced in the following manner:

Since (4.1) is zero, so is its integral when multiplied by non-divergent functions. In particular,

$$\int P(C, t) \mathcal{F}(C, t) \mathcal{J} [\hat{\Lambda}(C)] dC = 0 \quad (4.2)$$

with the function $\mathcal{F}(C, t)$ otherwise unspecified. Now one adds zero (in the form (4.2)) to the right hand side of the master equation (3.41) in the derivation of the Fokker-Planck equation of Section 3.4.1. (Actually, to be fully rigorous, one must

make sure that the full integral $\int P(\cdot)\widehat{\Lambda}$ over the entire master equation (3.41) plus all “zero” terms (4.2) is convergent, otherwise boundary term errors may result. This issue is considered in detail in Chapter 6, so for now let us defer this and assume that no boundary term errors are present.)

Adding (4.2) introduces an additional master equation term that can be written in the notation of Section 3.4.1 as

$$\mathcal{T}_{\mathcal{J}} [\widehat{\Lambda}] = \mathcal{F}\mathcal{J} [\widehat{\Lambda}]. \quad (4.3)$$

Following now the procedure of Sections 3.4.1 and (3.4.2), the result is that the coefficients in the Langevin equations (3.45) become modified to

$$A_j(C) \rightarrow A_j^{(\mathcal{F})}(C) = A_j(C) + \mathcal{F}(C)J_j^{(1)}(C) \quad (4.4a)$$

$$D_{jk}(C) \rightarrow D_{jk}^{(\mathcal{F})}(C) = D_{jk}(C) + \mathcal{F}(C)J_{jk}^{(2)}(C). \quad (4.4b)$$

In this manner an arbitrary function \mathcal{F} has been introduced into the stochastic equations. The number of these is unlimited provided we have at least one identity of the form (4.1), although there will be at most as many gauges with distinct effects on the equations as we have distinct identities.

The whole procedure is based on properties (4.1) of the kernel, and so the kernel gauges that can be used in a particular simulation will depend on the local subsystem bases used to expand the density matrix, and on how the kernel is constructed out of these.

Some (though certainly not all) broad classes of gauges that are possible with various kernels include:

- **ANALYTIC:** The procedure used in Section 3.4.3 to ensure a positive diffusive propagator for representations based on kernels analytic in complex variables z_j can be interpreted as using a gauge of the form

$$\left(\frac{\partial}{\partial \text{Re}\{z_j\}} + i \frac{\partial}{\partial \text{Im}\{z_j\}} \right) \widehat{\Lambda}(z_j) = 0. \quad (4.5)$$

This applies for any kernel analytic in z_j .

- **WEIGHTING:** If constant terms are present during the derivation of an FPE as in (3.43) in Section 3.4.1, they can be converted to deterministic evolution

of global weights, provided the kernel contains a global factor Ω , using an identity of the form

$$\left(\Omega \frac{\partial}{\partial \Omega} - 1\right) \widehat{\Lambda} = 0. \quad (4.6)$$

See Section 4.2, below.

- **DRIFT:** One can modify the deterministic part of the stochastic equations (3.45) in principle at will, provided compensation is made in the form of appropriate global trajectory weights. Building on the weighting gauge identity (4.6) for kernels with global weights Ω , the identities

$$\left(\Omega \frac{\partial}{\partial \Omega} - 1\right) \frac{\partial}{\partial C_j} \widehat{\Lambda} = 0. \quad (4.7)$$

follow for any variable C_j , and can be used for this purpose. See Section 4.3.

- **REDUCTION:** For some kernels, identities can be found that reduce second order partial derivatives of the kernel to first order. These can be used at the level of expressions (3.40) for each part \mathcal{T}_l of the master equation to reduce partial derivative terms obtained with the operator identities (3.39) to lower order. This will result in a change from diffusion behavior in the FPE to deterministic drift, or even better, in transformation of third- or higher-order partial derivative terms (which prevent one from obtaining any FPE) to lower-order terms, which can form part of an FPE. The identities required are of the general form

$$\left(\frac{\partial}{\partial C_j} \frac{\partial}{\partial C_k} - f(C) \frac{\partial}{\partial C_p}\right) \widehat{\Lambda} = 0. \quad (4.8)$$

for some variables C_j , C_k , and C_p , and some function $f(C)$. An example can be found in the recent work of Drummond and Corney on boson and fermion phase-space distributions[67, 68]

4.2 Weighting stochastic gauges

Consider the situation where during the derivation of the FPE in Section 3.4.1, one encounters nonzero constant terms $\sum_l T_l^{(0)}(C)P(C)$. Most typically, this situation

occurs with un-normalized density matrices e.g. in the thermodynamic equation (2.29). As it stands, such an equation of the form (3.43) does not lead immediately to stochastic equations for the variables C . However, if the kernel contains a global weight factor Ω , or by a simple modification is given one, this constant term can be converted to a deterministic evolution of this weight.

What is needed is a kernel of the form

$$\widehat{\Lambda}(C) = \Omega \widehat{\underline{\Lambda}}(\underline{C}). \quad (4.9)$$

where the “base” kernel $\widehat{\underline{\Lambda}}(\underline{C})$, containing all the local basis operators for subsystems does not depend on the global weight Ω . For notational convenience let us define the logarithm of the weight by

$$\Omega = e^{C_0}, \quad (4.10)$$

then the full parameter set is $C = \{C_0, \underline{C}\}$. Directly from (4.9), one can see that the gauge identity

$$\left(\frac{\partial}{\partial C_0} - 1 \right) \widehat{\Lambda}(C) = 0. \quad (4.11)$$

applies. If, now, one adds “zero” defined as

$$\sum_l \int P(C) T_l^{(0)}(\underline{C}) \left(\frac{\partial}{\partial C_0} - 1 \right) \widehat{\Lambda}(C) dC = 0 \quad (4.12)$$

to the master equation in the form (3.41), the original terms in $T_l^{(0)}$ vanish, to be replaced by first order differential terms

$$\sum_l \int P(C) T_l^{(0)}(\underline{C}) \frac{\partial}{\partial C_0} \widehat{\Lambda}(C) dC. \quad (4.13)$$

This can be incorporated into the formalism of the derivation in Section 3.4.1 by including labels $j, k = 0$ in the sums \sum_j, \sum_k , and noting that now

$$T_{l0}^{(1)}(\underline{C}) = T_l^{(0)}(\underline{C}) \quad (4.14a)$$

$$T_{l0k}^{(2)} = T_{lj0}^{(2)} = T_{l00}^{(2)} = 0. \quad (4.14b)$$

This then follows through to additional drift and diffusion coefficients (3.44) in the FPE (3.37)

$$A_0(\underline{C}) = \sum_l T_l^{(0)}(\underline{C}) \quad (4.15)$$

$$D_{j0} = D_{0j} = D_{00} = 0. \quad (4.16)$$

The new equation for $C_0 = \log \Omega$ becomes

$$dC_0(t) = A_0(\underline{C}, t) dt \quad (4.17)$$

(with no noise), while the stochastic equations for all the remaining variables C_j remain unchanged.

Lastly, the global weight can be made complex with no formal change to the above, if one needs to deal with complex constant terms $T_l^{(0)}$.

4.3 Drift stochastic gauges

4.3.1 Mechanism

Consider again a kernel with global weight $\Omega = e^{C_0}$, defined as (4.9) in the previous section. The weighting gauge identity (4.11) implies further gauge identities involving each of the variables in C (including C_0 itself):

$$0 = \left(\frac{\partial}{\partial C_0} - 1 \right) \frac{\partial}{\partial C_j} \hat{\Lambda}(C) \quad (4.18a)$$

$$= \left(\frac{1}{2} \frac{\partial}{\partial C_0} \frac{\partial}{\partial C_j} + \frac{1}{2} \frac{\partial}{\partial C_j} \frac{\partial}{\partial C_0} - \frac{\partial}{\partial C_j} \right) \hat{\Lambda}(C) \quad (4.18b)$$

$$= \mathcal{J}_j \left[\hat{\Lambda}(C) \right]. \quad (4.18c)$$

While (4.11) was used to convert constant terms in the ‘‘FPE-like’’ expression (3.43) to first order derivative terms (and hence deterministic evolution in the weight), the new identities can be used to convert first to second order terms. This allows conversion of deterministic evolution in a variable C_j to stochastic changes in the weight. Such a conversion can be highly desirable as will be shown in Chapters 6 and later.

The procedure is similar to that employed in Section 4.2, however some modifications are made to make the final expression for the stochastic equations in C_j more revealing. If one adds ‘‘zero’’, defined this time as

$$\sum_j \int P(C) \mathcal{F}_j(C) \mathcal{J}_j \left[\hat{\Lambda}(C) \right] dC = 0, \quad (4.19)$$

with arbitrary gauges \mathcal{F}_j for each variable C_j , then in the formalism of Section 3.4.1 additional terms $\int P(C) \mathcal{T}_0 \left[\widehat{\Lambda}(C) \right] dC = 0$ are introduced into the master equation, with the superoperator \mathcal{T}_0 containing coefficients

$$T_{0j}^{(1)}(C) = -\mathcal{F}_j(C) \quad (4.20a)$$

$$T_{0\underline{j}0}^{(2)}(C) = T_{00\underline{j}}^{(2)}(C) = \mathcal{F}_{\underline{j}}(C) \quad (4.20b)$$

$$T_{000}^{(2)}(C) = 2\mathcal{F}_0(C), \quad (4.20c)$$

where underlined indices have been used to indicate labels for the “base” variables such that $\underline{j} = 1, 2, \dots$, etc. Following the procedure through, the modified FPE coefficients

$$A_j = \underline{A}_j - \mathcal{F}_j \quad (4.21a)$$

$$D_{\underline{j}k} = \underline{D}_{\underline{j}k} \quad (4.21b)$$

$$D_{\underline{j}0} = \underline{D}_{\underline{j}0} + \mathcal{F}_{\underline{j}} = \underline{D}_{0\underline{j}} \quad (4.21c)$$

$$D_{00} = 2\mathcal{F}_0 \quad (4.21d)$$

are obtained (the original coefficients when $\mathcal{F}_j = 0$ are the underlined). The modifications to the deterministic evolution of variables are seen to be $-\mathcal{F}_j dt$.

A difficulty with the forms (4.21) is that in general the form of the diffusion matrix elements B_{jk} that actually go into the stochastic equations will depend in some complicated manner on all the \mathcal{F}_j . If the aim is to modify only the deterministic (drift) evolution of the variables $C_{\underline{j}}$ (as will be the case in later chapters), one may end up with the side-effect of additional changes in the noises $B_{\underline{j}k} dW_k$. This can prevent a clear assessment of what practical effect the change in the equations due a particular gauge will have.

A convenient form to have the noise matrices B in would be to have no change in the base variables $C_{\underline{j}}$ apart from the modification of the drift, and define new (also arbitrary) gauge functions $\mathcal{G}_{\underline{j}}(\mathcal{F})$ that give the exact stochastic terms for the weight variable C_0 , and do not introduce any more independent noises than were had in the original formulation. That is, if the prior weight evolution had no diffusion, one

can try the ansatz

$$B_{\underline{j}\underline{k}} = \underline{B}_{\underline{j}\underline{k}} \quad (4.22a)$$

$$B_{0\underline{k}} = \underline{\mathcal{G}}_{\underline{k}}. \quad (4.22b)$$

Using $D = BB^T$, the diffusion matrix elements that would result from (4.22) are

$$D_{\underline{j}\underline{k}} = \sum_{\underline{p}} \underline{B}_{\underline{j}\underline{p}} \underline{B}_{\underline{k}\underline{p}} = \underline{D}_{\underline{j}\underline{k}} \quad (4.23a)$$

$$D_{\underline{j}0} = \sum_{\underline{p}} \underline{B}_{\underline{j}\underline{p}} \underline{\mathcal{G}}_{\underline{p}} = D_{0\underline{j}} \quad (4.23b)$$

$$D_{00} = \sum_{\underline{p}} \underline{\mathcal{G}}_{\underline{p}}^2. \quad (4.23c)$$

Comparing these with (4.21), one identifies

$$\mathcal{F}_0 = \frac{1}{2} \sum_{\underline{p}} \underline{\mathcal{G}}_{\underline{p}}^2 \quad (4.24a)$$

$$\mathcal{F}_{\underline{j}} = \sum_{\underline{p}} \underline{B}_{\underline{j}\underline{p}} \underline{\mathcal{G}}_{\underline{p}}. \quad (4.24b)$$

Note that there has to be a small restriction of the gauge freedoms from (4.21) to (4.24) (one less arbitrary function due to (4.24a)) to achieve the convenient noise matrix B given by the ansatz (4.22).

The final stochastic Langevin equations with gauges $\underline{\mathcal{G}}_{\underline{k}}$ then are (noting that $\underline{B}_{\underline{j}\underline{k}} = B_{\underline{j}\underline{k}}$)

$$dC_{\underline{j}} = \underline{A}_{\underline{j}} dt + \sum_{\underline{k}} \underline{B}_{\underline{j}\underline{k}} (dW_{\underline{k}} - \underline{\mathcal{G}}_{\underline{k}} dt), \quad (4.25a)$$

$$dC_0 = \underline{A}_0 dt + \sum_{\underline{k}} \underline{\mathcal{G}}_{\underline{k}} \left(dW_{\underline{k}} - \frac{1}{2} \underline{\mathcal{G}}_{\underline{k}} dt \right). \quad (4.25b)$$

The (Ito) equation for the actual weight $\Omega = e^{C_0}$ takes on a simpler form

$$d\Omega = \Omega \left\{ \underline{A}_o dt + \sum_{\underline{k}} \underline{\mathcal{G}}_{\underline{k}} dW_{\underline{k}} \right\}. \quad (4.26)$$

with no new drift terms. Recapping, several assumptions beyond the identities (4.18) have been made to obtain (4.25), and these are:

1. The integral $\int P(\cdot)\widehat{\Lambda}dC$ for the master equation in form (3.41), including terms containing the \mathcal{F}_j , is convergent.
2. The function \mathcal{F}_0 is determined by (4.24a), rather than being an independent arbitrary function.
3. The prior evolution of the weight contains no stochastic terms ($\underline{D}_{j0} = \underline{D}_{00} = 0$).

Lastly, as with weighting gauges, if one is dealing with a kernel analytic in constant variables z_j , the above derivation goes through for z_j instead of C_j with no formal change, provided the weight variable $C_0 = z_0 = \log \Omega$ is now also complex. The gauges \mathcal{G}_k would then be arbitrary *complex* functions.

4.3.2 Real drift gauges, and their conceptual basis

If one arbitrarily adjusts the drift behavior of the stochastic equations, there must be some kind of compensation to ensure continued correspondence to the original quantum master equation. When the gauges \mathcal{G}_j are real, the modification of the weight can be understood in terms of compensation for the increments $\underline{B}_{jk}(dW_k - \mathcal{G}_k)$ no longer being described a Gaussian noise of mean zero (in the many infinitesimal timesteps limit when the CLT applies). Let us first look at the simple case of one subsystem with one complex variable z , with the gauge-less weight is constant, and the drift gauge \mathcal{G} real. The equations for the real and imaginary parts of the variables $z = z' + iz''$ and $z_0 = \log \Omega = z'_0 + iz''_0$ are

$$dz' = \underline{A}'dt + \underline{B}'dW - \underline{B}'\mathcal{G}dt \quad (4.27a)$$

$$dz'' = \underline{A}''dt + \underline{B}''dW - \underline{B}''\mathcal{G}dt \quad (4.27b)$$

$$dz'_0 = -\frac{1}{2}\mathcal{G}^2dt + \mathcal{G}dW \quad (4.27c)$$

$$dz''_0 = 0, \quad (4.27d)$$

where we have used the shorthand $\underline{B}' = \text{Re}\{\underline{B}\}$ and $\underline{B}'' = \text{Im}\{\underline{B}\}$ etc. for all quantities, and omitted writing subscripts $j, k = 1$.

Consider a very small time step dt in which the variables change from $z(t)$ to $z(t + dt)$ etc. This time step is small enough so that the stochastic equations for

infinitesimal dt apply, while large enough that the Wiener increment dW is Gaussian distributed due to the CLT. Suppose also that we consider just one particular trajectory so that the initial distributions of z and z_0 are delta functions around their initial values. The probability distribution of the Gaussian noise is

$$\Pr(dW) = e^{-dW^2/2dt}, \quad (4.28)$$

and if there is no gauge ($\mathcal{G} = 0$), the dependence of $z'(t + dt)$ on the realization of the noise dW can be inverted to give

$$dW = \frac{z'(t + dt) - z'(t) - \underline{A}'dt}{\underline{B}'}. \quad (4.29)$$

Substituting this into (4.28), we obtain an expression for the probability distribution² of $z(t + dt)$:

$$P(z(t + dt)) = \exp \left[-\frac{(z'(t + dt) - z'(t) - \underline{A}'dt)^2}{2(\underline{B}')^2dt} \right]. \quad (4.30)$$

If we now arbitrarily modify the drift with a nonzero gauge, (4.28) still applies, but the inverted relation $dW(z'(t + dt))$ now becomes

$$dW = \frac{z'(t + dt) - z'(t) - \underline{A}'dt + \mathcal{G}\underline{B}'dt}{\underline{B}'}, \quad (4.31)$$

leading to a different probability distribution $P_{\mathcal{G}}(z(t + dt))$ than the “correct” gaugeless (4.30). Precisely:

$$P_{\mathcal{G}}(z(t + dt)) = \exp \left[-\frac{(z'(t + dt) - z'(t) - \underline{A}'dt + \mathcal{G}\underline{B}'dt)^2}{2(\underline{B}')^2dt} \right] \quad (4.32)$$

$$= P(z(t + dt)) \exp \left[-\frac{\mathcal{G}}{2\underline{B}'} \{ \mathcal{G}\underline{B}'dt + 2(z'(t + dt) - z'(t) - \underline{A}'dt) \} \right]. \quad (4.33)$$

However, we can recover from this by introducing a global compensating weight Ω for the trajectory, which is always applied in all observable averages so that

$$\Omega P_{\mathcal{G}}(z(t + dt)) = P(z(t + dt)). \quad (4.34)$$

²Note that we use only expressions involving z' to characterize the probability distribution of the entire complex variable z . This is because z' and z'' are completely correlated with each other, since only the one real noise dW determines them both. The imaginary part, z'' , could have been used just as well.

Using (4.31)³ we can write this compensating weight from (4.33) as

$$\Omega = \exp \left[\mathcal{G} \left(-\frac{1}{2} \mathcal{G} dt + dW \right) \right]. \quad (4.35)$$

Identifying Ω from the definition of the kernel (4.9) one sees that the drift gauge equations for dz_0 (4.27c) and (4.27d) produce exactly the right global weight to compensate for the arbitrary drift corrections introduced by nonzero \mathcal{G} .

For the many-variable case with N_z base variables z_j plus the global complex log-weight z_0 , the gauged stochastic equations (4.25) can be written in vector form as

$$dz = \underline{\mathbf{A}} dt + \underline{\mathbf{B}} d\mathbf{W} - \underline{\mathbf{B}} \mathcal{G} dt \quad (4.36a)$$

$$dz_0 = -\frac{1}{2} \mathcal{G}^T \mathcal{G} dt + \mathcal{G}^T d\mathbf{W}, \quad (4.36b)$$

where all bold quantities denote column vectors of N_z elements labeled by $j = 1, \dots, N_z$ etc. The probability distribution of a noise realization is

$$\Pr(d\mathbf{W}) = \exp \left(-\frac{d\mathbf{W}^T d\mathbf{W}}{2dt} \right), \quad (4.37)$$

since all noise elements dW_j are independent. With no gauge, $d\mathbf{W} = (\underline{\mathbf{B}}')^{-1} [dz' - \underline{\mathbf{A}}' dt]$, leading to the probability distribution

$$P(\mathbf{z}(t+dt)) = \exp \left[-\frac{1}{2dt} (dz' - \underline{\mathbf{A}}' dt)^T [(\underline{\mathbf{B}}')^{-1}]^T (\underline{\mathbf{B}}')^{-1} (dz' - \underline{\mathbf{A}}' dt) \right], \quad (4.38)$$

while with a gauge the noise can be written

$$d\mathbf{W} = (\underline{\mathbf{B}}')^{-1} [dz' - \underline{\mathbf{A}}' dt + \underline{\mathbf{B}}' \mathcal{G} dt]. \quad (4.39)$$

This leads to a new probability distribution

$$\begin{aligned} P_{\mathcal{G}}(\mathbf{z}(t+dt)) &= \exp \left[-\frac{(dz' - \underline{\mathbf{A}}' dt + \underline{\mathbf{B}}' \mathcal{G} dt)^T [(\underline{\mathbf{B}}')^{-1}]^T (\underline{\mathbf{B}}')^{-1} (dz' - \underline{\mathbf{A}}' dt + \underline{\mathbf{B}}' \mathcal{G} dt)}{2dt} \right] \\ &= P_{\mathcal{G}}(\mathbf{z}(t+dt)) \\ &\quad \times \exp \left[-\frac{\mathcal{G}^T (\underline{\mathbf{B}}')^{-1} (dz' - \underline{\mathbf{A}}' dt) + (dz' - \underline{\mathbf{A}}' dt)^T [(\underline{\mathbf{B}}')^{-1}]^T \mathcal{G} + \mathcal{G}^T \mathcal{G}}{2} \right] \\ &= P_{\mathcal{G}}(\mathbf{z}(t+dt)) \exp \left[-\mathcal{G}^T d\mathbf{W} + \frac{1}{2} \mathcal{G}^T \mathcal{G} dt \right] \end{aligned} \quad (4.40)$$

³Not (4.29), since that applied only for $\mathcal{G} = 0$.

Where the last line was arrived at using (4.39). This requires a compensating weight

$$\Omega = \exp \left[\mathcal{G}^T \left(-\frac{1}{2} \mathcal{G} dt + d\mathbf{W} \right) \right], \quad (4.41)$$

which is given by the gauged equations if $\Omega = e^{dz_0}$.

As corollaries of this:

1. If there was no noise in the original equations ($\underline{B} = 0$), or the noise matrix was singular, then there would be no way a global weight could compensate for an arbitrary change in the drift equations — the new drift would just be plain wrong. For the case of very small noise in the \mathbf{z} equations, even small changes with respect to the original drift \underline{A} require very large weight compensation. These are the reasons why the gauges \mathcal{G} are multiplied by the coefficients \underline{B} in the $d\mathbf{z}$ equations, but not in the weight equation for dz_0 . Thus

If there is no random component to the evolution of a variable z , its drift cannot be modified using a drift gauge. If the noise matrix is singular, *no* variable drift can be modified using drift gauges.

2. While the weight equation (4.26) appears disordered and random due to the presence of noises dW_k , the evolution of Ω is actually strictly tied to the random walk of the other variables z_j , and acts to exactly compensate for any gauge modifications of the original drift.
3. If we consider the ($2N_z$ -dimensional) phase space of all base variables in $\underline{C} = \{z_j\}$, then a single real drift gauge \mathcal{G}_k changes the drift only in the phase-space direction specified by the k th column \underline{B}_k of the noise matrix⁴, associated with the k th noise dW_k .

4.3.3 Complex gauges

The change in the drift due to complex gauges $\mathcal{G} = \mathcal{G}' + i\mathcal{G}''$ is always

$$\underline{A} = \underline{A} dt - \underline{B} \mathcal{G}' dt - i \underline{B} \mathcal{G}'' dt. \quad (4.42)$$

⁴So that $\underline{B} = [\underline{B}_1 \ \underline{B}_2 \ \dots]$.

Table 4.1: Tally of drift gauge freedoms and comparison to phase space size.

	Number	Kind
Size of base phase-space $\underline{C} = \{z_j\}$ (degrees of freedom)	$2N_z$	real
Size of full phase-space $C = \{z_0, \underline{C}\}$	$2N_z + 2$	real
Number of standard drift gauges \mathcal{G}_k	N_z	complex
Number of noises dW_k in standard drift gauge scheme	N_z	real
Number of broadening drift gauges $\check{\mathcal{G}}_j$	N_z	complex
Number of noises in broadening drift gauge scheme (4.84)	$3N_z$	real

Imaginary gauges \mathcal{G}_k'' , in contrast to real ones, lead to changes in the drift only in a particular direction $i\underline{B}_k$ orthogonal to that affected by a real gauge. Note however, that this is a *particular* direction, and not any of the infinitely many orthogonal to \underline{B}_k . This kind of drift modification orthogonal to noise direction is not compensated for by changing the weight of the trajectory $|\Omega|$, but by modification of its phase $e^{iz_0''}$, leading to an interference effect between trajectories. It appears harder to grasp intuitively that this fully compensates, but in simulations such gauges are also seen to preserve observable averages (see Part B).

Similarly to purely real gauges, no drift modification can be made if the noise matrix \underline{B} is singular, as seen from (4.42). Conversely, with non-singular \underline{B} one has, in theory, full freedom to modify the deterministic evolution of the complex z_j to any arbitrary functional form. The weight Ω will keep exactly track of these modifications by virtue of using the same noises $d\mathbf{W}$. In fact, the weight evolution can be written as a deterministic function of the evolution of the remaining variables:

$$dz_0 = \mathcal{G}^T \left\{ \underline{B}^{-1} (dz - \underline{A} dt) + \frac{1}{2} \mathcal{G} dt \right\}. \quad (4.43)$$

A summary of drift gauge freedoms can be found in Table 4.1.

4.3.4 Weight spread

While the global weight e^{z_0} is completely determined by the evolution of the base variables z_j , in observable calculations (3.14) it appears (particularly in the denom-

inator) as effectively a random variable. Other things being equal, it is certainly desirable to make the spread of the weights small during a simulation.

Variations to consider

In general there are several properties of the weights that may be of relevance. $\text{Re}\{\Omega\}$ appears directly in the denominator average of observable estimates (3.14) via $\text{Tr}[\widehat{\Lambda}] \propto \Omega$. On the other hand, for general observables, both the phase and magnitude of the weight may be important to calculate the numerator average in an observable estimate (3.14) — the phase may be correlated in important ways with the other variables in the average.

For some observables and kernels, the numerator expression of (3.14) depends mostly only on $\text{Re}\{\Omega\}$. This is particularly so at short times when starting from a state well described as a classical (and separable) mixture of the kernels. In this case, both the numerator and denominator of the observable estimator (3.14) depend mostly on $\text{Re}\{\Omega\}$, and the variance of $\text{Re}\{\Omega\}$ is the most relevant to consider. In more general cases, the variance of $|\Omega|$ will be more relevant.

Log-weight spread estimate

A common hindrance when choosing the gauge is the difficulty of accurately assessing the size of the weight spread that will be produced by a given gauge. Here approximate expressions for the variance of z_0 at small times (4.47) will be derived.

Assuming there is no non-gauge weight drift ($\underline{A}_0 = 0$), the evolution in the variance of $z'_0 = \text{Re}\{z_0\}$ is given by

$$d\text{var}[z'_0] = d\langle (z'_0)^2 \rangle_{\text{stoch}} - 2\langle z'_0 \rangle_{\text{stoch}} d\langle z'_0 \rangle_{\text{stoch}}. \quad (4.44)$$

Using properties of the Ito calculus, one can evaluate the time increments of these

quantities from the stochastic equation (4.36b) as

$$d\langle z'_0 \rangle_{\text{stoch}} = \langle \text{Re} \{A_0\} \rangle_{\text{stoch}} dt = \frac{1}{2} \sum_k \langle (\mathcal{G}''_k)^2 - (\mathcal{G}'_k)^2 \rangle_{\text{stoch}} dt \quad (4.45a)$$

$$d\langle (z'_0)^2 \rangle_{\text{stoch}} = \langle 2z'_0 \text{Re} \{A_0\} \rangle_{\text{stoch}} dt + \sum_k \langle (B'_{0k})^2 dW_k \rangle_{\text{stoch}} \quad (4.45b)$$

$$= \sum_k \langle z'_0 (\mathcal{G}''_k)^2 - z'_0 (\mathcal{G}'_k)^2 + (\mathcal{G}'_k)^2 \rangle_{\text{stoch}} dt, \quad (4.45c)$$

and so in terms of covariances

$$d\text{var} [z'_0] = \sum_k \langle (\mathcal{G}'_k)^2 + \text{covar} [(\mathcal{G}''_k)^2, z'_0] - \text{covar} [(\mathcal{G}'_k)^2, z'_0] \rangle_{\text{stoch}} dt. \quad (4.46a)$$

Similarly one finds

$$d\text{var} [z''_0] = \sum_k \langle (\mathcal{G}''_k)^2 - 2 \text{covar} [\mathcal{G}'_k \mathcal{G}''_k, z''_0] \rangle_{\text{stoch}} dt. \quad (4.46b)$$

A practical approximation to the variance at short times can be gained by assuming lack of correlations between z_0 and \mathcal{G} (i.e. that the covariances are negligible). These approximations then would be

$$\text{var} [z'_0(t)] \approx V'_0(t) = \text{var} [z'_0(0)] + \int_0^t \sum_k \langle \mathcal{G}'_k(t')^2 \rangle_{\text{stoch}} dt' \quad (4.47a)$$

$$\text{var} [z''_0(t)] \approx V''_0(t) = \text{var} [z''_0(0)] + \int_0^t \sum_k \langle \mathcal{G}''_k(t')^2 \rangle_{\text{stoch}} dt'. \quad (4.47b)$$

Of course this is never exact because z_0 *does* depend on the gauges due to its evolution, however for small times, equal starting weights ($\text{var} [z'_0] = \text{var} [z''_0] = 0$), and *autonomous* gauges (i.e. \mathcal{G} does not depend explicitly on z_0) the expressions (4.47) are good approximations. They can be used to assess the amount of statistical noise that will be introduced by a particular gauge.

Small time regime for log-weights

How small is “small time” in this context? Clearly times when the covariances are much less than the $\sum_k (\mathcal{G}_k)^2$ terms. From (4.46), this will occur for autonomous gauges and equal starting weights if

$$z'_0(t) \sum_k |\mathcal{G}_k|^2 \ll \sum_k (\mathcal{G}'_k)^2, \quad (4.48a)$$

and

$$2z_0''(t) \sum_k \mathcal{G}'_k \mathcal{G}''_k \ll \sum_k (\mathcal{G}''_k)^2. \quad (4.48b)$$

To obtain limits on these, one needs an approximation to z_0 . Since $dW_k \approx \mathcal{O}(\sqrt{dt})$, then $\mathcal{G}^T d\mathbf{W}$ is the dominant term in the evolution of z_0 at short enough times. This implies

$$z'_0(t) \approx \mathcal{O} \left(\sqrt{dt} \sqrt{\sum_k (\mathcal{G}'_k)^2} \right) \quad (4.49)$$

$$z''_0(t) \approx \mathcal{O} \left(\sqrt{dt} \sqrt{\sum_k (\mathcal{G}''_k)^2} \right). \quad (4.50)$$

The root of sum of squares is due to independence of noises dW_k . For the approximations to z'_0 to apply one needs

$$\begin{aligned} \sqrt{dt \sum_k (\mathcal{G}'_k)^2} &\gg |\operatorname{Re} \{A_0\} dt| = \frac{dt}{2} \left| \sum_k [(\mathcal{G}'_k)^2 - (\mathcal{G}''_k)^2] \right| \\ \sum_k (\mathcal{G}'_k)^2 &\gg \frac{dt}{4} \left\{ \sum_k [(\mathcal{G}'_k)^2 - (\mathcal{G}''_k)^2] \right\}^2 \leq \frac{dt}{4} \left\{ \sum_k |\mathcal{G}_k|^2 \right\}^2 \\ dt &\ll \frac{4 \sum_k (\mathcal{G}'_k)^2}{\{\sum_k |\mathcal{G}_k|^2\}^2}. \end{aligned} \quad (4.51)$$

Given this is the case, z'_0 is given by (4.49), and so the condition (4.48a) for (4.47a) to hold becomes

$$\begin{aligned} \sum_k (\mathcal{G}'_k)^2 &\gg \sqrt{dt} \sqrt{\sum_k (\mathcal{G}'_k)^2} \sum_k |\mathcal{G}_k|^2 \\ dt &\ll \frac{\sum_k (\mathcal{G}'_k)^2}{\{\sum_k |\mathcal{G}_k|^2\}^2}. \end{aligned} \quad (4.52)$$

Since this agrees with the initial assumption of noise term dominance (4.51), it is the short time condition under which (4.47a) is accurate.

For dz''_0 to be given by (4.49), one needs

$$\begin{aligned} \sqrt{dt \sum_k (\mathcal{G}''_k)^2} &\gg |\operatorname{Im} \{A_0\} dt| = dt \left| \sum_k \mathcal{G}'_k \mathcal{G}''_k \right| \\ \sum_k (\mathcal{G}''_k)^2 &\gg dt \left\{ \sum_k \mathcal{G}'_k \mathcal{G}''_k \right\}^2 \leq dt \sum_j (\mathcal{G}'_j)^2 \sum_k (\mathcal{G}''_k)^2 \\ dt &\ll \frac{1}{\sum_k (\mathcal{G}'_k)^2}. \end{aligned} \quad (4.53)$$

The condition for (4.47b) is then

$$\begin{aligned} \sum_k (\mathcal{G}''_k)^2 &\gg 2\sqrt{dt} \sqrt{\sum_k (\mathcal{G}''_k)^2} \sum_k \mathcal{G}'_k \mathcal{G}''_k \\ \sum_k (\mathcal{G}''_k)^2 &\gg 4 dt \left\{ \sum_k \mathcal{G}'_k \mathcal{G}''_k \right\}^2 \leq 4 dt \sum_j (\mathcal{G}'_j)^2 \sum_k (\mathcal{G}''_k)^2 \\ dt &\ll \frac{1}{4 \sum_k (\mathcal{G}'_k)^2}. \end{aligned} \quad (4.54)$$

Again, this agrees with the initial assumption of noise term dominance (4.53), and is the short time condition under which (4.47b) is accurate.

Log-weight variance limits

As will be discussed in Section 7.4 and Appendix A, the variance of z'_0 should be $\lesssim \mathcal{O}(10)$ for the simulation to give results with any useful precision. Basically when z'_0 of a variance $\mathcal{O}(10)$ or more, the high positive z'_0 tail of the distribution contains few samples, but much weight, leading to possible bias. Also if z''_0 has a standard deviation of $\mathcal{O}(\pi)$ or more, mutual cancelling of trajectories with opposite phases will dominate the average, concealing any average in noise for reasonable sample sizes $\mathcal{O}(\lesssim 10^5)$.

This imposes a limit on how long a gauged simulation can last, and so it is extremely desirable to keep the magnitude of the gauges as small as possible.

As a corollary to this point, it is desirable to ensure that all drift gauges are zero at any attractors in phase space, to avoid accumulating unnecessary randomness in the weights when no significant evolution is occurring.

Direct weight spread estimate

One can also investigate the evolution of the variance of the weight Ω directly, and analytic estimates can be obtained in some situations. To better understand the effect of real or imaginary gauges on the weight, let us consider the case where gauges \mathcal{G}_k and weight Ω are decorrelated, there is no base weight drift ($\underline{A}_0 = 0$), and $\Omega = 1$ initially for all trajectories. Let $\Omega = \Omega' + i\Omega''$, then from (4.26)

$$d\Omega' = \sum_k (\Omega' \mathcal{G}'_k - \Omega'' \mathcal{G}''_k) dW_k \quad (4.55a)$$

$$d\Omega'' = \sum_k (\Omega' \mathcal{G}''_k + \Omega'' \mathcal{G}'_k) dW_k. \quad (4.55b)$$

If we consider the evolution of the second order moments of the weights, then under the uncorrelated Ω and \mathcal{G} assumption, it can be written as a closed system of equations

$$d \left\langle \begin{bmatrix} [\Omega']^2 \\ [\Omega'']^2 \\ \Omega' \Omega'' \end{bmatrix} \right\rangle_{\text{stoch}} = \begin{bmatrix} c_1 & c_2 & -2c_3 \\ c_2 & c_1 & 2c_3 \\ c_3 & -c_3 & c_1 - c_2 \end{bmatrix} \left\langle \begin{bmatrix} [\Omega']^2 \\ [\Omega'']^2 \\ \Omega' \Omega'' \end{bmatrix} \right\rangle_{\text{stoch}} dt, \quad (4.56)$$

where

$$c_1(t) = \sum_k \langle [\mathcal{G}'_k(t)]^2 \rangle_{\text{stoch}} \quad (4.57a)$$

$$c_2(t) = \sum_k \langle [\mathcal{G}''_k(t)]^2 \rangle_{\text{stoch}} \quad (4.57b)$$

$$c_3(t) = \sum_k \langle \mathcal{G}'_k(t) \mathcal{G}''_k(t) \rangle_{\text{stoch}}. \quad (4.57c)$$

This system can be solved, and remembering that $\langle \Omega(t) \rangle_{\text{stoch}} = 1$ (from (4.55)), and $\Omega(0) = 1$, one obtains the solutions

$$\begin{aligned} \text{var} [\Omega'] &= \frac{1}{2} e^{\int_0^t c_1(t') dt'} \left[e^{\int_0^t c_2(t') dt'} + e^{-\int_0^t c_2(t') dt'} \cos \left(2 \int_0^t c_3(t') dt' \right) \right] - 1 \\ \text{var} [\Omega''] &= \frac{1}{2} e^{\int_0^t c_1(t') dt'} \left[e^{\int_0^t c_2(t') dt'} - e^{-\int_0^t c_2(t') dt'} \cos \left(2 \int_0^t c_3(t') dt' \right) \right] - 1 \\ \langle \Omega' \Omega'' \rangle_{\text{stoch}} &= \frac{1}{2} e^{\int_0^t [c_1(t') - c_2(t')] dt'} \sin \left(2 \int_0^t c_3(t') dt' \right). \end{aligned} \quad (4.58)$$

If the gauge averages $c_j(t)$ are approximately constant with time, then at short times, when the condition that Ω and \mathcal{G}_k are uncorrelated holds, the variance of the

weight is

$$\text{var} [\Omega'] = t \sum_k \langle (\mathcal{G}'_k)^2 \rangle_{\text{stoch}} + \frac{t^2}{2} \sum_k \langle (\mathcal{G}'_k - \mathcal{G}''_k)^2 \rangle_{\text{stoch}} + \mathcal{O}(t^3). \quad (4.59)$$

In similar manner, one can also obtain

$$\text{var} [|\Omega|] = \exp \left[\int_0^t \langle |\mathcal{G}_k(t')|^2 \rangle_{\text{stoch}} dt' \right] - 1. \quad (4.60)$$

Some more conclusions about drift gauge forms

For most observables it is the variance of Ω' that is most relevant for uncertainties in the finite sample estimates, because it appears both in the numerator and denominator of (3.14). However, for some like the local quadratures $(\hat{a}_k^\dagger - \hat{a}_k)/2i$ the modulus of the weight is more relevant in the numerator of the observable expression (3.14). One can see from (4.59) that at short times imaginary gauges lead to smaller real weight spreads because the variance grows only quadratically, most of the gauge noise going into Ω'' . This makes imaginary drift gauges more convenient for short time estimates of most observables

4.4 Diffusion stochastic gauges

The Fokker-Planck equation specifies directly only the diffusion matrix $D(C)$, which is then decomposed via

$$D = BB^T \quad (4.61)$$

into noise matrices B , however these are not specified completely. This freedom in the choice of B , leads to a different kind of stochastic gauge than considered in the previous sections, which will be termed “diffusion gauges” here. No weights are required, and only the noise terms are modified. As with kernel stochastic gauges, the diffusion gauges may in some cases be used to choose a set of equations with the most convenient stochastic properties. The non-uniqueness of B has always been known, but has usually been considered to simply relabel the noises without any useful consequences. It has, however, been recently shown by Plimak *et al*[2]

that using a different B than the obvious “square root” form $B = B_0 = D^{1/2}$ leads to impressive improvement in the efficiency of positive P simulations of the Kerr oscillator in quantum optics. (This has a similar form of the Hamiltonian to the nonlinear term in (2.17).) This somewhat surprising result leads us to try to quantify the amount of freedom of choice available in the noise matrices. In the process, it will also become apparent why some extra properties of the kernel beyond the most general case are needed for diffusion gauges to be useful for simulations.

4.4.1 Noise matrix freedoms with general and complex analytic kernels

At first glance there appears to be a great deal of freedom in the choice of noise matrix B . Consider that the relationship $D = BB^T$ for N_v real variables can be satisfied by a $N_v \times N_W$ noise matrix, subject to the $N_v(N_v + 1)/2$ real constraints $D_{jk} = \sum_p B_{jp}B_{kp}$. The number (N_W) of independent real Wiener increments is formally unconstrained. It turns out, however, that much (if not all) of this freedom is simply freedom of labeling and splitting up a single Wiener increment into formally separate parts having no new statistical properties. More on this in Section 4.4.3.

Recalling Section 3.4.2, the Ito Langevin equations (3.45), (3.56), (4.25), or (4.36) can also be written (Whether the C_j are real or complex) as

$$dC_j(t) = A_j(C, t)dt + dX_j(C, t), \quad (4.62)$$

where the $\sum_k B_{jk}dW_k$ terms have been coalesced into a single stochastic increment dX_j . Only the means and variances of the Wiener increments dW_k are specified, and so by the properties of Ito stochastic calculus, the only binding relationships for the stochastic terms dX_j are (3.50) — mutual variances specified by the stochastic average of the diffusion matrix, and zero means.

Let us ask the question “do the relationships (3.50) completely specify the statistical properties of the dX_j ?”.

Consider firstly that because all the increments are taken to be “practically” infinitesimal in any simulation, then over any significant timescale the noise due

to the Wiener increments dW_k will be effectively Gaussian, whatever the actual distribution of dW_k used⁵. This is due to the central limit theorem. Also, (3.50a) is guaranteed by the independence of the B_{jk} and dW_k in the Ito calculus, combined with $\langle dW_k(t) \rangle_{\text{stoch}} = 0$. These constraints then imply that to check if there is any freedom in the statistical properties of the dX_j it suffices to compare the number of covariance conditions (3.50b) with the number of covariance relations for the dX_j .

Seemingly this is trivial: for N_v real variables C_j , there are $N_v(N_v + 1)/2$ conditions (3.50b), and the same number of possible covariance relationships $\langle dX_j dX_k \rangle_{\text{stoch}}$. This means that

In the general case of a kernel with no additional symmetry properties, all the (formally different) possible choices of the noise matrix B have exactly the same effect on the statistical behavior of the Langevin equations.

This is despite the noise matrix having formally many free parameters. These then are largely freedoms to relabel and split up the Wiener increments without introducing any new statistical behavior, as will be discussed in Section 4.4.

This simple comparison of conditions and relationships also explains why trying various forms of the noise matrices B was for a long time thought not to have any useful effect — under general conditions it doesn't.

It has been found, however, that a positive P simulation can be optimized by particular choices of B [2]. Let us consider the case when the kernel $\hat{\Lambda}$ can be written as an analytic function of N_z complex variables $C = \{z_j\}$. Now, there are $N_z(N_z + 1)/2$ complex conditions (3.50b). As for statistical properties of the dX_j , we have to consider independently both the real and imaginary parts, dX'_j and dX''_j respectively. Their covariance relations are $\langle dX'_j dX'_k \rangle_{\text{stoch}}$, $\langle dX''_j dX''_k \rangle_{\text{stoch}}$, and $\langle dX'_j dX''_k \rangle_{\text{stoch}}$, numbering $N_z(2N_z + 1)$ real relations in all. Since there are two real conditions per one complex, this means that there are left over N_z^2 degrees of freedom in the covariance relations between elements of the complex quantities dX_j . These might be used to tailor these N_z^2 free variances or covariances to our needs.

⁵Provided that the variance of dW_k is finite.

This begs the obvious question of “why do the analytic complex variable kernels seem allow more freedom than a general real variable kernel”, since they can also be written in terms of real variables – we just split each complex variable into real and imaginary parts.

The key lies in the fact that kernels analytic in complex variables have special symmetry properties that give us freedom to choose the complex derivatives of the kernel as (3.51), or indeed as

$$\frac{\partial \widehat{\Lambda}(C)}{\partial z_j} = \mathcal{F} \frac{\partial \widehat{\Lambda}(C)}{\partial \text{Re}\{z_j\}} - i(1 - \mathcal{F}) \frac{\partial \widehat{\Lambda}(C)}{\partial \text{Im}\{z_j\}} \quad (4.63)$$

using an analytic kernel stochastic gauge \mathcal{F} . This “analytic” symmetry then enters into the Fokker-Planck equation, and finds its way into the Langevin equations as freedom of choice of noise increments dX_j . As with all stochastic gauges, another way of looking at this is that because the kernel has certain symmetry properties, there is a whole family of distributions $P(C)$ of such kernels corresponding to the same quantum density matrix $\widehat{\rho}$. Choosing gauges chooses between these different distributions.

In the following subsections, several specific types of freedoms (or “stochastic gauges”) available in the noise matrix B will be considered, and a tally will be made at the end in Table 4.3.

4.4.2 Standard form of diffusion gauges for analytic kernels

Let us consider the analytic kernel case, where it was seen above that there may be useful noise matrix freedoms. Such kernels tend also to be the most convenient generally because a positive propagator, and so a stochastic interpretation, is always guaranteed by the method in Section 3.4.3.

Since $D = D^T$ is square and can be made symmetric⁶, its matrix square root is also symmetric $\sqrt{D} = \sqrt{D}^T$, and can be used as a noise matrix B_0

$$B_0 = \sqrt{D} \quad (4.64)$$

$$D = B_0 B_0^T. \quad (4.65)$$

⁶Since $\partial^2/\partial v_1 \partial v_2 = \partial^2/\partial v_2 \partial v_1$ for any variables v_1 and v_2 can be used in the FPE to achieve $D_{jk} = D_{kj}$.

This square root form can be considered as the “obvious” choice of decomposition, unique apart from the N_z signs of the diagonal terms⁷. However, for any complex orthogonal O such that $OO^T = I$, if B_0 is a valid decomposition of D , then so is the more general matrix $B = B_0O$. So, any matrix in the whole orthogonal family

$$B = B_0O \quad (4.66)$$

is a valid decomposition. A general complex orthogonal matrix can be written explicitly using an antisymmetric matrix basis $\sigma^{(jk)}$, ($j \neq k = 1, \dots, N_z$) having matrix elements

$$\sigma_{lp}^{(jk)} = \delta_{jl}\delta_{kp} - \delta_{jp}\delta_{kl}. \quad (4.67)$$

With these $(N_z - 1)N_z/2$ independent $N_z \times N_z$ matrices $\sigma^{(jk)}$ the general form is

$$O = \exp \left(\sum_{j < k} g_{jk}(C, C^*, t) \sigma^{(jk)} \right). \quad (4.68)$$

The g_{jk} are the $N_z(N_z - 1)/2$ complex diffusion gauge functions, which can in principle be *completely arbitrary*, including dependence on all variables in C (not necessarily analytic), and the time variable t , without affecting the correspondence between the Langevin stochastic equations, and the FPE.

As an example, in the case of two complex variables, there is one complex gauge function g_{12} , and the resulting transformation is

$$\begin{aligned} O &= \exp(g_{12}\sigma^{(12)}) \\ &= \cos(g_{12}) + \sigma^{(12)} \sin(g_{12}), \end{aligned} \quad (4.69)$$

where the anti-symmetric matrix $\sigma^{(12)}$ is proportional to a Pauli matrix:

$$\sigma^{(12)} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}. \quad (4.70)$$

Hence, e.g. if the diffusion matrix is diagonal, B_0 is also, and the transformed (but equivalent) noise matrix becomes:

$$B = \begin{bmatrix} \sqrt{D_{11}} \cos(g_{12}) & \sqrt{D_{11}} \sin(g_{12}) \\ -\sqrt{D_{22}} \sin(g_{12}) & \sqrt{D_{22}} \cos(g_{12}) \end{bmatrix}. \quad (4.71)$$

⁷One for each complex variable.

The square form of B (4.66) with O given by (4.68) will be termed here the “standard” diffusion gauge, to distinguish it from some other forms that will be discussed in later subsections.

4.4.3 Real standard diffusion gauges and noise mixing

Not all the canonical diffusion gauges g_{jk} are useful. One class of useless gauges to avoid are those that seemingly change B , but effectively only swap around the linear combinations of noises dW_k , without affecting dX_j or hence any statistical properties of the equations. Their main effect is usually to complicate the math.

Consider, for example, the two-complex variable case given by (4.69). For a purely real diffusion gauge $g_{12} = g'_{12}$, the effect of the noise terms in the Langevin equations becomes

$$d\mathbf{X} = B(g'_{12}) d\mathbf{W} = B_0 \begin{bmatrix} \cos g'_{12} & \sin g'_{12} \\ -\sin g'_{12} & \cos g'_{12} \end{bmatrix} \begin{bmatrix} dW_1 \\ dW_2 \end{bmatrix} = B_0 d\mathbf{W}'. \quad (4.72)$$

This is just a rotation of the noises $d\mathbf{W}$, leading to $d\mathbf{W}'$ being a linear combination of the noises dW_1 and dW_2 . The new noises dW'_j have the same statistical properties as the old dW_j . Thus $d\mathbf{X}(g'_{12}) = B(g'_{12}) d\mathbf{W}$ has exactly the same effect in the equations as the un-gauged $d\mathbf{X}(g'_{12} = 0)$, regardless of any complicated form of g'_{12} we choose to try.

More generally, with N_z complex variables, there are at least $N_z(N_z - 1)/2$ such useless noise rotations available, one for each pair of variables. Looking at the standard orthogonal matrix form (4.68), one sees that the terms in the exponential proportional to real parts of gauges g_{jk} represent simply all these useless rotations.

In a formalism with only real variables C_j , and no extra symmetries, diffusion gauges arise formally in the same manner as in Section 4.4.2, although all the matrices D, B, B_0 , and O must now be real, making g_{jk} also only real. This means that in such a case *all* the resulting real gauge functions in $O(g_{jk})$ are useless noise mixers, and no useful modifications of the stochastic equations can be achieved by choosing them. This explains why the potential of noise matrix choice to improve the stochastic equations went unnoticed for many years.

This subsection can be summarized as

Only the imaginary parts of the standard diffusion gauges (4.68) can affect the statistical properties of the stochastic equations.

Lastly, it may be worth pointing out that real gauges can achieve the same useless noise rotations in non-square noise matrices as well. For example, for any arbitrary real function f the noise matrix

$$B^{(f)} = \begin{bmatrix} B \cos f & B \sin f \end{bmatrix} \quad (4.73)$$

satisfies $B^{(f)}[B^{(f)}]^T = D$ just as well as the original B , whatever the form of the arbitrary complex function f . The stochastic increments are (for N_z complex variables)

$$\begin{aligned} dX_j &= \sum_{k=1}^{2N_z} B_{jk}^{(f)} dW_k \\ &= \sum_{k=1}^{N_z} B_{jk} (\cos f dW_k + \sin f dW_{N_z+k}) \\ &= \sum_{k=1}^{N_z} B_{jk} dW'_k, \end{aligned} \quad (4.74)$$

with the combined noises dW'_k having the same statistical properties irrespective of the choice of f .

4.4.4 Imaginary standard diffusion gauges and statistics

The imaginary part of standard diffusion gauges $g''_{jk} = \text{Im} \{g_{jk}\}$ do affect the statistics. Again consider the two-variable case of diagonal D of expression (4.69), this time with imaginary gauge $g_{12} = ig''_{12}$.

$$d\mathbf{X} = B_0 \begin{bmatrix} \cosh g''_{12} & i \sinh g''_{12} \\ -i \sinh g''_{12} & \cosh g''_{12} \end{bmatrix} \begin{bmatrix} dW_1 \\ dW_2 \end{bmatrix}. \quad (4.75)$$

For example, in the simplest case of diagonal positive real D , one has (using the notation $dX_j = dX'_j + idX''_j$)

$$dX'_1 = \sqrt{D_{11}} \cosh g''_{12} dW_1 \quad (4.76a)$$

$$dX'_2 = \sqrt{D_{22}} \cosh g''_{12} dW_2 \quad (4.76b)$$

$$dX''_1 = \sqrt{D_{11}} \sinh g''_{12} dW_2 \quad (4.76c)$$

$$dX''_2 = -\sqrt{D_{22}} \sinh g''_{12} dW_1. \quad (4.76d)$$

While the covariance requirements (3.50b) are satisfied, the variances of the dX_j components may vary depending on g''_{12} . The nonzero correlations are

$$\langle (dX'_1)^2 \rangle_{\text{stoch}} = \cosh^2 g''_{12} \langle D_{11} \rangle_{\text{stoch}} dt \quad (4.77a)$$

$$\langle (dX'_2)^2 \rangle_{\text{stoch}} = \cosh^2 g''_{12} \langle D_{22} \rangle_{\text{stoch}} dt \quad (4.77b)$$

$$\langle (dX''_1)^2 \rangle_{\text{stoch}} = \sinh^2 g''_{12} \langle D_{11} \rangle_{\text{stoch}} dt \quad (4.77c)$$

$$\langle (dX''_2)^2 \rangle_{\text{stoch}} = \sinh^2 g''_{12} \langle D_{22} \rangle_{\text{stoch}} dt \quad (4.77d)$$

$$\langle dX'_1 dX''_2 \rangle_{\text{stoch}} = -\frac{1}{2} \sinh 2g''_{12} \left\langle \sqrt{D_{11} D_{22}} \right\rangle_{\text{stoch}} dt \quad (4.77e)$$

$$\langle dX'_2 dX''_1 \rangle_{\text{stoch}} = \frac{1}{2} \sinh 2g''_{12} \left\langle \sqrt{D_{11} D_{22}} \right\rangle_{\text{stoch}} dt. \quad (4.77f)$$

Such imaginary gauges only apply when the model is expressed in complex variables z_j with an analytic kernel. A glance at the canonical gauge form (4.68), indicates that there is then one such imaginary gauge $g''_{jk} = \text{Im} \{g_{jk}\}$ for each possible pair of variables.

4.4.5 Non-standard diffusion gauges and further freedoms

In Section 4.4.1 it was seen that there are N_z^2 degrees of freedom⁸ for the statistics of the dX_j , but later, in Sections 4.4.3 and 4.4.4 it was shown that there are at most $N_z(N_z - 1)/2$ useful (imaginary) standard gauges $g_{jk} = ig''_{jk}$. Conclusion: there are some more freedoms not included in the standard square noise matrix expression (4.66) and (4.68). The simplest example occurs when there is only one complex kernel parameter z_1 : there is one degree of freedom in dX_1 , but zero standard gauges.

⁸The case of a kernel analytic in complex variables is being considered all the time.

To see how the extra non-square gauges can enter the picture, let us consider the simplest $N_z = 1$ case. Here there will be three covariance relations for dX_1' and dX_1'' , but two conditions (3.50) on these. Firstly, note that while the standard gauged noise matrix $B = B_0 O$ is square in the complex variables z_j , in the real variables z_j' and z_j'' it is a $2N_z \times N_z$ — in this case a 2×1 . This leaves no room for a third degree of freedom in $dX_1 = B d\mathbf{W}$. There is, however, no particular limit on the number of columns of B or the number of noises in $d\mathbf{W}$, provided $BB^T = D$, so let us instead consider a 2×2 noise matrix,

$$B = \begin{bmatrix} b_{rr} & b_{ri} \\ b_{ir} & b_{ii} \end{bmatrix}, \quad (4.78)$$

where subscripts r and i denote elements relating to the real and imaginary parts of the complex variable z_1 . Since there are only three degrees of freedom, we can set one of the elements to zero with no restrictions on the statistical properties of dX_1 (say $b_{ri} = 0$). The remaining elements must satisfy the conditions (from (3.50))

$$\langle \text{Re} \{D\} \rangle_{\text{stoch}} dt = \langle (dX_1')^2 - (dX_1'')^2 \rangle_{\text{stoch}} = dt \langle b_{rr}^2 - b_{ir}^2 - b_{ii}^2 \rangle_{\text{stoch}} \quad (4.79a)$$

$$\langle \text{Im} \{D\} \rangle_{\text{stoch}} dt = 2 \langle dX_1' dX_1'' \rangle_{\text{stoch}} = dt \langle b_{rr} b_{ir} \rangle_{\text{stoch}}, \quad (4.79b)$$

leaving room for one arbitrary function (i.e. gauge) out of the three nonzero elements b_{rr} , b_{ri} , or b_{ii} . In the standard noise matrix form from (3.54) one had $b_{ii} = 0$, so all the noise matrix elements were then completely specified by (4.79), with no room for a gauge.

4.4.6 Distribution broadening gauges

Consider the case of no diffusion $D = 0$. When the kernel is analytic in complex variables, there can actually be nonzero noise matrices, leading to nonzero noise despite no diffusion in the FPE. All that is required is that the complex B obeys $BB^T = 0$, or in terms of stochastic increments $d\mathbf{X} = B d\mathbf{W}$ that

$$\langle d\mathbf{X} \cdot d\mathbf{X}^T \rangle_{\text{stoch}} = D = 0 \quad (4.80a)$$

$$\langle d\mathbf{X} \rangle_{\text{stoch}} = 0. \quad (4.80b)$$

A viable noise matrix is, for example,

$$\check{B} = \begin{bmatrix} \check{g} & i\check{g} \end{bmatrix}, \quad (4.81)$$

where \check{g} is a $N_z \times N_W$ complex matrix, all of whose elements can be any arbitrary functions (gauges) we like without affecting $\check{B}\check{B}^T = D = 0$. There is no limit on the width N_W . For nonzero diffusion, the same kind of gauges can be attached to an existing noise matrix ($BB^T = D$) at one's leisure via

$$\tilde{B} = \begin{bmatrix} B & \check{B} \end{bmatrix}. \quad (4.82)$$

The new noise matrix also obeys $\tilde{B}\tilde{B}^T = D$. Nonzero noises corresponding to zero diffusion in the FPE are possible because the correspondence between stochastic equations and the FPE is exact only in the limit of infinite samples. Thus the conditions (4.80) apply only in the limit of infinitely many trajectories, and $B = 0$ is just one special case in which these conditions are satisfied for every trajectory on its own.

The diffusion gauges \check{g} will be termed “distribution broadening” here because their effect is to make the distribution of the complex variables z_j broader, while preserving their complex means and mutual correlations. Generally such broadening gauges are not of much use, simply making everything more noisy and reducing precision of observable averages, but there are some situations when this is advantageous.

One example occurs when the noise matrix is singular, or there is no native diffusion in the FPE (i.e. $dX_j = 0$) for a variable z_j . In this situation drift gauges are unable to make modifications to the deterministic evolution. If one, however, uses a broadening gauge to force a nonzero noise matrix, then the usual drift gauge formalism can be used and arbitrary modifications to the deterministic evolution made. The idea of creating additional drift gauges in this manner was first proposed (in a $N_z = 2$ case) by Dowling[69]. Let us see how this proceeds in a general case:

For clarity in the resultant equations it is best to define a diagonal broadening gauge matrix \check{g} with elements

$$\check{g}_{jk} = \check{g}_j \delta_{jk}. \quad (4.83)$$

Adding the broadening gauge to a pre-existing noise matrix B as in (4.81) and (4.82), one obtains a $N_z \times 3N_z$ noise matrix and $3N_z$ independent real noises. Each of these noises can now have a drift kernel gauge attached to it in the manner described in Section 4.3. For the purpose of arbitrary drift manipulation it suffices to introduce N_z complex drift gauges $\check{\mathcal{G}}_j$ on just N_z of these noises — say on the noise matrix elements \check{g}_j . This leads then to the stochastic equations

$$dz_j = A_j dt + \sum_k B_{jk} dW_k + \check{g}_j (d\check{W}_j + id\check{\check{W}}_j - \check{\mathcal{G}}_j dt) \quad (4.84a)$$

$$dz_0 = A_0 dt - \frac{1}{2} \sum_j \check{\mathcal{G}}_j^2 dt + \sum_j \check{\mathcal{G}}_j d\check{W}_j. \quad (4.84b)$$

The “broadening noises” $d\check{W}_j$ and $d\check{\check{W}}_j$ are independent real Wiener increments just like the dW_k .

The arbitrary gauge modifications to the drift of a variable z_j are

$$-\check{g}_j \check{\mathcal{G}}_j dt. \quad (4.85)$$

A related situation occurs when the native stochastic increment dX_j for a variable z_j is small. In such a situation, making a significant change in the drift of z_j would require large compensation in the log-weight z_0 , leading to rapidly increasing statistical uncertainty or even bias from widely varying trajectory weights. However, if one introduces a broadening gauge such that

$$\begin{aligned} \text{var} \left[\text{Re} \{ \check{g}_j \} d\check{W}_j \right] &\gg \text{var} \left[\text{Re} \{ dX_j \} \right] \\ \langle \text{Re} \{ \check{g}_j \}^2 \rangle_{\text{stoch}} &\gg \sum_k \langle \text{Re} \{ B_{jk} \}^2 \rangle_{\text{stoch}}, \end{aligned} \quad (4.86)$$

then the drift of $\text{Re} \{ z_j \}$ can be modified at smaller weighting cost than if one had used standard drift gauges. Analogously for $\text{Im} \{ z_j \}$. In fact, there appears to be a tradeoff here between the amount of noise introduced into the log-weight z_0 (proportional to $\check{\mathcal{G}}_j$), and the amount of noise introduced directly into the variable whose drift is being modified (proportional to \check{g}_j).

4.4.7 Diffusion from different physical processes

A commonly occurring situation is that several physically distinct processes give separate contributions to the diffusion matrix, e.g.

$$D = \sum_l D^{(l)}. \quad (4.87)$$

Calculating the square-root noise matrix B_0 may, in some cases, give a very complicated expression in this situation. If this is a hindrance in calculations, a much more transparent noise matrix decomposition is possible. One decomposes each diffusion contribution $D^{(l)}$ separately into its own noise matrix as

$$D^{(l)} = B^{(l)}(B^{(l)})^T. \quad (4.88)$$

and then combines them as

$$B = \begin{bmatrix} B^{(1)} & B^{(2)} & \dots \end{bmatrix}. \quad (4.89)$$

This results in separate noise processes for each diffusion contribution $D^{(l)}$, and formally separate drift gauges (one complex drift gauge per real noise). The benefit of doing this is that the resulting stochastic equations have stochastic terms of a relatively simple form. On the other hand, a possible benefit of doing things the hard way with B_0 a direct square root of the full diffusion matrix D is that diffusion contributions from different processes l may partly cancel, leading to a less noisy simulation.

4.5 Summary of standard gauges

The standard gauge choices that will be used in subsequent chapters can be summarized in vector form by the equations

$$dz = \underline{A} dt + \underline{B}_0 O(\{g\}) (d\mathbf{W} - \underline{G} dt), \quad (4.90a)$$

$$dz_0 = \underline{A}_0 dt + \underline{G}^T \left(d\mathbf{W} - \frac{1}{2} \underline{G} dt \right). \quad (4.90b)$$

This uses a kernel of the form (4.9) proportional to a global weight $\Omega = e^{z_0}$, and analytic in $N_z + 1$ complex variables z_j, z_0 . A weighting kernel gauge, standard drift

kernel gauges, and standard imaginary diffusion gauges have been used ($g_{jk} = ig''_{jk}$ forming the set $\{g\}$ of arbitrary real-valued gauge functions). Underlined quantities are those obtained from the FPE before introduction of drift gauges, while the bold quantities denote column vectors with an element per base configuration variable $z_{j \neq 0}$. And so, $\underline{\mathbf{A}}$ and $\underline{\mathbf{A}}_0$ are drift coefficients from the un-gauged FPE, $d\mathbf{W}$ are real Wiener increments (with each element dW_j statistically independent for each j and each timestep, of zero mean and variance dt), and \mathcal{G} are arbitrary complex drift gauge functions. $\underline{\mathbf{B}}_0 = \sqrt{\underline{\mathbf{D}}}$ is the square-root noise matrix form, O is the $N_z \times N_z$ orthogonal matrix given by (4.68), but in this case dependent only on imaginary gauges ig''_{jk} . The gauge freedoms in this standard formulation are summarized in Table 4.2.

Table 4.2: **Tally of diffusion and noise matrix freedoms in the standard formulation** (4.90): Kernel analytic in complex phase-space variables (Section 3.4.3), standard imaginary diffusion gauges (Section 4.4.2), standard drift gauges (Section 4.3).

	Number	Kind
Base phase-space variables $\underline{\mathcal{C}} = \{z_j\}$	N_z	complex
All variables $\mathcal{C} = \{z_0, \underline{\mathcal{C}}\}$	$N_z + 1$	complex
Noises (Wiener increments dW_k)	N_z	real
Drift gauges \mathcal{G}_k	N_z	complex
Imaginary diffusion gauges $g_{jk} = ig''_{jk}$	$N_z(N_z - 1)/2$	imaginary

Table 4.3: **Tally of diffusion gauge and noise matrix freedoms.** The degrees of freedom counted are always real, not complex.

Number of:	Kernel analytic in N_z complex variables	General kernel in N_v real variables
Real variables	$2N_z$	N_v
Covariance relations between dX_j	$N_z(2N_z + 1)$	$N_v(N_v + 1)/2$
Covariance constraints on dX_j	$N_z(N_z + 1)$	$N_v(N_v + 1)/2$
Potentially useful statistical freedoms between dX_j	N_z^2	0
maximum Wiener increments dW_k	∞	∞
standard Wiener increments dW_k	N_z	N_v
Elements in standard square root noise matrix $B_0 = \sqrt{D}$	$2N_z^2$	N_v^2
Elements in extended noise matrix with $2N_z$ real columns	$4N_z^2$	N_v^2
Canonical gauges g_{jk} in (4.68)	$N_z(N_z - 1)$	$N_v(N_v - 1)/2$
— useless real gauges g'_{jk}	$N_z(N_z - 1)/2$	$N_v(N_v - 1)/2$
— useful imaginary gauges g''_{jk}	$N_z(N_z - 1)/2$	0
Potentially useful non-standard gauges	$N_z(N_z + 1)/2$	0
Useless potential noise spawning gauges	∞	∞
Potential broadening gauges in \check{g}	∞	not applicable
Broadening gauges \check{g}_j in drift gauge scheme (4.84).	N_z	not applicable

Chapter 5

The gauge P representation

The gauge P representation, which will be used in subsequent chapters, is explained and its properties investigated here. It is based on the positive P representation of quantum optics, which uses an off-diagonal coherent state kernel, but due to the inclusion of a global phase can be used to introduce weighting and drift gauges. The original distribution concept is due to P. D. Drummond, and its basic derivation is given in published work by Drummond and Deuar[66, 56, 61].

The positive P distribution is a promising starting point because it has already been successfully used in some many-body problems in quantum optics (e.g. squeezing in optical solitons [12, 13]) and with Bose atoms as well (evaporative cooling[14, 15, 46]). The coherent-state mode-based approach gives simple equations similar to the mean-field Gross-Pitaevskii (GP) equations, and automatically applies to open systems. However, stability and systematic error problems can occur[70, 71, 64, 66] with these simulations, which hinders effective use of the method. The gauge P representation inherits all the useful features of the positive P but allows drift gauges, which can be used to tailor the equations so that stability problems and systematic (boundary term) errors are removed. Weighting gauges allow one to also perform thermodynamic simulations of grand canonical ensembles by including an evolving trajectory weight.

Stochastic equations for the interacting Bose gas model are found in Sections 5.3, and 5.6. Extension to non-local interparticle interactions is derived in Section 5.5.

5.1 Properties of the representation

The representation uses an un-normalized (Bargmann) coherent state basis. On a subsystem j this basis is

$$||\alpha_j\rangle_j = \exp\left(\alpha_j \widehat{a}_j^\dagger\right) |0\rangle = \exp\left(\frac{|\alpha_j|^2}{2}\right) |\alpha_j\rangle_j, \quad (5.1)$$

with the complex amplitude α_j , and mean particle occupation $\langle \widehat{n}_j \rangle = |\alpha_j|^2$. As usual, the boson annihilation operators at subsystem j are \widehat{a}_j , and obey the commutation relations

$$\left[\widehat{a}_j, \widehat{a}_k^\dagger \right] = \delta_{jk}. \quad (5.2)$$

The basis states are mutually non-orthogonal

$$\langle \beta_j^* ||_j | | \alpha_k \rangle_k = \delta_{jk} e^{\alpha_j \beta_j}, \quad (5.3)$$

(from (3.30)) and overcomplete (from (3.31)).

The kernel (on N separable subsystems) is chosen to be

$$\begin{aligned} \widehat{\Lambda} &= e^{z_0} \otimes_{j=1}^N \widehat{\Lambda}_j, \\ \widehat{\Lambda}_j &= ||\alpha_j\rangle_j \langle \beta_j^* ||_j \exp(-\alpha_j \beta_j). \end{aligned} \quad (5.4)$$

If one defines coherent amplitude vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ to contain all N elements α_j and β_j , respectively, then the full variable set is $C = \{z_0, \boldsymbol{\alpha}, \boldsymbol{\beta}\}$, containing $2N + 1$ complex variables. The non-orthogonality of the basis coherent states allows normalization $\text{Tr} \left[\widehat{\Lambda} \right] = e^{z_0}$ apart from the complex global weight $\Omega = e^{z_0}$. The kernel is also seen to be analytic in all complex variables α_j , β_j , and z_0 , so the procedure of Section 3.4.3 can be used to ensure a stochastic realization for any FPE.

As per expression (3.1), the density matrix (possibly un-normalized) is expanded as

$$\widehat{\rho}_u = \int P_G(C) \widehat{\Lambda}(C) d^{2N} \boldsymbol{\alpha} d^{2N} \boldsymbol{\beta} d^2 z_0 \quad (5.5)$$

with the gauge P distribution $P_G(C)$.

Now, it can be verified by expansion that

$$\widehat{a}_k ||\alpha_j\rangle_j = \delta_{jk} \alpha_j ||\alpha_j\rangle_j \quad (5.6a)$$

$$\widehat{a}_k^\dagger ||\alpha_j\rangle_j = \frac{\partial}{\partial \alpha_k} ||\alpha_j\rangle_j. \quad (5.6b)$$

This then leads to the basic kernel operator correspondences

$$\widehat{a}_j \widehat{\Lambda} = \alpha_j \widehat{\Lambda}, \quad (5.7a)$$

$$\widehat{a}_j^\dagger \widehat{\Lambda} = \left(\beta_j + \frac{\partial}{\partial \alpha_j} \right) \widehat{\Lambda}, \quad (5.7b)$$

$$\widehat{\Lambda} \widehat{a}_j = \left(\alpha_j + \frac{\partial}{\partial \beta_j} \right) \widehat{\Lambda}, \quad (5.7c)$$

$$\widehat{\Lambda} \widehat{a}_j^\dagger = \beta_j \widehat{\Lambda} \quad (5.7d)$$

in the form (3.39) Together with

$$\frac{\partial}{\partial z_0} \widehat{\Lambda} = \widehat{\Lambda}, \quad (5.8)$$

which is seen to apply by inspection of (5.4), these operator equations can be used to obtain observable moment estimates and stochastic equations using the procedures of Sections 3.3.1, 3.4.1, 3.4.3.

The similarity of the gauge P to the positive P representation allows one to adopt some exact results obtained for the latter. Any state that has the positive P representation $P_+(\boldsymbol{\alpha}, \boldsymbol{\beta})$ can be represented by the gauge P representation

$$P_G(z_0, \boldsymbol{\alpha}, \boldsymbol{\beta}) = \delta^2(z_0) P_+(\boldsymbol{\alpha}, \boldsymbol{\beta}), \quad (5.9)$$

by inspection of the kernels (5.4) and (3.6). (This is not a unique correspondence, but the simplest of many.)

Since it has been shown that any quantum state $\widehat{\rho}$ has a positive P representation[11], it follows from (5.9) that all quantum states must also have gauge P representations. The constructive expression (3.10) for a positive P representation of arbitrary $\widehat{\rho}$ can be substituted into (5.9) to obtain a similar expression for the gauge P.

5.2 Observables

Calculation of observables proceeds in a very similar manner to the positive P representation in Section 3.3.5. All operators on supported states can be written as linear combinations of the moments of the local annihilation and creation operators $\widehat{a}_j, \widehat{a}_j^\dagger$. Thus, to evaluate any observable it suffices to know how to evaluate an expectation

value of a Hermitian operator with two adjoint separable terms of the form

$$\widehat{Q}(\{L_j, L'_k\}, \theta) = \frac{e^{i\theta}}{2} \otimes_j \widehat{a}_{L_j}^\dagger \otimes_k \widehat{a}_{L'_k} + \frac{e^{-i\theta}}{2} \otimes_k \widehat{a}_{L'_k}^\dagger \otimes_j \widehat{a}_{L_j}. \quad (5.10)$$

θ is a phase, the L_j and L'_k are subsystem labels (not necessarily unique), while the j and k are “subsystem label counters”. For example the particle number operator for subsystem p has $j = k = \{1\}$, $L_1 = L'_1 = p$, $\theta = 0$; while a quadrature operator $\widehat{q}(\theta) = \frac{1}{2}(\widehat{a}^\dagger e^{i\theta} + \widehat{a} e^{-i\theta})$ on subsystem p has $j = \{1\}$, $k = \{\}$, $L_1 = p$.

Using (5.7) and $\text{Tr} [\widehat{\Lambda}] = e^{z_0}$, one then obtains

$$\begin{aligned} \text{Tr} [\widehat{Q}\widehat{\Lambda}] &= \frac{e^{i\theta}}{2} \prod_k \alpha_{L'_k} \prod_j \left(\beta_{L_j} + \frac{\partial}{\partial \alpha_{L_j}} \right) \text{Tr} [\widehat{\Lambda}] + \frac{e^{-i\theta}}{2} \prod_j \alpha_{L_j} \prod_k \left(\beta_{L'_k} + \frac{\partial}{\partial \alpha_{L'_k}} \right) \text{Tr} [\widehat{\Lambda}] \\ &= \frac{e^{i\theta+z_0}}{2} \prod_j \beta_{L_j} \prod_k \alpha_{L'_k} + \frac{e^{-i\theta+z_0}}{2} \prod_k \beta_{L'_k} \prod_j \alpha_{L_j}. \end{aligned} \quad (5.11)$$

Similarly,

$$\text{Tr} [\widehat{Q}\widehat{\Lambda}^\dagger] = \frac{e^{i\theta+z_0^*}}{2} \prod_j \alpha_{L_j}^* \prod_k \beta_{L'_k}^* + \frac{e^{-i\theta+z_0^*}}{2} \prod_k \alpha_{L'_k}^* \prod_j \beta_{L_j}^*. \quad (5.12)$$

These can be entered directly into the general observable estimate expression (3.14) giving

$$\begin{aligned} \overline{Q}(\{L_j, L'_k\}, \theta) &= \frac{\langle \text{Re} \{ \frac{1}{2} e^{z_0} (e^{i\theta} \prod_{jk} \beta_{L_j} \alpha_{L'_k} + e^{-i\theta} \prod_{jk} \beta_{L'_k} \alpha_{L_j}) \} \rangle_{\text{stoch}}}{\langle \text{Re} \{ e^{z_0} \} \rangle_{\text{stoch}}} \\ \langle \widehat{Q}(\{L_j, L'_k\}, \theta) \rangle &= \lim_{S \rightarrow \infty} \overline{Q}(\{L_j, L'_k\}, \theta). \end{aligned} \quad (5.13)$$

Comparing (5.10) and (5.13), one sees that the same form appears in both, and so an algorithm to determine the finite-sample estimate of the expectation value of an arbitrary observable $\langle \widehat{O} \rangle$ can be given:

1. Normally-order \widehat{O} , using (5.2) by placing all annihilation operators to the right.
2. The numerator of the stochastic estimate \overline{O} is formed by replacing \widehat{a}_k , \widehat{a}_j^\dagger , and \otimes in the normally-ordered expression for \widehat{O} by α_k , β_j , and \prod , respectively, and finally multiplying by the global weight $e^{z_0} = \Omega$.
3. Take the average of the real part of the numerator terms, and divide by the average of the real part of the weight $\langle \text{Re} \{ e^{z_0} \} \rangle_{\text{stoch}}$.

In cases where the normalization of $\hat{\rho}$ is known to be conserved (e.g. dynamical master equations (2.20), but not thermodynamic (2.29)), the average in the denominator can be dropped, because it is known to always be equal to one in the limit of many trajectories. Its variations about unity are an indication of the sampling error in the calculation, but are not necessary to obtain expectation values.

The above algorithm applies also when \hat{O} is an infinite sum of moments, such as for example the parity operator at subsystem k , which can be written as

$$\hat{\pi}_k = \sum_{n=0}^{\infty} (-1)^n |n\rangle_k \langle n|_k \quad (5.14)$$

in terms of Fock number states $|n\rangle_k$ with occupation n at subsystem k . Since $|n\rangle_k = (1/\sqrt{n!})(\hat{a}_k^\dagger)^n |0\rangle$, the one obtains

$$\langle \hat{\pi}_k \rangle = \lim_{S \rightarrow \infty} \frac{\langle \text{Re} \{ \exp(z_0 - 2\alpha_k \beta_k) \} \rangle_{\text{stoch}}}{\langle \text{Re} \{ e^{z_0} \} \rangle_{\text{stoch}}}. \quad (5.15)$$

With such operators of infinite order in annihilation and creation operators, one should, however, be wary of boundary term errors (of the second kind), as explained in Section 6.2.1.

Finally, for estimates of fidelity using expression (3.22), one requires the trace of kernel products. Using (5.4) and the properties of Bargmann states (5.3), one obtains

$$\text{Tr} \left[\hat{\Lambda}(\boldsymbol{\alpha}_1, \boldsymbol{\beta}_1, \Omega_1) \hat{\Lambda}(\boldsymbol{\alpha}_2, \boldsymbol{\beta}_2, \Omega_2) \right] = \Omega_1 \Omega_2 \exp [-(\boldsymbol{\alpha}_1 - \boldsymbol{\alpha}_2) \cdot (\boldsymbol{\beta}_1 - \boldsymbol{\beta}_2)] \quad (5.16a)$$

$$\text{Tr} \left[\hat{\Lambda}(\boldsymbol{\alpha}_1, \boldsymbol{\beta}_1, \Omega_1) \hat{\Lambda}^\dagger(\boldsymbol{\alpha}_2^*, \boldsymbol{\beta}_2^*, \Omega_2^*) \right] = \Omega_1 \Omega_2^* \exp [-(\boldsymbol{\alpha}_1 - \boldsymbol{\beta}_2^*) \cdot (\boldsymbol{\beta}_1 - \boldsymbol{\alpha}_2^*)] \quad (5.16b)$$

5.3 Dynamics of locally-interacting Bose gas

The master equation for the dynamics of a locally interacting Bose gas on a lattice is (2.20) with the Hamiltonian given by (2.17). Using the operator correspondences (5.7), and the methods of Sections 3.4.1, 3.4.3, one can obtain an FPE for this system. Subsequently using the methods of Section 3.4.2, the square root form of the noise matrix $\underline{B} = \underline{B}_0 = \sqrt{\underline{D}}$, and standard drift gauges, the stochastic equations given below are obtained from (4.25) and (4.26) (or more directly, from

(4.90), setting $g_{jk} = 0$). Diffusion gauges have been omitted at this stage (so, $O(g_{jk}) = I$) for clarity. Note that the M spatial modes are now labeled by the lattice labels \mathbf{n} , or \mathbf{m} as defined in Section 2.3. Each mode is a “subsystem”.

It was chosen to separate the noise contributions from each process, as discussed in Section 4.4.7. This is done so that the noise terms take on a simple form.

With no environment interactions, the purely Hamiltonian evolution is found to lead to the Ito stochastic equations

$$d\alpha_{\mathbf{n}} = -i \sum_{\mathbf{m}} \omega_{\mathbf{nm}} \alpha_{\mathbf{m}} dt - 2i\chi \alpha_{\mathbf{n}}^2 \beta_{\mathbf{n}} dt + i\alpha_{\mathbf{n}} \sqrt{2i\chi} (dW_{\mathbf{n}} - \mathcal{G}_{\mathbf{n}}), \quad (5.17a)$$

$$d\beta_{\mathbf{n}} = i \sum_{\mathbf{m}} \omega_{\mathbf{mn}} \beta_{\mathbf{m}} dt + 2i\chi \alpha_{\mathbf{n}} \beta_{\mathbf{n}}^2 dt + \beta_{\mathbf{n}} \sqrt{2i\chi} (d\widetilde{W}_{\mathbf{n}} - \widetilde{\mathcal{G}}_{\mathbf{n}}), \quad (5.17b)$$

$$d\Omega = \Omega \sum_{\mathbf{n}} \left[\mathcal{G}_{\mathbf{n}} dW_{\mathbf{n}} + \widetilde{\mathcal{G}}_{\mathbf{n}} d\widetilde{W}_{\mathbf{n}} \right], \quad (5.17c)$$

with $2M$ independent real Wiener increments $dW_{\mathbf{n}}$ and $d\widetilde{W}_{\mathbf{n}}$. The corresponding complex drift gauges are $\mathcal{G}_{\mathbf{n}}$ and $\widetilde{\mathcal{G}}_{\mathbf{n}}$.

Addition of a heat bath at temperature T results in the following additions:

$$d\alpha_{\mathbf{n}} = \dots - \frac{\gamma_{\mathbf{n}}}{2} \alpha_{\mathbf{n}} dt + \sqrt{\gamma_{\mathbf{n}} \bar{n}_{\text{bath}}(T)} (d\eta_{\mathbf{n}} - \mathcal{G}_{\mathbf{n}}^{(1)}) \quad (5.18a)$$

$$d\beta_{\mathbf{n}} = \dots - \frac{\gamma_{\mathbf{n}}}{2} \beta_{\mathbf{n}} dt + \sqrt{\gamma_{\mathbf{n}} \bar{n}_{\text{bath}}(T)} (d\eta_{\mathbf{n}}^* - \widetilde{\mathcal{G}}_{\mathbf{n}}^{(1)}) \quad (5.18b)$$

$$d\Omega = \dots + \Omega \sum_{\mathbf{n}} \left[d\eta_{\mathbf{n}} \widetilde{\mathcal{G}}_{\mathbf{n}}^{(1)} + d\eta_{\mathbf{n}}^* \mathcal{G}_{\mathbf{n}}^{(1)} \right], \quad (5.18c)$$

where \bar{n}_{bath} is given by the Bose-Einstein distribution (2.23). Here the *complex* noises $d\eta_{\mathbf{n}}$ are independent, and satisfy

$$\langle d\eta_{\mathbf{n}} d\eta_{\mathbf{m}}^* \rangle_{\text{stoch}} = \delta_{\mathbf{nm}} dt, \quad (5.19a)$$

$$\langle d\eta_{\mathbf{n}} d\eta_{\mathbf{m}} \rangle_{\text{stoch}} = 0, \quad (5.19b)$$

$$\langle d\eta_{\mathbf{n}} \rangle_{\text{stoch}} = 0. \quad (5.19c)$$

The drift gauges $\mathcal{G}_{\mathbf{n}}^{(1)}$ and $\widetilde{\mathcal{G}}_{\mathbf{n}}^{(1)}$ are also complex.

Two-particle losses to a zero temperature heat bath result in the following addi-

tions to the dynamical equations:

$$d\alpha_{\mathbf{n}} = \cdots - \gamma_{\mathbf{n}}^{(2)} \alpha_{\mathbf{n}}^2 \beta_{\mathbf{n}} dt + i\alpha_{\mathbf{n}} \sqrt{\gamma_{\mathbf{n}}^{(2)}} (dW_{\mathbf{n}}^{(2)} - \mathcal{G}_{\mathbf{n}}^{(2)}), \quad (5.20a)$$

$$d\beta_{\mathbf{n}} = \cdots - \gamma_{\mathbf{n}}^{(2)} \alpha_{\mathbf{n}} \beta_{\mathbf{n}}^2 dt + i\beta_{\mathbf{n}} \sqrt{\gamma_{\mathbf{n}}^{(2)}} (d\widetilde{W}_{\mathbf{n}}^{(2)} - \widetilde{\mathcal{G}}_{\mathbf{n}}^{(2)}), \quad (5.20b)$$

$$d\Omega = \cdots + \Omega \sum_{\mathbf{n}} \left[\mathcal{G}_{\mathbf{n}}^{(2)} dW_{\mathbf{n}}^{(2)} + \widetilde{\mathcal{G}}_{\mathbf{n}}^{(2)} d\widetilde{W}_{\mathbf{n}}^{(2)} \right]. \quad (5.20c)$$

Here, again, the noises $dW_{\mathbf{n}}^{(2)}$ and $d\widetilde{W}_{\mathbf{n}}^{(2)}$ are independent Wiener increments, and the drift gauges $\mathcal{G}_{\mathbf{n}}^{(2)}$ and $\widetilde{\mathcal{G}}_{\mathbf{n}}^{(2)}$ are complex.

Lastly, A coherent driving field leads to the deterministic corrections

$$d\alpha_{\mathbf{n}} = \cdots + \varepsilon dt, \quad (5.21a)$$

$$d\beta_{\mathbf{n}} = \cdots + \varepsilon^* dt. \quad (5.21b)$$

The numerical simulation strategy is (briefly)

1. Sample a trajectory according to the known initial condition $P_G(0) = P_G(\widehat{\rho}(0))$
2. Evolve according to the stochastic equations, calculating moments of interest, and accumulating appropriate sums of them.
3. Repeat for $\mathcal{S} \gg 1$ independent trajectories.

One could equally well remain with $2M \times 2M$ complex noise matrix, which would use only $2M$ real noises and $2M$ drift gauges, instead of the $6M$ real noises and $6M$ complex drift gauges above. In general, however, the stochastic terms would then have a complicated dependence on the parameters χ , \bar{n}_{bath} , $\gamma_{\mathbf{n}}$, and $\gamma_{\mathbf{n}}^{(2)}$ — possibly undesirable. On the other hand, the amount of noise in the simulation might be reduced, because the effects of one process may partly cancel the effects of another. This may be particularly so if both scattering $\propto \chi$ and two-boson heat bath interactions $\propto \gamma_{\mathbf{n}}^{(2)}$ are present. These two processes lead to similar terms in the equations (apart from factors of i , etc.), and may be expected to cancel some noise without causing excessively complicated equations.

Some brief comments about the behavior of the above equations:

- The (un-gauged) equations for $d\beta_{\mathbf{n}}$ are simply the complex conjugates of $d\alpha_{\mathbf{n}}$, apart from possessing independent noises.

- The Hamiltonian evolution (5.17) leads to a nonlinear equation in the coherent amplitudes, with obligatory noise.
- All interactions with a heat bath (irrespective of the temperature) cause a deterministic exponential loss of particles.
- Finite temperature thermal interactions cause a directionless randomization of the coherent amplitudes α and β , leading also to a mean growth of boson numbers.
- Two-particle losses also lead to nonlinear equations and noise.
- The noise from scattering and two-particle losses is directional in phase space, as opposed to thermal noise.
- Coherent gain causes no noise in the equations.

5.4 Comparison to Gross-Pitaevskii semiclassical equations

The lossless equations (5.17) are similar in form to the Gross-Pitaevskii (GP) semiclassical equations ubiquitous in calculations on Bose-Einstein Condensates at temperatures well below condensation. (For details of these gases and results that are obtained with the GP equations see e.g. the review by Dalfovo *et al* [31].) Derivation of the GP equations basically assumes the particles all coherently occupy a single orbital, and can be described by its wavefunction.

In fact, if one

1. Ignores stochastic and drift gauge terms in (5.17).
2. Makes the assumption that the field is coherent — i.e. the local kernels are coherent state projectors $\hat{\Lambda}_{\mathbf{n}} \rightarrow |\alpha_{\mathbf{n}}\rangle_{\mathbf{n}} \langle\alpha_{\mathbf{n}}|_{\mathbf{n}}$, implying

$$\beta \rightarrow \alpha^*. \tag{5.22}$$

Note that this is consistent with the first assumption of no stochastic terms, since $d\beta_{\mathbf{n}}$ differs from $d\alpha_{\mathbf{n}}^*$ only by having independent noises.

3. Makes a variable change

$$\psi_{\mathbf{n}} = \psi(\mathbf{x}_{\mathbf{n}}) = \frac{\alpha_{\mathbf{n}}}{\sqrt{\prod_d \Delta x_d}} = \frac{\beta_{\mathbf{n}}^*}{\sqrt{\prod_d \Delta x_d}}, \quad (5.23)$$

such that ψ is the order parameter (i.e. the wavefunction normalized to $\bar{N} = \int \psi(\mathbf{x}) d^{\mathcal{D}}\mathbf{x}$, where \bar{N} is the mean particle number):

then precisely the (lattice) GP equations for $\psi(\mathbf{x}_{\mathbf{n}})$ are obtained. From this, it can be surmised that in regimes where the GP equations are a reasonably good approximation, the noise will be relatively small and calculations precise.

It is quite remarkable that just by the addition of simple stochastic terms, full first-principles quantum evolution is recovered from a mean field theory. This convenient property suggests that full quantum simulations using this method have the potential to remain numerically tractable.

5.5 Extended interparticle interactions

Equations for the case of extended interparticle interactions as in the Hamiltonian (2.12) will be derived here. Proceeding as in Section 5.3, the complex diffusion matrix (before introducing any drift gauges) in the FPE can now be written

$$\underline{D} = \begin{bmatrix} D^{(\alpha)} & 0 \\ 0 & D^{(\beta)} \end{bmatrix}, \quad (5.24)$$

where $D^{(\alpha)}$ appears in the FPE as $\frac{1}{2}D_{\mathbf{nm}}^{(\alpha)}\partial^2/\partial\alpha_{\mathbf{n}}\partial\alpha_{\mathbf{m}}$, and $D^{(\beta)}$ as $\frac{1}{2}D_{\mathbf{nm}}^{(\beta)}\partial^2/\partial\beta_{\mathbf{n}}\partial\beta_{\mathbf{m}}$. Their matrix elements are given by

$$D_{\mathbf{nm}}^{(\alpha)} = -i\frac{u_{\mathbf{nm}}}{\hbar}\alpha_{\mathbf{n}}\alpha_{\mathbf{m}}, \quad (5.25a)$$

$$D_{\mathbf{nm}}^{(\beta)} = i\frac{u_{\mathbf{nm}}}{\hbar}\beta_{\mathbf{n}}\beta_{\mathbf{m}}. \quad (5.25b)$$

Since the interaction potential $u_{\mathbf{nm}}$ is symmetric, then it could be orthogonally decomposed by its matrix square root v , which satisfies

$$vv = u = vv^T. \quad (5.26)$$

The matrix v could be calculated at the beginning of a simulation and subsequently used in noise matrices satisfying $B^{(z)}[B^{(z)}]^T = D^{(z)}$ (with z taking on the labels α or β), where

$$B_{\mathbf{nm}}^{(\alpha)} = \sqrt{\frac{-i}{\hbar}} \alpha_{\mathbf{n}} v_{\mathbf{nm}} \quad (5.27a)$$

$$B_{\mathbf{nm}}^{(\beta)} = \sqrt{\frac{i}{\hbar}} \beta_{\mathbf{n}} v_{\mathbf{nm}}. \quad (5.27b)$$

Unfortunately finding the matrix square root would usually need to be done numerically, which would require storing $\mathcal{O}(M^2)$ matrix elements $v_{\mathbf{nm}}$ and take a time $\mathcal{O}(M^4)$ to compute — not efficient for large lattices.

A much more efficient, though involved, approach is possible. One would like to obtain some orthogonal decomposition of \underline{D} dependent directly on $U_{\mathbf{n}}$ that depends only on the interparticle spacing. This potential has M elements, rather than the $M \times M$ of $u_{\mathbf{nm}}$. Writing the Langevin equations in terms of stochastic increments $dX_{\mathbf{n}}^{(z)}$ (See Section 3.4.2) as

$$dz_{\mathbf{n}} = A_{\mathbf{n}}^{(z)} dt + dX_{\mathbf{n}}^{(z)}, \quad (5.28)$$

the stochastic increments must obey

$$\langle dX_{\mathbf{n}}^{(z)} \rangle_{\text{stoch}} = 0 \quad (5.29a)$$

$$\langle dX_{\mathbf{n}}^{(z)} dX_{\mathbf{m}}^{(z)} \rangle_{\text{stoch}} = D_{\mathbf{nm}}^{(z)} dt \quad (5.29b)$$

$$\langle dX_{\mathbf{n}}^{(\alpha)} dX_{\mathbf{m}}^{(\beta)} \rangle_{\text{stoch}} = 0. \quad (5.29c)$$

Firstly, we can define new stochastic increments $dY_{\mathbf{n}}^{(\alpha)}$ and $dY_{\mathbf{n}}^{(\beta)}$, such that

$$dX_{\mathbf{n}}^{(\alpha)} = \alpha_{\mathbf{n}} \sqrt{\frac{-i}{\hbar}} dY_{\mathbf{n}}^{(\alpha)} \quad (5.30a)$$

$$dX_{\mathbf{n}}^{(\beta)} = \beta_{\mathbf{n}} \sqrt{\frac{i}{\hbar}} dY_{\mathbf{n}}^{(\beta)}, \quad (5.30b)$$

and these must obey relationships (Remembering from (2.15) that $u_{\mathbf{nm}} = u_{\mathbf{mn}} = U_{|\mathbf{n}-\mathbf{m}|}$),

$$\langle dY_{\mathbf{n}}^{(z)} \rangle_{\text{stoch}} = 0 \quad (5.31a)$$

$$\langle dY_{\mathbf{n}}^{(z)} dY_{\mathbf{m}}^{(z)} \rangle_{\text{stoch}} = U_{|\mathbf{n}-\mathbf{m}|} dt, \quad (5.31b)$$

which now depend only on the *a-priori* potential $U_{\mathbf{n}}$, not on the dynamically evolving mode amplitudes. Note that while the variances of the $dY_{\mathbf{n}}^{(z)}$ are the same for both choices of z , the actual realizations of the increments must be independent.

It is useful to now consider the \mathcal{D} -dimensional discrete Fourier transform of the interparticle potential:

$$\tilde{U}_{\tilde{\mathbf{n}}} = \tilde{U}'_{\tilde{\mathbf{n}}} + i\tilde{U}''_{\tilde{\mathbf{n}}} = \frac{1}{C_{\text{norm}}} \sum_{\mathbf{n}} U_{\mathbf{n}} e^{-i\mathbf{k}_{\tilde{\mathbf{n}}}\cdot\mathbf{x}_{\mathbf{n}}}, \quad (5.32)$$

with the normalization constant $C_{\text{norm}} = (2\pi)^{\mathcal{D}/2} / \prod_d \Delta x_d$. All the lattice notation used from here on has been defined in Section 2.3. Note now that since $U_{\mathbf{n}}$ is real, and $\mathbf{k}_{-\tilde{\mathbf{n}}} = -\mathbf{k}_{\tilde{\mathbf{n}}}$, then (5.32) implies

$$\tilde{U}'_{\tilde{\mathbf{n}}} = \tilde{U}'_{-\tilde{\mathbf{n}}} \quad (5.33a)$$

$$\tilde{U}''_{\tilde{\mathbf{n}}} = -\tilde{U}''_{-\tilde{\mathbf{n}}}. \quad (5.33b)$$

The inverse transform is

$$U_{\mathbf{n}} = \frac{C_{\text{norm}}}{M} \sum_{\tilde{\mathbf{n}}} \tilde{U}_{\tilde{\mathbf{n}}} e^{i\mathbf{k}_{\tilde{\mathbf{n}}}\cdot\mathbf{x}_{\mathbf{n}}}. \quad (5.34)$$

Expanding out the elements of lattice coordinate vectors, one can write $|\mathbf{x}_{\mathbf{n}} - \mathbf{x}_{\mathbf{m}}| = \{\vartheta_1 \Delta x_1 (n_1 - m_1), \dots, \vartheta_{\mathcal{D}} \Delta x_{\mathcal{D}} (n_{\mathcal{D}} - m_{\mathcal{D}})\}$, where the quantities ϑ_d can take on the values $+1$ or -1 , depending on what is required to take the modulus. Recalling the symmetry property of the interparticle potential that was assumed in Section 2.3 ($U_{\{n_1, \dots, n_d, \dots\}} = U_{\{n_1, \dots, M_d - n_d, \dots\}}$ for any dimension d), one obtains

$$\begin{aligned} U_{\{n_1, \dots, n_d, \dots\}} &= \frac{C_{\text{norm}}}{M} \sum_{\tilde{\mathbf{n}}} \tilde{U}_{\tilde{\mathbf{n}}} \prod_d e^{ik_d(\tilde{n}_d)x_d(n_d)} \\ &= \frac{C_{\text{norm}}}{M} \sum_{\tilde{\mathbf{n}}} \tilde{U}_{\tilde{\mathbf{n}}} e^{-ik_d(\tilde{n}_d)x_d(n_d)} \prod_{d' \neq d} e^{ik_{d'}(\tilde{n}_{d'})x_{d'}(n_{d'})}, \end{aligned} \quad (5.35)$$

since $\exp[ik_d M_d \Delta x_d] = 1$ for any k_d . So, the phase for dimension d can have either sign. This result can then be applied to each dimension where $\vartheta_d = -1$ to obtain (using also $\mathbf{x}_{\mathbf{n}-\mathbf{m}} = \mathbf{x}_{\mathbf{n}} - \mathbf{x}_{\mathbf{m}}$) the expression (subtly different from (5.34))

$$U_{|\mathbf{n}-\mathbf{m}|} = \frac{C_{\text{norm}}}{M} \sum_{\tilde{\mathbf{n}}} \tilde{U}_{\tilde{\mathbf{n}}} e^{i\mathbf{k}_{\tilde{\mathbf{n}}}\cdot(\mathbf{x}_{\mathbf{n}} - \mathbf{x}_{\mathbf{m}})}. \quad (5.36)$$

If one introduces new stochastic increments $dZ_{\mathbf{n}}^{(v)}$, then using (5.36), condition (5.31b) is equivalent to

$$\langle dY_{\mathbf{n}}^{(z)} dY_{\mathbf{m}}^{(z)} \rangle_{\text{stoch}} = \frac{C_{\text{norm}}}{M} \sum_{\tilde{\mathbf{n}}\tilde{\mathbf{m}}} e^{i\mathbf{k}_{\tilde{\mathbf{n}}}\cdot\mathbf{x}_{\mathbf{n}}} e^{-i\mathbf{k}_{\tilde{\mathbf{m}}}\cdot\mathbf{x}_{\mathbf{m}}} \sqrt{\tilde{U}_{\tilde{\mathbf{n}}}\tilde{U}_{\tilde{\mathbf{m}}}} \langle dZ_{\tilde{\mathbf{n}}}^{(v)} dZ_{\tilde{\mathbf{m}}}^{(v)} \rangle_{\text{stoch}} \quad (5.37)$$

provided that

$$\langle dZ_{\tilde{\mathbf{n}}}^{(z)} \rangle_{\text{stoch}} = 0 \quad (5.38a)$$

$$\langle dZ_{\tilde{\mathbf{n}}}^{(z)} dZ_{\tilde{\mathbf{m}}}^{(z)} \rangle_{\text{stoch}} = \delta_{\tilde{\mathbf{n}}\tilde{\mathbf{m}}} dt. \quad (5.38b)$$

This can be checked by substitution. This is not yet quite what one wants to be able to decompose into $dY_{\mathbf{n}}^{(z)}$ because the second phase in (5.37) has the wrong sign. Again using $-\mathbf{k}_{\tilde{\mathbf{n}}} = \mathbf{k}_{-\tilde{\mathbf{n}}}$, to relabel $\tilde{\mathbf{m}} \rightarrow -\tilde{\mathbf{m}}$, and applying (5.33) one has

$$\langle dY_{\mathbf{n}}^{(z)} dY_{\mathbf{m}}^{(z)} \rangle_{\text{stoch}} = \frac{C_{\text{norm}}}{M} \sum_{\tilde{\mathbf{n}}\tilde{\mathbf{m}}} e^{i\mathbf{k}_{\tilde{\mathbf{n}}}\cdot\mathbf{x}_{\mathbf{n}}} e^{i\mathbf{k}_{\tilde{\mathbf{m}}}\cdot\mathbf{x}_{\mathbf{m}}} \sqrt{\tilde{U}_{\tilde{\mathbf{n}}}\tilde{U}_{\tilde{\mathbf{m}}}^*} \delta_{\mathbf{n},-\mathbf{m}} dt. \quad (5.39)$$

The phase factor is now fine, but the complex conjugate $U_{\tilde{\mathbf{m}}}^*$ spoils the potential decomposition. What is needed are some stochastic increments $d\tilde{Z}_{\tilde{\mathbf{n}}}^{(z)}$ that will satisfy

$$\langle d\tilde{Z}_{\tilde{\mathbf{n}}}^{(z)} \rangle_{\text{stoch}} = 0 \quad (5.40a)$$

$$\langle d\tilde{Z}_{\tilde{\mathbf{n}}}^{(z)} d\tilde{Z}_{\tilde{\mathbf{m}}}^{(z)} \rangle_{\text{stoch}} = \delta_{\tilde{\mathbf{n}},-\tilde{\mathbf{m}}} \tilde{U}_{\tilde{\mathbf{n}}} dt, \quad (5.40b)$$

to allow then a decomposition

$$dY_{\mathbf{n}}^{(z)} = \sqrt{\frac{C_{\text{norm}}}{M}} \sum_{\tilde{\mathbf{n}}} e^{i\mathbf{k}_{\tilde{\mathbf{n}}}\cdot\mathbf{x}_{\mathbf{n}}} d\tilde{Z}_{\tilde{\mathbf{n}}}^{(z)}. \quad (5.41)$$

This can be achieved by separating out the real and imaginary parts of \tilde{U} as

$$d\tilde{Z}_{\tilde{\mathbf{n}}}^{(z)} = \sqrt{\tilde{U}'_{\tilde{\mathbf{n}}}} d\zeta_{\tilde{\mathbf{n}}}^{(z)} + \sqrt{\tilde{U}''_{\tilde{\mathbf{n}}}} d\tilde{\zeta}_{\tilde{\mathbf{n}}}^{(z)}, \quad (5.42)$$

with the new stochastic increments

$$\langle d\zeta_{\tilde{\mathbf{n}}}^{(z)} \rangle_{\text{stoch}} = \langle d\tilde{\zeta}_{\tilde{\mathbf{n}}}^{(z)} \rangle_{\text{stoch}} = 0 \quad (5.43a)$$

$$\langle d\zeta_{\tilde{\mathbf{n}}}^{(z)} d\zeta_{\tilde{\mathbf{m}}}^{(z)} \rangle_{\text{stoch}} = \delta_{\tilde{\mathbf{n}},-\tilde{\mathbf{m}}} dt \quad (5.43b)$$

$$\langle d\tilde{\zeta}_{\tilde{\mathbf{n}}}^{(z)} d\tilde{\zeta}_{\tilde{\mathbf{m}}}^{(z)} \rangle_{\text{stoch}} = \delta_{\tilde{\mathbf{n}},-\tilde{\mathbf{m}}} dt \quad (5.43c)$$

$$\langle d\zeta_{\tilde{\mathbf{n}}}^{(z)} d\tilde{\zeta}_{\tilde{\mathbf{m}}}^{(z)} \rangle_{\text{stoch}} = 0. \quad (5.43d)$$

Note that $\sqrt{\widetilde{U}_{\tilde{\mathbf{n}}}''\widetilde{U}_{-\tilde{\mathbf{n}}}''} = i\widetilde{U}_{\tilde{\mathbf{n}}}''$ by (5.33). Let us divide the $\tilde{\mathbf{n}} \neq \mathbf{0}$ momentum mode space into two symmetric halves \mathcal{R} and $\widetilde{\mathcal{R}}$, such that e.g. when $\mathcal{D} = 3$:

$$\tilde{\mathbf{n}} \in \mathcal{R} \text{ if } \begin{cases} \tilde{n}_1 > 0 \\ \text{or, } \tilde{n}_1 = 0 & \text{and } \tilde{n}_2 > 0 \\ \text{or, } \tilde{n}_1 = \tilde{n}_2 = 0 & \text{and } \tilde{n}_3 > 0 \end{cases} \quad (5.44)$$

while $\tilde{\mathbf{n}} \in \widetilde{\mathcal{R}}$ if $\tilde{\mathbf{n}} \notin \{\mathcal{R}, \mathbf{0}\}$. The $d\zeta_{\tilde{\mathbf{n}}}^{(z)}$ noises can now be realized by the construction

$$d\zeta_{\tilde{\mathbf{n}}}^{(z)} = \begin{cases} (dW_{\tilde{\mathbf{n}},1} + idW_{\tilde{\mathbf{n}},2})/\sqrt{2} & \text{if } \tilde{\mathbf{n}} \in \mathcal{R} \\ dW_{\mathbf{0},1} & \text{if } \tilde{\mathbf{n}} = \mathbf{0} \\ (dW_{-\tilde{\mathbf{n}},1} - idW_{-\tilde{\mathbf{n}},2})/\sqrt{2} & \text{if } \tilde{\mathbf{n}} \in \widetilde{\mathcal{R}}. \end{cases} \quad (5.45)$$

in terms of real independent Wiener increments $dW_{\tilde{\mathbf{n}},j}$, for all $\tilde{\mathbf{n}} \in \{\mathcal{R}, \mathbf{0}\}$, numbering M in total¹. $d\zeta_{\tilde{\mathbf{n}}}^{(z)}$ requires separate M independent noises $dW_{\tilde{\mathbf{n}},3}$ and $dW_{\tilde{\mathbf{n}},4}$.

Collecting all this together, one obtains

$$dX_{\mathbf{n}}^{(\alpha)} = \alpha_{\mathbf{n}} \sqrt{\frac{-iC_{\text{norm}}}{\hbar M}} \sum_{\tilde{\mathbf{n}}} e^{i\mathbf{k}_{\tilde{\mathbf{n}}}\cdot\mathbf{x}_{\mathbf{n}}} \left\{ \sqrt{\widetilde{U}_{\tilde{\mathbf{n}}}'} d\zeta_{\tilde{\mathbf{n}}}^{(\alpha)} + \sqrt{\widetilde{U}_{\tilde{\mathbf{n}}}''} d\tilde{\zeta}_{\tilde{\mathbf{n}}}^{(\alpha)} \right\} \quad (5.46a)$$

$$dX_{\mathbf{n}}^{(\beta)} = \beta_{\mathbf{n}} \sqrt{\frac{iC_{\text{norm}}}{\hbar M}} \sum_{\tilde{\mathbf{n}}} e^{i\mathbf{k}_{\tilde{\mathbf{n}}}\cdot\mathbf{x}_{\mathbf{n}}} \left\{ \sqrt{\widetilde{U}_{\tilde{\mathbf{n}}}'} d\zeta_{\tilde{\mathbf{n}}}^{(\beta)} + \sqrt{\widetilde{U}_{\tilde{\mathbf{n}}}''} d\tilde{\zeta}_{\tilde{\mathbf{n}}}^{(\beta)} \right\}. \quad (5.46b)$$

All four $\zeta_{\tilde{\mathbf{n}}}^{(z)}$ and $\tilde{\zeta}_{\tilde{\mathbf{n}}}^{(z)}$ complex noises per (Fourier space) lattice point are independent of each other, and of the noises at all other (momentum) lattice points and times. Explicitly they have the form (5.45) in terms of the $4M$ real Wiener increments² $dW_{\tilde{\mathbf{n}},j}$. Storage space for M complex variables $\widetilde{U}_{\tilde{\mathbf{n}}}$ is required, and calculation of these³ takes a time $\propto M \log M$ — much more tractable than the calculations of $v_{\mathbf{nm}}$. The $dX_{\mathbf{n}}^{(z)}$ reduce to the noise terms of (5.17) under local interactions $U_{\mathbf{n}} = 2\hbar\chi\delta_{\mathbf{n},\mathbf{0}}$.

With no drift gauges, the two-particle interaction terms of the Ito stochastic equations become (the terms due to other processes are unchanged)

$$d\alpha_{\mathbf{n}} = \dots - i \sum_{\mathbf{m}} \frac{U_{|\mathbf{n}-\mathbf{m}|}}{\hbar} \alpha_{\mathbf{n}} \alpha_{\mathbf{m}} \beta_{\mathbf{m}} + dX_{\mathbf{n}}^{(\alpha)}, \quad (5.47a)$$

$$d\beta_{\mathbf{n}} = \dots + i \sum_{\mathbf{m}} \frac{U_{|\mathbf{n}-\mathbf{m}|}}{\hbar} \beta_{\mathbf{n}} \alpha_{\mathbf{m}} \beta_{\mathbf{m}} + dX_{\mathbf{n}}^{(\beta)}. \quad (5.47b)$$

¹ $M + 1$ if M is even.

²Or $4(M + 1)$ if M is even.

³Using the “fast Fourier transform” algorithm.

Drift gauges can be introduced by making the replacements

$$dW_{\mathbf{n},j} \rightarrow dW_{\mathbf{n},j} - \mathcal{G}_{\mathbf{n},j} dt, \quad (5.48)$$

in (5.45) (this follows straight from the standard form (4.90)) and

$$d\Omega = \cdots + \Omega \sum_{\mathbf{n}} \sum_j \mathcal{G}_{\mathbf{n},j} dW_{\mathbf{n},j}. \quad (5.49)$$

Simulations of such models pose no fundamental problem (see e.g. Sections 10.5 and 10.6), however the required computer time scales more steeply with M . There are now $2M$ complex terms to calculate in the equation for each variable rather than the 2 for a local interaction model (5.17).

5.6 Thermodynamics of interacting Bose gas

The grand canonical thermodynamics of a system with Hamiltonian \hat{H} in thermal and diffusive contact with a reservoir at temperature T and chemical potential μ can be simulated using the master equation (2.29).

With the locally-interacting Hamiltonian (2.17), using the operator correspondences (5.7), and the methods of Sections 3.4.1, 3.4.3, one can obtain the FPE, and then directly by (4.90) (not using diffusion gauges here, so $g_{jk} = 0$) the stochastic equations. Using the same notation as in Section 5.3, these can be written

$$d\alpha_{\mathbf{n}} = -\hbar \sum_{\mathbf{m}} \omega_{\mathbf{nm}} \alpha_{\mathbf{m}} d\tau + (\mu_e - 2\hbar\chi\alpha_{\mathbf{n}}\beta_{\mathbf{n}}) \alpha_{\mathbf{n}} d\tau + i\alpha_{\mathbf{n}} \sqrt{2\hbar\chi} (dW_{\mathbf{n}} - \mathcal{G}_{\mathbf{n}}) \quad (5.50a)$$

$$d\beta_{\mathbf{n}} = 0, \quad (5.50b)$$

$$d\Omega = \Omega \left[-\hbar \sum_{\mathbf{nm}} \omega_{\mathbf{nm}} \alpha_{\mathbf{m}} \beta_{\mathbf{n}} d\tau + \sum_{\mathbf{n}} (\mu_e - \hbar\chi\alpha_{\mathbf{n}}\beta_{\mathbf{n}}) \alpha_{\mathbf{n}} \beta_{\mathbf{n}} d\tau + \sum_{\mathbf{n}} \mathcal{G}_{\mathbf{n}} dW_{\mathbf{n}} \right] \quad (5.50c)$$

There are M real Wiener increments $dW_{\mathbf{n}}$, and hence the same number of complex drift gauges $\mathcal{G}_{\mathbf{n}}$.

The asymmetric form (no $\beta_{\mathbf{n}}$ evolution) arises because the Hamiltonian acts only from the left on the density matrix in (2.27). The initial $\beta_{\mathbf{n}}$ takes on a range of random initial values (see (5.59) below), which then remain constant. A symmetric set of equations is also possible by starting from the middle (anticommutator) term

of (2.27), however this appears to serve no useful purpose but needs more noises (another set of M for the $\beta_{\mathbf{n}}$ evolution).

The initial condition (2.31) is $\hat{\rho}_u(0) = \exp[-\lambda_n \hat{N}] = \otimes_{\mathbf{n}} \hat{\rho}_{\mathbf{n}}(0)$, where λ_n is given by (2.32), and with

$$\hat{\rho}_{\mathbf{n}}(0) = \exp[-\lambda_n \hat{a}_{\mathbf{n}}^\dagger \hat{a}_{\mathbf{n}}]. \quad (5.51)$$

One could use (5.9) and (3.10) to obtain a gauge P distribution corresponding to this initial state, but a more compact distribution can be found as follows:

Since $\hat{\rho}_u(0)$ is separable, let us just consider the initial conditions in a single mode \mathbf{n} (all modes are in this same state initially). In a local Fock number state complete orthogonal basis $|n\rangle$, the initial state can be written

$$\begin{aligned} \hat{\rho}_{\mathbf{n}}(0) = \exp[-\lambda_n \hat{n}_{\mathbf{n}}] \hat{I}_{\mathbf{n}} &= \sum_p \frac{(-\lambda_n)^p}{p!} \hat{n}_{\mathbf{n}}^p \sum_n |n\rangle \langle n| \\ &= \sum_n e^{-\lambda_n n} |n\rangle \langle n|. \end{aligned} \quad (5.52)$$

The local kernel (5.4), on the other hand, expanding the Bargmann states, is

$$\begin{aligned} \hat{\Lambda}_{\mathbf{n}} &= e^{\alpha_{\mathbf{n}} \hat{a}_{\mathbf{n}}^\dagger} |0\rangle \langle 0| e^{\beta_{\mathbf{n}} \hat{a}_{\mathbf{n}}} e^{-\alpha_{\mathbf{n}} \beta_{\mathbf{n}}} \\ &= \sum_{n\tilde{n}} \frac{\alpha_{\mathbf{n}}^n \beta_{\mathbf{n}}^{\tilde{n}}}{\sqrt{n! \tilde{n}!}} |n\rangle \langle \tilde{n}| e^{-\alpha_{\mathbf{n}} \beta_{\mathbf{n}}}. \end{aligned} \quad (5.53)$$

Since the high temperature state should be classical let us try a gauge P distribution over just diagonal coherent states and so postulate $\boldsymbol{\alpha} = \boldsymbol{\beta}^*$. Also, at high temperature the modes should be separable, so let us try a distribution where the amplitudes at each mode are independent. From (5.51), the state of each mode should be identical, and lastly, for simplicity, let us choose the initial weight of each trajectory to be equal: $\Omega(0) = e^{z_0(0)} = 1$ Let us start, then, with the Gaussian ansatz

$$P_G^{\text{try}}(\boldsymbol{\alpha}, \boldsymbol{\beta}, z_0) = \delta^2(z_0) \delta^{2M}(\boldsymbol{\beta} - \boldsymbol{\alpha}^*) \prod_{\mathbf{n}} \frac{1}{2\pi\sigma^2} \exp\left(\frac{-|\alpha_{\mathbf{n}}|^2}{2\sigma^2}\right). \quad (5.54)$$

To see whether this is sufficient to represent the initial state, and to find the value of σ , let us substitute into (5.5) and see if (5.52) can be satisfied. Integrating over Ω and $\boldsymbol{\beta}$, and separating modes, one has (for each mode \mathbf{n}):

$$\hat{\rho}_{\mathbf{n}}^{\text{try}}(0) = \int \frac{1}{2\pi\sigma^2} \exp\left(\frac{-|\alpha_{\mathbf{n}}|^2}{2\sigma^2}\right) \sum_{n\tilde{n}} \frac{\alpha_{\mathbf{n}}^n (\alpha_{\mathbf{n}}^*)^{\tilde{n}}}{\sqrt{n! \tilde{n}!}} |n\rangle \langle \tilde{n}| e^{-|\alpha_{\mathbf{n}}|^2} d^2\alpha_{\mathbf{n}}. \quad (5.55)$$

Writing $\alpha_{\mathbf{n}} = r e^{i\theta}$, one has

$$\begin{aligned}\widehat{\rho}_{\mathbf{n}}^{\text{try}}(0) &= \frac{1}{2\pi\sigma^2} \sum_{n\tilde{n}} \frac{1}{\sqrt{n!\tilde{n}!}} \int_{-\pi}^{\pi} d\theta e^{i\theta(n-\tilde{n})} \int_0^{\infty} dr r^{n+\tilde{n}+1} \exp\left(-\frac{r^2}{1+1/2\sigma^2}\right) |n\rangle \langle\tilde{n}| \\ &= \frac{1}{\sigma^2} \sum_n \frac{1}{n!} \int_0^{\infty} dr r^{2n+1} \exp\left(-\frac{r^2}{1+1/2\sigma^2}\right) |n\rangle \langle n|,\end{aligned}\quad (5.56)$$

and using[72] $\int_0^{\infty} r^{2n+1} e^{-fr^2} dr = n!/2f^{n+1}$,

$$\widehat{\rho}_{\mathbf{n}}^{\text{try}}(0) = \frac{1}{1+2\sigma^2} \sum_n \left(\frac{\sigma^2}{1+2\sigma^2}\right)^n |n\rangle \langle n|. \quad (5.57)$$

One wants to have $\widehat{\rho}_{\mathbf{n}}^{\text{try}}(0) = \widehat{\rho}_{\mathbf{n}}(0)/\text{Tr}[\widehat{\rho}_{\mathbf{n}}(0)]$. Since $\text{Tr}[\widehat{\rho}_{\mathbf{n}}(0)] = 1/[1 - e^{-\lambda_n}]$, this implies

$$\sigma = \frac{1}{\sqrt{2(e^{\lambda_n} - 1)}} = \sqrt{\frac{\bar{n}_0}{2}}, \quad (5.58)$$

($\bar{n}_0 = 1/[e^{\lambda_n} - 1]$ is the mean occupation per mode). So then, it has been verified that a gauge P distribution for $\widehat{\rho}_u(0)$ is just a Gaussian in $\boldsymbol{\alpha}$.

$$P_G(\boldsymbol{\alpha}, \boldsymbol{\beta}, z_0) = \delta^2(z_0) \delta^{2M}(\boldsymbol{\beta} - \boldsymbol{\alpha}^*) \prod_{\mathbf{n}} \frac{1}{\pi \bar{n}_0} \exp\left(\frac{-|\alpha_{\mathbf{n}}|^2}{\bar{n}_0}\right). \quad (5.59)$$

This is easily sampled to obtain initial values of $\alpha_{\mathbf{n}} = \beta_{\mathbf{n}}^*$, and $z_0 = 0$ for each trajectory.

For the case of extended interparticle interactions $u_{\mathbf{nm}} = U_{|\mathbf{n}-\mathbf{m}|}$, a similar procedure can be followed as was done for the dynamics in Section 5.5. The diffusion matrix in the FPE is now

$$D_{\mathbf{nm}}^{(\alpha)} = -u_{\mathbf{nm}} \alpha_{\mathbf{n}} \alpha_{\mathbf{m}}, \quad (5.60a)$$

$$D_{\mathbf{nm}}^{(\beta)} = 0, \quad (5.60b)$$

rather than (5.25). With no drift gauges, the Ito stochastic equations are

$$d\alpha_{\mathbf{n}} = -\hbar \sum_{\mathbf{m}} \omega_{\mathbf{nm}} \alpha_{\mathbf{m}} d\tau + \mu_e \alpha_{\mathbf{n}} d\tau - \sum_{\mathbf{m}} U_{|\mathbf{n}-\mathbf{m}|} \alpha_{\mathbf{n}} \alpha_{\mathbf{m}} \beta_{\mathbf{m}} d\tau + dX_{\mathbf{n}}^{(\alpha)} \quad (5.61a)$$

$$d\beta_{\mathbf{n}} = 0, \quad (5.61b)$$

$$d\Omega = \Omega \left[-\hbar \sum_{\mathbf{nm}} \omega_{\mathbf{nm}} \alpha_{\mathbf{m}} \beta_{\mathbf{n}} + \sum_{\mathbf{n}} \left(\mu_e - \sum_{\mathbf{m}} \frac{U_{|\mathbf{n}-\mathbf{m}|}}{2} \alpha_{\mathbf{m}} \beta_{\mathbf{m}} \right) \alpha_{\mathbf{n}} \beta_{\mathbf{n}} \right] d\tau \quad (5.61c)$$

With the stochastic increments given by

$$dX_{\mathbf{n}}^{(\alpha)} = i\alpha_{\mathbf{n}} \sqrt{\frac{C_{\text{norm}}}{M}} \sum_{\tilde{\mathbf{n}}} e^{i\mathbf{k}_{\tilde{\mathbf{n}}} \cdot \mathbf{x}_{\mathbf{n}}} \left\{ \sqrt{\tilde{U}'_{\tilde{\mathbf{n}}}} d\zeta_{\tilde{\mathbf{n}}}^{(\alpha)} + \sqrt{\tilde{U}''_{\tilde{\mathbf{n}}}} d\tilde{\zeta}_{\tilde{\mathbf{n}}}^{(\alpha)} \right\}. \quad (5.62)$$

Here the noises $d\zeta_{\tilde{\mathbf{n}}}^{(\alpha)}$ and $d\tilde{\zeta}_{\tilde{\mathbf{n}}}^{(\alpha)}$ are given again by (5.45), but the $2M$ real Wiener increments $dW_{\tilde{\mathbf{n}},j}$ now have variance $d\tau$ instead of dt . Drift gauges can again be introduced by making the replacements

$$dW_{\mathbf{n},j} \rightarrow dW_{\mathbf{n},j} - \mathcal{G}_{\mathbf{n},j} d\tau, \quad (5.63)$$

in (5.45) and

$$d\Omega = \dots + \Omega \sum_{\mathbf{n}} \sum_j \mathcal{G}_{\mathbf{n},j} dW_{\mathbf{n},j}. \quad (5.64)$$

5.7 Comparison with historical distributions

To put the gauge P representation in perspective, let us compare to the more commonly used phase-space representations from the field of quantum optics where phase-space distributions have arguably been most used. This section is based on Section 2 of the published article by Drummond and Deuar[56]. The concepts and general layout of this section are due to P. D. Drummond.

To understand the reasons for development of the various distributions, it is useful to peruse the requirements listed in Section 3.6 that they have to satisfy to result in stochastic simulations of many-body models.

Phase-space mappings were first introduced by Wigner as the famous Wigner representation[57]. Historically, phase-space distributions have developed in three stages:

- **Stage one (classical-like phase space):** A classical-like phase space was used in which the number of real configuration variables in C was the same as the number of classical degrees of freedom. Typically the kernel is of the form of a diagonal projector, and not all quantum states can be represented by positive nonsingular P . This usually manifests itself as either a non positive-definite propagator or 3rd order terms in the FPE in master equations involving several-body processes. Either way, a quantum-equivalent stochastic

process is not recovered. The Wigner[57], Husimi Q[73], Glauber-Sudarshan P[58, 74, 59], and the Poisson[75, 76, 62] representations, among others, all fall into this category.

- **Stage two (doubled phase space):** By the use of off-diagonal kernels, representations in a higher-dimensional phase space were developed, for which a non-singular distribution P exists for all quantum states. These typically have at least two real configuration variables for each classical degree of freedom (one for the ket and one for the bra in the kernel). Examples of such representations are the Glauber R[58, 74] and the positive P[10, 11] representations. While these work very well for highly damped systems, in models with several-body processes and low damping, they typically develop unstable trajectories. This leads to large sampling uncertainties or even systematic “boundary term errors” [70, 71].
- **Stage three (global amplitude and stochastic gauges):** Addition of a global weight to the kernel allows the introduction of drift stochastic gauges as in Section 4.3, and (by appropriate gauge choice) modification of the stochastic equations to remove the instabilities. The gauge P representation introduced above is of this type, as is the stochastic wavefunction method of Carusotto *et al*[1, 65]. This work of Carusotto *et al* is in some ways complementary to that presented in this thesis, and has been developed approximately in parallel. One fundamental difference between the two representations is that the gauge P representation allows the particle number to vary, allowing open behavior such as lasing or evaporative cooling, whereas the stochastic wavefunction method is hardwired to a constant particle number N .

How do these various representations compare? A check of their applicability to interacting Bose gas simulations with Hamiltonians (2.17) or (2.12) is shown in Table 5.1. One can see that the early distributions were often hindered in obtaining a many-body simulation by a whole variety of problems.

For completeness, it should be mentioned that there are other distributions of the general form (3.1) that are more suited to non-stochastic calculations. For example

Table 5.1: Check list of required representation properties for the more commonly used phase-space representations, when applied to a lattice interacting Bose gas Hamiltonian of the forms (2.17) or (2.12), with no external coupling.

distribution type	positive real	complete non-singular	FPE ^a	positive propagator	stable unbiased	UV convergent	open systems
Wigner	no	yes	no	varies ^b	–	no	yes
Q	yes	yes	yes	no	–	no	yes
P	yes	no	yes	no	–	yes	yes
R	no	yes	yes	no	–	yes	yes
positive P	yes	yes	yes	yes	no	yes	yes
sw. ^c Fock	yes	yes	yes	yes	yes ^d	yes	no
sw. ^b coherent	yes	yes	yes	yes	yes ^c	yes	no
gauge P	yes	yes	yes	yes	yes ^c	yes	yes

^ai.e. only 1st and 2nd partial derivative terms in $\partial P/\partial t$.

^bThere are no second order terms in the FPE for dynamics, while for thermodynamics the propagator may or may not be positive semi-definite, depending on occupation of modes.

^cStochastic wavefunction

^dGiven an appropriate choice of gauge — see Chapter 7

the symplectic tomography scheme of Mancini *et al*[77, 78], which expresses the quantum state as a probability distribution of a quadrature observable depending on a range of lab parameters. This has been used to investigate quantum entanglement and failure of local realism, but apparently has not led to many-mode quantum simulations, presumably due to the lack of a positive propagator in nonlinear evolution. The complex P representation[11] allows one to derive exact results for certain problems, but does not lead to stochastic equations, since the distribution is neither real nor positive.

Chapter 6

Boundary term errors and their removal

As has been mentioned previously, when one attempts to simulate non-separable physical models using stochastic equations derived via phase-space distributions, instabilities in the equations are common. These can give rise to systematic biases in the observable estimates (3.14) using a finite number of samples. Sometimes these biases abate at large sample number (Such a situation is considered in detail in Appendix A), while in other cases the biases remain even in the limit $\mathcal{S} \rightarrow \infty$.

The latter have been called *boundary term errors* because the root of the biasing lies in discarding nonzero boundary terms in integrations involving the distribution $P(C)$ (and some factors) over C . Rigorous results regarding how and when these biases arise in practical simulations are few, and many researchers while aware of phase-space distribution methods, are either 1) unaware of the potential for biases, or 2) know that biases exist and because of this avoid using phase-space distribution methods altogether. With the aim of reducing this confusion, in Sections 6.1 and 6.2, the mechanisms by which these biases come about are investigated and classified for general phase-space distributions. To clarify the conditions under which boundary term errors may be a problem, some symptoms that can be checked for by inspection of the stochastic equations and observable estimators are pointed out.

It will be seen in Chapters 7 and 9 that simulations of interacting Bose gases

can contain boundary term errors (at least when using coherent state kernels), and a method to overcome this problem is developed in Section 6.3. Subsequently, this un-biasing procedure is demonstrated on the two rather different examples from the literature where boundary term errors have been seen when using a positive P representation: Two-boson absorption (Section 6.4), and a single-mode laser model (Section 6.5). Removal of possible or confirmed biases in the interacting Bose gas is carried out in Sections 7.3 and 9.2, respectively.

6.1 Boundary term errors of the first kind

6.1.1 Origin and general form

During the derivation of the Fokker-Planck equation (from the original master equation), there is a point between (3.41) and (3.42) where integration by parts is performed, and boundary terms discarded. The correspondence between full quantum evolution and FPE (and hence the final stochastic equations) occurs *only* if these discarded boundary terms are zero. Otherwise, there is a discrepancy between quantum mechanics and what is calculated by the stochastic equations in the limit of $\mathcal{S} \rightarrow \infty$. This has been termed a boundary term error. Let us see what exact form these discarded boundary terms take.

The partial integration is made on a master equation that already contains all kernel gauge expressions (4.19), (4.12) and/or (4.2) as well as the basic nonzero terms arising from directly applying the operator correspondences (3.39) to the master equation. In any case, all these terms can be incorporated into the formalism of (3.41). With the real variable set $C = \{C_j\}$, when going from (3.41) to (3.42) the discarded boundary terms are found to be (after some algebra)

$$\widehat{\mathcal{B}}(C) = \sum_{l_j} \int \left[\widehat{\mathcal{B}}_j^{\text{int}}(C) \right]_{\min[C_j]}^{\max[C_j]} \prod_{p \neq j} dC_p \quad (6.1a)$$

where

$$\widehat{\mathcal{B}}_j^{\text{int}}(C) = P(C)T_{lj}^{(1)}(C)\widehat{\Lambda}(C) + \frac{1}{2}P(C)\sum_k T_{ljk}^{(2)}\frac{\partial\widehat{\Lambda}(C)}{\partial C_k} - \frac{1}{2}\widehat{\Lambda}(C)\sum_k \frac{\partial}{\partial C_k} \left(P(C)T_{lkj}^{(2)}(C) \right). \quad (6.1b)$$

The usual definite evaluation notation $[f(v)]_a^b = \lim_{v \rightarrow b} f(v) - \lim_{v \rightarrow a} f(v)$ has been used. Note that the kernel is still an operator, and hence so is the boundary term expression $\widehat{\mathcal{B}}$. Every matrix element of the boundary term operator should be zero for the correspondence between quantum mechanics and the FPE to be exact. Such matrix elements could in principle be evaluated by inserting the elements of $\widehat{\Lambda}$ and $\partial\widehat{\Lambda}/\partial C_k$, which can usually be worked out exactly, into (6.1). This will be attempted for the gauge P representation in Section 6.1.5, and a toy example of the calculation of $\widehat{\mathcal{B}}$ is given in Section 6.2.2.

While (6.1) is nontrivial, there are generally a variety of ways in which it could turn out to be zero.

1. Especially for an open phase space, one expects that $P(C)$ will tend towards zero at the boundary, and if this rate is faster than the growth of the multiplying factors ($T^{(\bullet)}(C)$ times the matrix elements of $\widehat{\Lambda}$), then the terms in (6.1) will be zero.
2. Even if individually the terms in (6.1) do not tend to zero, but the distribution $P(C)$ times multiplying factors is an even function of C_j , then the values evaluated at $\max[C_j]$ and $\min[C_j]$ may cancel each other. A typical example of this is when a particular $C_j = \theta$ is the phase of a complex variable — any function of θ evaluated at the limits of its domain $[-\pi, \pi]$ is the same at both limits, and any boundary terms in θ sum to zero.
3. Also, there may be a cancellation between separate terms in the expression (6.1).

Conversely, the boundary terms can easily turn out to be nonzero in an open phase-space if $P(C)$ drops off too slowly as $|C_k|$ grows (typically — if the tails of $P(C)$ obey a power law), or, in a partly or fully closed phase-space, when $P(C) \neq 0$ at the boundary.

6.1.2 Symptoms indicating boundary term errors

Unfortunately, it is not generally possible to directly evaluate the boundary terms (6.1) and check correctness, because this involves knowing at least the far tail behavior (in an open phase space) of $P(C)$ as a function of time. In a fully or partly closed phase-space one needs to know the exact behavior near the boundary. These require detailed knowledge about aspects of the full solution to the master equation — generally not achievable in non-trivial systems, although power law tails of $P(C)$ can be considered a bad sign.

What is really needed are some symptoms of likely boundary term errors that could be looked for in the stochastic equations before simulation. Since boundary term errors arise when the distribution falls off too slowly as the (possibly open) boundaries of phase space are approached, at least four (there may be more) broad classes of features come to mind that indicate the probable presence of boundary term errors of the first kind.

1. The first such symptom, fairly well known, is the presence of so-called MOVING SINGULARITIES. In a set of deterministic equations, these manifest themselves as solutions (usually a set of measure zero) that diverge in a finite time. Previous studies [64, 66] have shown that when moving singularities are present in the deterministic parts of stochastic equations, boundary term errors usually result. In particular, it was seen that
 - In a number of studied systems, moving singularities in equations derived using the positive P representation occur whenever systematic errors are seen in observables calculated with estimators polynomial in the system variables.
 - In these systems, when the moving singularities are removed with the aid of stochastic gauges (as explained in Section 6.3), the systematic errors are removed as well[66].

The presence of pathological behaviour when moving singularities are present can be grasped intuitively: Due to the stochasticity of the equations, and

the continuous nature of the noise, one expects that after small times, the distribution of trajectories (at least the far tails of it) will extend over all of phase space. In particular into those phase-space regions where the solutions that diverge in a finite time occur. While the divergent trajectories are usually a set of measure zero, an infinitesimal proportion of divergent trajectories may have a finite effect on the observable means, on the general basis that $0 \times \infty$ does not necessarily equal zero. However, because the number of such trajectories is infinitesimal in comparison with the rest, this effect can never be estimated by a finite sample of them, hence the systematic errors.

Moving singularities may arise either in a dynamic fashion (e.g. in an equation $dv = v^2 dt$, $v(t) = v(0)/[1 - tv(0)]$, and a divergence occurs at $t = 1/v(0)$), or due to divergent drift terms for finite phase-space variables (e.g. $dv = dt/(1 - v)$ at $v = 1$, or $dv = dt/(1 - t)$ at $t = 1$). In the latter situation the divergent trajectories may even be a set of nonzero measure.

2. Secondly, one should also be wary of NOISE DIVERGENCES — i.e. instabilities in the noise matrices B . For some small set of trajectories the Wiener increments $dW(t)$ will conspire to act approximately as a constant $(\Delta W) \approx \mathcal{O}(\sqrt{\Delta t})$ over a time interval $\Delta t \gg dt$. This results in approximately an effective drift term proportional to B , and if such “drift terms” give rise to what effectively are moving singularities, then systematic errors will also result by an analogous mechanism as for instabilities of the drift terms.
3. Thirdly, Another indicator of possible boundary term errors is a DISCONTINUOUS DRIFT AND/OR NOISE MATRIX B . This may cause an effect that can not be simulated stochastically, as the discontinuity is only ever sampled by a set of trajectories of measure zero. In particular, there may be a region near such a boundary that contains only an infinitesimal portion of total trajectories, but those trajectories have large impact on the observable estimates.
4. Fourthly, if the INITIAL DISTRIBUTION IS TOO BROAD, boundary term errors may be present from the outset. An example would be a Lorentzian ini-

tial distribution of variable $v(0)$ if the Fokker-Planck equation contains terms quadratic in v . Another example is the single-mode laser model of Section 6.5 with non-delta-function initial conditions.

If any of the above symptoms is present, boundary term errors may be suspected. Conversely, if the symptoms are absent, it *appears likely* that no boundary term errors occur, but unfortunately rigorous results for general distributions are hard to come by. No boundary term errors were seen for any simulation that did not suffer from one or more of these symptoms.

There have also been developed several numerical indicators (described in detail in [64]) that one can use to check for the presence of boundary term errors from a finite set of trajectories.

6.1.3 Moving singularities in a complex phase-space

For a set of stochastic equations (e.g. such as (3.45) or (4.90)) it is relatively easy to rule out boundary term symptoms 3 and 4 from the previous section (by inspection). Moving singularities and noise divergences (symptoms 1 and 2) are a little more slippery, but practical conditions for their absence can be stated. Simulations in this thesis are with complex variable sets on an open phase space, so here a practical condition for absence of moving singularities will be derived only for this particular (though commonly occurring) kind of phase-space.

In an open complex phase space, barring symptoms 3 and 4, moving singularities cannot occur if the both the drift and noise matrices of the stochastic equations contain no radial components that lead to super-exponential growth. This is because then all variables remain finite at finite times, and so at finite times never reach the “boundary” at modulus infinity. Since $z = ze^{2i\pi}$ for any complex z , tangential evolution does not lead to divergences, because boundary terms evaluated at the boundaries of the polar variable ($\angle z = 0$ or 2π) are equal, and thus cancel in (6.1a).

More precisely, for a set of stochastic equations

$$dz_j = A_j(C) dt + \sum_k B_{jk}(C) dW_k(t), \quad (6.2)$$

in complex variables z_j ($C = \{z_j\}$, including, possibly z_0), this gives the *Condition*:

- Moving singularities or noise divergences will not occur provided that the limits

$$\lim_{|z_j| \rightarrow \infty} \frac{A_j}{|z_j|} \quad \text{and} \quad \lim_{|z_j| \rightarrow \infty} \frac{B_{jk}}{|z_j|} \quad (6.3)$$

converge for all $z_j \in C$ and all k , and that symptoms 3 and 4 of Section 6.1.2 are ruled out (by inspection).

6.1.4 Generic presence in many-mode simulations

Off-diagonal kernels appear necessary to allow all quantum states to be described by a positive real distribution that can be sampled stochastically (see Section 3.2.3). In such a case, the independent “ket” and “bra” parameters of the kernel will (if complex) almost always enter the kernel in an analytic way. Then, to ensure a positive propagator, requires in turn that the complex derivatives be chosen as in Section 3.4.3, preserving the analytic nature of the formulation into the stochastic equations.

The presence of moving singularities in (deterministic) dynamical equations for an analytically continued phase-space is a well-studied problem for both ordinary and partial differential nonlinear evolution equations. According to the Painleve conjecture (see, e.g. [79, 80, 81]), this is equivalent to non-integrability - and is therefore generic to many (in fact, all but a set of measure zero) complex dynamical equations with large numbers of degrees of freedom. Just the sort one obtains when simulating many-body systems.

Kernel gauges are, however, not restricted by the above mechanisms to be analytic. This is because they arise as arbitrary (hence, possibly non-analytic) multipliers to operators on the kernel, rather than as a result of applying analytic operator identities (3.39) to it. Because of this freedom of gauge choice, it has been conjectured by P. D. Drummond[56] that breaking the analyticity of the complex equations using stochastic gauges is a necessary (though not sufficient just by itself) condition to remove boundary term errors. This agrees with previous work[66, 1], and also with the gauges used to remove boundary term errors in this thesis.

Diffusion gauges, while possibly non-analytic, do not appear useful for removal of boundary term errors of the first kind because they appear only at the FPE-Langevin stage, after any nonzero boundary term operators have been already discarded.

6.1.5 Boundary terms in the gauge \mathbf{P} representation

Let us consider the gauge \mathbf{P} representation. The kernel (5.4), expanded in the complete Fock number state basis for each subsystem is $\widehat{\Lambda} = \Omega \otimes_k \widehat{\Lambda}_k$, with

$$\widehat{\Lambda}_k = \sum_{n\tilde{n}} \frac{\alpha_k^n \beta_k^{\tilde{n}}}{\sqrt{n!\tilde{n}!}} |n\rangle_k \langle \tilde{n}|_k e^{-\alpha_k \beta_k}, \quad (6.4)$$

as in (5.53). A master equation leads via (5.7), without invoking gauges, to terms $T_{lj}^{(1)}$ and $T_{ljk}^{(2)}$ polynomial (of low order) in α_k and β_k . Use of a weighting gauge will lead to some terms $T^{(\bullet)} \propto \Omega$. Drift gauges could generally lead to any form of the corresponding coefficients $T^{(\bullet)}$, but let us *assume* that the drift gauges \mathcal{G}_k , etc. are polynomial in $\alpha_k, \beta_k, \alpha_k^*, \beta_k^*$, and autonomous (i.e. not dependent on Ω). Under this assumption, the terms $T^{(\bullet)}$ due to drift gauges will be polynomial in all the variables, and at most $\propto \Omega^2$ (Diffusion coefficient in Ω due to standard drift gauge introduction is $\propto \Omega^2 \mathcal{G}_k^2$).

Combining this together, the (matrix elements of the) multipliers of $P(C)$ appearing in the boundary term expression (6.1) will at most scale as :

$$\leq \propto \Omega^3 \prod_k \alpha_k^{a_k} \beta_k^{b_k} e^{-\alpha_k \beta_k}, \quad (6.5)$$

where the coefficients take on all values $a_k \geq 0$ and $b_k \geq 0$ for various matrix elements.

In an open phase space of complex variables z_j (as here), the boundary lies at $|z_j| \rightarrow \infty$. The multipliers of $P(C)$ in (6.1) scale in these far tails as

$$\leq \propto |\Omega^3| \prod_k |\alpha_k|^{a_k} |\beta_k|^{b_k} e^{c_k |\alpha_k| |\beta_k|}, \quad (6.6)$$

where $c_k = -e^{i\angle\alpha_k} e^{i\angle\beta_k}$, can have positive real part. This implies that if one wants to be sure that no boundary term errors are present, (assuming any drift gauges \mathcal{G}_k

are polynomial in α_k , β_k , or their conjugates, and independent of Ω), then: $P(C)$ should decay in the far tails no slower than

$$P(C) < \infty \quad 1/|\Omega^{4+\tilde{b}_k}|, \quad (6.7a)$$

$$P(C) < \infty \quad \exp(-\tilde{c}_k|\alpha_k|^{\tilde{a}_k}), \quad (6.7b)$$

$$P(C) < \infty \quad \exp(-\tilde{c}_k|\beta_k|^{\tilde{a}_k}), \quad (6.7c)$$

where $\tilde{b}_k > 0$, $\tilde{c}_k > 0$ and $\tilde{a}_k > 1$. (And there should be no time-dependent discontinuities in the drift or noise equations).

This does not preclude that by some canceling of terms, a less favorable scaling of the tails may also become acceptable, but certainly if P scales like (6.7) or better in the far tails, boundary term errors (of the first kind) will be absent in a gauge P representation simulation. Note in particular that from (6.7):

1. Power law tails of $P(C)$ in the α, β part of phase-space will invariably lead to boundary term errors. For some matrix elements of $\hat{\mathcal{B}}$ for which the coefficient of $|n\rangle_k \langle \tilde{n}|_k$ in (6.4) is of high enough order, the boundary terms will become nonzero. This is consistent with the conclusions drawn by Gilchrist *et al*[64] that power law tails in a positive P distribution are an indicator of boundary term errors.
2. Exponential decay of $P(C)$ with $|\alpha_k|$ or $|\beta_k|$ is also not good enough in general. This is because the coefficient of (say) $|\alpha_k|$ in the exponents of (6.7) is not constant, but $\propto |\beta_k|$ — which may be any arbitrarily large value. Only a faster-than-exponential decay will be able to overcome this when $|\beta_k|$ is large. However, Gilchrist *et al*[64] found no boundary term errors in a single-mode anharmonic oscillator (6.9) simulation for which the far tail behaviour of P_+ appeared to be exponential, and no indications of boundary term errors in this system are seen here in Chapter 7 either. This may be because the single-mode system is a special case — certainly in Chapter 8 it is seen that when acting on several coupled modes, a similar Hamiltonian *can* lead to boundary term errors.

6.2 Boundary term errors of the second kind

Biases that do not abate for $\mathcal{S} \rightarrow \infty$ may also arise via a different mechanism when the estimator functions to be averaged to obtain the expectation value of a given observable grow too rapidly in phase space.

6.2.1 Mechanism and overview

Working at the level of the distribution $P(C)$, expectation values of observables are given by the real part of (3.13). When one takes a number \mathcal{S} of samples of this distribution, the estimator (3.14) is used, but if the real parts of the integrals in (3.13) do not converge then this finite sample estimate does not approach the quantum mechanical value even in the limit $\mathcal{S} \rightarrow \infty$. Systematic errors arising from this source will be called here boundary term errors of the second kind, and have very different properties to those of the first kind. Explicitly, from (3.13) and by imposing Hermiticity, the integrals that must converge are

$$\int P(C) \operatorname{Re} \left\{ \operatorname{Tr} \left[\widehat{O} \left(\widehat{\Lambda}(C) + \widehat{\Lambda}^\dagger(C^*) \right) \right] \right\} dC \quad (6.8a)$$

$$\int P(C) \operatorname{Re} \left\{ \operatorname{Tr} \left[\widehat{\Lambda}(C) \right] \right\} dC, \quad (6.8b)$$

for any particular observable \widehat{O} .

One can see that even when (6.8b) is convergent, errors may occur for some observables for which only (6.8a) diverges. Given a particular distribution, divergences may come about due to the matrix elements of either the kernel $\widehat{\Lambda}$, the operator \widehat{O} , or both growing too rapidly as the boundaries of phase space are approached. Observables that involve an infinite sum of moments polynomial in the system variables (such as the parity in the gauge P representation, whose estimate is given by the sum (5.15)) may be particularly prone to these boundary term errors, because the dependence of $\operatorname{Tr} \left[\widehat{O} \widehat{\Lambda} \right]$ on the phase-space variables will grow particularly rapidly.

Boundary term errors of the second kind depend on the equations of motion only by their effect on the distribution of samples $P(C)$, and so do not show up as moving singularities or noise divergences in the equations.

A sobering point is that, in theory, one could always come up with some observables that grow fast enough as the boundary of phase space is approached to overcome whatever far tail behaviour occurs in $\widehat{\Lambda}(C)P(C)$. Such an observable could be defined as an infinite sum of polynomial moments whose terms grow appropriately rapidly. Thus, boundary term errors of the second kind are likely to occur for *some* observable estimates for any phase-space distribution simulation method, for any model. In practice, fortunately, it suffices to make the estimators of only all observables *of interest* convergent.

6.2.2 Example: un-normalized positive P representation

A good example of boundary term errors of the second kind occurs for the “anharmonic” oscillator single-mode Hamiltonian, when using an un-normalized coherent state kernel. This Hamiltonian¹ is

$$\widehat{H} = \hbar \widehat{a}^\dagger \widehat{a}^2, \quad (6.9)$$

the single-mode analogue of a repulsive two-body potential in the interaction picture². Let us consider the kernel

$$\widehat{\Lambda} = ||\alpha\rangle\langle\beta^*||, \quad (6.10)$$

off-diagonal in the un-normalized Bargmann coherent states (5.1). It is convenient to change to logarithmic variables $n_L = \log(\alpha\beta)$, and $m_L = \log(\alpha/\beta)$, then the basic operator correspondences are, from (5.6),

$$\widehat{a}\widehat{\Lambda} = \alpha\widehat{\Lambda} = e^{(n_L+m_L)/2}\widehat{\Lambda} \quad (6.11a)$$

$$\widehat{a}^\dagger\widehat{\Lambda} = \frac{\partial}{\partial\alpha}\widehat{\Lambda} = e^{-(n_L+m_L)/2}\left(\frac{\partial}{\partial n_L} + \frac{\partial}{\partial m_L}\right)\widehat{\Lambda} \quad (6.11b)$$

$$\widehat{\Lambda}\widehat{a}^\dagger = \beta\widehat{\Lambda} = e^{(n_L-m_L)/2}\widehat{\Lambda} \quad (6.11c)$$

$$\widehat{\Lambda}\widehat{a} = \frac{\partial}{\partial\beta}\widehat{\Lambda} = e^{(m_L-n_L)/2}\left(\frac{\partial}{\partial n_L} - \frac{\partial}{\partial m_L}\right)\widehat{\Lambda}. \quad (6.11d)$$

It is also convenient to split the variables into real and imaginary parts: $n_L = n'_L + in''_L$, and $m_L = m'_L + im''_L$. For a dynamical evolution model with no coupling to the

¹In some appropriate time units.

²That is, with terms linear in \widehat{a} and \widehat{a}^\dagger transformed away. Compare with (2.17).

environment, utilizing the analytic nature of the kernel as explained in Section 3.4.3, one obtains the FPE (provided no boundary term errors of the first kind are present)

$$\frac{\partial P}{\partial t} = \left\{ 2 \frac{\partial}{\partial m_L''} + \frac{\partial^2}{\partial n_L'^2} + \frac{\partial^2}{\partial n_L''^2} + \frac{\partial^2}{\partial m_L'^2} + \frac{\partial^2}{\partial m_L''^2} - \frac{\partial^2}{\partial n_L' \partial m_L''} - \frac{\partial^2}{\partial m_L'' \partial n_L'} - \frac{\partial^2}{\partial n_L'' \partial m_L'} - \frac{\partial^2}{\partial m_L' \partial n_L''} \right\} P. \quad (6.12)$$

This gives the Ito stochastic equations (with standard square root form of $B_0 = \sqrt{D}$)

$$dn_L = i\sqrt{2i} [dW(t) - id\widetilde{W}(t)] = 2d\eta(t) \quad (6.13a)$$

$$dm_L = i\sqrt{2i} [dW(t) + id\widetilde{W}(t)] - 2i dt = 2i[d\eta^*(t) - dt]. \quad (6.13b)$$

$dW(t)$ and $d\widetilde{W}(t)$ are the usual independent Wiener increments, and the complex noise $d\eta(t)$ defined here obeys

$$\langle d\eta(t) \rangle_{\text{stoch}} = 0 \quad (6.14a)$$

$$\langle d\eta^*(t) d\eta(t') \rangle_{\text{stoch}} = \delta(t - t') dt^2 \quad (6.14b)$$

$$\langle d\eta(t) d\eta(t') \rangle_{\text{stoch}} = 0. \quad (6.14c)$$

The FPE (or stochastic equations) can be solved, and in the case of coherent state initial conditions (with initial amplitude α_0)

$$P(n_L, m_L, 0) = \delta^2(m_L - 2i\angle\alpha_0) \delta^2(n_L - 2\log|\alpha_0|), \quad (6.15)$$

the distribution at time t is

$$P(n_L, m_L, t) = \delta^2\left(m_L - \left[i(n_L^* - 2\log|\alpha_0|) - 2it + 2i\angle\alpha_0\right]\right) \times \frac{1}{2\pi t} \exp\left(-\frac{|n_L - 2\log|\alpha_0||^2}{4t}\right). \quad (6.16)$$

In this toy model, one can actually check if there really were no boundary terms of the first kind, by substituting (6.16), and (6.10) into the expression for $\widehat{\mathcal{B}}$. The boundary term operator $\widehat{\mathcal{B}}$ is given by the integrals (6.1) of operators $\widehat{\mathcal{B}}_j^{\text{int}}$ composed of P , $\widehat{\Lambda}$ and coefficients $T^{(\bullet)}$, with individual terms of $\widehat{\mathcal{B}}$ evaluated at the boundaries of phase space. In our case these boundaries are $C_j \rightarrow -\infty$ and $C_j \rightarrow \infty$ for all $j = 1, \dots, 4$, and one requires all the matrix elements of $\widehat{\mathcal{B}}$ to be zero. By comparison of (6.12) to (3.43), and labeling variables as $n_L' = C_1$, $n_L'' = C_2$, $m_L' = C_3$, and

$m_L'' = C_4$, and also setting the term counter $l = 1$, the nonzero $T^{(\bullet)}$ coefficients in the $\widehat{B}_j^{\text{int}}$ are $T_{14}^{(1)} = -2$, $T_{1jj}^{(2)} = 2$ for all j , and $T_{114}^{(2)} = T_{141}^{(2)} = T_{123}^{(2)} = T_{132}^{(2)} = -2$. Also, the matrix elements of $\widehat{\Lambda}$ (in a Fock basis $|n\rangle$) can be written (from the definition (5.1) of $|\alpha\rangle$) as

$$\langle \tilde{n} | \widehat{\Lambda} | n \rangle = \frac{1}{\sqrt{n! \tilde{n}!}} \exp \left[n_L(n + \tilde{n})/2 + m_L(n - \tilde{n})/2 \right]. \quad (6.17)$$

So:

1. The $T^{(\bullet)}$ coefficients are constants.
2. The matrix elements of $\widehat{\Lambda}$ and $\partial \widehat{\Lambda} / \partial z$ are exponential in both variables $z = n_L$ or m_L , with some constant positive or negative coefficient.
3. The distribution P is Gaussian in real and imaginary parts of both variables (n_L around mean $2 \log |\alpha_0|$), m_L around mean $2i(\angle \alpha_0 - t)$).

One can see that all matrix elements of all the integrands ($\langle \tilde{n} | \mathcal{B}_j^{\text{int}} | n \rangle$) will be dominated by the Gaussian decay of $P(C)$ in the far tails for all n and \tilde{n} . For $j = 1, 2$, after integration over m_L and a (real or imaginary) component of n_L , the expression to be evaluated at large n_L will be Gaussian, hence zero in the far tail limit. Also, for the terms ($j = 3, 4$) in which one integrates over n_L and a component of m_L , the expression to be evaluated will be a Dirac delta function — also zero in the far tail limit. Hence,

$$\widehat{\mathcal{B}} = 0. \quad (6.18)$$

Thus it has been confirmed that there are no boundary term errors of the first kind.

So far it seems like this might be a very good kernel to try even for many-body locally interacting Bose gases, which have the same kind of interparticle potential term. The advantages would be that there are no boundary term errors of the first kind, no moving singularities (compare to Section 6.1.4) and even no nonlinear terms in the stochastic equations (compare to (5.17)).

Having the luxury of knowing $P(C)$ for this toy problem, let us see what happens when one tries to calculate an observable. Directly from (6.10) and (5.3):

$$\text{Tr} \left[\widehat{\Lambda} \right] = \exp(\alpha\beta) = \exp[e^{n_L}]. \quad (6.19)$$

The essential denominator term (6.8b) in any observable calculation, is then (using (6.16) and defining $n_L = n'_L + in''_L$),

$$\begin{aligned} & \int \operatorname{Re} \{ \exp [e^{n_L}] \} \frac{1}{4\pi t} e^{-|n_L - 2 \log |\alpha_0|^2 / 4t} d^2 n_L \\ \propto & \int \exp \left[e^{n'_L} \cos n''_L \right] \cos \left(e^{n'_L} \sin n''_L \right) e^{-(n'_L - 2 \log |\alpha_0|^2) / 4t} e^{-(n''_L)^2 / 4t} d^2 n_L \end{aligned} \quad (6.20)$$

For n_L with large positive real part, the factor from $\operatorname{Tr} [\widehat{\Lambda}]$ dominates (it's an exponential of an exponential!), and since the integrand is symmetric in n''_L and nonzero, it must become unbounded, and the integral diverge.

A plot of the integrand $P \operatorname{Tr} [\widehat{\Lambda}]$ for several time values is shown in Figure 6.1. The unbounded behaviour at high n'_L values indicates that (6.8b) will diverge and boundary term errors (of the second kind) will be present. A comparison to the probability distribution of n'_L is also made there to indicate which contributions to the integral (6.8b) are sampled by a stochastic simulation.

This means that: *All observable estimates using (3.14) are suspect with this representation!*

6.2.3 In the gauge P representation

For the gauge P representation, $\operatorname{Tr} [\widehat{\Lambda}] = e^{z_0}$, and so the denominator integral (6.8b) becomes simply

$$\int P(C) e^{z'_0} \cos z''_0 dC \quad (6.21)$$

in terms of $z_0 = z'_0 + iz''_0$. This will be convergent provided the far tails of the marginal probability distribution of z'_0 decay as $e^{-z'_0}$, or faster in the limit $|z'_0| \rightarrow \infty$.

From the estimator expression (5.13) for an arbitrary observable that is a polynomial function of \widehat{a} and \widehat{a}^\dagger , one can see that the numerator integral (6.8a) will be of the form

$$\int P(C) \sum_j f_j(C) dC, \quad (6.22)$$

where the functions for each term f_j are

$$f_j(C) = \frac{1}{2} e^{z'_0} \left(\prod_k |\alpha_k|^{a_k} |\beta_k|^{b_k} \right) \cos \left(z''_0 \pm \theta + \sum_k [a_k \angle \alpha_k + b_k \angle \beta_k] \right), \quad (6.23)$$

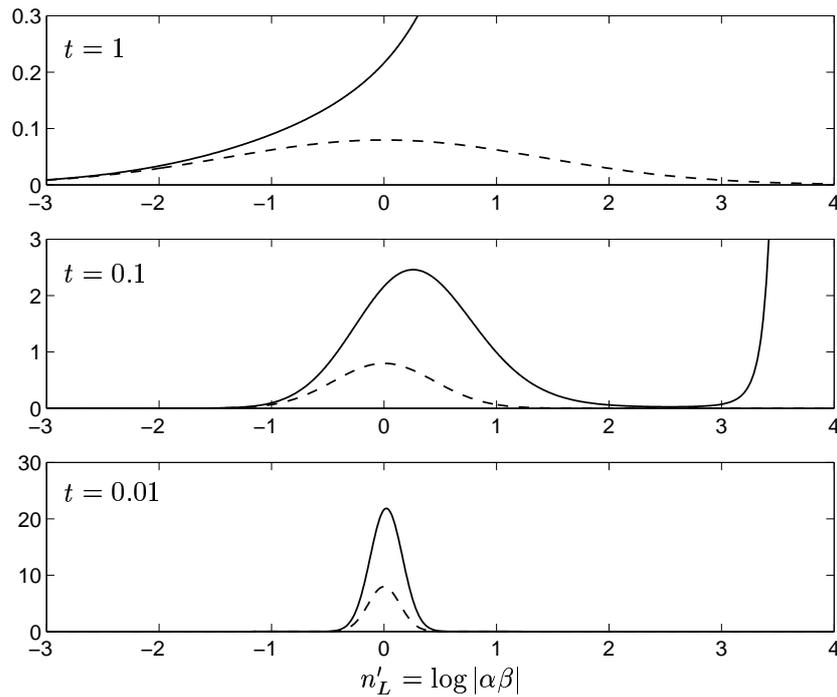


Figure 6.1: **Integrands appearing in the denominator** of observable estimates for anharmonic oscillator dynamics governed by the Hamiltonian (6.9). Values plotted are (SOLID LINE): $\int \text{Tr} [\widehat{\Lambda}(C)] P(C) d^2 m_L$, evaluated (using (6.19) and (6.16)) at $n''_L = 0$. Coherent state initial conditions (6.15) were used with mean occupation number unity ($|\alpha_0| = 1$). To show what parts of the integral (6.8b) would be sampled by a stochastic simulation, the probability distribution of n_L evaluated at $n''_L = 0$ (i.e. $\int P(C) d^2 m_L$ evaluated at $n''_L = 0$) is also plotted as the DASHED LINE. Note the unbounded behaviour of the integrand at large positive n'_L values (which are not sampled).

with some non-negative integer coefficients a_k and b_k , which take on different values for different f_j . The behaviour of each such term as one approaches the boundaries of phase space, where the modulus of at least one variable tends to $+\infty$, is exponential in z'_0 (with unity coefficient), and polynomial in $|\alpha_k|$ and $|\beta_k|$.

This means that

- Boundary term errors can be expected for observables polynomial in \widehat{a}_k and/or \widehat{a}_k^\dagger if the tails of $P(C)$ decay as any power law in the $|\alpha_k|$, or $|\beta_k|$, or as $e^{-z'_0}$ or slower in z'_0 .

Conversely, if the tails of $P(C)$ decay faster than power law in all $|\alpha_k|$ and $|\beta_k|$ and

faster than $e^{-z'_0}$ in the modulus of the weight, then boundary term errors (*at least of the second kind*) will not occur for any observable polynomial in the annihilation and creation operators on subsystems. That is, any observable given by (5.10), or a finite linear combination of such expressions.

Power law tails of $P(C)$ in $|\alpha_k|$ or $|\beta_k|$ do not immediately exclude all observables, only those whose estimators grow too rapidly. By inspection of (5.10), (5.13), one can see that an observable that is an a_k th polynomial of \hat{a}_k will lead to a term $f_j \propto e^{z'_0} |\alpha_k|^{a_k}$. If $P(C) \propto e^{-cz'_0} |\alpha_k|^{-\tilde{a}_k}$ in the far tails, then the indefinite integral $\int P(C) f_j(C) dC$ will be $\propto e^{(1-c)z'_0} |\alpha_k|^{a_k - \tilde{a}_k + 1}$ (if $a_k - \tilde{a}_k + 1 \neq 0$), or $\propto e^{(1-c)z'_0} \log |\alpha_k|$ (otherwise) in this far tail region. When this is evaluated in the far tail limit, the integral will convergent if the decay of $P(C)$ is fast enough. Similarly for the relationship between \hat{a}_k^\dagger and β_k . Hence, one can state the conditions for expectation value estimates of a particular operator \hat{A} to not suffer from boundary term errors of the second kind: When ϵ_j are infinitesimal positive real constants, one requires

$$\hat{A} \propto \hat{a}_k^{a_k} \implies \lim_{|\alpha_k| \rightarrow \infty} P(C) < \propto \frac{1}{|\alpha_k|^{-(1+a_k+\epsilon_1)}} \quad (6.24a)$$

$$\hat{A} \propto \hat{a}_k^{\dagger b_k} \implies \lim_{|\beta_k| \rightarrow \infty} P(C) < \propto \frac{1}{|\beta_k|^{-(1+b_k+\epsilon_2)}} \quad (6.24b)$$

as well as

$$\lim_{z'_0 \rightarrow \infty} P(C) < \propto e^{-(1+\epsilon_3)z'_0}. \quad (6.24c)$$

Non-polynomial observables defined as infinite sums (e.g. the parity (5.14)) can be more troublesome, because the functions $\text{Tr} \left[\hat{\Lambda} \hat{O} \right]$ etc. that are to be sampled by (6.8a) can grow faster than polynomially as $|\alpha_k|, |\beta_k| \rightarrow \infty$. Conditions of the kind (6.24) can also be derived, but will not scale as a power law. For example since the parity estimator (5.15) has a numerator expression that scales $\propto e^{-2\text{Re}\{\alpha_k \beta_k\}}$, one requires $P(C)$ to decay faster than $e^{2\text{Re}\{\alpha_k \beta_k\}}$ in the far tails for an unbiased estimate of parity expectation values (remember that $\alpha_k \beta_k$ may have negative real part in highly non-classical parts of phase-space).

6.2.4 Locally-interacting Bose gas calculations

Here the likelihood of boundary term errors of the second kind is considered for a many-mode interacting Bose gas model using the gauge P representation as described in Sections 5.3 and 5.6. By inspection of the dynamics and thermodynamics gauge P equations for the locally interacting Bose gas (5.17)–(5.21), and (5.50), one sees that the stochastic terms (assuming no gauges) take the forms

$$d\alpha_{\mathbf{n}} \cdots + c \alpha_{\mathbf{n}} dW_{\mathbf{n}} \quad (6.25a)$$

$$d\beta_{\mathbf{n}} \cdots + \tilde{c} \beta_{\mathbf{n}} d\tilde{W}_{\mathbf{n}}, \quad (6.25b)$$

with some complex constants c and \tilde{c} , apart from possibly additional terms caused by finite temperature heat bath interactions (5.18)

$$d\alpha_{\mathbf{n}} \cdots + \sqrt{\gamma_{\mathbf{n}} \bar{n}_{\text{bath}}(T)} d\eta_{\mathbf{n}} \quad (6.26a)$$

$$d\beta_{\mathbf{n}} \cdots + \sqrt{\gamma_{\mathbf{n}} \bar{n}_{\text{bath}}(T)} d\eta_{\mathbf{n}}^*. \quad (6.26b)$$

In this Subsection, the far-tail regime of phase space where $|\alpha_{\mathbf{n}}|$ and/or $|\beta_{\mathbf{n}}|$ are large will be of interest, and there the terms (6.26) are negligible, and hence are ignored here.

Upon change of variables to $z_{\mathbf{n}} = \log \alpha_{\mathbf{n}}$ and $\tilde{z}_{\mathbf{n}} = \log \beta_{\mathbf{n}}^*$, one finds by the rules of Ito calculus that these logarithmic variables will have stochastic evolution

$$dz_{\mathbf{n}} \cdots + c dW_{\mathbf{n}} \quad (6.27a)$$

$$d\tilde{z}_{\mathbf{n}} \cdots + \tilde{c} d\tilde{W}_{\mathbf{n}}. \quad (6.27b)$$

This indicates that the *short time* evolution of $z_{\mathbf{n}}$ and $\tilde{z}_{\mathbf{n}}$ is dominated by Gaussian noise. Thus at short times, (defining $z_{\mathbf{n}} = z'_{\mathbf{n}} + iz''_{\mathbf{n}}$, and $\tilde{z}_{\mathbf{n}} = \tilde{z}'_{\mathbf{n}} + i\tilde{z}''_{\mathbf{n}}$)

$$P(C, t) \prod_{\mathbf{n}} d^2 z_{\mathbf{n}} d^2 \tilde{z}_{\mathbf{n}} \propto P(C, 0) \prod_{\mathbf{n}} \exp \left(-\frac{[z'_{\mathbf{n}}(t) - z'_{\mathbf{n}}(0)]^2}{2\text{Re}\{c\}^2 t} - \frac{[\tilde{z}'_{\mathbf{n}}(t) - \tilde{z}'_{\mathbf{n}}(0)]^2}{2\text{Re}\{\tilde{c}\}^2 t} \right) \quad (6.28)$$

Changing back to the coherent state amplitude variables,

$$P(C, t) \prod_{\mathbf{n}} d^2 \alpha_{\mathbf{n}} d^2 \beta_{\mathbf{n}} \propto P(C, 0) \prod_{\mathbf{n}} \frac{1}{|\alpha_{\mathbf{n}} \beta_{\mathbf{n}}|} \exp \left\{ -\frac{\left(\log \left| \frac{\alpha_{\mathbf{n}}(t)}{\alpha_{\mathbf{n}}(0)} \right| \right)^2}{2\text{Re}\{c\}^2 t} - \frac{\left(\log \left| \frac{\beta_{\mathbf{n}}(t)}{\beta_{\mathbf{n}}(0)} \right| \right)^2}{2\text{Re}\{\tilde{c}\}^2 t} \right\} \quad (6.29)$$

The far tail decay as a Gaussian of a logarithm of $|\alpha_{\mathbf{n}}|$ or $|\beta_{\mathbf{n}}|$ is strong enough to overcome any polynomial of these variables if one moves far enough into the tails. (polynomials will be only exponential in the logarithm). So, there is no immediate reason to suspect boundary term errors of the second kind for polynomial observables here. (They cannot be ruled out by the above discussion, however, because at longer times the evolution of $P(C)$ is much more complicated and no proof for such a regime has been given).

On the other hand, (6.29) does not decay strongly enough to make an average over a parity estimator converge, since these latter can grow as fast as $\exp(2e^{z'_{\mathbf{n}}+\bar{z}'_{\mathbf{n}}})$ (when $\alpha_{\mathbf{n}}\beta_{\mathbf{n}}$ is negative real), which grows much faster than (6.29) decays. One thus suspects that parity expectation values will not be estimated by (5.15) in an unbiased way for a gauge P representation of the locally interacting Bose gas model.

6.2.5 Unresolved issues

Lest one think that boundary term errors are fully understood, a variety of issues are still unclear (to the best of the author's knowledge). Some of these have to do with tradeoffs between the two kinds of error. For example:

- **The role of normalization:** In Section 6.2.2 it was seen that for a (single-mode) interacting Bose gas, the un-normalized positive P representation did not suffer boundary term errors of the first kind, but had severe problems with errors of the second kind when $\text{Re}\{\alpha\beta\} \rightarrow +\infty$. On the other hand, the normalized positive P representation or the gauge P do not suffer from the second kind of error for most observables, but (as will be seen in Chapter 8) does have problems with moving singularities, and hence presumably the first kind of boundary term error, when $\text{Re}\{\alpha\beta\} \rightarrow -\infty$.

There appears to be a tradeoff between the two processes causing boundary term errors, which can involve the normalization of the kernel. Nevertheless, the normalization does not parameterize the tradeoff in a straightforward way, since the two kinds of boundary term error occur in different regions of phase-space.

- **The role of moving singularities:** Only the first kind of boundary term error appears during derivation of the FPE and stochastic equations while on the other hand, errors of the second kind can occur simply on a static initial distribution, which does not evolve in any way. For these reasons it appears that moving singularities in the deterministic part of the stochastic equations should be related to boundary term errors of the first kind.

On the other hand, moving singularities do lead to pathological tail behaviour in $P(C)$ after a certain time t_{sing} (even without any noise terms, if the initial distribution is nonzero on any point of a trajectory that escapes to infinity at finite time.) Because the far tail behaviour of $P(C)$ changes in this qualitative way after t_{sing} , one should be able (at least in some cases) to find some observables that would be estimated in an unbiased way for $t < t_{\text{sing}}$, but suffer from boundary term errors of the second kind for $t > t_{\text{sing}}$. So it appears that moving singularities can also influence the second kind of boundary term error in some way, presumably by causing slowly decaying distribution tails.

A better understanding of the relationship between the two kinds of error might lead to a way to make an advantageous tradeoff that avoids both, at least in some cases.

6.3 Removal of the errors using kernel gauges

6.3.1 Conceptual justification

If one introduces a kernel gauge \mathcal{F} (e.g. a drift gauge \mathcal{G}_k) using a gauge identity $\mathcal{J}[\widehat{\Lambda}] = 0$, defined by (4.1) as in Section 4.1, then zero is added to the master equation in the form $\int P(C)\mathcal{F}(C)\widehat{J}[\widehat{\Lambda}]dC$, as in (4.2). Introducing such a kernel gauge can have an effect on the presence of boundary term errors in two ways:

1. **Modification of the distribution:** A change in the FPE due to addition of extra gauge terms will cause a direct change to its solution $P(C, t)$. A good choice of gauge can lead to faster decaying tails, and hence directly to $\widehat{\mathcal{B}} = 0$ from evaluation of expression (6.1), and no boundary term errors of the first

kind. This is the easiest approach in practice, and will be considered in some detail in the next Subsection 6.3.2.

Less directly, this approach also has the potential to remove boundary term errors of the *second* kind. This is again because $P(C)$ changes, and if the far tail behaviour is made to decay faster, the observable expressions (6.8a) and (6.8b) may become convergent.

2. **Renormalization of boundary terms:** The expression (6.1b) gains extra gauge terms, which can have an effect of the final value of $\widehat{\mathcal{B}}$. During the derivation of the FPE in Section 3.4.1, these extra added terms to the master equation can be written in the $T^{(\bullet)}$ notation of (3.41) and also (6.1) for $\widehat{\mathcal{B}}$ as

$$T_0^{(0)} = \mathcal{F}J^{(0)} \quad (6.30a)$$

$$T_{0j}^{(1)} = \mathcal{F}J_j^{(1)} \quad (6.30b)$$

$$T_{0jk}^{(2)} = \mathcal{F}J_{jk}^{(2)}. \quad (6.30c)$$

One finds that the additional boundary terms introduced by this kernel gauge are (using (6.1)) are of the same general form as (6.1a), but with integration over some additional terms $\widehat{\mathcal{B}}_{\mathcal{J},j}^{\text{int}}$ on top of the usual $\widehat{\mathcal{B}}_j^{\text{int}}$. Using (6.1b), these new terms are

$$\widehat{\mathcal{B}}_{\mathcal{J},j}^{\text{int}}(C) = \mathcal{F}(C) \left\{ P(C)J_j^{(1)}(C) + \frac{1}{2}P(C) \sum_k T_{jk}^{(2)} \frac{\partial}{\partial C_k} - \frac{1}{2} \sum_k \frac{\partial}{\partial C_k} \left(P(C)T_{kj}^{(2)}(C) \right) \right\} \widehat{\Lambda}(C). \quad (6.31)$$

The point to be made is that while the identity $\mathcal{J}[\widehat{\Lambda}]$, which can be expressed in differential form as (4.1), is zero, the integrands $\widehat{\mathcal{B}}_{\mathcal{J},j}^{\text{int}}$ for the boundary term expression $\widehat{\mathcal{B}}$ have a different form, and are not necessarily zero at all. This then may open the possibility that if originally (with no gauges), the boundary terms were nonzero, an appropriate choice of gauge can cancel these, again leading to $\widehat{\mathcal{B}} = 0$, and no boundary term errors of the first kind.

6.3.2 Practical implementation using drift gauges

Let us take the direct approach of changing $P(C, t)$ to make the boundary terms $\widehat{\mathcal{B}} = 0$. A difficulty is that the solution of $P(C, t)$ is not known for non-trivial cases, so one cannot work with it directly. However, as pointed out in Section 6.1.2, moving singularities in the noiseless equations are very closely tied to boundary term errors of the first kind. Certainly in the sense that their presence indicates weakly decaying distribution tails, and hence boundary term errors. The drift part of stochastic equations can be modified practically at will using drift kernel gauges \mathcal{G}_k (see Section 4.3), and in particular, the moving singularities can be removed in this way. As will be shown in examples of Sections 6.4 and 6.5, removal of these moving singularities does indeed appear to result in disappearance of the boundary term errors.

Given a kernel amenable to drift gauges (i.e. possessing a global weight Ω), and keeping in mind the symptoms of boundary term errors of the first kind outlined in Section 6.1.2 one wishes to choose drift gauges according to the following heuristic criteria:

1. Moving singularities that were present in the original (ungauged) stochastic equations are removed. This can be assessed by analysis of the deterministic parts of the stochastic equations. See examples in subsequent sections.
2. No new moving singularities are introduced.
3. No noise divergences are introduced (in the $d\Omega = \dots + \Omega \sum_k \mathcal{G}_k dW_k$ evolution).
4. No discontinuities in the drift or diffusion of any variables are introduced.
5. Given the above criteria 1–4 are satisfied (which can be assessed by inspection or analysis of the gauged stochastic equations), one wants to minimize the weight spread in the resulting simulation, to optimize efficiency.

A broadly applicable rule of thumb is to **keep the gauges as small as possible while achieving the goal of removing boundary term errors.**

This can be aimed for using several complementary approaches:

- (a) Make the gauge functions zero in regions of phase space where no correction to the un-gauged drift is necessary.
 - (b) Minimize a relevant variance of the weights. This weight spread has been considered in detail in Section 4.3.4. Depending on the situation, the quantity $\sum_k |\mathcal{G}_k|^2$ (more generally) or $\sum_k \text{Re} \{\mathcal{G}\}_k^2$ (for short time calculations of some observables) can be minimized, since the short-time growth of the variance of $|\Omega|$ or $\text{Re} \{\Omega\}$ (respectively) has been shown in Section 4.3.4 to be proportional to these quantities.
 - (c) Ensure that drift gauges are zero at deterministic attractors in phase space, so that no weight spread is accumulated when no significant evolution is taking place.
 - (d) Minimize the variances of estimators of observables of interest. That is, the variance of quantities $\text{Tr} [\widehat{O}\widehat{\Lambda}]$ and $\text{Tr} [\widehat{O}\widehat{\Lambda}^\dagger]$ appearing in the numerator of observable estimates (3.14). In this manner, one can optimize the simulation for the estimation of a particular observable \widehat{O} , rather than more generally as in point (a) above.
6. One should also avoid gauges with the following features:
- (a) Gauges that are nonzero only in a rarely visited region of phase space. The problem is that good sampling of the gauged region is needed to make the correct change to observable estimates — otherwise, if the gauged region is not visited by any trajectories, one has practically the same simulation as without gauges! In a less extreme case where only a tiny (but nonzero) fraction of trajectories visit the gauged region, the correction from biased-ungauged to unbiased-gauged observable estimates is badly resolved, and the simulation results become very noisy. Furthermore, if the bias to be corrected is significant, then the change in the few gauged trajectories needs to be large to compensate for the bias, and hence the weight correction on those trajectories will become large as well. This leads to a large weight spread. It is much better to make small changes in many trajectories contribute to removal of the bias.

- (b) Gauges rapidly changing in phase-space. These typically lead to stiff equations, and a gauged region of phase space that is badly sampled by reasonable-sized ensembles. This bad sampling arises when no single trajectory stays in the gauged region long, to such a degree that at a given time few (or, on average, possibly less than one) trajectories are sampling the gauged region, i.e. sampling the resulting observable estimate corrections. This results in similar problems as in the previous point (a).
- (c) Non autonomous gauges (i.e. depending directly on the weight Ω). The problem with these is that they are usually much more difficult to analyze to check whether or not new moving singularities or noise divergences have been introduced.
- (d) Gauges analytic in complex variables if the kernel is also analytic in these. As discussed in Section 6.1.4, it is desirable to break the analyticity of the stochastic equations to avoid moving singularities.

Note that an explicit variational minimization according to 5(b) or 5(d) could be in conflict with the recommendations 6(a) or 6(b), in which case it is probably best to try the simulations and see which works better.

This technique appears to be broadly applicable, and only requires the recognition of what instabilities in the stochastic equations could lead to problems. It does not require detailed knowledge of what the boundary terms $\widehat{\mathcal{B}}$ are, nor of the behaviour of the far tails of $P(C)$, provided instabilities are removed.

6.3.3 Ito vs. Stratonovich moving singularities

There is a subtle ambiguity that can arise when using the method just presented: Should one remove moving singularities from the Ito or the Stratonovich form of the stochastic equations?

As indicated in Appendix B, when one numerically integrates the stochastic differential equations using a finite but small time step Δt , the exact form of the drift that is integrated can depend on details of the integration algorithm chosen. For

example, if one evaluates the derivatives using the variable values at the beginning of a time step (i.e. simply those calculated using the previous step), then one integrates the standard Ito form of the equations. However, with a semi-implicit method, where one first estimates the variables in the middle of the time step, before using *these* to calculate the derivative and actually advance the simulation, one must integrate the *Stratonovich* form of the stochastic differential equations. The difference between these is given by the Stratonovich correction to the drift (B.7). In practice, multiplicative³ stochastic equations, such as those considered in this thesis, are more stable (and hence, efficient) using the semi-implicit method[82].

In some situations it may occur that moving singularities occur in the deterministic parts of the Stratonovich equations but not the Ito equations (or vice-versa). To arrive at practical guidelines on this issue, one can argue in the following way:

- Since both the explicit (Ito) and semi-implicit (Stratonovich) algorithm simulate the same underlying equations, then to be sure that no bias is being introduced, moving singularities should not occur in *either* algorithm. In fact, a whole family of hybrid Ito-Stratonovich algorithms are possible where derivatives are evaluated at arbitrary time values during the time step Δt , and the difference in the drift terms between two such algorithms is a multiple $\in [0, 2]$ of the Stratonovich correction (B.7) — See Appendix B. *None* of these algorithms should contain moving singularities, otherwise some kind of bias may be suspected.
- The Stratonovich correction for variable C_j (B.7) is highly dependent on the form of the noise matrix elements ($\propto \sum_{kl} B_{kl} \partial B_{jl} / \partial C_k$). If one finds that moving singularities are present for one algorithm but not for another, this indicates that the Stratonovich correction is growing very rapidly as one heads for the boundaries of phase space. This in turn indicates that the noise matrix elements must be growing in some appropriately rapid fashion as well — possibly rapidly enough to cause noise divergences.

³i.e. equations in which the noise terms are proportional to the variables, and hence of non-constant variance.

Taking the above points into consideration, it is conjectured here that:

1. In a simulation that does not suffer from boundary term errors, there must be no moving singularities in the deterministic part of either the Ito, Stratonovich, or any of the hybrid Ito-Stratonovich forms of the stochastic equations.
2. If a simulation has moving singularities for one of the family of algorithms (e.g. Ito), but not for some other member of this family, then noise divergences may be suspected as an additional cause of boundary term errors.
3. To eliminate biases caused by moving singularities using a drift gauge, one must remove moving singularities from *all* hybrid Ito-Stratonovich algorithms with the same gauge choice.

6.4 Removal example 1: Two-particle absorber

Two examples will now be worked through — those systems in which boundary term errors in positive P simulations have been reported in published work. Removal of boundary term errors using the guidelines of Section 6.3.2 will be demonstrated, and the procedure considered in some detail. The first example will be a single-mode two-boson absorber (in this section), while the second example of a single-mode laser will be treated in Section 6.5. These two examples demonstrate somewhat different aspects of boundary term errors and their removal.

This section is closely based on Section IV of the published article by Deuar and Drummond[66].

6.4.1 The single-mode model

Consider a single optical cavity mode driven by coherent radiation, and damped by a zero temperature bath that causes both one- and two-photon losses. For convenience, units of time are chosen so that the rate of two-photon loss is unity. The scaled one-photon loss rate is γ , and ε is the scaled (complex) driving field amplitude. Following

standard techniques, the system can be described with the master equation

$$\begin{aligned} \frac{\partial \widehat{\rho}}{\partial t} = & \left[\varepsilon \widehat{a}^\dagger - \varepsilon^* \widehat{a}, \widehat{\rho} \right] + \frac{\gamma}{2} (2 \widehat{a} \widehat{\rho} \widehat{a}^\dagger - \widehat{a}^\dagger \widehat{a} \widehat{\rho} - \widehat{\rho} \widehat{a}^\dagger \widehat{a}) \\ & + \frac{1}{2} (2 \widehat{a}^2 \widehat{\rho} \widehat{a}^{\dagger 2} - \widehat{a}^{\dagger 2} \widehat{a}^2 \widehat{\rho} - \widehat{\rho} \widehat{a}^{\dagger 2} \widehat{a}^2) . \end{aligned} \quad (6.32)$$

This is also the master equation of two-particle absorption in an interaction picture where kinetic processes have been relegated to the Heisenberg picture evolution of the operators.

Let us use the gauge P representation to convert this master equation into a set of stochastic equations. The single-mode kernel is written as

$$\widehat{\Lambda}(C = \{\alpha, \beta, \Omega\}) = \Omega ||\alpha\rangle \langle \beta^* || e^{-\alpha\beta}. \quad (6.33)$$

Following the treatment of Sections 3.4, 4.3, and 5.3, and converting to Stratonovich form using (B.7), one arrives at the Stratonovich stochastic equations

$$d\alpha = [\varepsilon - \alpha(\alpha\beta + i\mathcal{G} + \{\gamma - 1\}/2)] dt + i\alpha dW \quad (6.34a)$$

$$d\beta = [\varepsilon^* - \beta(\alpha\beta + i\widetilde{\mathcal{G}} + \{\gamma - 1\}/2)] dt + i\beta d\widetilde{W} \quad (6.34b)$$

$$d\Omega = S_\Omega dt + \Omega [\mathcal{G}dW + \widetilde{\mathcal{G}}d\widetilde{W}] . \quad (6.34c)$$

Here $S_\Omega dt$ is the appropriate Stratonovich correction term given by (B.7), which depends on the functional form of the particular gauges chosen.

With no gauge ($\mathcal{G} = \widetilde{\mathcal{G}} = 0$), the positive P Stratonovich equations are recovered:

$$d\alpha = [\varepsilon - \alpha(\alpha\beta + \{\gamma - 1\}/2)] dt + i\alpha dW \quad (6.35a)$$

$$d\beta = [\varepsilon^* - \beta(\alpha\beta + \{\gamma - 1\}/2)] dt + i\beta d\widetilde{W} . \quad (6.35b)$$

This two-particle loss model can give either correct or incorrect results when treated with the usual positive P representation methods. Generally, problems only arise when any accompanying single-particle losses have small rates or when the number of bosons per mode is small (see Figure 6.3). It is a well-studied case, and a detailed treatment of the boundary term errors that appear can be found in [64]. It also has the merit that exact solutions can be readily found using other means. Analyzing this example gives a good check that the moving singularity

removal procedure of Section 6.3.2 does in fact eliminate boundary terms and leads to correct results. Several simplifications of this model will now be concentrated on, which correspond to existing literature.

6.4.2 Two-boson absorber

In the simplest form of this model, corresponding to $\gamma = \varepsilon = 0$, only two-boson absorption takes place. One expects that for some state $|\psi\rangle = \sum_{n=0}^{\infty} c_n |n\rangle$ in a Fock number state basis, all even boson number components will decay to vacuum, and all odd-numbered components will decay to $|1\rangle$, leaving a mixture of vacuum and one-boson states at long times.

The positive P representation has been found to give erroneous results [70, 83, 84, 85] accompanied by moving singularities [64]. There are power-law tails in the distribution that lead to boundary term errors — consistently with the conditions (6.7). The observable usually concentrated on in this system is the occupation number $\hat{n} = \hat{a}^\dagger \hat{a}$, which is estimated by $\bar{n} = \langle \alpha\beta = \check{n} \rangle_{\text{stoch}}$ in the positive P representation. The variable $\check{n} = \alpha\beta$ has a convenient closed equation (Stratonovich),

$$d\check{n} = -2\check{n}(\check{n} + i\mathcal{G}^+ - 1/2) dt + i\check{n} dW^+ \quad (6.36)$$

with $dW^+ = (dW + d\widetilde{W})$, and $\mathcal{G}^+ = (\mathcal{G} + \widetilde{\mathcal{G}})/2$.

Let us examine the behavior of the above equation when $\mathcal{G}^+ = 0$, i.e., in the standard, un-gauged formulation. The deterministic part of the evolution has a repeller at $\check{n} = 0$, and an attractor at $\check{n} = 1/2$. The noise is real, finite, and of standard deviation $\sqrt{dt/2}$ at the attractor. The deterministic part of the evolution has a single trajectory of measure zero that can escape to infinity along the negative real axis,

$$\alpha(t) = -\beta(t) = \frac{1}{\sqrt{2(t_{\text{sing}} - t)}}, \quad (6.37)$$

where $t_{\text{sing}} = 1/2\alpha(0)^2 = -1/2\check{n}(0)$. This moving singularity is a symptom of likely boundary term errors.

Indeed, in the steady-state limit, all trajectories in a simulation will head toward $\check{n} = 1/2$, with some remnant noise around this value being kept up by the stochastic

term. This noise does not affect the average, however, and the estimator of $\langle \hat{n} \rangle$ becomes $\lim_{t \rightarrow \infty} \bar{n} = \lim_{t \rightarrow \infty} \langle \check{n} \rangle_{\text{stoch}} = 1/2$ in the many trajectories limit. Quantum mechanics, however, predicts that if we start from a state $\hat{\rho}(0)$, the steady state will be

$$\lim_{t \rightarrow \infty} \langle \hat{n} \rangle = \sum_{j=0}^{\infty} \langle 1 + 2j | \hat{\rho}(0) | 1 + 2j \rangle , \quad (6.38)$$

since all even occupations decay to vacuum, and all odd to $|1\rangle$. For a coherent state $|\alpha_0\rangle$ input, say, this will be

$$\lim_{t \rightarrow \infty} \langle \hat{n} \rangle = \frac{1}{2} \left(1 - e^{-2|\alpha_0|^2} \right) . \quad (6.39)$$

Thus one expects that the positive P simulation will give correct results only when $e^{|\alpha_0|^2} \gg 1$.

To correct the problem one has to change the phase-space topology in some way to prevent the occurrence of such moving singularities. It was found that a good gauge for a two-boson absorber nonlinearity in general is

$$\mathcal{G} = \tilde{\mathcal{G}} = \mathcal{G}^+ = i(\check{n} - |\check{n}|) . \quad (6.40)$$

This will be dubbed the ‘‘circular’’ gauge due to the creation of an attractor at $|\check{n}| = 1/2$.

This gauge has the effect of replacing the $-2\check{n}^2$ term in (6.36) with $-2\check{n}|\check{n}|$, which is always a restoring force, and so never leads to any escape to infinity in finite time. With the gauge (6.40), the Stratonovich equations become

$$d\check{n} = \check{n}(1 - 2|\check{n}|) dt + i\check{n}dW^+ , \quad (6.41a)$$

$$d\alpha = -\alpha(|\check{n}| - 1/2) dt + i\alpha dW , \quad (6.41b)$$

$$d\Omega = \Omega \left\{ [\check{n} + (\check{n} - |\check{n}|)^2] dt + i(\check{n} - |\check{n}|) dW^+ \right\} . \quad (6.41c)$$

How do these new equations measure up to the four symptoms of boundary term errors of the first kind outlined in Section 6.1.2? In logarithmic variables $\alpha_L = \log \alpha$, $n_L = \log \check{n}$, and $z_0 = \log \Omega$, the Stratonovich system equations are

$$dn_L = (1 - 2|\check{n}|) dt + idW^+ \quad (6.42a)$$

$$d\alpha_L = -(|\check{n}| - 1/2) dt + idW \quad (6.42b)$$

$$dz_0 = [\check{n} + (\check{n} - |\check{n}|)^2] dt + i(\check{n} - |\check{n}|) dW^+ . \quad (6.42c)$$

One now has an attractor on the circle $|\check{n}| = 1/2$, and a repellor at $\check{n} = 0$, with free phase diffusion of n in the tangential direction. Once trajectories reach the attractor, only phase diffusion occurs. So:

1. *Moving singularities:* There are no moving singularities in n_L , since $|\check{n}|$ always remains bounded. Also, since the deterministic evolution of the remaining independent variables α_L and z_0 depends only on \check{n} or $|\check{n}|$ in such a way that $d\alpha_L$ and dz_0 remain finite for finite \check{n} , then these variables never escape to infinity in finite time (being just a time integral of finite values of \check{n}). Thus, no moving singularities occur.
2. *Noise divergences:* Similarly because the noise coefficient of n_L is constant, while the noise coefficients of α_L and z_0 depend only on \check{n} , and in such a way that for finite \check{n} the noise magnitude remains finite — then one concludes that noise divergences are not present either.
3. *Discontinuous drift or noise:* Not present, by inspection.
4. *Initial distribution:* Initial coherent state has $P(0) = \delta^2(\alpha - \alpha_0)\delta^2(\beta - \alpha^*)\delta^2(\Omega - 1)$. There are no boundary terms at the outset with such an infinitely-peaked distribution.

How does this change in the Ito formulation? Using (B.7) on (6.42), one finds that the Ito and Stratonovich equations are identical for n_L and α_L , while the Ito weight equation is $dz_0 = (\check{n} - |\check{n}|)^2 dt + i(\check{n} - |\check{n}|) dW^+$. The above arguments apply without change to Ito, and to hybrid Ito-Stratonovich algorithms. So — no symptoms of boundary term errors remain in any formulation.

Let us see how the gauge P simulation compares with the positive P in practice — i.e. are boundary term errors *really* removed along with the moving singularities? Figure 6.2 compares results for an (exact) truncated number-state basis calculation, a positive P calculation, and a gauge calculation using the circular gauge (6.40) for an initial coherent state of $\alpha_0 = 1/\sqrt{2}$. Figure 6.3 compares steady-state values for exact, positive P, and gauge calculations for a wide range of initial coherent states. It is seen that the gauge calculation is correct to within the small errors due to

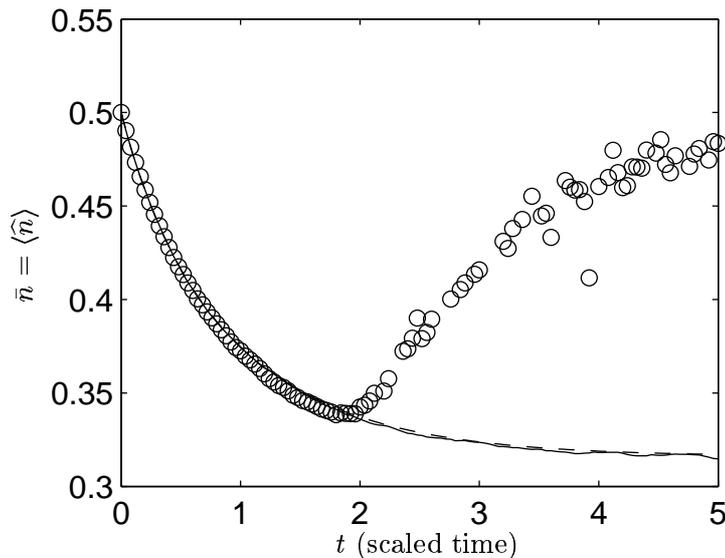


Figure 6.2: **Comparison of two-boson absorption simulations:** time-varying mode occupation. CIRCLES: positive P simulation; SOLID LINE: gauge simulation using circular gauge (6.40); DASHED LINE: exact calculation (truncated number-state basis). Simulation parameters: $\mathcal{S} = 4 \times 10^4$ trajectories; initial coherent state.

finite sample size, whereas the positive P calculation leads to severe errors at low occupations.

6.4.3 One- and two-boson absorber

If one now considers single-boson decay as well, but still with no driving, one expects that all states will decay to the vacuum on a time scale $1/\gamma$ on top of the two-particle losses, which produce decay of particle numbers on a timescale of 1 (although not necessarily right down to the vacuum.) When $\gamma \gg 1$, the single-particle decay dominates and nothing interesting is seen beyond an exponential decay of particle number. However, if $\gamma \lesssim 1$, one should first see a rapid decay to a mixture of vacuum and one-boson states due to the two-boson process, and then a slow decay of the one-boson state to the vacuum on a time scale of $t \approx 1/\gamma$.

In this case the positive P (Stratonovich) equations display different behavior depending on whether γ is above or below the threshold $\gamma = 1$. Below threshold, there is an attractor at $\check{n} = (1 - \gamma)/2$, and a repeller at $\check{n} = 0$, while above threshold, the attractor is at $\check{n} = 0$, and the repeller at $\check{n} = -(\gamma - 1)/2$. In either case, there is a

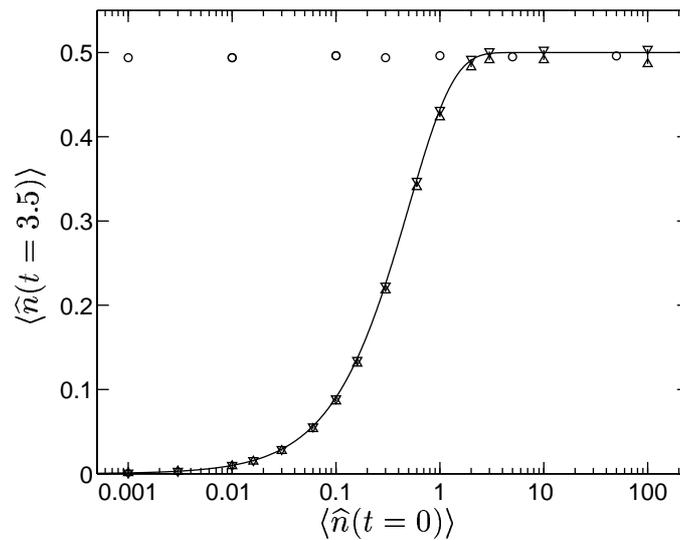


Figure 6.3: **Steady state expectation values of boson number $\langle \hat{n} \rangle$** obtained by gauge simulations (DOUBLE TRIANGLES) compared to exact analytic results (6.39) (SOLID LINE) and positive P simulations (CIRCLES) for a wide range of initial coherent states. Size of uncertainty in gauge results due to finite sample size is indicated by vertical extent of ‘double-triangle’ symbol. Steady state was observed to have been reached in all simulations by $t = 3.5$ or earlier (compare with Figure 6.2 and 6.4), hence this is the time for which the simulation data is plotted. $\mathcal{S} = 10^5$ trajectories

divergent trajectory along the negative real axis, which is again a moving singularity. It turns out that the steady state calculated with the positive P is erroneous while $\gamma < 1$, and that there are transient boundary term errors while $\gamma < 2$ [70]. The false steady state below threshold lies at the location of the attractor: $(1 - \gamma)/2$.

Let us try to fix this problem using the same circular gauge (6.40) as before. The gauged equation for n_L is now

$$dn_L = (1 - \gamma - 2|\check{n}|) dt + idW^+ , \quad (6.43)$$

while the α and Ω evolution is unchanged. So, above threshold one is left with only an attractor at $\check{n} = 0$, while below threshold one has a repellor at $\check{n} = 0$ surrounded by an attracting circle at $|\check{n}| = (1 - \gamma)/2$. This phase space again has no moving singularities or noise divergences.

The results of simulations for the parameter $\gamma = 0.1$ are shown in Figure 6.4 ($\gamma \ll 1$ was chosen so that a system with two widely differing time scales is tested.) The gauge simulation tracks the exact results and avoids the false results of the positive P simulation. Note also that the gauge simulation remains efficient for a wide range of occupation numbers — from $\langle \hat{n} \rangle \approx 100 \gg 1$, where the positive P is also accurate, to $\langle \hat{n} \rangle \approx 0.1 \ll 1$ where it is totally incorrect.

6.4.4 Driven two-boson absorber

The other type of situation to consider is when there is a driving field as well as two-boson damping. In this Subsection the coherent driving ε is nonzero, but the single-particle loss rate is assumed negligible ($\gamma = 0$), since this process never causes any of the simulation problems anyway, but leaving it out simplifies analysis. Failure of the positive P representation method has been found in this limit as well [71], and is evident in Figure 6.5. The equation for \check{n} is no longer stand-alone in this case, and all three complex variables must be simulated as in (6.34c), the Ω equation being the same as in the undriven case (6.41).

A treatment of the singular trajectory problem with the same circular gauge (6.40) leads again to correct results, as seen in Figure 6.5.

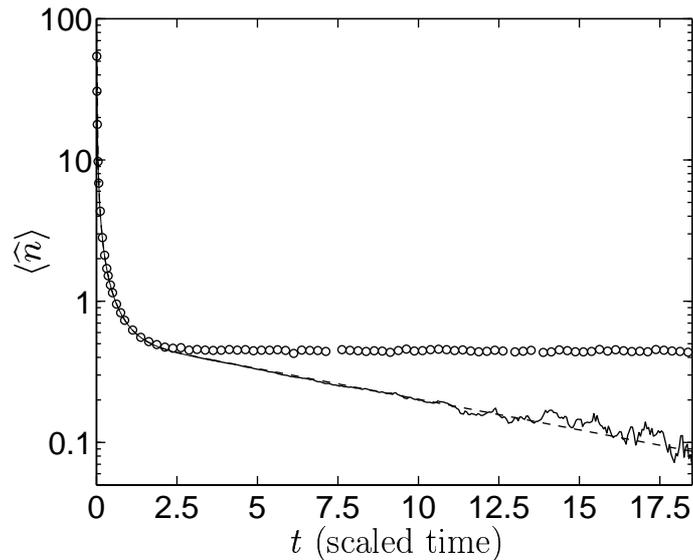


Figure 6.4: **Comparison of simulations for system with both one- and two-boson damping.** Relative single-boson damping strength $\gamma = 0.1$; CIRCLES: positive P simulation. SOLID LINE: simulation using the circular gauge (6.40); DASHED LINE: exact calculation (truncated number-state basis). Gauge simulation parameters: 100 000 trajectories; initial coherent state $|10\rangle$ with $\langle \hat{n} \rangle = 100$ bosons.

6.4.5 Relevance to interacting Bose gas

The master equation (6.32) was first considered in the context of quantum optics. There, single-photon losses are usually large relative to two-photon processes, and mode occupations are high, so the boundary term errors have not been a practical problem in their original quantum optics regime. However, for other physical systems such as interacting Bose gases (which this thesis is most interested in), two-particle losses can be stronger, one expects a large number of modes to have low or practically zero occupation, and so the boundary term errors may be significant. For example, collisional many-body processes that remove particles in pairs from the modeled field $\hat{\Psi}$, will lead to similar two-particle loss terms (2.24) in the master equation.

6.5 Removal example 2: Single-mode laser

Let us now consider the second quantum system for which systematic errors have been seen with the positive P representation. The mechanism by which boundary

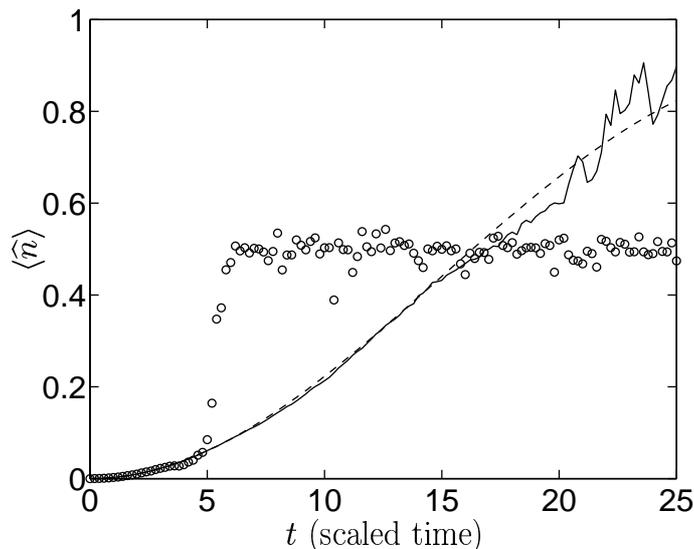


Figure 6.5: **Driven two-boson absorber** with $\varepsilon = 0.05$. CIRCLES: positive P simulation (1000 trajectories); SOLID LINE: circular gauge (6.40) simulation (100 000 trajectories); DASHED LINE: exact calculation (truncated number-state basis). Initial vacuum state.

term errors occur here is somewhat different than in the two-boson absorber. For two-boson damping, boundary term errors occur even when one chooses an optimal (i.e., compact) initial distribution to represent the starting state $|\alpha_0\rangle\langle\alpha_0|$, whereas for this laser model the systematic errors occur only if one starts with unusually broad initial distributions. Nevertheless, such broader distributions may arise during time-evolution from some other state, so consideration of what occurs for significant distribution broadness is also relevant to practical simulations.

This section is closely based on Section IV of the published article by Deuar and Drummond[66].

6.5.1 The laser model

For a simple model of a photonic or atomic laser, using the single-mode positive P representation,

$$\widehat{\Lambda}(C = \{\alpha, \beta\}) = \|\alpha\rangle\langle\beta^*|e^{-\alpha\beta}, \quad (6.44)$$

one can derive[71, 64] the the Fokker-Planck equation

$$\frac{\partial P}{\partial t} = \left\{ \frac{\partial}{\partial \alpha} \left[\frac{\alpha\beta}{\mathcal{N}} - G \right] \alpha + \frac{\partial}{\partial \beta} \left[\frac{\alpha\beta}{\mathcal{N}} - G \right] \beta + 2Q\mathcal{N} \frac{\partial^2}{\partial \alpha \partial \beta} \right\} P, \quad (6.45)$$

where the average number of photons in the mode is $\bar{n} = \langle \alpha\beta \rangle_{\text{stoch}}$, and \mathcal{N} is a scaling parameter that equals the number of atoms in the gain medium. Both G , the gain parameter, and $Q \geq G/\mathcal{N}$, the noise parameter, are real and positive. One usually considers the scaled variables

$$\tilde{\alpha} = \alpha/\sqrt{\mathcal{N}} \quad (6.46a)$$

$$\tilde{\beta} = \beta/\sqrt{\mathcal{N}}, \quad (6.46b)$$

The stochastic Ito equations are then [71, 64]

$$d\tilde{\alpha} = (G - \tilde{\alpha}\tilde{\beta})\tilde{\alpha} dt + \sqrt{2Q} d\eta \quad (6.47a)$$

$$d\tilde{\beta} = (G - \tilde{\alpha}\tilde{\beta})\tilde{\beta} dt + \sqrt{2Q} d\eta^*, \quad (6.47b)$$

where $d\eta$ is a complex Wiener increment obeying $\langle d\eta^* d\eta \rangle_{\text{stoch}} = dt$, $\langle d\eta \rangle_{\text{stoch}} = \langle d\eta^2 \rangle_{\text{stoch}} = 0$.

One is again usually interested in the (scaled) occupation of the lasing mode

$$\bar{n}_{\text{sc}} = \left\langle n_{\text{sc}} = \tilde{\alpha}\tilde{\beta} \right\rangle_{\text{stoch}} = \langle \hat{a}^\dagger \hat{a} \rangle / \mathcal{N}, \quad (6.48)$$

i.e. the number of bosons in the lasing mode per each atom in the gain medium. Changing variables to n_{sc} and an auxiliary ‘‘phase-like’’ variable $z_2 = \tilde{\alpha}/\tilde{\beta}$, the FPE becomes

$$\frac{\partial P}{\partial t} = \left\{ 2 \frac{\partial}{\partial n_{\text{sc}}} [Q + n_{\text{sc}}(n_{\text{sc}} - G)] + 2Q \frac{\partial}{\partial z_2} \frac{n_{\text{sc}}}{z_2} + 2Q \frac{\partial^2}{\partial n_{\text{sc}}^2} n_{\text{sc}} - 2Q \frac{\partial^2}{\partial z_2^2} \frac{z_2^2}{n_{\text{sc}}} \right\} P. \quad (6.49)$$

The diffusion matrix is now decoupled

$$4Q \begin{bmatrix} n_{\text{sc}} & 0 \\ 0 & -z_2^2/n_{\text{sc}} \end{bmatrix}, \quad (6.50)$$

and so the two complex variables can evolve under independent noises. In fact the entire n_{sc} evolution decouples from z_2 , and can be written as the closed Ito equation

$$dn_{\text{sc}} = -2(n_{\text{sc}} - a)(n_{\text{sc}} - b)dt + 2\sqrt{Qn_{\text{sc}}}dW, \quad (6.51)$$

where now the *real* Wiener increments (implemented as Gaussian noises) have variance $\langle dW^2 \rangle_{\text{stoch}} = dt$, and the deterministic stationary points are

$$a = \frac{1}{2} \left(G + \sqrt{G^2 + 4Q} \right) \geq 0 \quad (6.52a)$$

$$b = \frac{1}{2} \left(G - \sqrt{G^2 + 4Q} \right) \leq 0. \quad (6.52b)$$

(In the Stratonovich calculus, $[a/b] = \frac{1}{2}(G \pm \sqrt{G^2 + 2Q})$). One finds that the stationary point at a is an attractor, and at b there is a repellor. Defining $n_{\text{neg}} = b - n_{\text{sc}}$, as the distance (in the negative direction) from the repellor, one obtains

$$\frac{dn_{\text{neg}}}{dt} = 2n_{\text{neg}}(n_{\text{neg}} + \sqrt{G^2 + 4Q}) + \text{noise} , \quad (6.53)$$

which shows that there is a singular trajectory escaping to infinity in finite time along the negative real axis $n_{\text{sc}} < b$ — i.e. a moving singularity.

6.5.2 Initial conditions

Let us consider the case of vacuum initial conditions. A vacuum can be represented by the positive P representation

$$P_+(\tilde{\alpha}, \tilde{\beta}) = \delta^2(\tilde{\alpha}) \delta^2(\tilde{\beta}) , \quad (6.54)$$

but also by Gaussian distributions of any variance σ_0^2 , around the above,

$$P^{(+)}(\tilde{\alpha}, \tilde{\beta}) = \frac{1}{4\pi^2\sigma_0^4} \exp \left\{ -\frac{|\tilde{\alpha}|^2 + |\tilde{\beta}|^2}{2\sigma_0^2} \right\} . \quad (6.55)$$

It can be checked that this satisfies the vacuum positive P representation condition (3.34), giving another family of equivalent vacuum distributions different to (3.35), which was found in Section 3.3.6. Note: the distribution of n_{sc} is non-Gaussian, but has a standard deviation of $\sigma_{n_{\text{sc}}} \approx \sqrt{2}\sigma_0^2$ in both the real and imaginary directions.

It has been found by Schack and Schenzle [71] that for this single-mode laser model, a positive P simulation of pumping from a vacuum will give correct answers if the usual δ -function initial condition (6.54) is used, but will have systematic errors if an initial condition (6.55) has a sufficiently large variance (see Figure 6.6).

This can be understood by inspection of the stochastic equation (6.51). If one has a sufficiently broad initial distribution, that some trajectories start with

$\text{Re}\{n_{\text{sc}}\} < b$, then the singular trajectory will be explored by the distribution. From this it appears that for all $\sigma_0 > 0$, some boundary term errors may be present. In practice, no errors were seen by Schack and Schenzle for small enough σ_0 , presumably because their magnitude was below the signal-to-noise resolution obtained in those simulations — an interesting effect in itself. Additionally, even if initially no trajectories fall on the singular trajectory (e.g. some initial distribution that is exactly zero for $\text{Re}\{n_{\text{sc}}\} < b$), the region $\text{Re}\{n_{\text{sc}}\} < b$ may be subsequently explored due to the presence of the noise terms.

Apart from the obvious δ -function initial condition, one might want to try the canonical distribution (3.10) applicable generally to any density matrix. For optical laser models, this will not cause practical problems as then the variance is $\sigma_0^2 = 1/\mathcal{N}$, which for any realistic case will be very small (i.e., $\sigma_{n_{\text{sc}}} \ll |b|$). Hence, boundary term errors would be negligible for a wide range of simulation times, based on the evidence that they were not seen [71] in this regime. Nevertheless, broader distributions may arise during time-evolution of some other state *to* a vacuum, so consideration of what occurs for significant σ_0 is also relevant to practical simulations.

Incidentally, the anomalous results were discovered in [71] when $\sigma_0^2 = 1$ was chosen by the erroneous procedure of scaling the equations while not simultaneously scaling the canonical initial condition in α .

6.5.3 Drift gauged equations

Let us introduce the single-mode gauge P kernel (5.4) (compare with (6.44), which differs only by lack of the global weight factor e^{z_0}). Proceeding to introduce drift gauges as in Section 4.3 (for n_{sc} but not for z_2 , as averages of this variable will not figure in what follows), one obtains the Ito stochastic equations

$$dn_{\text{sc}} = 2n_{\text{sc}}(G - n_{\text{sc}}) dt + 2Q dt + 2\sqrt{Qn_{\text{sc}}}(dW - \mathcal{G} dt) \quad (6.56a)$$

$$dz_0 = -\frac{1}{2}\mathcal{G}^2 dt + \mathcal{G}dW, \quad (6.56b)$$

as per the standard gauge formulation (4.90). It is convenient to define a transformed gauge function

$$\mathcal{G}_{(n)} = \mathcal{G} \sqrt{\frac{Q}{n_{\text{sc}}}} \quad (6.57)$$

, which is also arbitrary and enters into the simulation as

$$dn_{\text{sc}} = 2n_{\text{sc}}(G - n_{\text{sc}} - \mathcal{G}_{(n)}) dt + 2Q dt + 2\sqrt{Qn_{\text{sc}}} dW \quad (6.58a)$$

$$dz_0 = -\frac{n_{\text{sc}}\mathcal{G}_{(n)}^2}{2Q} dt + \mathcal{G}_{(n)}\sqrt{\frac{n_{\text{sc}}}{Q}} dW. \quad (6.58b)$$

6.5.4 Correcting for the moving singularity

Consider the (Ito) deterministic evolution of, n'_{sc} , the real part of $n_{\text{sc}} = n'_{\text{sc}} + in''_{\text{sc}}$,

$$\frac{dn'_{\text{sc}}}{dt} = -2n_{\text{sc}}'^2 + 2Gn'_{\text{sc}} + 2Q + 2n_{\text{sc}}''^2 - 2\text{Re}\{n_{\text{sc}}\mathcal{G}_{(n)}\}. \quad (6.59)$$

The moving singularity is due to the $-2n_{\text{sc}}'^2$ leading term for negative values of n'_{sc} . The drift gauge is now chosen according to the criteria below (as in Section 6.3.2, but in a different order):

1. As pointed out in Sections 4.3.4 and 6.3.2, it is desirable to keep the gauge terms to a minimum because whenever they act the weights of trajectories become more randomized. Thus, let us restrict ourselves to functions $\mathcal{G}_{(n)}$ that are only nonzero for $n'_{\text{sc}} < 0$.
2. There should not be any discontinuities in the drift equations, as this might lead to bad sampling of some parts of the distribution and thus possibly to systematic biases. A further problem with discontinuous gauges is that the Stratonovich correction term for dz_0 cannot be calculated at the discontinuity, preventing use of the more stable semi-implicit numerical algorithm (See Appendix B).

Hence, in particular, in light of the previous point 1., one should have $\lim_{n'_{\text{sc}} \rightarrow 0}(\mathcal{G}_{(n)}) = 0$. For ease of analysis, let us start with a simple form for the gauge, $\mathcal{G}_{(n)} = c - \lambda n'_{\text{sc}} + \lambda_y n''_{\text{sc}}$. Continuity immediately implies $c = \lambda_y = 0$, hence

$$\mathcal{G}_{(n)} = \begin{cases} -\lambda n'_{\text{sc}} & \text{if } n'_{\text{sc}} < 0 \\ 0 & \text{if } n'_{\text{sc}} \geq 0 \end{cases}, \quad (6.60)$$

3. The next necessary (and most important) condition, to remove moving singularities, is that the $-2n'_{\text{sc}}{}^2$ term is canceled, hence:

$$\lambda \geq 1. \quad (6.61)$$

Using (B.7), the Stratonovich correction for dn_{sc} is $-Q dt$, and so the family of hybrid Ito-Stratonovich equations differ only in having this term a multiple of $Q dt$ in the range $[0, 2Q dt]$. This has no effect on the removal or not of the moving singularity.

4. Now, if $\lambda = 1$ there are no systematic errors, but the sampling error very quickly obscures any observable estimates because n'_{sc} still heads to $-\infty$ exponentially fast due to the $2Gn'_{\text{sc}}$ term. This takes it into regions of ever increasing $|\mathcal{G}_{(n)}|$, and weights quickly become randomized. For slightly larger parameters λ , the n'_{sc} evolution takes trajectories to a point lying far into the negative n'_{sc} region where the two leading terms balance. Here the trajectories sit, and rapidly accumulate weight noise. It is clear that for an optimum simulation all stationary points of n'_{sc} in the nonzero gauge region must be removed. In this system the condition for this is

$$\lambda > 1 + \frac{G^2}{2Q}, \quad (6.62)$$

for the Stratonovich calculus, and $\lambda > 1 + G^2/4Q$ for Ito.

5. Are any new moving singularities or noise divergences introduced? The gauged equation for the imaginary part of n_{sc} is

$$dn''_{\text{sc}} = 2n''_{\text{sc}}[G + n'_{\text{sc}}(\lambda - 2)] dt + 2\sqrt{Q|n_{\text{sc}}|} \sin(\frac{1}{2}\angle n_{\text{sc}}) dW, \quad (6.63)$$

which is attractive towards $n''_{\text{sc}} = 0$ when $\lambda > 2$ in the gauged ($n'_{\text{sc}} < 0$) region, and exponentially growing otherwise. In either case, no super-exponential escape occurs, and no moving singularities. The evolution of z_0 depends only on finite expressions of finite n_{sc} , is thus an integration over the history of a trajectory, and so does not diverge in finite time. The Stratonovich correction for dz_0 is $\lambda(n'_{\text{sc}} + n_{\text{sc}} + |n_{\text{sc}}|)/2$ in the gauged region, zero otherwise. Introduction

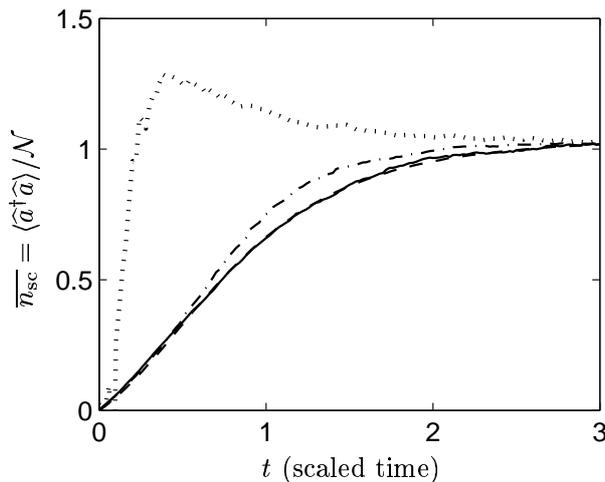


Figure 6.6: One-mode laser $G = 1$, $Q = 0.25$. DASHED LINE: (correct) positive P simulation with δ -function initial conditions (6.54) $\sigma_0^2 = 0$, and 100000 trajectories. DOTTED-DASHED LINE: erroneous positive P simulation with Gaussian initial conditions (6.55) $\sigma_0^2 = 0.1$ initially, and 100000 trajectories. DOTTED LINE: erroneous positive P simulation with $\sigma_0^2 = 1$, and 10000 trajectories. SOLID LINE: gauge calculation for $\sigma_0^2 = 0.1$ with $\lambda = 4$, which corrects the systematic error of the positive P. Only 4000 trajectories, so as not to obscure other data.

of such a term or its multiple does not lead to any moving singularities for z_0 evolution either. One concludes that no new moving singularities or noise divergences are introduced by the gauge (6.60).

The results for an example have been plotted in Figure 6.6. The parameters there were $G = 1$, $Q = 0.25$ (leading to $a \approx 1.1124$ and $b \approx -0.1124$ in Stratonovich formulation). The initial condition in the example was $\sigma_0^2 = 0.1$, which is already much larger than the canonical variance for physically likely parameters in a photonic laser. Typical values of $|n_{sc}|$ initially will be of order $\sigma_{n_{sc}} \approx 0.14 \gtrsim |b|$ here. A good choice of gauge was found to be $\lambda = 4$. The use of this gauge restores the correct results.

6.5.5 Non-optimal initial conditions

As one increases the spread of the initial distribution beyond $\sigma_{n_{sc}} \approx |b|$, it becomes increasingly difficult to find a gauge that will give reasonable simulations. For example a wide variety of what seemed like promising gauges for $\sigma_0^2 = 0.3$ have been

tried (with the same values of parameters Q and G as in Figure 6.6), and none have come close to success. The problem is that while systematic errors are removed, large random noise appears and obscures whatever one is trying to calculate.

Trajectories that start off at a value of n'_{sc} lying significantly beyond b require a lot of modification to their subsequent evolution to (1) stop them from escaping to $-\infty$ and (2) move them out of the gauged region of phase space so that they do not accumulate excessive weight noise. If there are many of these, the trade-off between the gauge size and length of time spent in the gauged region does not give much benefit anymore. Some consolation is gained in knowing that results will at worst be noisy and unusable, rather than being systematically incorrect.

6.6 Summary of main boundary term results

The two mechanisms that lead to boundary term errors have been specified and investigated for general phase-space distribution methods: 1) Errors of the first kind, which are caused by discarding nonzero boundary terms in the partial integration step between the master and Fokker-Planck equations, and 2) Errors of the second kind, which can arise if the estimator of a particular observable for a given kernel grows too rapidly as the boundaries of phase-space C are approached, and may be present or absent depending on the moment chosen.

The exact expression for the first kind of boundary terms has been found: (6.1), and the integrals that must converge for moment calculations to avoid the second kind are given by (6.8). While these cannot in general be explicitly evaluated for non-toy models because the exact form of the distribution $P(C)$ at the boundaries of phase space is required, a simple example has been solved for instructive purposes in Section 6.2.2, and these boundary term expressions explicitly evaluated. For the gauge P representation, conditions on $P(C)$ (6.7) and (6.24) are found, which are sufficient to avoid boundary term errors of the first and second kind (respectively). In particular, for the gauge P representation:

1. The simulation will be free of boundary terms of the first kind when: A) Kernel drift gauges are polynomial in the system variables or their conjugates, and B)

the far tails of $P(C)$ decay faster than exponential in the coherent amplitudes and faster than power law in Ω .

2. Estimators of observables polynomial in the \hat{a}_k and \hat{a}_k^\dagger operators are found to be free of boundary term errors of the second kind provided that the far tails of $P(C)$ decay faster than a power law.

This is consistent with the results found by Gilchrist *et al*[64]. No indication that such biases will occur for interacting Bose gas simulations is found in Section 6.2.4.

In Section 6.1.2, four symptoms of boundary term errors of the first kind are pointed out. These are moving singularities, noise divergences, discontinuities, and excessively broad initial conditions. All four can be checked for by inspection and some analysis of the stochastic equations and initial condition, before any simulation. Explicit conditions (6.3) have been found for simulations on open complex phase-spaces, which are sufficient to ensure that neither moving singularities nor noise divergences occur. It is also noted that by the Painleve conjecture, boundary term errors of the first kind are likely to be generically present in ungauged many-mode simulations based on analytic kernels.

All the above results have been collected here with the aim of systematizing and reducing the confusion on the topic of boundary term errors. Subsequently, Section 6.3.2 introduces a broadly-applicable heuristic method by which the symptoms of boundary term errors of the first kind can be removed using drift kernel gauges (and hence, presumably, also the boundary term errors as such). This removal method is applied to the two cases in the literature where systematic bias has been reported when using the positive P representation: two-boson absorption, and a single-mode laser model. It is found that the systematic bias is indeed removed from simulation results by this method for these examples.

Part B

Development of gauges for interacting Bose gases

Chapter 7

Gauges for single-mode interacting Bose gas dynamics

7.1 Motivation

Simulations of the dynamics of multi-mode (locally) interacting Bose gases (with lattice Hamiltonian (2.17)) using the positive P representation suffer two major technical problems caused by instabilities in the stochastic equations: 1) Exponential growth of distribution broadness leading to rapid loss of any useful accuracy. 2) Probable moving singularities when two or more modes are coupled (See Section 7.2.5). These issues prevent simulation of all but short time behaviour. Gases with non-local interactions (Hamiltonian (2.12)) are expected to suffer from similar problems on the basis that the locally-interacting lattice model is a special case of these more general models.

The instabilities arise from the two-body interaction terms in the Hamiltonian, and so for a locally-interacting model the unstable processes decouple and are local to each spatial lattice mode of the model. Hence, if the instabilities are brought under control for each mode on its own, then simulations of the full many-mode model should benefit as well.

In this chapter, a single-boson mode with two-body interactions is considered, and gauges (kernel drift and diffusion) are introduced to make improvements. A

drift gauge can remove the offending instability, while the diffusion gauge makes a tradeoff between noise in the phase-space (i.e. α and β) equations and noise in the global weight z_0 , and will be optimized to improve simulation times. In Chapter 8, the performance of the method developed here will be investigated in a coupled two-mode system as a prelude to multi-mode simulations.

7.2 The single-mode interacting Bose gas

7.2.1 Physical model

Consider a single mode extracted from the open multi-mode locally-interacting lattice model of Section 2.5. Annihilation and creation operators for the mode are \hat{a} and \hat{a}^\dagger , and obey $[\hat{a}, \hat{a}^\dagger] = 1$. From (2.17), and with possible coherent gain added as in (2.25), the Hamiltonian is

$$\hat{H} = \hbar\hat{n}(\omega + \chi[\hat{n} - 1]) + i\hbar(\varepsilon\hat{a}^\dagger - \varepsilon^*\hat{a}), \quad (7.1)$$

The master equation of this system interacting with an environment is given by the usual Linblad form (2.20), with Linblad operators

$$\hat{L}_1 = \hat{a}\sqrt{\gamma(1 + \bar{n}_{\text{bath}})} \quad (7.2a)$$

$$\hat{L}_2 = \hat{a}^\dagger\sqrt{\gamma\bar{n}_{\text{bath}}} \quad (7.2b)$$

modeling single-particle interactions with a standard boson heat bath with a mean number of particles $\bar{n}_{\text{bath}}(T)$ per bath mode, as in Section 2.5.

Physically this model approximates a single mode of interest \mathbf{n} of a multi-mode system with two-body scattering, where evolution of the other modes $\mathbf{m} \neq \mathbf{n}$ is assumed negligible. Roughly, the linear self-energy term becomes $\omega = \omega_{\mathbf{nn}}$, highly occupied coherent modes \mathbf{m} can be collected into $\varepsilon = -i \sum_{\mathbf{m}} \omega_{\mathbf{nm}} \langle \hat{a}_{\mathbf{m}} \rangle$, while the remainder of modes can become the heat bath. The single-mode model can also sometimes represent an approximation to a single orbital of a more complex system, if the constants $\omega, \varepsilon, \bar{n}_{\text{bath}}$ and χ are chosen appropriately.

7.2.2 Stochastic gauge P equations

A single-mode gauge P kernel (6.33) is used, where the complex configuration variables are coherent state amplitudes α and β , and a global weight $\Omega = e^{z_0}$. From the multi-mode equations (5.17), (5.18), and (5.21) of Section 5.3, the gauge P Ito equations for this model are

$$d\alpha = -i\omega\alpha dt - 2i\chi\alpha^2\beta dt - \frac{\gamma}{2}\alpha dt + \epsilon dt + \sum_k \underline{B}_{1k}(dW_k - \mathcal{G}_k dt) \quad (7.3a)$$

$$d\beta = i\omega\beta dt + 2i\chi\beta^2\alpha dt - \frac{\gamma}{2}\beta dt + \epsilon^* dt + \sum_k \underline{B}_{2k}(dW_k - \mathcal{G}_k dt) \quad (7.3b)$$

$$d\Omega = \Omega \sum_k \mathcal{G}_k dW_k, \quad (7.3c)$$

where the pre-drift-gauge noise matrices \underline{B}_{jk} defined as in Section 4.3.1 and drift gauges \mathcal{G}_k have not been specified yet. The dW_k are independent Wiener increments, which can be implemented by Gaussian noises of variance dt . The elements of the $2 \times N_W$ complex noise matrices \underline{B} satisfy

$$\underline{D}_{11} = \sum_k \underline{B}_{1k}^2 = -2i\chi\alpha^2 \quad (7.4a)$$

$$\underline{D}_{22} = \sum_k \underline{B}_{2k}^2 = 2i\chi\beta^2 \quad (7.4b)$$

$$\underline{D}_{12} = \sum_k \underline{B}_{1k}\underline{B}_{2k} = \gamma\bar{n}_{\text{bath}}(T). \quad (7.4c)$$

The N_W complex drift gauge functions \mathcal{G}_k are arbitrary in principle.

The standard gauge formulation (4.90) will be used, with the proviso that noises related to the bath interaction and the interparticle interactions are chosen to be independent to simplify the equations as outlined in Section 4.4.7. That is, using diffusion matrix $\underline{D} = \underline{B}\underline{B}^T = \sum_l \underline{D}^{(l)}$ with

$$\underline{D}^{(1)} = 2i\chi \begin{bmatrix} -\alpha^2 & 0 \\ 0 & \beta^2 \end{bmatrix} \quad ; \quad \underline{D}^{(2)} = \gamma\bar{n}_{\text{bath}}(T) \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad (7.5)$$

one obtains the square root noise matrix forms

$$\underline{B}_0^{(1)} = \sqrt{2i\chi} \begin{bmatrix} i\alpha & 0 \\ 0 & \beta \end{bmatrix} \quad ; \quad \underline{B}_0^{(2)} = \sqrt{\frac{\gamma\bar{n}_{\text{bath}}(T)}{2}} \begin{bmatrix} 1 & i \\ 1 & -i \end{bmatrix}. \quad (7.6)$$

In Section 7.2.4 it will be shown that the process responsible for instabilities is the two-particle collisions (parameterized by χ), and so improvements should be

searched for by gauging this process, not the bath interactions (parameterized by γ). Thus, the standard gauges are applied as:

1. Imaginary diffusion gauges g''_{jk} as in Section 4.5. Recall from Sections 4.4.3 and 4.4.4 that only imaginary standard diffusion gauges can alter the stochastic behaviour of the simulation, so only these will be considered. One obtains

$$\underline{B} = \left[\underline{B}_0^{(1)} O(g''_{12}) \quad \underline{B}_0^{(2)} \right], \quad (7.7)$$

with an orthogonal matrix O of the form (4.68) dependent on a single imaginary diffusion gauge $g_{12} = ig''$ with real g'' .

2. Complex drift gauges \mathcal{G}_1 and \mathcal{G}_2 , with the remaining $\mathcal{G}_{k>2} = 0$.

The two-variable diagonal diffusion gauge expression (4.71) applies here, and the Ito stochastic equations become

$$\begin{aligned} d\alpha &= -i\omega\alpha dt - 2i\chi\alpha^2\beta dt - \frac{\gamma}{2}\alpha dt + \epsilon dt + \sqrt{\gamma\bar{n}_{\text{bath}}}d\eta_{\text{bath}} \\ &\quad + i\alpha\sqrt{2i\chi}[(dW_1 - \mathcal{G}_1 dt)\cosh g'' + i(dW_2 - \mathcal{G}_2 dt)\sinh g''] \end{aligned} \quad (7.8a)$$

$$\begin{aligned} d\beta &= i\omega\beta dt + 2i\chi\beta^2\alpha dt - \frac{\gamma}{2}\beta dt + \epsilon^* dt + \sqrt{\gamma\bar{n}_{\text{bath}}}d\eta_{\text{bath}}^* \\ &\quad + \beta\sqrt{2i\chi}[-i(dW_1 - \mathcal{G}_1 dt)\sinh g'' + (dW_2 - \mathcal{G}_2 dt)\cosh g''] \end{aligned} \quad (7.8b)$$

$$d\Omega = \Omega \{ \mathcal{G}_1 dW_1 + \mathcal{G}_2 dW_2 \} \quad (7.8c)$$

in terms of the real Wiener increments dW_1, dW_2 implemented by Gaussian noises of variance dt , and the complex stochastic Wiener-like increment $d\eta_{\text{bath}}$ obeying $\langle d\eta_{\text{bath}} d\eta_{\text{bath}}^* \rangle_{\text{stoch}} = dt$ and $\langle d\eta_{\text{bath}}^2 \rangle_{\text{stoch}} = 0$.

7.2.3 Anharmonic oscillator

Consider the gain-less system ($\bar{n}_{\text{bath}}(T) = \varepsilon = 0$), which contains all the essential features of the two-body interactions. This system is known in the literature as the damped anharmonic oscillator.

For this toy system, the observables can be solved analytically, and comparing simulations with these will be of great use here to optimize diffusion gauge choices. In particular, consider coherent state projector initial conditions $|\alpha_0\rangle\langle\beta_0^*|$, which

correspond to the configuration of a single trajectory out of the \mathcal{S} system samples that together can represent any arbitrary quantum state. Let us concentrate on the mean particle number

$$\langle \hat{n} \rangle = \langle \hat{a}^\dagger \hat{a} \rangle, \quad (7.9)$$

and the first-order time correlation function¹

$$G^{(1)}(0, t) = \beta_0 \langle \hat{a} \rangle = \alpha_0^* \langle \hat{a}^\dagger \rangle^*, \quad (7.10)$$

which contains phase coherence information. Normalizing by $\langle \alpha_0 | \beta_0^* \rangle = \text{Tr} [|\alpha_0\rangle \langle \beta_0^* |]$, their expectation values are found to be

$$\langle \hat{n} \rangle = n_0 e^{-\gamma t}, \quad (7.11a)$$

$$G^{(1)}(0, t) = n_0 e^{-\gamma t/2} e^{-i\omega t} \exp \left\{ \frac{n_0}{1 - i\gamma/2\chi} (e^{-2i\chi t - \gamma t} - 1) \right\}, \quad (7.11b)$$

where $n_0 = \alpha_0 \beta_0$.

When the damping is negligible, n_0 real, and the number of particles is $n_0 \gg 1$, one sees that the initial phase oscillation period (ignoring ω) is

$$t_{\text{osc}} = \frac{1}{2\chi} \sin^{-1} \left(\frac{2\pi}{n_0} \right) \approx \frac{\pi}{\chi n_0}, \quad (7.12)$$

and the phase coherence time, over which $|G^{(1)}(0, t)|$ decays is

$$t_{\text{coh}} = \frac{1}{2\chi} \cos^{-1} \left(1 - \frac{1}{2n_0} \right) \approx \frac{1}{2\chi \sqrt{n_0}}. \quad (7.13)$$

The first quantum revival occurs at

$$t_{\text{revival}} = \frac{\pi}{\chi}. \quad (7.14)$$

¹This first order correlation function is usually written in the Heisenberg picture as $G^{(1)}(0, t) = \langle \hat{a}^\dagger(0) \hat{a}(t) \rangle$. Moving to the Schrödinger picture in which the gauge P representation is defined, one has $G^{(1)} = \text{Tr} [e^{i\hat{H}t/\hbar} \hat{a} e^{-i\hat{H}t/\hbar} \hat{\rho}(0) \hat{a}^\dagger]$. Because of the coherent state conditions used here, $\hat{\rho}(0) = \hat{\Lambda}(\alpha_0, \beta_0, \Omega = 1)$. Thus, $\hat{\rho}(0) \hat{a}^\dagger = \beta_0 \hat{\rho}(0)$ using (5.7d). This gives $G^{(1)} = \beta_0 \text{Tr} [\hat{a} e^{-i\hat{H}t/\hbar} \hat{\rho}(0) e^{i\hat{H}t/\hbar}]$, and identifying the Schrödinger picture density matrix as $\hat{\rho}(t) = e^{-i\hat{H}t/\hbar} \hat{\rho}(0) e^{i\hat{H}t/\hbar}$, this gives the desired expression $G^{(1)}(0, t) = \beta_0 \langle \hat{a} \rangle$. To obtain the right-hand side expression in (7.10), note that $G^{(1)}$ is a complex quantity composed of two observables (see (7.20)) — one for the real, and one for the imaginary part. Taking the adjoint, $G^{(1)}(0, t)^* = \langle \hat{a}^\dagger(t) \hat{a}(0) \rangle$ in the Heisenberg picture, and following the same procedure as above, one obtains $G^{(1)}(0, t)^* = \alpha_0 \langle \hat{a}^\dagger \rangle$.

The stochastic equations for this gain-less system are more convenient in terms of the logarithmic number-phase variables

$$n_L = \log(\alpha\beta) \quad (7.15a)$$

$$m_L = \log\left(\frac{\alpha}{\beta}\right). \quad (7.15b)$$

For coherent states where $\beta = \alpha^*$, n_L is the logarithm of the mean number of particles $|\alpha|^2$, while $m_L = 2i\angle\alpha$, and characterizes the phase of the coherent amplitude.

The Ito equations in the new variables are

$$dn_L = -\gamma dt + 2i\sqrt{i\chi}e^{-g''}(d\eta - \mathcal{G}_{(n)} dt) \quad (7.16a)$$

$$dm_L = -2i(\omega - \chi + 2\chi e^{n_L}) dt + 2i\sqrt{i\chi}e^{g''}(d\eta^* - \mathcal{G}_{(m)} dt) \quad (7.16b)$$

$$dz_0 = \mathcal{G}_{(m)}d\eta + \mathcal{G}_{(n)}d\eta^* - \mathcal{G}_{(n)}\mathcal{G}_{(m)} dt = d(\log \Omega), \quad (7.16c)$$

where the complex Wiener-like increment $d\eta = (dW_1 - idW_2)/\sqrt{2}$ has variances

$$\langle d\eta d\eta^* \rangle_{\text{stoch}} = dt \quad (7.17a)$$

$$\langle d\eta^2 \rangle_{\text{stoch}} = 0, \quad (7.17b)$$

and the independent transformed complex drift gauge functions are

$$\mathcal{G}_{(n)} = \frac{\mathcal{G}_1 - i\mathcal{G}_2}{\sqrt{2}} \quad ; \quad \mathcal{G}_{(m)} = \frac{\mathcal{G}_1 + i\mathcal{G}_2}{\sqrt{2}}. \quad (7.18a)$$

Note: these are not necessarily complex conjugates since the original \mathcal{G}_1 and \mathcal{G}_2 are in general complex.

7.2.4 Behaviour of the positive P simulation

Lack of moving singularities

Consider the anharmonic oscillator equations (7.16) in the positive P representation, where $\mathcal{G}_k = 0$, and $g'' = 0$. By comparison with the condition (6.3), one can check by inspection that moving singularities do not occur : dn_L does not depend on any variables and so is stable, while dm_L depends only on the other variable n_L . Thus the deterministic behaviour of m_L is just integration of a finite function of n_L (which itself remains finite). The noise in m_L is of constant magnitude. z_0 does

not evolve. All variables remain finite, no moving singularities or noise divergences occur. Therefore, none of the symptoms of boundary term errors from Section 6.1.2 are present in the equations.

Indeed, the investigations of Gilchrist *et al*[64] found that boundary term errors do not occur for the special case of the single-mode anharmonic oscillator.

Rapid growth of statistical error

If the observable estimators converge to the quantum mechanical expectation values in the limit of many trajectories so that *accuracy* of the simulation is not a problem, there still remains the issue of how large a sample of trajectories has to be to give a *precise* result.

While the evolution of n_L is well-behaved, being simply constant isotropic diffusion and possibly some decay, the deterministic evolution of m_L can be very sensitive to n_L due to the exponential drift term.

Particularly $\text{Re}\{m_L\}$ has a great influence on observable estimates of phase-dependent observables, such as $G^{(1)}(0, t)$. Consider the quadrature observables

$$\widehat{q}(\theta) = \frac{\widehat{a}e^{i\theta} + \widehat{a}^\dagger e^{-i\theta}}{2}, \quad (7.19)$$

which are closely related to the first-order correlation function

$$G(0, t) = \beta_0 \langle \widehat{q}(0) - i\widehat{q}(\pi/2) \rangle = (\alpha_0 \langle \widehat{q}(0) + i\widehat{q}(\pi/2) \rangle)^*. \quad (7.20)$$

Comparing to (5.10) and (5.13), the estimator for $\langle \widehat{q} \rangle$ is

$$\bar{q}(\theta) = \frac{\langle \text{Re} \left\{ \frac{\Omega}{2} (\alpha e^{i\theta} + \beta e^{-i\theta}) \right\} \rangle_{\text{stoch}}}{\langle \text{Re} \{ \Omega \} \rangle_{\text{stoch}}} \quad (7.21)$$

$$= \frac{\langle \text{Re} \left\{ \exp \left(z_0 + \frac{n_L}{2} \right) \cosh \left(i\theta + \frac{m_L}{2} \right) \right\} \rangle_{\text{stoch}}}{\langle \text{Re} \{ e^{z_0} \} \rangle_{\text{stoch}}}. \quad (7.22)$$

We can see that \bar{q} is very sensitive to $\text{Re}\{m_L\}$ due to the cosh factor. This real part of m_L evolves as

$$d\text{Re}\{m_L\} = 4\chi e^{\text{Re}\{n_L\}} \sin(\text{Im}\{n_L\}) dt + \dots, \quad (7.23)$$

which can be a very rapid growth even for moderate values of $\text{Re}\{n_L\}$. Worse, even a moderate spread in $\text{Re}\{n_L\}$ leads rapidly to a very wide *spread* of $\text{Re}\{m_L\}$,

not to mention the spread in the factor $\cosh(m_L/2)$ that appears in the observable estimate. The behaviour of this spread will be considered in detail in Section 7.5.3, but for now it suffices to point out that there arises a characteristic time scale t_{sim} beyond which the uncertainty in phase-dependent observable estimates grows faster than exponentially with time, and any simulations become effectively useless no matter how many trajectories are calculated.

Numerical investigations (See Figure 7.3 and Table 7.1) found that a positive P simulation lasts at most for times

$$t_{\text{sim}} \approx \frac{(1.27 \pm 0.08)}{\chi n_0^{2/3}}. \quad (7.24)$$

For large mode occupation $n_0 \gg 1$, this is not enough time for significant phase decoherence to occur (compare to t_{coh} in (7.13)). This unfavorable scaling is known to be a major unresolved stumbling block for positive P simulations of the interacting Bose gas model (2.1) in many physical regimes. For example, evaporative cooling simulations encounter such sampling problems upon the onset of Bose condensation[15]. The expression (7.24) indicates that the simulation time for a many-mode system is likely to be limited by the simulation time of the most highly occupied mode (this is confirmed by the simulations in Chapter 10), which makes it especially worthwhile to improve simulation times at large mode occupations.

Greater stability when damped

Consider the evolution of the log-occupation n_L in (7.16). In a positive P simulation (with $g'' = 0$), the noise term produces a spread in both the real and imaginary parts of n_L of standard deviation $\sigma = \sqrt{2\chi t}$. To a good approximation, all trajectories in a reasonable-sized sample will lie within about 4σ of the mean value (which is $\log(n_0) - \gamma t$). If the values of $\text{Re}\{n_L\}$ for all trajectories are $\ll -\log 2\chi$, then the nonlinear term in the dm_L evolution equation becomes negligible, and stability problems abate. This will occur if the damping γ is large enough in comparison with the two-particle collisions χ . In Section 7.5.4, this behaviour will be found to depend on the ‘‘damping strength’’ parameter (7.70), which will quantify what is meant by ‘‘ γ big enough’’. The increased stability at higher damping is a well-known

feature of positive P simulations of the anharmonic oscillator, and/or an interacting Bose gas[46].

7.2.5 Coupling to other modes and moving singularities

Non phase-dependent observables such as $\hat{n} = \hat{a}^\dagger \hat{a}$ can still be calculated at long times for the anharmonic oscillator with the positive P method, because the evolution of their estimators (e.g. $\langle e^{n_L} \rangle_{\text{stoch}}$) does not depend on the unstable m_L . This is no consolation, however, because we are *ultimately* interested in simulating many-mode systems. (Single-mode toy problems will fall even to a brute force truncated number-state basis calculation, so all these involved stochastic schemes are only ultimately justified for many-mode models). When many modes are present, the convenient separation of n_L and m_L evolution seen in (7.16) is no longer present due to mode-mixing terms (from kinetics and external potentials). This occurs even for a single-mode model with nonzero pumping ε , or a finite-temperature heat bath $\bar{n}_{\text{bath}} > 0$, which also model underlying mode-coupling processes. If the n_L and m_L equations are coupled, the growth of the spread of $\text{Re}\{m_L\}$ will also feed the growth of the spread of n_L , making *all* moment calculations intractable after some relatively short time.

Furthermore, moving singularities may appear. Reverting back to considering the α and β equations (7.8), the Ito positive P evolution of α will be of the form

$$d\alpha = -2i\chi\alpha^2\beta dt + \epsilon dt + \dots, \quad (7.25)$$

and in particular

$$d|\alpha| = 2\chi|\alpha|^2|\beta| \sin(\angle\alpha + \angle\beta) dt + \dots \quad (7.26)$$

When $|\beta| \sin(\angle\alpha + \angle\beta)$ is positive, this violates the no-moving-singularities conditions (6.3), $|\alpha|$ grows faster than exponentially, and moving singularities may be possible, leading to boundary term errors. This did not occur for the closed anharmonic oscillator only because of its special symmetry properties that allowed the convenient decoupled logarithmic form of the equations (7.16).

Thus it can be seen that when mode mixing occurs, there are two potential problems that can arise after short times:

1. Rapid appearance of massive statistical errors, masking observable estimates in noise.
2. Possible systematic biases caused by moving singularities in the equations.

Ways to deal with these issues will be investigated in this chapter for a single-mode system, and for coupled-mode systems in Chapter 8.

7.3 Drift gauges: Removal of instability

As was seen in (7.26), the instabilities in the equations arise from the nonlinear two-body drift terms, and in particular from the part that leads to super exponential growth of $|\alpha|$ or $|\beta|$. Defining the real and imaginary parts of the number variable

$$\check{n} = \alpha\beta = e^{nL} = n' + in'' \quad (7.27)$$

for convenience, the offending terms are:

$$d\alpha = 2\chi\alpha n'' dt + \dots \quad (7.28a)$$

$$d\beta = -2\chi\beta n'' dt + \dots \quad (7.28b)$$

whereas the terms

$$d\alpha = -2i\chi\alpha n' dt + \dots \quad (7.29a)$$

$$d\beta = 2i\chi\beta n' dt + \dots \quad (7.29b)$$

affect only the phase of α or β , and are harmless. Furthermore we can see that any nonzero value of $n'' = |\alpha|\text{Im}\{\beta e^{i\angle\alpha}\}$ can lead to moving singularities (in α evolution if $n'' > 0$, or in β evolution if $n'' < 0$). Note that nonzero n'' are characteristic of “non-classical” states that cannot be represented by non-singular Glauber P distributions — i.e. as a mixture of coherent states. This is because for all coherent state trajectories $\beta = \alpha^*$, and so $\check{n} = n'$ only.

Let us follow the heuristic procedure of Section 6.3.2 to remove the instabilities. Considering the points there, in order:

1. REMOVE INSTABILITY: Drift gauges \mathcal{G}_k are chosen so that the terms causing the instability (7.28) are canceled by the drift correction. Assuming no other changes to the drift takes place, then from (7.8), after some algebra, one obtains the required gauges:

$$\mathcal{G}_1 = -\sqrt{2i\chi} n'' e^{-g''} \quad (7.30a)$$

$$\mathcal{G}_2 = i\sqrt{2i\chi} n'' e^{-g''} = -i\mathcal{G}_1 \quad (7.30b)$$

These effectively cause the replacement in the stochastic equations

$$-2i\chi\alpha\check{n} dt \rightarrow -2i\chi\alpha n' dt \quad (7.31a)$$

$$2i\chi\beta\check{n} dt \rightarrow 2i\chi\beta n' dt. \quad (7.31b)$$

In terms of the logarithmic variables, the scaled gauges (7.18) are

$$\mathcal{G}_{(n)} = 0 \quad (7.32a)$$

$$\mathcal{G}_{(m)} = -2\sqrt{i\chi} n'' e^{-g''} = \mathcal{G}_1 \sqrt{2}. \quad (7.32b)$$

For the anharmonic oscillator, dn_L is unchanged from the positive P simulation, and the full equations are

$$dn_L = -\gamma dt + 2i\sqrt{i\chi} e^{-g''} d\eta \quad (7.33a)$$

$$dm_L = -2i(\omega - \chi + 2\chi \operatorname{Re}\{e^{n_L}\}) dt + 2i\sqrt{i\chi} e^{g''} d\eta^* \quad (7.33b)$$

$$dz_0 = -2\sqrt{i\chi} \operatorname{Im}\{e^{n_L}\} e^{-g''} d\eta. \quad (7.33c)$$

2. CHECK AGAIN FOR MOVING SINGULARITIES: (Ignoring, for the time being, any dependence on g'' since the diffusion gauge has not been chosen yet)

- Comparing to conditions (6.3), it is seen that the linear drift terms in parameters ε , ω , γ , and \bar{n}_{bath} in (7.8) will not lead to moving singularities.
- The new nonlinear drift terms (7.31) have lost their radial component due to the drift gauges, and so do not lead to any moving singularities either.

- From (B.7), the Stratonovich corrections for $d\alpha$ and $d\beta$ are $i\chi\alpha dt$ and $-i\chi\beta dt$, respectively. These extra terms satisfy (6.3), and so do not contribute any moving singularities in any of the Ito-Stratonovich hybrid family of algorithms.
- That leaves the new dz_0 evolution. The weight deterministically tracks the n_L evolution, and from (7.33)

$$dz_0 = i \operatorname{Im} \{e^{n_L}\} (dn_L + \gamma dt), \quad (7.34)$$

and so if $n_L = \log(\alpha\beta)$ remains finite for all trajectories (as is seen from the lack of moving singularities in α and β evolution), then so will z_0 .

Thus it is seen that no new moving singularities are introduced provided g'' is well-enough behaved. The dependence on g'' is considered in Section 7.5.7.

3. The noise terms in $d\alpha$ and $d\beta$ are exponential in α , and so also satisfy (6.3) together with the noise terms of dz_0 . thus noise divergences are absent, given a well behaved g'' .
4. As desired, there are no discontinuities in equations, by inspection (provided g'' is well behaved).
5. GAUGE EFFICIENCY. Comparing to corresponding sub-points from Section 6.3.2:
 - (a) Corrections to the drift are necessary in all of phase space apart from the subspaces $n'' = 0$ of measure zero, so generally $\mathcal{G}_k \neq 0$ apart from this special region.
 - (b) For general purpose calculations, one wishes to minimize the quantity $\sum_k |\mathcal{G}_k|^2$, given that the instabilities are removed. This is achieved by the gauges (7.30) as they are only just large enough to remove the instabilities, and do not introduce any other modifications to the $d\alpha$, or $d\beta$ equations.
 - (c) Attractors? For the single-mode anharmonic oscillator when $\varepsilon = 0$, there are deterministic attractors in the phase space at vacuum ($n = 0$), and

$n = (\chi - \omega)/2\chi$ (Stratonovich) or $n = -\omega/2\chi$ (Ito). At all these, $n'' = 0$, so the gauge is also zero. This is the desired situation.

In the more complicated case of nonzero ε , there will be some stationary points elsewhere in phase space. One could try to construct some gauges that would behave as (7.30) in the far tails of phase space i.e. as $|\alpha|, |\beta| \rightarrow \infty$, but would be zero at these stationary points. In the broad picture, however, there seems little point to do this for the single-mode case, because in a many-mode simulation all modes will be coupled together by the kinetic interaction $\omega_{\mathbf{n} \neq \mathbf{m}}$, and ε will be different or absent.

- (d) Rather than tailoring the simulation for a single observable \widehat{O} , the quantity $\sum_k |\mathcal{G}_k|^2$ was minimized here to keep the gauge applicable for a general case.

6. None of the features to be avoided occur.

- (a) The gauge is nonzero over a wide range of phase-space, including regions often visited.
- (b) The gauge does not change in a particularly rapid fashion in phase-space
- (c) The gauge is autonomous.
- (d) The drift gauges break the analyticity of the equations, as suggested by the conjecture of Section 6.1.4.

7.4 Exponentials of Gaussian random variables

From the single-mode equations for the locally-interacting Bose gas (7.8) (or indeed from the many-mode equations (5.17) or (5.50)), it is seen that noise terms due to the interparticle interactions are generally of the multiplicative form

$$d\alpha \propto \alpha \sqrt{\chi} dW_k + \dots \quad ; \quad d\beta \propto \beta \sqrt{\chi} dW_k + \dots \quad (7.35)$$

This leads to Brownian-like motion in the *logarithmic* variables n_L or m_L , as e.g. in (7.16). On the other hand, the observable estimates (e.g. $\langle \widehat{n} \rangle$ or $G^{(1)}(0, t)$) typically

involve quantities such as α or $\alpha\beta$ — in the original non-logarithmic variables. Clearly there will be a lot of averaging over random variables that are of similar form to the exponential of a Gaussian. Let us investigate the behaviour of such random variables.

Let ξ be a Gaussian random variable of mean zero and variance unity, thus its distribution is

$$\Pr(\xi) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{\xi^2}{2}\right). \quad (7.36)$$

The moments of ξ can be found by integration of (7.36) to be

$$\langle \xi^k \rangle_{\text{stoch}} = \begin{cases} \frac{k!}{2^{k/2}(k/2)!} & \text{if } k > 0 \text{ is even} \\ 0 & \text{if } k > 0 \text{ is odd} \end{cases} \quad (7.37)$$

Now let us define the exponential random variable

$$v_\sigma = v_0 e^{\sigma\xi} = e^{vL} \quad (7.38)$$

with positive real σ and v_0 . Using (7.37), the exponential variable's mean is

$$\langle v_\sigma \rangle_{\text{stoch}} = \bar{v}_\sigma = \sum_{k=0}^{\infty} \frac{\sigma^k \langle \xi^k \rangle_{\text{stoch}}}{k!} = v_0 \exp\left(\frac{\sigma^2}{2}\right). \quad (7.39)$$

and so, also

$$\text{var}[v_\sigma] = \sigma_v^2 = v_0^2 e^{\sigma^2} (e^{\sigma^2} - 1) = \left[\bar{v}_\sigma \sqrt{\left(\frac{\bar{v}_\sigma}{v_0}\right)^2 - 1} \right]^2. \quad (7.40)$$

If one is interested in estimating quantities such as \bar{v}_σ using \mathcal{S} samples, then by the Central Limit Theorem, the relative uncertainty in such a finite-sample estimate will be

$$\Delta \bar{v}_\sigma = \frac{\sigma_v}{\bar{v}_\sigma \sqrt{\mathcal{S}}}. \quad (7.41)$$

To obtain only a single significant digit of accuracy, the number of trajectories needed is then

$$\mathcal{S}_{\min} = \frac{100\sigma_v}{\bar{v}_\sigma} = 100\sqrt{e^{\sigma^2} - 1}. \quad (7.42)$$

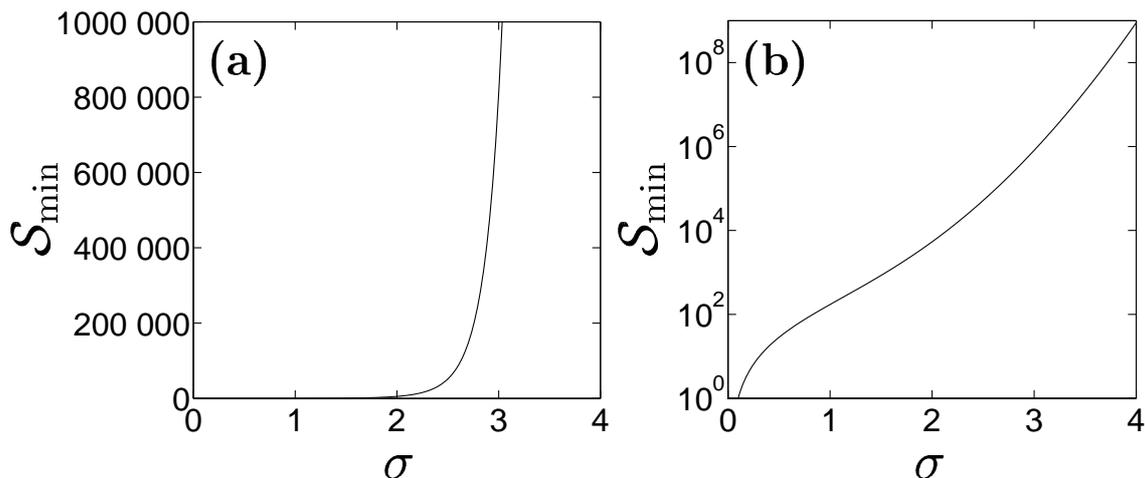


Figure 7.1: **Number of samples** S_{\min} required to obtain a single significant digit of precision in the mean of a random variable v_σ that is an exponential of a Gaussian random variable with standard deviation σ . Scales: **(a)** Linear, **(b)** logarithmic.

This minimum number of required samples (independent of v_0) is plotted in Figure 7.1. It can be seen there that for $\sigma \gtrsim 3$ the number of samples required becomes intractable, and this sets a practical limit on the variance of the logarithmic variable $v_L = \sigma\xi$:

$$\sigma^2 \lesssim \mathcal{O}(10), \quad (7.43)$$

which will be referred to numerous times in this and following chapters.

At small standard deviations $\sigma \lesssim 1$, on the other hand, $\sigma_v \approx \sigma$, and the variance of the logarithmic and exponential variables ($\text{var}[v_L]$ and σ_v^2 , respectively) is approximately the same — a result that will also be useful.

Lastly, this sampling problem can also lead to systematic biases (but not boundary term errors) in averages of v with finite sample number \mathcal{S} once $\sigma \gtrsim 3$, when (7.43) is violated. This is considered in detail in Appendix A.

Means of some combinations of Gaussian noises

Some expressions that will be used in Sections 7.5, and 7.8.2 are given below. Since they contain only odd powers of ξ ,

$$\langle \sin(\sigma\xi) \rangle_{\text{stoch}} = 0 \quad (7.44)$$

$$\langle \xi \cos(\sigma\xi) \rangle_{\text{stoch}} = 0. \quad (7.45)$$

And, by the same series expansion approach as for (7.39)

$$\langle \xi e^{\sigma \xi} \rangle_{\text{stoch}} = \sigma \exp\left(\frac{\sigma^2}{2}\right) \quad (7.46)$$

$$\langle \cos(\sigma \xi) \rangle_{\text{stoch}} = \exp\left(-\frac{\sigma^2}{2}\right) \quad (7.47)$$

$$\langle \xi \sin(\sigma \xi) \rangle_{\text{stoch}} = \sigma \exp\left(-\frac{\sigma^2}{2}\right). \quad (7.48)$$

7.5 Optimization of diffusion gauges

7.5.1 Aims

Upon choice of the drift gauges (7.30) to remove the offending moving singularities, diffusion gauges can be chosen to vastly (as it turns out) improve the sampling behaviour of the simulation. The aim here will be to arrive at a diffusion gauge g'' for the one-mode system that satisfies the following:

1. Improves useful simulation times in the low-damped high boson occupation regime (this is the regime where simulations without diffusion gauges give very unsatisfactory results – see Figure 7.3 and Section 7.2.4).
2. Is easily generalized to many-mode situations. This means, in particular, that the gauge choice should allow for the possibility of mode occupations changing dynamically due to inter-mode coupling (although this is not actually seen in the one-mode toy system).
3. Is expressed as an exact expression, or does not require excessively involved calculations to evaluate. This is important, since in a simulation the gauge should be evaluated at each time step if it is to adapt to dynamically changing mode occupations.
4. Applies on timescales of the order of the coherence time (and shorter timescales as well). This is a relevant timescale only when mode occupations are larger than order unity, and is then smaller than the quantum revival time $t_{\text{revival}} = \pi/\chi$. At lower mean occupations (order unity or smaller), we require that the

simulation remains stable for a time $\mathcal{O}(1/\chi)$, which should be sufficient not to prematurely destabilize any coupled highly-occupied modes in multi-mode systems.

5. Depends only on variables in the current trajectory — i.e. on parameters of the current coherent-state projector $|\alpha\rangle\langle\beta^*|$ component of the full state $\hat{\rho}$. This is necessary if we want to be able to parallelize the calculation, which enormously improves calculation times. Also, if the evolution of all possible coherent-state projectors can be stabilized to restrict distribution size and statistical error, then so will the evolution of all possible states since these are always expressed as a distribution, and estimated as an appropriate sample, of such projectors.
6. We are especially interested in the case of low or absent damping $\gamma \ll 2\chi$, since this is the regime where quantum effects are strongest, and also where the simulation is most unstable (see e.g. Sections 7.2.4 and 7.5.4). The highly-damped regime is a lesser priority, since there a mean-field, or other approximate simulation would be sufficient for most purposes.

This may seem like a lot of conditions for one quantity, but it helps to remember that g'' is in principle an *arbitrary* function.

7.5.2 Variables to be optimized

Before proceeding directly to searching for advantageous values of the diffusion gauge, there remain several more issues to address:

Which moment to optimize?

While, strictly speaking, this depends on which moment we might be interested in, in general we should concentrate on occupation number $\langle\hat{n}\rangle$ or phase correlations $G^{(1)}$. If these basic low-order observables are badly calculated then higher-order observables will not do any better, because they are more sensitive to the distribution broadness.

Consider that estimators for these two observables are (by comparing their definitions (7.9) and (7.10) to observable estimators in the gauge P representation (5.10) and (5.13)) are

$$\langle \hat{n} \rangle \propto \text{Re} \{ \langle \Omega \check{n} \rangle_{\text{stoch}} \} \quad (7.49)$$

$$G^{(1)}(0, t) \propto \beta_0 \langle \Omega \alpha \rangle_{\text{stoch}} \quad \text{or} \quad \propto \alpha_0^* \langle \Omega^* \beta^* \rangle_{\text{stoch}} \quad (7.50)$$

(the normalization by denominators $\text{Re} \{ \langle \Omega \rangle_{\text{stoch}} \}$ has been omitted). One finds from (7.16) or (7.8) that both $d\check{n}$ and $d\Omega$ are proportional to $e^{-g''}$, so varying the diffusion gauge g'' acts to scale noise in the number observable $\langle \hat{n} \rangle \sim \langle \Omega \check{n} \rangle_{\text{stoch}}$ smoothly from very large at negative g'' to very small at positive g'' . This does not suggest any “optimal” values of g'' . On the other hand, the evolution of phase-dependent expectation values such as $G^{(1)}(0, t)$ displays high noise at both large negative and positive values of g'' but via different processes. (Large noise in $\Omega = e^{z_0}$ at negative g'' , and large noise directly in α or β at positive values). This suggests that g'' parameterizes some tradeoff between phase-space and weight noise, and that there is some intermediate value of g'' at which the resulting uncertainty in phase observables is minimized. Thus optimization should be based on such phase variables.

Logarithmic variables

Ideally one would like to optimize for the variance of variables like $(\alpha\Omega)$, since these appear directly in the calculation of the observable $G^{(1)}$. Unfortunately, exact expressions for such variances are difficult to obtain in a closed form², and would probably be very complicated if obtained. A complicated form makes it difficult to arrive at an expression for the optimum g'' that would satisfy condition 3. in Section 7.5.1 (i.e. simple to evaluate during each simulation step).

It turns out, however, that exact expressions for the variances of the logarithms of phase-dependent observables such as $\text{var} [\log(\alpha\Omega)] = \text{var} [z_0 + n_L/2 + m_L/2]$ can

²The reason for this becomes clear after proceeding to calculate some logarithmic variances in (7.65). These require the evaluation of quantities such as $\langle n''(t')n''(t'') \rangle_{\text{stoch}}$ given in (7.64). To evaluate $\text{var} [\Omega\alpha] = \text{var} [e^{G_L}]$, on the other hand, requires calculating averages of the form $\langle n''(t_1)n''(t_2)n''(t_3), \dots \rangle_{\text{stoch}}$ to all orders, which becomes increasingly involved as the number of factors grow.

be found with (relative) convenience and ease, and we will look for values of g'' to minimize this logarithmic variance instead of $\text{var}[\alpha\Omega]$.

As further justification of this choice, one can make an analogy between the behaviour of $\log \alpha$ and α , etc. versus the behaviour of the Gaussian random variable v_L of Section 7.4 and $v_\sigma = e^{v_L}$. In both cases the logarithmic variable is generated by Brownian motion, although in the $\log \alpha$ case, there is additional drift. One sees from (7.40) that for $\sigma \lesssim \mathcal{O}(1)$, the variances of both the logarithmic and “normal” variables are the same. This corresponds to short time evolution of the anharmonic oscillator in the above analogy. Furthermore, because of the rapid rise of variance with time, the range of σ for which $\sigma \not\approx \sigma_v$ but the simulation gives any useful accuracy is fairly narrow and occurs in the more noisy part of the simulation.

In summary, it can be expected that a optimization of g'' in logarithmic variances will still give good results. This is borne out by the massive improvement in simulation times seen in Figure 7.3. Certainly, however, some further improvement could be obtained by considering non-logarithmic variances (particularly since in the full simulation, $\log \alpha\Omega$ etc. are not exactly Gaussian distributed due to the effect of the drift terms), although it is not clear whether the difference would be significant or not.

Which phase-dependent variable to optimize

The first-order correlation function $G^{(1)}(0, t)$ can be estimated in two ways, as seen in (7.50). Hence, it is best to optimize for the average variance of the logarithm of the two random variables $\Omega\alpha\beta_0$ and $(\Omega\alpha_0\beta)^*$ corresponding to the $G^{(1)}$ estimate. Also, since general coherent state initial conditions to be optimized will have arbitrary phases, we should optimize a variance of variables related to $|G^{(1)}(0, t)|$ rather than the complex $G^{(1)}$.

7.5.3 Optimization of g''

$|G^{(1)}(0, t)|$ is estimated by either e^{G_L} or $e^{\tilde{G}_L}$, where

$$G_L(t) = \operatorname{Re} \{ \log \beta_0 + \log \alpha(t) + z_0(t) \} \quad (7.51a)$$

$$\tilde{G}_L(t) = \operatorname{Re} \{ \log \alpha_0 + \log \beta(t) + z_0(t) \}, \quad (7.51b)$$

while the uncertainty in the estimate will be proportional to $\operatorname{var} [e^{G_L}]$ or $\operatorname{var} [e^{\tilde{G}_L}]$. Taking the considerations of the previous Section 7.5.2 into account, let us look for such $g'' = g''_{\text{opt}}$ that the mean variance of G_L and \tilde{G}_L is minimized for the anharmonic oscillator system. This minimum occurs when

$$\frac{\partial \left(\operatorname{var} [G_L(t_{\text{opt}})] + \operatorname{var} [\tilde{G}_L(t_{\text{opt}})] \right)}{\partial g''} = 0. \quad (7.52)$$

Note that the optimal value g''_{opt} will in general depend upon the target time $t_{\text{opt}} = t$ at which the variances are considered.

The formal solution of the gauged anharmonic oscillator equations (7.33), assuming uniform initial weight $\Omega(0) = 1$, is straightforward to find:

$$n_L(t) = \log(n_0) - \gamma t + \sqrt{\chi} e^{-g''} [i\zeta^+(t) - \zeta^-(t)] \quad (7.53a)$$

$$m_L(t) = \log(\alpha_0/\beta_0) - 2i(\omega - \chi)t - 4i\chi \int_0^t n'(t') dt' + \sqrt{\chi} e^{g''} [i\zeta^-(t) - \zeta^+(t)] \quad (7.53b)$$

$$z_0(t) = -\sqrt{\chi} e^{-g''} \int_0^t n''(t') [\eta^+(t') + i\eta^-(t')] dt', \quad (7.53c)$$

where using the definition (7.27).

$$\eta^\pm(t) = \frac{dW_1(t)}{dt} \pm \frac{dW_2(t)}{dt} \quad (7.54)$$

are independent real Gaussian noises obeying

$$\langle \eta^\pm(t) \rangle_{\text{stoch}} = 0 \quad (7.55a)$$

$$\langle \eta^\pm(t) \eta^\pm(t') \rangle_{\text{stoch}} = 2\delta(t - t') \quad (7.55b)$$

$$\langle \eta^\pm(t) \eta^\mp(t') \rangle_{\text{stoch}} = 0, \quad (7.55c)$$

and so $d\eta = \sqrt{i}(\eta^- - \eta^+) dt/2$. Also,

$$\zeta^\pm(t) = \int_0^t \eta^\pm(t') dt'. \quad (7.56)$$

These ζ^\pm are time-correlated Gaussian random variables. Using (7.55), they obey the relationships

$$\langle \zeta^\pm(t) \rangle_{\text{stoch}} = 0 \quad (7.57a)$$

$$\langle \zeta^\pm(t) \zeta^\pm(t') \rangle_{\text{stoch}} = 2 \min[t, t'] \quad (7.57b)$$

$$\langle \zeta^\pm(t) \zeta^\mp(t') \rangle_{\text{stoch}} = 0. \quad (7.57c)$$

The variables to optimize are

$$\begin{bmatrix} G_L \\ \tilde{G}_L \end{bmatrix} = \log |n_0| - \frac{\gamma}{2} t - \frac{\sqrt{\chi}}{2} \left(e^{-g''} \zeta^-(t) \begin{bmatrix} + \\ - \end{bmatrix} e^{g''} \zeta^+(t) \right) - \sqrt{\chi} e^{-g''} \int_0^t n''(t') \eta^+(t') dt'. \quad (7.58)$$

In the Ito calculus the noises at t' are independent of any variables at t' , so one finds

$$\begin{bmatrix} \langle G_L \rangle_{\text{stoch}} \\ \langle \tilde{G}_L \rangle_{\text{stoch}} \end{bmatrix} = \log |n_0| - \frac{\gamma}{2} t, \quad (7.59)$$

and using (7.57b)

$$\begin{aligned} \begin{bmatrix} \langle G_L^2 \rangle_{\text{stoch}} \\ \langle \tilde{G}_L^2 \rangle_{\text{stoch}} \end{bmatrix} &= \begin{bmatrix} \langle G_L \rangle_{\text{stoch}}^2 \\ \langle \tilde{G}_L \rangle_{\text{stoch}}^2 \end{bmatrix} + \chi t \cosh 2g'' \begin{bmatrix} + \\ - \end{bmatrix} \chi \int_0^t \langle n''(t') \eta^+(t') \zeta^\pm(t') \rangle_{\text{stoch}} dt' \\ &+ \chi e^{-2g''} \int_0^t dt' \int_0^t dt'' \langle n''(t') n''(t'') \rangle_{\text{stoch}} \langle \eta^+(t') \eta^+(t'') \rangle_{\text{stoch}}, \end{aligned} \quad (7.60)$$

where use has also been made of the independence of $\eta^\pm(t')$ and $\zeta^\mp(t)$.

To further evaluate (7.60), one needs to calculate the averages containing n'' . Define $c = \sqrt{\chi} e^{-g''}$, and $n_0 = n'_0 + i n''_0$. Firstly,

$$\begin{aligned} \langle n''(t) \rangle_{\text{stoch}} &= e^{-\gamma t} \left\langle e^{-c \zeta^-} \right\rangle_{\text{stoch}} \left[n'_0 \langle \sin c \zeta^+ \rangle_{\text{stoch}} + n''_0 \langle \cos c \zeta^+ \rangle_{\text{stoch}} \right]. \\ &= n''_0 e^{-\gamma t} \end{aligned} \quad (7.61)$$

where we have first used the independence of the ζ^\pm to separate the stochastic averages, and then applied expressions (7.39), (7.44), and (7.47) to evaluate them.

Now, since noises at times $t > t'$ are uncorrelated with those at t' or earlier,

$$\begin{aligned}
\langle n''(t')\eta^+(t')\zeta^+(t) \rangle_{\text{stoch}} &= \langle n''(t')\eta^+(t')\zeta^+(t') \rangle_{\text{stoch}} \\
&= \int_0^{t'} \langle n''(t')\eta^+(t')\eta^+(t'') \rangle_{\text{stoch}} dt'' \\
&= \int_0^{t'} \langle n''(t') \rangle_{\text{stoch}} \langle \eta^+(t')\eta^+(t'') \rangle_{\text{stoch}} dt'' \\
&= 2 \langle n''(t') \rangle_{\text{stoch}}, \tag{7.62}
\end{aligned}$$

using (7.55b).

Secondly, if $t'' > t'$, noting that

$$\zeta^\pm(t'') = \zeta^\pm(t') + \tilde{\zeta}^\pm(t'' - t'),$$

where the $\tilde{\zeta}^\pm(t)$ are independent of the $\zeta^\pm(t)$, but have the same properties (7.57), one finds that

$$\begin{aligned}
\langle n''(t')n''(t'') \rangle_{\text{stoch}} &\tag{7.63} \\
&= e^{-\gamma(t'+t'')} \left\langle e^{-2c\zeta^-(t')} \right\rangle_{\text{stoch}} \left\langle e^{-c\tilde{\zeta}^-(t''-t')} \right\rangle_{\text{stoch}} \left\langle \cos c\tilde{\zeta}^+(t''-t') \right\rangle_{\text{stoch}} \\
&\quad \times \left\{ [n'_0]^2 \langle \sin^2 c\zeta^+(t') \rangle_{\text{stoch}} [n''_0]^2 \langle \cos^2 c\zeta^+(t') \rangle_{\text{stoch}} + n'_0 n''_0 \langle \sin 2c\zeta^+(t') \rangle_{\text{stoch}} \right\}.
\end{aligned}$$

Terms containing $\langle \sin c\tilde{\zeta} \rangle_{\text{stoch}} = 0$ have already been discarded. Evaluating the averages using (7.39) and (7.47), one obtains

$$\langle n''(t')n''(t'') \rangle_{\text{stoch}} = \frac{e^{-\gamma(t'+t'')}}{2} \left\{ |n_0|^2 e^{4t'c^2} - (n'_0)^2 + (n''_0)^2 \right\}. \tag{7.64}$$

Due to symmetry

$$\int_0^t dt' \int_0^t dt'' n''(t')n''(t'') = 2 \int_0^t dt'' \int_0^{t''} dt' n''(t')n''(t''),$$

and substituting (7.64), (7.62), and (7.61) into (7.60) and (7.59), then integrating, one obtains

$$\begin{aligned}
\begin{bmatrix} \text{var}[G_L(t)] \\ \text{var}[\tilde{G}_L(t)] \end{bmatrix} &= \chi t \cosh(2g'') \begin{bmatrix} + \\ - \end{bmatrix} 2\chi n''_0 \left(\frac{1 - e^{-\gamma t}}{\gamma} \right) \\
&\quad + \chi e^{-2g''} |n_0|^2 \left(\frac{e^{4\chi e^{-2g''} t} e^{-2\gamma t} - 1}{4\chi e^{-2g''} - 2\gamma} \right) - \chi e^{-2g''} [(n'_0)^2 - (n''_0)^2] \left(\frac{1 - e^{-2\gamma t}}{2\gamma} \right). \tag{7.65}
\end{aligned}$$

The optimum $g'' = g''_{\text{opt}}$ can now in theory be calculated by imposing (7.52).

Since this would involve the solution of a transcendental equation, it would be cumbersome to use in a numerical simulation, going against requirement 3 of Section 7.5.1, because one would have to execute a time-consuming algorithm at each time step. By considering, below, some important special cases, an approximation to the optimum g''_{opt} applicable under the broad conditions aimed for in Section 7.5.1 will be found, which can then be easily used in actual simulations.

7.5.4 Important special cases

1. Perhaps the most important special case is when damping is absent ($\gamma = 0$).

This is also the worst case in terms of simulation stability. At relatively short times one has

$$4\chi t_{\text{opt}} e^{-2g''_{\text{opt}}} \ll 1, \quad (7.66)$$

and the optimum is given (from (7.52) and (7.65)) by the roots of the cubic in $V_g = e^{-2g''_{\text{opt}}}$:

$$8\chi t_{\text{opt}} |n_0|^2 V_g^3 + [4(n_0'')^2 + 1] V_g^2 - 1 = 0. \quad (7.67)$$

In the usual case of simulations with sizeable occupation numbers in modes, and times up to coherence time (which then is much shorter than the quantum revival time π/χ) this short time condition is satisfied. A very useful expression

$$g''_{\text{opt}} \approx \frac{1}{3} \log (|n_0| \sqrt{8\chi t_{\text{opt}}}). \quad (7.68)$$

applies when the V_g^2 term is negligible. This occurs at long enough times when n_0 is large enough and mostly real: i.e. when $1 + 4(n_0'')^2 \ll (8\chi t_{\text{opt}} |n_0|^2)^{2/3}$. So χt_{opt} must be at least $\gg 1/8|n_0|^2$ (a higher limit applies if $n_0'' \neq 0$).

The opposite case when n_0 is either too small, too imaginary, or the time is too short has the V_g^3 term negligible and leads to

$$g''_{\text{opt}} \approx \frac{1}{4} \log [1 + 4(n_0'')^2]. \quad (7.69)$$

2. It can be seen that the long time behaviour of the variances depends on the “damping strength” parameter

$$q = 2\gamma - 4\chi e^{-2g''}. \quad (7.70)$$

When q is positive, the long time (i.e. $qt \gg 1$) behaviour of (7.65) is asymptotic to a linear increase

$$\begin{bmatrix} \text{var}[G_L] \\ \text{var}[\tilde{G}_L] \end{bmatrix} \rightarrow \chi t \cosh 2g'' - b \quad (7.71)$$

where $b = \chi e^{-2g''} \{[(n'_0)^2 - (n''_0)^2]/2\gamma - |n_0|^2/q\} \mp 2\chi n''_0/\gamma$ is a constant. This means that the sampling uncertainty grows relatively slowly, and long simulation times are possible. When q is negative, on the other hand, the long time behaviour is $\chi e^{-2g''} |n_0|^2 e^{|q|t}/|q|$, and we expect a rapid appearance of intractable sampling error after some time $|q|t \gg 1$.

The parameter q depends on the relative strengths of the nonlinearity and the damping, and determines the long time behavior of the statistical error. Note though that if damping is present, then for large enough diffusion gauge g'' , q can always be made positive, and the linear (in t) variance regime can be reached with a choice $g'' \geq \log \sqrt{2\chi/\gamma}$.

3. An immediate extension to nonzero damping of expressions (7.67), (7.68) and (7.69) can be derived for times $|qt_{\text{opt}}| \ll 1$ to give the cubic

$$8\chi t_{\text{opt}} |n_0|^2 V_g^3 + a_2 V_g^2 - 1 = 0 \quad (7.72a)$$

where

$$\begin{aligned} a_2(n_0, \gamma t_{\text{opt}}) = & 1 + 4(n''_0)^2 \left(\frac{1 - e^{-2\gamma t_{\text{opt}}}}{2\gamma t_{\text{opt}}} \right) \\ & - 2|n_0|^2 \left(\frac{1 - 2\gamma t_{\text{opt}} + 2(\gamma t_{\text{opt}})^2 - e^{-2\gamma t_{\text{opt}}}}{\gamma t_{\text{opt}}} \right). \end{aligned} \quad (7.72b)$$

The expression (7.68) still applies at large enough and real enough n_0 : i.e. when $a_2 \ll (8\chi t_{\text{opt}} |n_0|^2)^{2/3}$. In the opposite case of $a_2 \gg (8\chi t_{\text{opt}} |n_0|^2)^{2/3}$, one has

$$g''_{\text{opt}} \approx \frac{1}{4} \log(a_2). \quad (7.73)$$

7.5.5 Suggested approximate form of diffusion gauge

Guided by conditions 2, 4, 5, and 6, in Section 7.5.1, we can see that the choice of g''_{approx} for the anharmonic oscillator might depend on the four real parameters n'_0 , n''_0 , χt_{opt} and γ . Because of conditions 4 and 5, we are most interested in the regime of small γ , and either $\chi t_{\text{opt}} \lesssim \mathcal{O}\left(1/2\sqrt{|n_0|}\right)$ when $|n_0| \gtrsim \mathcal{O}(1)$, or $\chi t_{\text{opt}} \lesssim \mathcal{O}(1)$ when $n_0 \lesssim \mathcal{O}(1)$. These cases are covered by expression (7.72).

This can be most easily seen in the two limits $|n_0| \gg 1$ and $|n_0| \ll 1$ where for small γ , we have $qt_{\text{opt}} \approx -2(\chi t_{\text{opt}}/|n_0|)^{2/3}$ and $\approx -4\chi t_{\text{opt}}/\sqrt{1+4(n''_0)^2}$ respectively. Hence the condition $|q|t_{\text{opt}} \ll 1$ applies for target times $\chi t_{\text{opt}} \ll |n_0|/2\sqrt{2}$ and $\ll \approx (1+2|n''_0|)/4$ respectively, which is roughly sufficient for low occupations, and more than sufficient for high occupations.

To obtain an explicit estimate for g''_{opt} , one can either evaluate the roots of the polynomials (7.67) or (7.72) by standard expressions, which can be still quite complicated although reasonably rapid, or use the approximation

$$g''_{\text{approx}} = \frac{1}{6} \log \left\{ 8|n_0|^2 \chi t_{\text{opt}} + a_2^{3/2} \right\}, \quad (7.74)$$

which reduces to (7.68) and (7.73) in their limits of applicability, and works very well in practice (see figures 7.3, 7.6, and 7.5).

The discrepancy Δ between (7.74) and the exact optimization obtained by solving (7.52) with (7.65) is shown for real n_0 for a wide range of parameters in Figure 7.2. Note that the ubiquitous $e^{-2g''}$ factor is, for small discrepancy, $e^{-2g''_{\text{opt}}} \approx (1-2\Delta)e^{-2g''_{\text{approx}}}$. It can be seen that for occupations $\gtrsim \mathcal{O}(10)$ and/or for times shorter than, or of the order of, singly-occupied coherence time the approximation is very good.

The diffusion gauge choice (7.74) will be used from here on.

7.5.6 Relationship between target time and mode occupation

The expression (7.74) was worked out under the conditions that the mean occupation of the mode is conserved. In coupled-mode simulations this is no longer the case, and

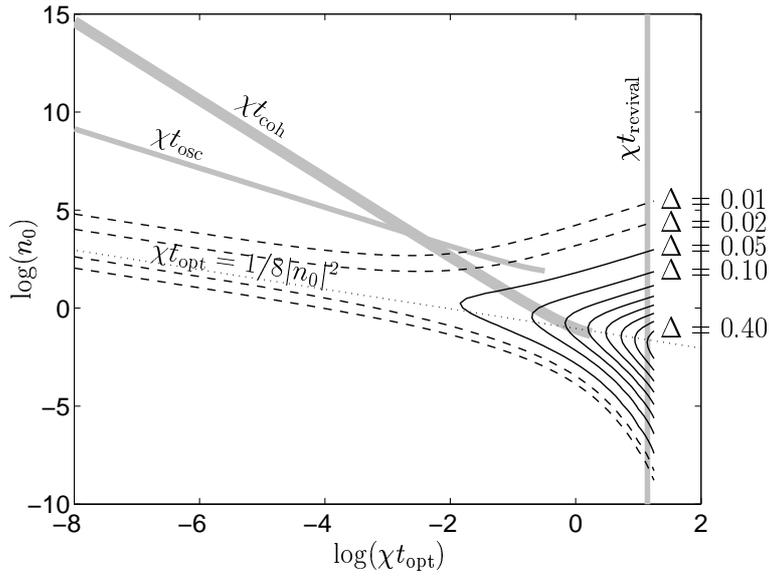


Figure 7.2: **Discrepancy** $\Delta = g''_{\text{opt}} - g''_{\text{approx}}$ between g''_{opt} (the exact optimization of g'' by solving (7.52) using (7.65)), and the approximate expression (7.74). Displayed is the case of no damping ($\gamma = 0$) and classical initial occupation ($n_0 = n'_0$), shown as a function of t_{opt} and n_0 . Discrepancy values Δ are shown as SOLID CONTOURS with spacing 0.05. Additional DASHED CONTOURS shown at very low discrepancy. DOTTED LINE approximates region of greatest discrepancy $\chi t_{\text{opt}} \approx 1/8n_0'^2$. For comparison, several physical timescales are also shown in GREY: time of first quantum revival t_{revival} , phase coherence time t_{coh} and phase oscillation period t_{osc} .

could be adapted for by replacing n_0 by $\check{n}(t)$, which explicitly assumes independence of trajectories, and a Markovian process. Although the mean over all trajectories $\langle \Omega(t)\check{n}(t) \rangle$ would be a better estimator of the mean boson occupation, this would be in conflict with aim 5 in Section 7.5.1, and might lead to biases due to complicated feedback mechanisms between trajectories.

Another assumption used to arrive at (7.74) was that g'' would be constant in time, which is now no longer the case. This raises the question of how to include the target time in g''_{opt} . Two ways that quickly come to mind is either to calculate $g''_{\text{opt}}(t)$

1. Always optimizing for a time t_{opt} forward from the present t (choosing then, explicitly, $g''_{\text{opt}}(t_{\text{opt}})$), or:
2. Optimizing for only the remaining time to the absolute target time $t_{\text{opt}} \geq t$

(choosing, then, $g''_{\text{opt}}(\max[t_{\text{opt}} - t, 0])$).

In combination, this gives rise to the four possible gauge forms (in the (7.74) approximation, using (7.72b))

$$g''_{\text{approx}}(t) = \frac{1}{6} \log \{8|n_0|^2 \chi t_{\text{opt}} + a_2(n_0, \gamma t_{\text{opt}})^{3/2}\}, \quad (7.75a)$$

$$g''_{\text{approx}}(t) = \frac{1}{6} \log \{8|n_0|^2 \chi t_{\text{rem}} + a_2(n_0, \gamma t_{\text{rem}})^{3/2}\}, \quad (7.75b)$$

$$g''_{\text{approx}}(t) = \frac{1}{6} \log \{8|\check{n}(t)|^2 \chi t_{\text{opt}} + a_2(\check{n}(t), \gamma t_{\text{opt}})^{3/2}\}, \quad (7.75c)$$

$$g''_{\text{approx}} = \frac{1}{6} \log \{8|\check{n}(t)|^2 \chi t_{\text{rem}} + a_2(\check{n}(t), \gamma t_{\text{rem}})^{3/2}\}(t), \quad (7.75d)$$

where the “remaining time to target” is

$$t_{\text{rem}} = \max[t_{\text{opt}} - t, 0]. \quad (7.76)$$

These strategies have been numerically investigated for the undamped ($\gamma = 0$) case with real n_0 starting conditions. It turns out that form (7.75d) has some advantage over the others, particularly at high occupations. Details are discussed in Section 7.7.1 and shown in Table 7.1.

7.5.7 Boundary term issues

When considering the use of an adaptive $g''(t, \check{n}(t))$, one should be careful that the g'' dependent terms do not introduce new noise divergences that were not present when the standard $g'' = 0$ square root noise matrix B_0 was used³. The condition to avoid noise divergence symptoms (and so, presumably, boundary term errors) is as always (6.3): The stochastic equations should contain no radial component that grows faster than exponentially as large radial values of the phase-space variables are reached.

Since the suggested forms (7.75) depend directly only on the complex occupation variable $\check{n} = e^{n_L}$, and not on either of the other independent variables (m_L , and z_0), then it suffices to only check that the \check{n} evolution equation contains no radial super-exponential growth. This is because the g'' dependent terms in the m_L and z_0

³Incidentally, any boundary term errors introduced by such noise divergences would have to be of the second kind, since the choice of diffusion gauge is made at the level of the FPE-stochastic equation correspondence, long after any boundary terms $\hat{\mathcal{B}}$ has been discarded.

evolutions simply accumulate integrals of functions of n_L , and if \check{n} remains finite, then so will the other variables.

The \check{n} evolution is now

$$\begin{aligned} d\check{n} &= 2i\check{n}\sqrt{i\chi}e^{-g''}d\eta + \gamma(\bar{n}_{\text{bath}} - \check{n})dt \\ &+ (\varepsilon dt + \sqrt{\gamma\bar{n}_{\text{bath}}}d\eta_{\text{bath}})\beta + (\varepsilon dt + \sqrt{\gamma\bar{n}_{\text{bath}}}d\eta_{\text{bath}})^*\alpha. \end{aligned} \quad (7.77)$$

Provided the time dependence of parameters γ , \bar{n}_{bath} , ε , or χ is not pathological, the condition (6.3) means g'' must obey

$$\lim_{|\check{n}(t)| \rightarrow \infty} e^{-g''} \propto |n(t)|^a \text{ where } a \leq 0. \quad (7.78)$$

for no super-exponential growth to occur. That is, $e^{g''}$ must grow as a non-negative power law (or faster) as $|\check{n}|$ becomes large. This is seen to be satisfied by the suggested forms (7.75). Finally, non g'' -dependent terms were considered in Section 7.3, and found not to lead to any moving singularities.

It is concluded, then, that the diffusion gauge form (7.74) does not lead to any new noise divergences or moving singularities, and hence none of the usual boundary term error symptoms have been reintroduced.

7.5.8 Particle gain

External particle gain (rather than loss due to a zero-temperature heat bath) complicates the behaviour. Looking at (7.77), the bath interactions tend to equilibrate the mode occupation to the bath value \bar{n}_{bath} . If $\bar{n}_{\text{bath}} \ll \langle \hat{n} \rangle$, then one can expect that the optimum gauge calculated with nonzero γ will be largely unchanged. For highly occupied bath modes, or strong coherent particle gain ε , a new analysis of g'' optimization may give better results than (7.75), however this appears to be a very involved procedure. An obvious first try is to just see what happens with the gauge form (7.75).

7.6 Estimates of simulation times

The limit (7.43) on the variance of exponentials of Gaussian random variables can be used to estimate times of useful simulation t_{sim} from expressions for $\text{var}[G_L]$ — since G_L behaves much like a Gaussian random variable in analogy to $\sigma\xi$, and observables are estimated by e^{G_L} in analogy to $v_\sigma \propto e^{\sigma\xi}$. Imposing

$$(\text{var}[G_L(t_{\text{sim}})] + \text{var}[\tilde{G}_L(t_{\text{sim}})])/2 \approx 10, \quad (7.79)$$

one can solve for t_{sim} to obtain an estimate at least of the scaling of t_{sim} with system parameters.

Consider first the **gauged** simulation. For coherent state initial conditions $\beta_0 = \alpha_0^*$ at small damping $\gamma t \ll 1$, and short enough times $|q|t \ll 1$,

$$\text{var}[G_L] = \frac{\chi t}{2} \left(V_g + \frac{1}{V_g} \right) + 2(\chi V_g t n'_0)^2. \quad (7.80)$$

For large particle number $n'_0 \gg 1$, g''_{opt} takes the form (7.68), so $V_g \approx 1/2(n'_0 \chi t)^{1/3}$, the $\chi t V_g/2$ term is negligible, and $\text{var}[G_L] \approx 3(n'_0 \chi^2 t^2)^{2/3}/2$, and (7.79) leads to a useful simulation time

$$t_{\text{sim}} \approx \frac{(20/3)^{3/4}}{\chi \sqrt{n'_0}} \approx \mathcal{O}(10 t_{\text{coh}}). \quad (7.81)$$

Since in this regime $|q|t \approx 4\chi t V_g \approx 2(\chi t/n'_0)^{2/3}$, then $|q|t_{\text{sim}} \approx \mathcal{O}(5)/n'_0 \ll 1$, and so the expression (7.81) is consistent with the original short time assumption $|q|t \ll 1$ when $n'_0 \gg \mathcal{O}(5)$.

At small occupations $n'_0 \ll 1$ on the other hand, $g''_{\text{opt}} \rightarrow 0$, and at long times, $\text{var}[G_L] \approx \frac{1}{4}n'_0{}^2 e^{4\chi t}$. Thus useful simulation time scales very slowly with n'_0 , and is

$$t_{\text{sim}} \approx \frac{\mathcal{O}(1) - \frac{1}{2} \log n'_0}{\chi}. \quad (7.82)$$

The “long time” condition holds while $\frac{1}{4}n'_0{}^2 e^{4\chi t}$ is much greater than the lower order term χt . That is, $n'_0 \gg \mathcal{O}(10^{-8})$. At even smaller n'_0 , the χt term dominates (7.80), and

$$t_{\text{sim}} \approx \mathcal{O}(10)/\chi. \quad (7.83)$$

For comparison, consider the **positive P** behaviour, which can be estimated from the logarithmic variances (7.92) for the case with no drift gauge, which are worked out in Section 7.8.2. For coherent state initial conditions $n_0 = n'_0$ and small damping $\gamma t \ll 1$,

$$\text{var}[G_L] \approx \chi t + 2n'_0(\chi t)^2 + 2\chi^2 n'^2_0 \left\{ \frac{e^{4\chi t} - 1}{8\chi^2} - \frac{t}{2\chi} - t^2 \right\}. \quad (7.84)$$

At short times $4\chi t \ll 1$ the leading term in the $\{\bullet\}$ expression is $\frac{8}{3}(\chi t)^3 n'^2_0$. In such a case, for large mode occupation $n'_0 \gg \mathcal{O}(1)/\chi t$ this term dominates, and using (7.79), one obtains

$$t_{\text{sim}} \approx \frac{\mathcal{O}(1)}{\chi n'^{2/3}_0}. \quad (7.85)$$

Checking back, $1/\chi t \approx \mathcal{O}(1) n'^{2/3}_0$, so (7.85) is consistent with the short time assumption for $n'_0 \gg 1$. At long times, one again has (7.82) and (7.83). The expression (7.85) is also useful in estimating simulation times for many-mode models, as will be seen in Section 10.2.1.

Curiously, the “stable” drift-gauged simulation with **no diffusion gauge** does even worse than the positive P. In this case, $V_g = 1$, and using (7.80), at large n'_0

$$t_{\text{sim}} \approx \frac{\mathcal{O}(2)}{\chi n'_0} \approx t_{\text{osc}}, \quad (7.86)$$

which is consistent when $n'_0 \gg \mathcal{O}(\frac{1}{4})$.

The numerical results in Table 7.1 and Figure 7.3, are seen to agree with these estimates to within constant factors of $\mathcal{O}(1)$.

7.7 Numerical investigation of improvement

To unambiguously determine the improvements in simulation time that are achieved by the use of the proposed gauges, simulations of an undamped anharmonic oscillator were carried out for a wide range of mode occupations and a variety of gauges, testing various target times t_{opt} where appropriate. (The undamped system was chosen because this is the worst case.)

Figure 7.3 compares t_{sim} , the maximum time achieved with various methods at which useful precision in the phase correlations $G^{(1)}(0, t)$ can be obtained. Note

the logarithmic scale. Results at high boson occupation are tabulated in Table 7.1, which includes data for a larger set of gauge choices. Figure 7.6 gives examples of calculated values $|G^{(1)}(0, t)|$ along with error estimates. Table 7.2 gives some empirical fitting parameters to expression (7.89) for the useful simulation time t_{sim} . These may be useful to assess simulation times when the particle occupation is of $\mathcal{O}(10)$ or smaller, and the expressions (7.81) or (7.85) are not accurate.

Details of the sampling error behaviour are shown in Figure 7.5 for a range of gauge choices, while Figure 7.4 shows the dependence of useful simulation times on the target time parameter t_{opt} for a variety of gauge forms for two example mode occupations: $n_0 = 1$ and $n_0 = 10^4$.

7.7.1 Procedure

Simulations were carried out with coherent state initial conditions for a wide range of mean occupation number

$$n_0 = \{10^{-5}, 10^{-4}, 10^{-3}, 0.01, 0.1, 1, 10, 100, 1000, 10^4, 10^5, 10^6, 10^8, 10^{10}\}. \quad (7.87)$$

The following gauges, and their varieties as seen in Table 7.1, were tried for each n_0 value:

- i) No gauge: $\mathcal{G}_k = g'' = 0$. This is the standard positive P distribution technique.
- ii) Drift gauge only: \mathcal{G}_k given by (7.30), and $g'' = 0$.
- iii) Both drift gauge (7.30) and diffusion gauges of the four related forms (7.75).
- iv) Diffusion gauge (7.100) only, as described by Plimak *et al*[2] (See Section 7.9.2).
- v) Drift gauge (7.99) only, as described by Carusotto *et al*[1] (See Section 7.9.1)
- vi) Adaptive diffusion gauges (7.94) only (of the four forms (7.97)), and $\mathcal{G}_k = 0$.

For the diffusion gauges of iii), iv), and vi), which depend on an *a priori* target time parameter t_{opt} , a wide variety of target times were tried to investigate the dependence between t_{sim} and t_{opt} , and ascertain what are the longest simulation times achievable. The dependence of t_{sim} on t_{opt} is plotted in Figure 7.4, while the best times are tabulated in Tables 7.1, 7.2, and Figure 7.3.

The term *useful* precision has been taken to indicate a situation where the estimate \bar{O} of the expectation value of some observable \hat{O} is known to be precise to at least one significant digit when $\mathcal{S} = 10^6$ trajectories are used. That is, $\Delta\bar{O}(\mathcal{S} = 10^6) \leq |\bar{O}|/10$. Since for many-mode systems, the calculation of even one trajectory is reasonably time consuming, it is clear that $\mathcal{S} = 10^6$ trajectories is very unlikely to be exceeded for most non-trivial problems. Hence, meaningful results for non-trivial problems are at best likely to be obtained only in the parameter region where the above-defined “useful” precision condition is satisfied. If one calculates some lesser (than 10^6) number of trajectories, yet $\mathcal{S} \gg 1$, then the Central Limit theorem can be used to extrapolate $\Delta\bar{O}$ to the situation when $\mathcal{S} = 10^6$, since $\Delta\bar{O} \propto 1/\sqrt{\mathcal{S}}$. This gives finally, that the precision is taken to be “useful” when

$$\Delta\bar{O}(\mathcal{S})\sqrt{\frac{\mathcal{S}}{10^6}} \leq \frac{1}{10}|\bar{O}|. \quad (7.88)$$

In the current simulations, the observable in question is $|G^{(1)}(0, t)|$, and the “simulation time” t_{sim} that will be referred to here is the maximum time at which useful precision in $|G^{(1)}(0, t)/G^{(1)}(0, 0)|$ is retained. Each actual simulation was done with $\mathcal{S} = 10^4$ trajectories.

Uncertainties in the calculated useful simulated times arise because the $\Delta|G^{(1)}|$ are themselves estimated from the finite ensemble of \mathcal{S} trajectories. The uncertainty in $\Delta|G^{(1)}|$ was estimated by inspection of several (usually 10) independent runs with identical parameters. The range of values of t_{sim} seen was taken to be twice the uncertainty in t_{sim} .

The scalings of simulation time with particle number for $n'_0 \gg 1$ and $n'_0 \ll 1$ have been worked out in Section 7.6. Taking these results into account, Table 7.2 empirically fits the simulation time t_{sim} to

$$t_{\text{est}} = \frac{1}{\chi} \left\{ [c_1 n'^{-c_0}]^{-c_2} + \left[\log \left(\frac{e^{c_3}}{n'^{c_4}} + 1 \right) \right]^{-c_2} \right\}^{-1/c_2} \quad (7.89)$$

for intermediate times, as this may be useful for evaluation of simulation strategies in a many-mode case. c_0 characterizes the power-law scaling of t_{sim} at high occupation, c_1 the pre-factor for high n'_0 , c_3 a constant residual t_{sim} at near vacuum, c_4 characterizes the curvature at small n'_0 , while c_2 is related to the stiffness of the transition.

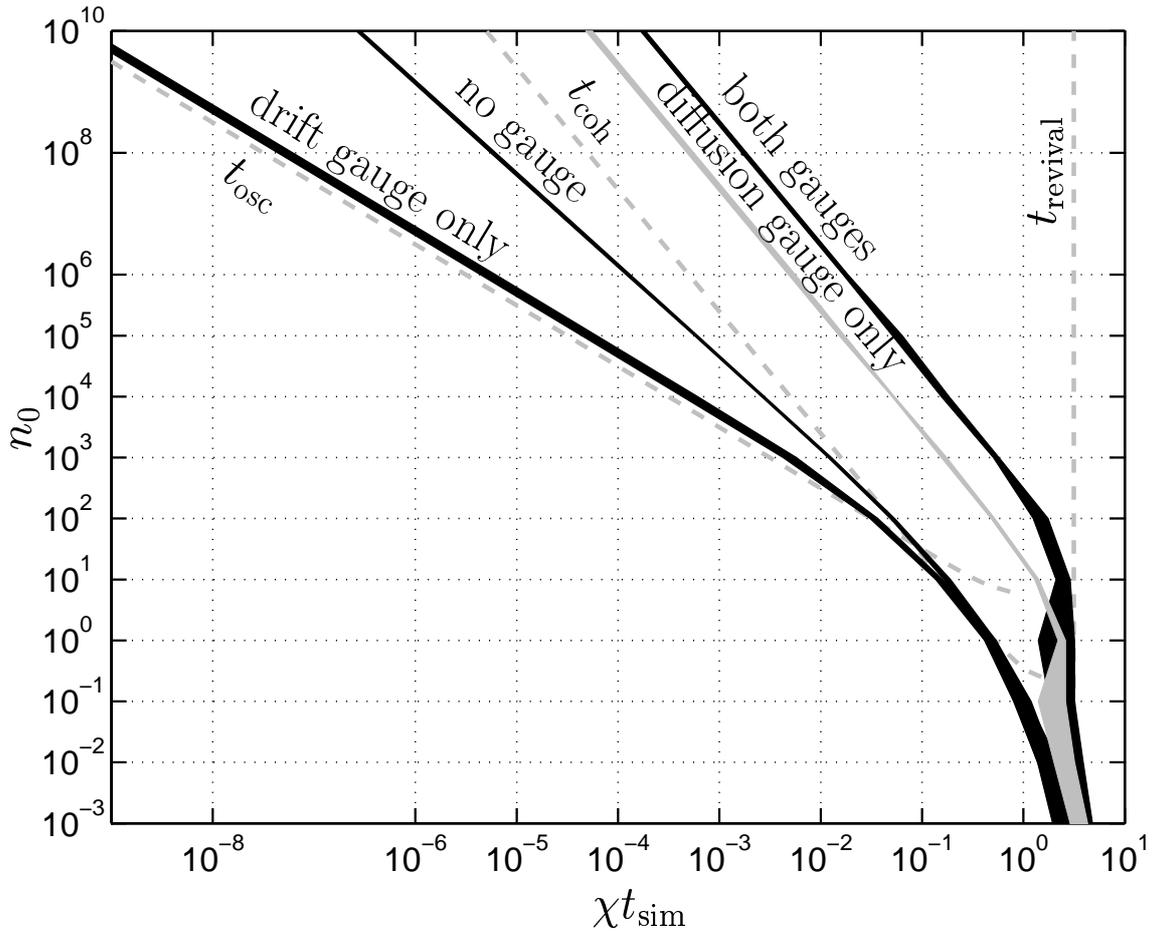


Figure 7.3: **Maximum useful simulation time** t_{sim} , of the one-mode undamped anharmonic oscillator with various gauge choices. Initial coherent state mean mode occupations $n_0 = n'_0$. Width of plotted lines shows estimated uncertainty in the values shown. The DRIFT GAUGE is (7.30), while the DIFFUSION GAUGE is (7.100) of Plimak *et al*[2] when on its own, or (7.75d) when with the drift gauge (7.30). For comparison, several timescales from Section 7.2.3 are also shown as broken lines: time of first quantum revival t_{revival} , phase coherence time t_{coh} and phase oscillation period t_{osc} .

The expression (7.89) reduces to $c_1 n_0'^{-c_0}/\chi$ and $(c_3 - c_4 \log n_0')/\chi$, when $n_0' \gg 1$ and $n_0' \ll 1$, respectively, in agreement with the limiting expressions (7.81) and (7.82). Uncertainty Δc_j in parameters c_j was worked out by requiring $\sum_{n_0} \{ [t_{\text{est}}(c_j \pm \Delta c_j, n_0) - t_{\text{sim}}(n_0)] / \Delta t_{\text{sim}} \}^2 = \sum_{n_0} \{ 1 + ([t_{\text{est}}(c_j, n_0) - t_{\text{sim}}(n_0)] / \Delta t_{\text{sim}})^2 \}$. In the range checked ($n_0 \in [10^{-5}, 10^{10}]$), the fit is good — i.e. there are no outlier data that would lie significantly beyond the range of t_{est} specified by parameters $c_j \pm \Delta c_j$.

Table 7.1: **Maximum simulation time**, at useful precision, of the one-mode system in the limit ($n_0 = n'_0 \gg 1$), achievable with various gauge choices. Calculations are for the undamped ($\gamma = 0$) system, which is the worst case in terms of sampling error. More details are given in Section 7.7.1. For comparison, several timescales from Section 7.2.3 are also given: time of first quantum revival t_{revival} , phase coherence time t_{coh} and phase oscillation period t_{osc} .

Drift gauge \mathcal{G}_k	Diffusion gauge g''	Useful simulation time t_{sim} when $n_0 = n'_0 \gg 1$	Maximum n_0 for which $\chi t_{\text{sim}} \geq 1$
$t_{\text{osc}} = \pi/\chi n'_0$			
(7.99)	0	$(1.06 \pm 0.16) t_{\text{osc}}$	$0.014 \begin{smallmatrix} + 0.016 \\ - 0.008 \end{smallmatrix}$
(7.30)	0	$(1.7 \pm 0.4) t_{\text{osc}}$	$0.08 \begin{smallmatrix} + 0.07 \\ - 0.05 \end{smallmatrix}$
0	0	$(1.27 \pm 0.08) / \chi n'_0{}^{2/3}$	0.11 ± 0.06
$t_{\text{coh}} = 1/2\chi\sqrt{n'_0}$			
0	(7.100) or (7.97a) or (7.97c)	$(8.2 \pm 0.4) t_{\text{coh}}$	12 ± 3
0	(7.97b) or (7.97d)	$(10.4 \pm 0.7) t_{\text{coh}}$	19 ± 4
(7.30)	(7.75b)	$(25.6 \pm 1.0) t_{\text{coh}}$	120 ± 30
(7.30)	(7.75a)	$(30 \pm 3) t_{\text{coh}}$	150 ± 40
(7.30)	(7.75c)	$(32 \pm 3) t_{\text{coh}}$	190 ± 15
(7.30)	(7.75d)	$(35 \pm 4) t_{\text{coh}}$	240 ± 70
$t_{\text{revival}} = \pi/\chi$			

Table 7.2: **Empirical fitting parameters** for maximum useful simulation time t_{sim} with several gauge choices. The fit is to expression (7.89).

	Positive P	Both gauges	Drift gauge	Diffusion gauge
Drift Gauge	0	(7.30)	(7.30)	0
Diffusion Gauge	0	(7.75d)	0	(7.97d)
c_0	$2/3$	$1/2$	1	$1/2$
c_1	1.27 ± 0.08	17.6 ± 1.7	5.5 ± 1.2	5.2 ± 0.4
c_2	$3.2 \pm \infty_{1.2}$	$3.6 \pm \infty_{2.3}$	$1.4 \pm \infty_{0.4}$	$2.7 \pm \infty_{1.0}$
c_3	-0.5 ± 0.3	2.8 ± 0.9	-0.5 ± 0.3	-2.4 ± 0.6
c_4	0.45 ± 0.07	0.23 ± 0.13	0.49 ± 0.08	0.23 ± 0.13

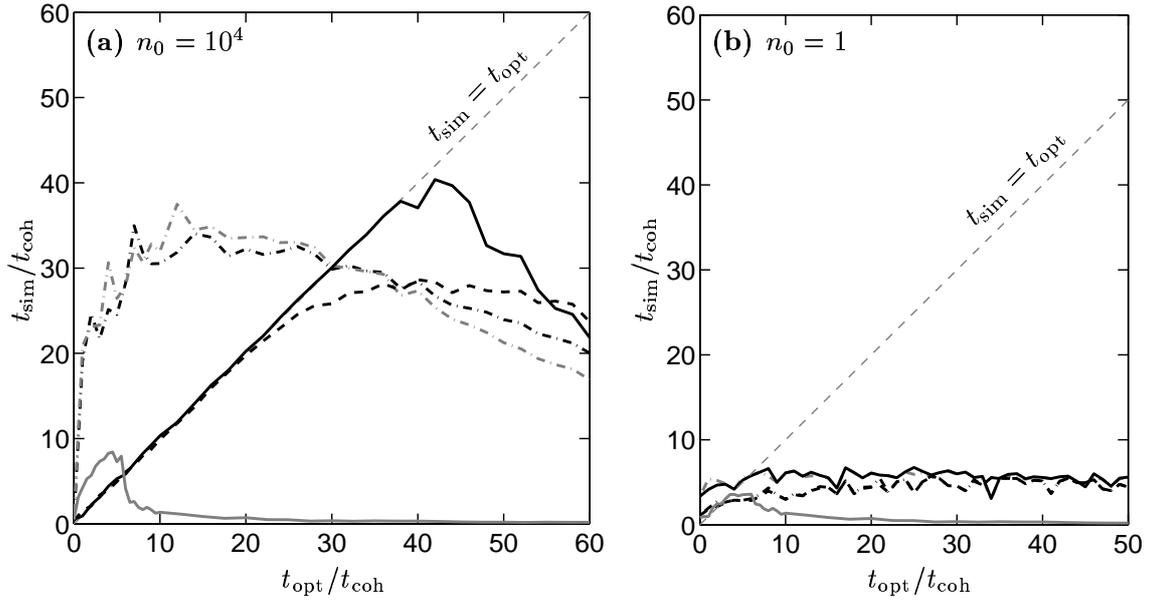


Figure 7.4: Comparison of *a priori* target time t_{opt} with actual useful simulation time t_{sim} for a variety of diffusion gauges: Results for drift gauges (7.30) with the four g'' forms (7.75) are shown as: (7.75d) – SOLID DARK; (7.75c) – DASH-DOTTED LIGHT; (7.75b) – DASHED; (7.75a) – DASH-DOTTED DARK. Relationship obtained using the diffusion gauge (7.100) of Plimak *et al*[2] (but no drift gauge) is also shown (SOLID LIGHT). The dashed line in the background shows, for reference $t_{\text{opt}} = t_{\text{sim}}$. Subplot (a): mean particle number $n_0 = 10^4$, (b): $n_0 = 1$.

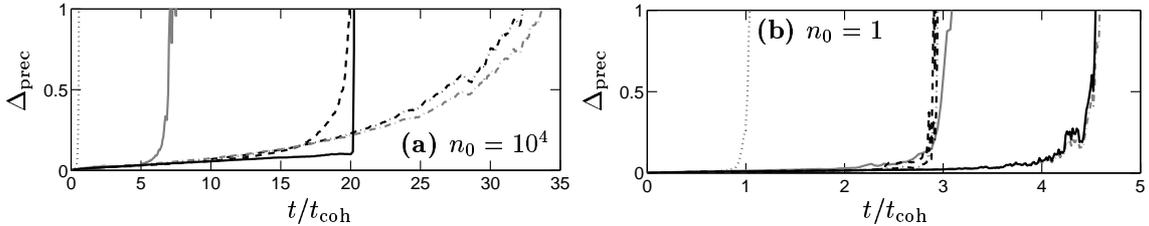


Figure 7.5: Uncertainty in $|G^{(1)}(0, t)|$, as a function of time. The quantity plotted is $\Delta_{\text{prec}} = \sqrt{\mathcal{S}/10^6} (10 \Delta |G^{(1)}(0, t)| / |G^{(1)}(0, 0)|)$, so that $\Delta_{\text{prec}} \leq 1$ corresponds to useful precision as defined in Section 7.7.1. Results are plotted for combined drift (7.30) and diffusion (7.75) gauges, where the four forms are shown as: (7.75d) – SOLID DARK; (7.75c) – DASH-DOTTED LIGHT; (7.75b) – DASHED; (7.75a) – DASH-DOTTED DARK. Data for the diffusion gauge (7.100) of Plimak *et al*[2] with $t_{\text{opt}} = 3t_{\text{coh}}$ as used therein is shown as a LIGHT SOLID line, while the ungauged positive P calculation is shown as a DOTTED line. Simulations were carried out with $\mathcal{S} = 10^4$ trajectories, starting with initial coherent state conditions: (a) $n_0 = 10^4$ (b) $n_0 = 1$. The combined drift-and-diffusion-gauge plots were calculated with target times of: (a) $t_{\text{opt}} = 20t_{\text{coh}}$ (b) $t_{\text{opt}} = 4t_{\text{coh}}$.

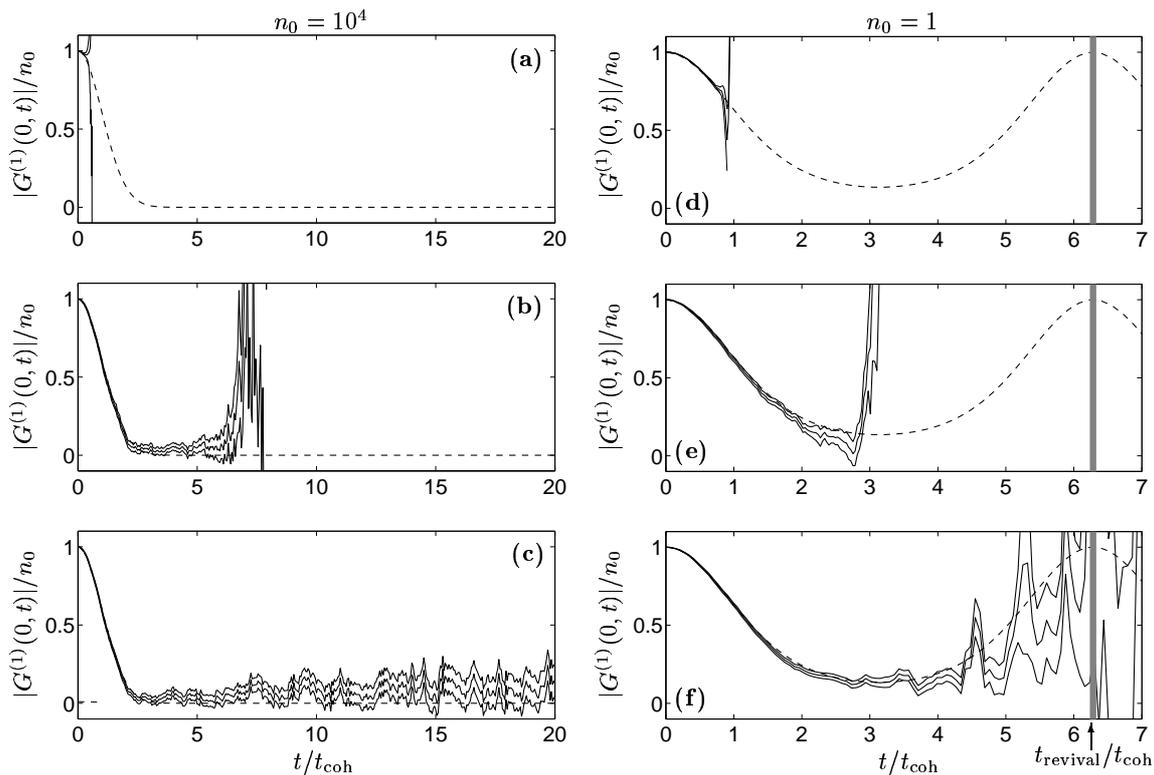


Figure 7.6: **Modulus of the phase correlation function** $G^{(1)}(0, t)$. Comparison of calculations with different gauges: Subplots (a) and (d): Ungauged positive P. Subplots (b) and (e): Diffusion gauge (7.100) of Plimak *et al*[2] with $t_{\text{opt}} = 3t_{\text{coh}}$. Subplots (c) and (f): Combined drift and diffusion gauges (7.30) and (7.75d) with the choice $t_{\text{opt}} = 20t_{\text{coh}}$. The initial conditions were a coherent state with $\langle \hat{n} \rangle = n_0$, where in subplots (a)-(c): $n_0 = 10^4$ particles, and (d)-(f): $n_0 = 1$. Triple SOLID lines indicate $G^{(1)}$ estimate with error bars. Exact values are also shown (single DASHED LINE). The quantum revival time is shown SHADED for the $n_0 = 1$ plots. $S = 10^4$ trajectories in all cases

7.7.2 Features seen

- Combining drift and diffusion gauges gives the longest useful simulation times. Such simulations give good precision well beyond the point at which all coherence has decayed away for highly-occupied modes - potentially up to about 35 coherence times.
- Diffusion-gauge-only simulations also give quite good statistical behaviour (although useful simulation times are about 4 times shorter at high occupation than with both gauges).

- Despite the efficient behaviour of combined gauge simulations, using only a drift-gauge gives even worse statistical error than no gauge at all (although, this does eliminate potential boundary term systematics in a multi-mode system, which are a problem with no gauge). Such simulations are restricted in time to about one phase oscillation.
- Plain positive P simulations at high occupation last for about $n_0'^{1/3}$ phase oscillation periods, which is much less than the time required for significant coherence to be lost if n_0 is large.
- The diffusion gauge forms (7.75b) and (7.75d) that optimize for the “remaining time to target” ($t_{\text{opt}} - t$) show statistical errors that are well controlled by the choice of the target time t_{opt} . Basically statistical error can be reliably expected to remain small up to the explicit target time t_{opt} , provided that this is within the useful simulation range given in Table 7.1 — see Figure 7.5. Thus, a heuristic approach to get the maximum time out of a simulation would be to
 1. Pick t_{opt} to be some time that one wants to be able to simulate to.
 2. Run a simulation, and see how long useful observable estimates occur: $t_{\text{sim}}(t_{\text{opt}})$.
 3. Choose:
 - (a) If $t_{\text{sim}}(t_{\text{opt}}) < t_{\text{opt}}$ then reduce t_{opt} to some value between $t_{\text{sim}}(t_{\text{opt}})$ and the present t_{opt} . Using this new value should give a new better simulation time also in that range. Iterate back to step 2. Or,
 - (b) if $t_{\text{sim}}(t_{\text{opt}}) \approx t_{\text{opt}}$ then either keep the simulation if one is happy with it, or one can try to increase t_{opt} to perhaps obtain more simulation time, going back to step 2.

The rest of the diffusion gauge forms (7.75) do not follow such a simple dependence and appear to require tedious parameter searching to find the best gauge parameter choice for given initial conditions.

- Simulations showing quantum revivals have not been achieved. Generally it appears that useful simulation times are proportional to coherence times at high occupations. In this context, however, it is worth mentioning that quantum revivals *have* been seen in this system by Dowling[69] using the gauge P representation, by imposing externally-induced time reversal in the mode at a time t_{reversal} , where $t_{\text{coh}} < t_{\text{reversal}} \ll t_{\text{revival}}$. The phase coherence was seen by Dowling[69] to decay and then revive at $t = 2t_{\text{reversal}}$. This indicates that the potential for quantum revivals in these stochastic simulations is there, just not realized in the free evolution analyzed in this chapter.
- At low occupation, i.e. of the order of one boson or less, combined-gauge methods still give the best results, but the magnitude of their advantage becomes smaller.
- Of the four gauge forms (7.75), the form (7.75d) gives the longest useful simulations.
- The diffusion gauge forms (7.75d) and (7.75c) optimizing depending on the occupation of the current trajectory do better than those optimizing on the basis of only the initial condition. This is not surprising, since this approach is simply better tailored to the predicted subsequent evolution of each trajectory.
- At low occupations, a broad range of t_{opt} choices give much the same statistical behaviour. This is most likely because the forms (7.75) are not always the best optimizations of (7.65) in this regime for two reasons: 1) Because the “small time” condition $|q|t \ll 1$ is not always satisfied at the long time end of the simulation, and 2) because as seen in Figure 7.2, the form (7.74) is not an accurate root of the cubic (7.72).
- The simulation times with diffusion gauges (whether accompanied by drift gauge (7.30) or not) not only have better scaling with n_0 when n_0 is large, but this power-law decay of simulation time starts much later, as seen in Figure 7.3 and the right hand column of Table 7.1.

7.8 Diffusion gauges on their own

7.8.1 Motivation

The previous Sections 7.5 to 7.7 considered optimization of the diffusion gauge given that instabilities in the equations have been removed using the drift gauge (7.30). An approach that could be an alternative is to only use a diffusion gauge g'' to make a tradeoff between noise in the number and phase variables as before, but without introducing any drift gauge.

Such an approach could have the advantage that no weight spread is introduced since $dz_0 = 0$. The weight spread is not a major issue in a single-mode system, and clearly from Table 7.1 the drift-gauged simulations last longer, but with many modes there may be a problem: *All* the gauges introduce modifications to the *same* single weight variable. A vivid example occurs if we have M identical but independent modes. Then, $dz_0 = \sum_{k=1}^{2M} \mathcal{G}_k dW_k$, and the standard deviation of dz_0 is \sqrt{M} times greater than for a single mode but with no new physics. Since the log-weight z_0 enters into the observable estimates in exponential fashion, the situation is again analogous to the average of an exponential of a Gaussian random variable, and there is the non-scalable useful simulation limit $\text{var}[z_0] \lesssim \mathcal{O}(10)$ by condition(7.43). Since $\text{var}[z_0(t)] \approx \propto Mt$ at short times, this translates to

$$t_{\text{sim}} \propto \frac{1}{M} \tag{7.90}$$

in a worst case.

The major disadvantage of the $\mathcal{G} = 0$ approach, is that boundary term errors cannot be ruled out, as in the plain positive P method. One can, however, try to monitor for their appearance using the indicators developed by Gilchrist *et al*[64]: Spiking in moment calculations, rapid onset of statistical error after some time, or power-law distribution tails. A caveat is that these indicators tend to emerge only for fairly large ensemble sizes \mathcal{S} (Typically what happens is that the indicators appear with a certain delay time if \mathcal{S} small).

This $\mathcal{G} = 0$ method variant, although less desirable due to the need for careful monitoring of the simulation, was found to be much more successful than the

drift-gauged method (in its present form) at increasing simulation times for the preliminary example many-mode models simulated in Chapter 10.

7.8.2 Optimization of g'' and comments

Proceeding as before in Section 7.5.3, the formal solution of the anharmonic oscillator equations with $\Omega(0) = 1$ are as (7.53), but with the changes

$$m_L = (7.53b) + 4\chi \int_0^t n''(t') dt \quad (7.91a)$$

$$z_0(t) = 0. \quad (7.91b)$$

Following the same procedure as before, one obtains

$$\begin{aligned} \begin{bmatrix} \text{var}[G_L] \\ \text{var}[\tilde{G}_L] \end{bmatrix} &= \chi t \cosh(2g'') - 4\chi^2 \left(n'_0 \begin{bmatrix} - \\ + \end{bmatrix} n''_0 e^{-2g''} \right) \left\{ \frac{1 - e^{-\gamma t}(1 + \gamma t)}{\gamma^2} \right\} \\ &+ 4\chi^2 |n_0|^2 \left\{ \frac{e^{-qt} - 1}{q(q - \gamma)} - \frac{1 - e^{-\gamma t}}{\gamma(\gamma - q)} - \frac{1}{2} \left(\frac{1 - e^{-\gamma t}}{\gamma} \right)^2 \right\}, \end{aligned} \quad (7.92)$$

where the “damping strength” q is given by (7.70). The variational condition (7.52) can now be used to optimize g'' .

Considering the same special cases as in Section 7.5.4,

1. With damping absent, and at relatively short times $|q|t \ll 1$, the optimized gauge is

$$g''_{\text{opt}} \approx g_{\text{approx}} = \frac{1}{4} \log \left[\frac{1}{3} (4\chi t_{\text{opt}} |n_0|)^2 + 1 \right]. \quad (7.93)$$

(no cubic polynomial this time.)

2. The long time behaviour when $|q|t \gg 1$: When $q > 0$, the increase in the variances of G_L and \tilde{G}_L is still linear just like in (7.71), but with the constant $b = 4\chi^2 |n_0|^2 (q/2\gamma - 1)/q\gamma - 4\chi^2 [n'_0 \mp n''_0 e^{-2g''}]/\gamma^2$. When $q < 0$ the long time behaviour is again exponential: $G_L \approx 4\chi^2 |n_0|^2 e^{|q|t} (q/2\gamma - 1)/\gamma q$.
3. For nonzero damping in the $|qt_{\text{opt}}| \ll 1$ regime, the optimized diffusion gauge is

$$g''_{\text{opt}} \approx g_{\text{approx}} = \frac{1}{4} \log \left[\frac{(4\chi t_{\text{opt}} |n_0|)^2}{3} c_2(\gamma t_{\text{opt}}) + 1 \right], \quad (7.94a)$$

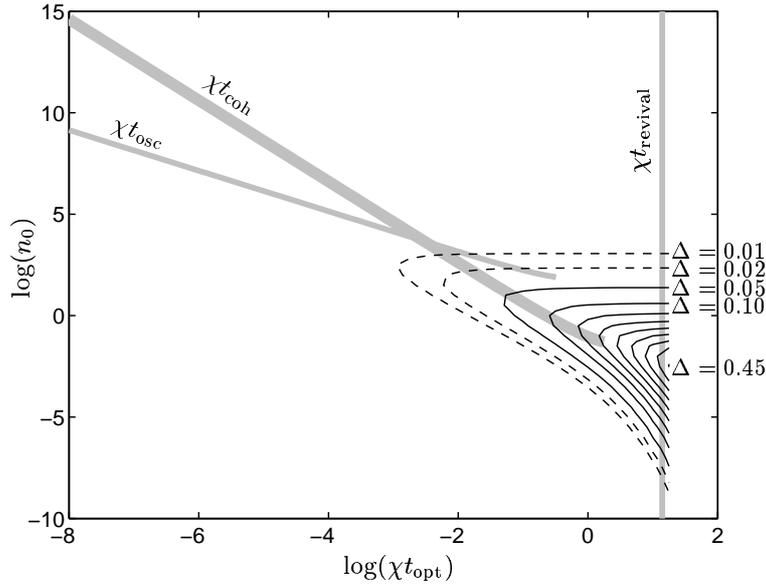


Figure 7.7: **Discrepancy** $\Delta = g''_{\text{opt}} - g''_{\text{approx}}$ between g''_{opt} (the exact optimization of g'' obtained by solving (7.52) using (7.92)) and the approximate expression (7.94). Displayed is the case of no damping ($\gamma = 0$) and classical initial occupation ($n_0 = n'_0$), shown as a function of t_{opt} and n_0 . Discrepancy values Δ are shown as SOLID CONTOURS with spacing 0.05. Additional DASHED CONTOURS shown at very low discrepancy. For comparison, several physical timescales are also shown in GREY: time of first quantum revival t_{revival} , phase coherence time t_{coh} and phase oscillation period t_{osc} .

where the coefficient is

$$c_2(v) = \frac{3}{2} \left(\frac{e^{-2v}(3 + 2v) + 1 - 4e^{-v}}{v^3} \right), \quad (7.94b)$$

which reduces to $c_2(0) = 1$ in the undamped case.

The discrepancy between (7.94) and the exact optimization obtained by solving (7.52) with (7.92) is shown for real n_0 for a wide range of parameters in Figure 7.7. It can be seen that for occupations $\gtrsim \mathcal{O}(10)$ and/or for times shorter than singly-occupied coherence time $1/2\chi$, the approximation is good. Compare with the analogous case with drift gauge shown in Figure 7.2.

For zero damping, and coherent state initial conditions $n_0 = n'_0$, and relatively short times $4\chi t e^{-2g''} \ll 1$ one finds

$$\text{var}[G_L(t)] = \text{var}[\tilde{G}_L(t)] = \chi t \cosh 2g'' - 2(\chi t)^2 n'_0 + \frac{8}{3} \chi t (n'_0)^2 e^{-2g''}. \quad (7.95)$$

At large occupation n'_0 , the gauge (7.94) gives $e^{g''} \approx (4\chi t_{\text{opt}} n'_0)^2/3$, and one finds that all three terms in (7.95) are of similar size. Using the condition (7.43) the simulation time is

$$t_{\text{sim}} \approx \mathcal{O}(10) t_{\text{coh}}. \quad (7.96)$$

The $4\chi t e^{-2g''_{\text{approx}}} \ll 1$ condition for $t < t_{\text{opt}}$ implies large occupation $n'_0 \gg \mathcal{O}(1)$. At low occupation on the other hand, $e^{g''_{\text{approx}}} \rightarrow 1$, the variances are $\approx e^{4\chi t} n'^2_0/4$, and one has again (7.82)

7.8.3 Numerical investigation of performance

The results of numerical simulations using such diffusion gauge only schemes are shown in Figures 7.8 and 7.9, in analogy with Figures 7.4 and 7.5 for the drift gauged schemes. Some data is also given in Tables 7.1 and 7.2. As before, simulations were made with $\mathcal{S} = 10^4$ for the mode occupations (7.87), and a wide range of target times t_{opt} .

The four gauges simulated were all of the general form (7.94), but with the four adaptive varieties forms

$$g''(t) = g''_{\text{approx}}(n_0, t_{\text{opt}}) \quad (7.97a)$$

$$g''(t) = g''_{\text{approx}}(n_0, t_{\text{rem}}) \quad (7.97b)$$

$$g''(t) = g''_{\text{approx}}(\check{n}(t), t_{\text{opt}}) \quad (7.97c)$$

$$g''(t) = g''_{\text{approx}}(\check{n}(t), t_{\text{rem}}) \quad (7.97d)$$

in analogy with (7.75), where t_{rem} is given by (7.76). Otherwise, the procedure was the same as described in Section 7.7. Additional features beyond what is mentioned there include:

- The diffusion-gauge-only simulations give improvement over the positive P, but give simulation times $\mathcal{O}(4)$ times shorter than when combined with drift gauge (7.30). For example, compare Figure 7.8 to Figure 7.4). t_{sim} is still $\gg t_{\text{coh}}$ at large n_0 , however.

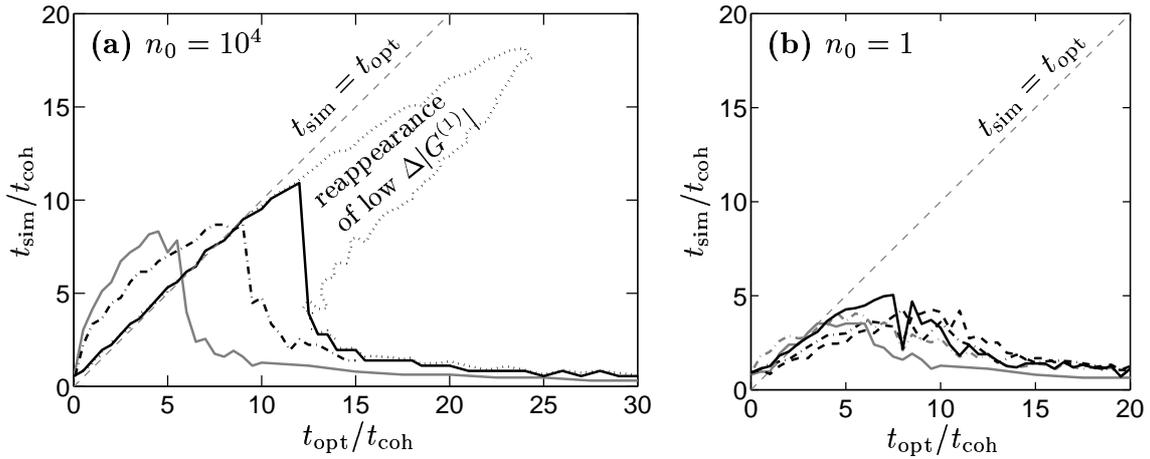


Figure 7.8: **Comparison of *a priori* target time t_{opt} with actual useful simulation time t_{sim}** for a variety of diffusion gauges in the $\mathcal{G} = 0$ schemes: Results for the four forms (7.97) are shown as: (7.97d) – SOLID DARK; (7.97c) – DASH-DOTTED LIGHT; (7.97b) – DASHED; (7.97a) – DASH-DOTTED DARK. Relationship obtained using the diffusion gauge (7.100) of Plimak *et al*[2] (but no drift gauge) is also shown (SOLID LIGHT). The dashed line in the background shows, for reference $t_{\text{opt}} = t_{\text{sim}}$. For the gauge (7.97d), the whole region where useful precision occurs is shown by the DOTTED line. Subplot (a): mean particle number $n_0 = 10^4$, (b): $n_0 = 1$.

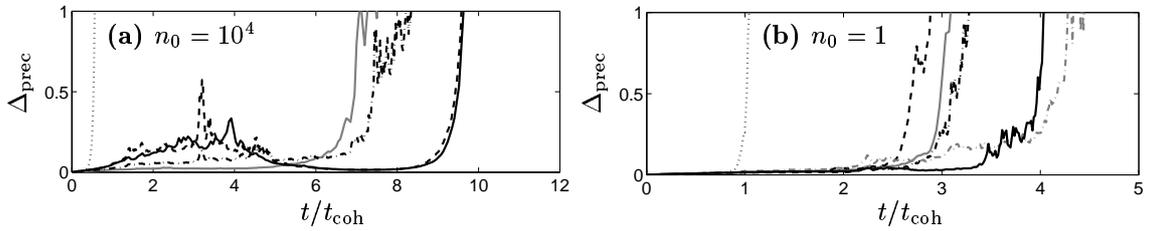


Figure 7.9: **Uncertainty in $|G^{(1)}(0, t)|$** , as a function of time for various $\mathcal{G} = 0$ schemes. The quantity plotted is $\Delta_{\text{prec}} = \sqrt{\mathcal{S}/10^6} (10 \Delta |G^{(1)}(0, t)| / |G^{(1)}(0, 0)|)$, so that $\Delta_{\text{prec}} \leq 1$ corresponds to useful precision as defined in Section 7.7.1. Results are plotted for the four diffusion gauge forms (7.97), shown as : (7.97d) – SOLID DARK; (7.97c) – DASH-DOTTED LIGHT; (7.97b) – DASHED; (7.97a) – DASH-DOTTED DARK. Data for the diffusion gauge (7.100) of Plimak *et al*[2] with $t_{\text{opt}} = 3t_{\text{coh}}$ (as used therein) is shown as a LIGHT SOLID line, while the ungauged positive P calculation is shown as a DOTTED line. Simulations were carried out with $\mathcal{S} = 10^4$ trajectories, starting with initial coherent state conditions: (a) $n_0 = 10^4$ (b) $n_0 = 1$. The diffusion-gauge plots in (a) were calculated with target times of $t_{\text{opt}} = 10t_{\text{coh}}$ for forms (7.97d) and (7.97b), and $t_{\text{opt}} = 7t_{\text{coh}}$ for forms (7.97c) and (7.97a). In subplot (b), the four forms (7.97) were simulated with $t_{\text{opt}} = 5t_{\text{coh}}$.

- The full adaptive gauge form (7.97d) gives the most predictable results, in analogy to the drift-gauged case. That is, $t_{\text{sim}} \approx t_{\text{opt}}$ for times up till those given in Table 7.1.
- At low n'_0 the simulation time appears more sensitive to the choice of t_{opt} than for the drift-gauged simulations.
- At high occupations $n'_0 \gg 1$, the pairs of adaptive gauges (7.97d)&(7.97b) and (7.97a)&(7.97c) behave identically, and are thus not shown in all plots. The simulation appears to be insensitive to whether one uses a gauge choice dependent on $\check{n}(t)$ or n_0 . This is probably a feature peculiar to the particle number conserving single-mode anharmonic oscillator model, because here $\check{n}(t) \approx n_0$ while useful precision is seen.
- For $n'_0 \gg 1$, the time-adaptive gauge forms (7.97b) and (7.97d) lead to a peculiar effect if the optimization time t_{opt} is chosen larger than the usual maximum t_{sim} given in Table 7.1. The statistical error in the $G^{(1)}$ estimate first rises rapidly, then falls again, and finally grows definitively. This is seen in Figure 7.9(a), and the parameter region for which this occurs is shown in Figure 7.8(a). In effect one has two time intervals when the simulation gives useful results: at short times, and later in a time interval around $t \approx \mathcal{O}(10t_{\text{coh}})$.

7.9 Comparison to recent related work

Improvements to the basic positive P simulation method for specific cases of interacting Bose gas systems have been tried with some success in several recent publications[1, 2, 3, 56]. Here we tie these together with the stochastic gauge formalism, and make some comparison to the results and analysis in the present chapter.

7.9.1 The work [1] of Carusotto, Castin, and Dalibard

The article *N-boson time-dependent problem: A reformulation with stochastic wavefunctions* considered an isolated (i.e. particle-conserving) system of \bar{N} interacting

bosons. The “coherent state simple scheme” described in Section III B 2 therein can be identified as using drift gauges of the form

$$\mathcal{G}_{\mathbf{n}} = i\sqrt{2i\chi}(\alpha_{\mathbf{n}}\beta_{\mathbf{n}} - |\alpha_{\mathbf{n}}|^2) \quad (7.98a)$$

$$\tilde{\mathcal{G}}_{\mathbf{n}} = \sqrt{2i\chi}(\alpha_{\mathbf{n}}\beta_{\mathbf{n}} - |\beta_{\mathbf{n}}|^2) \quad (7.98b)$$

when re-written in the notation of this thesis for a many-mode system as in equations (5.17), with mode labels \mathbf{n} . For the single-mode system considered in this chapter, this corresponds to

$$\mathcal{G}_1 = i\sqrt{2i\chi}(\check{n} - |\alpha|^2) \quad (7.99a)$$

$$\mathcal{G}_2 = \sqrt{2i\chi}(\check{n} - |\beta|^2). \quad (7.99b)$$

This gauge causes a full decoupling of the complementary α and β equations. Like (7.30) it is also successful in removing moving singularities, since the nonlinear terms in the radial equations for $d|\alpha|$ and $d|\beta|$ are removed.

There are two major differences between the coherent state wavefunction method and the gauge P method: Firstly, the former is hardwired to models that conserve particle number. This hardwiring to \bar{N} particles has both a major benefit and a disadvantage, as compared to the gauge P method considered in this thesis:

- **BENEFIT:** coherent state amplitudes are not free to explore the entire phase space, are bounded from above, and cannot form moving singularities.
- **DISADVANTAGE:** Losses or gains from external reservoirs are unable to be simulated, and so it is not applicable in its present form to simulations of e.g. evaporative cooling or a continuously loaded system such as an atom laser.

Another major difference is that diffusion gauges were not considered by Carusotto *et al*, and hence simulation times with this method for a single mode were very short ($\approx t_{\text{osc}}$).

The gauges (7.99) could also be applied to the present gauge P method, and their efficiency can be compared with the efficiency of gauges developed here: (7.30) when $g'' = 0$. The gauge (7.99) mediates the replacements $\check{n} \rightarrow |\alpha|^2$ or $|\beta|^2$ in the $d\alpha$ and

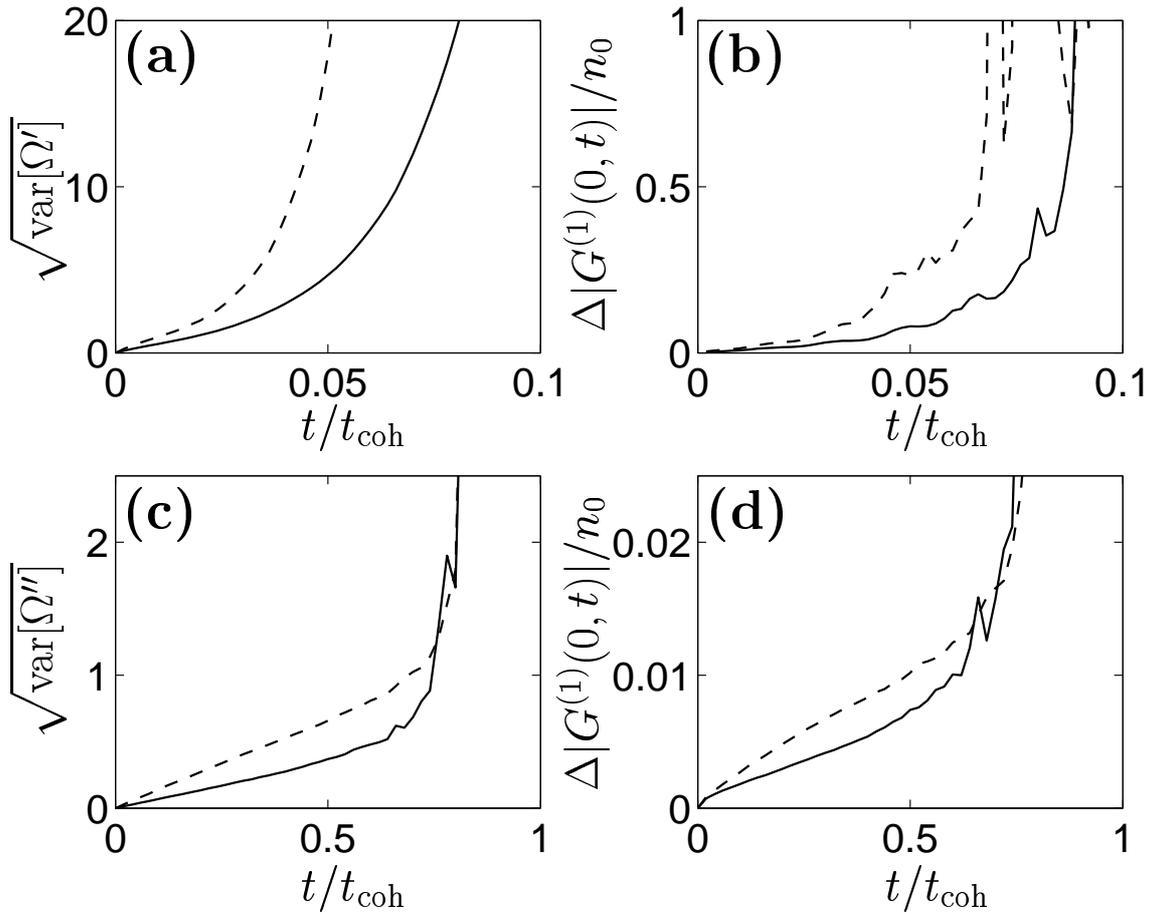


Figure 7.10: **Spread in trajectory weights** ((a) and (c)) and *phase correlation function uncertainties* ((b) and (d)), compared for the drift (only) gauges (7.99) – DASHED, and (7.30) (with no diffusion gauge: $g'' = 0$)– SOLID, in a one-mode, undamped, gainless system. Coherent state initial conditions with initial mean occupations (a) and (b): $n_0 = 10^4$, while in (c) and (d): $n_0 = 1$. All calculations are with 10^4 trajectories.

$d\beta$ equations (respectively), as opposed to (7.30), which only replaces $\check{n} \rightarrow \text{Re}\{\check{n}\}$ and does not decouple $d\beta$ from $d\alpha$. So, since the magnitude of the spread in Ω behaves approximately proportional to $\sqrt{\sum_k |\mathcal{G}_k|^2}$, the gauge (7.99) is expected to produce a broader distribution of weights. Numerical simulations were carried out, and the results are shown in Figure 7.10 for both high and low mode occupations, and Table 7.1 for high occupation. Useful simulation time with (7.99) is somewhat smaller than seen with (7.30).

In the same published work[1], a stochastic wavefunction method was developed based on Fock number states as well. This method was shown to give good results

and not be prone to boundary term errors but unfortunately does not appear extensible to open systems in any straightforward fashion because it is very strongly hardwired to \bar{N} total particles.

The stochastic wavefunction method was also adapted in [1] to some cases of non-local interactions, arriving at equations that use effectively a similar noise expansion to that in (5.46), but were only applicable to potentials having exactly real Fourier transforms $\tilde{U}_{\mathbf{n}} = \tilde{U}'_{\mathbf{n}}$.

7.9.2 The work [2] of Plimak, Olsen, and Collett

The article *Optimization of the positive- P representation for the anharmonic oscillator* considered a single-mode undamped, gainless system at high Bose occupation with coherent state initial conditions. The “noise optimization” scheme applied therein to greatly improve simulation times can be identified as an imaginary diffusion gauge of the form (rewritten in the present notation)

$$g_{12} = \frac{i}{2} \cosh^{-1} [2n_0\chi t_{\text{opt}}] = ig''_A \quad (7.100)$$

defined at high occupation or long times (i.e. while $2n_0\chi t_{\text{opt}} \geq 1$). This is dependent on a target time t_{opt} (which was taken to be $t_{\text{opt}} = 3t_{\text{coh}}$ in the calculations of Ref. [2]), and the initial Bose occupation $n_0 = n'_0 = |\alpha_0|^2$. The value of the gauge was held constant throughout the calculation, without allowance for a changing mode occupation.

The useful simulation times obtainable with this method are also shown in Figures 7.3 and 7.5, and Table 7.1. Their precise dependence on the target time parameter t_{opt} has been calculated here, and is shown in Figure 7.4.

Comparing (7.100) with the optimized diffusion gauge (7.97a) that is constant with time, one finds that in the $n_0 \gg 1$ limit the two gauges differ by a constant:

$$g''_A(n_0, t_{\text{opt}}) \approx g''_{\text{approx}}(n_0, t_{\text{opt}}) + \frac{\log 3}{4}. \quad (7.101)$$

While the simulation times achievable using g''_A are comparable with those obtained with the adaptive gauge $g''_{\text{approx}}(\check{n}(t), t_{\text{opt}})$ of (7.97a), the target time t_{opt} is no longer

a good indicator, and has a complicated relationship with t_{sim} , as seen in Figure 7.8. The simulation time is also much shorter than with the drift gauged simulation.

Extensions to multi-mode systems or low mode occupations were not considered, however since the system under consideration was only a single mode, the lack of drift gauges did not lead to any boundary term errors in this particular case.

7.9.3 The work [3] of Deuar and Drummond

Some initial research by Drummond and Deuar into gauges for the single-mode undamped anharmonic oscillator system was reported in[3]. Here the state was written as an expansion over a normalized, Hermitian, coherent state projector kernel

$$\widehat{\Lambda}(\alpha, \beta, \theta) = \frac{e^{i\theta} \|\alpha\rangle\langle\beta^*\| + e^{-i\theta} \|\beta^*\rangle\langle\alpha\|}{2e^{n'} \cos(\theta + n'')}, \quad (7.102)$$

where the phase variable θ was real. This is equivalent to imposing Hermiticity on the positive P kernel before normalizing it.

Drift gauges of the form

$$\mathcal{G}_1 = -\lambda\sqrt{\chi} [n'' - n' + |\alpha|^2] [\tan(\theta + n'') + i] \quad (7.103a)$$

$$\mathcal{G}_2 = \lambda\sqrt{\chi} [n'' + n' - |\beta|^2] [\tan(\theta + n'') + i] \quad (7.103b)$$

were used, where λ is a (real) strength parameter chosen *a priori*.

Simulation precision was greatly improved, but evident boundary term errors were present. This may have been caused by several factors, which are pointed out here because they are good examples of what should be avoided when choosing kernels and gauges:

1. The radial super-exponential behavior of $d\alpha$ and $d\beta$ is not removed in full, hence one can expect that moving singularities will still be present — as is ultimately seen as boundary term errors.
2. The magnitude of the kernel (7.102) diverges when $\theta \rightarrow -n''$. Hence, this parameter region might be unable to be sampled in an unbiased way. In particular, matrix elements of $\widehat{\Lambda}$ appearing in the expression (6.1) for $\widehat{\mathcal{B}}$ diverge when $\theta = \pm\pi/2$, and may lead to nonzero boundary terms once integrated.

3. The gauges were dependent on the trajectory weight $e^{i\theta}$, and thus non-autonomous. This causes feedback behavior between the phase-space variables α and β , and the weight, which is difficult to analyze to definitively determine whether moving singularities are present or not. (Compare to point 6(c) in Section 6.3.2).
4. The gauges diverge at $\theta \rightarrow \pm\pi/2$, against recommendation 6(b) of Section 6.3.2.

7.10 Gauge recommendation

Collecting together the analysis, and results of numerical calculations reported in this chapter, the following gauge choices appear to be advantageous for a single mode (labeled \mathbf{n}) of a two-body locally interacting Bose gas that is free to interact with its environment:

Defining

$$\check{n}_{\mathbf{n}} = \alpha_{\mathbf{n}}\beta_{\mathbf{n}}, \quad (7.104)$$

then

$$\mathcal{G}_{\mathbf{n}} = -i\tilde{\mathcal{G}}_{\mathbf{n}} = -\sqrt{2i\chi} \operatorname{Im} \{ \check{n}_{\mathbf{n}} \} e^{-g_{\mathbf{n}}''} \quad (7.105)$$

The local diffusion gauge dependent on the single “target time” parameter t_{opt} , or more precisely on the “remaining time to target”

$$t_{\text{rem}} = \begin{cases} t_{\text{opt}} - t & \text{if } t < t_{\text{opt}} \\ 0 & \text{if } t \geq t_{\text{opt}} \end{cases}. \quad (7.106)$$

is then given by

$$g_{\mathbf{n}}'' = \frac{1}{6} \log \left\{ 8|\check{n}_{\mathbf{n}}(t)|^2 \chi t_{\text{rem}} + a_2^{3/2} (\check{n}_{\mathbf{n}}(t), \gamma_{\mathbf{n}} t_{\text{rem}}) \right\}, \quad (7.107)$$

i.e. (7.74). The coefficient a_2 is

$$a_2(\check{n}, v) = 1 + 4\operatorname{Im} \{ \check{n} \}^2 \left(\frac{1 - e^{-2v}}{2v} \right) - 2|\check{n}|^2 \left(\frac{1 - 2v + 2v^2 - e^{-2v}}{v} \right). \quad (7.108)$$

7.11 Summary

The principal aim of this chapter has been to develop a “black box” form of drift and diffusion gauges. A form that can be expected to remove boundary term errors and extend simulation times beyond what is possible with the positive P for a single mode of a locally-interacting Bose gas with two-particle collisions that is open to the environment. The recommended form is given by (7.104)-(7.108). These gauges:

1. Extend simulation time for the anharmonic oscillator beyond what was possible with the positive P representation (i.e. when $g''_{jk} = 0$ and $\mathcal{G}_{\mathbf{n}} = \tilde{\mathcal{G}}_{\mathbf{n}}$). For high mode occupations, the simulation time is extended by a factor of $\mathcal{O}(15n_0^{1/6})$ (see Table 7.1) when starting from a coherent initial state, representative of a single trajectory. This simulation time is $\mathcal{O}(40)$ coherence times at high mode occupation, and all decoherence behavior is simulated. The improvement in numerical performance is best summarized by Figures 7.3, 7.6, and in Table 7.1.
2. Remove the instability responsible for moving singularities.
3. Apply for dynamically changing mode occupation.
4. Apply for simulations of open systems.

For the aim of using these gauges in many-mode simulations of open interacting Bose gases, it is crucial that the gauge choice be freely adaptable to any changes caused to the mode occupation (or other properties). Such changes will be caused by interactions with the rest of the modes, and an external environment.

A second possibly alternative method where only diffusion gauges are used has been considered, and appropriate optimized diffusion gauges derived. This method does not guarantee removal of systematic biases, but appears to have advantages in efficiency for many-mode calculations provided the simulation can be successfully monitored using the indicators of Gilchrist *et al*[64] to catch any boundary term errors if these form.

Lastly, the method developed has been tied in with and compared to some previous related work in the field[1, 2, 3].

Chapter 8

Dynamics of two coupled Bose gas modes

The next step up towards a many-mode simulation is to see how the statistical behavior of the gauge methods is affected once mode mixing occurs. In this chapter, the effect of the combined drift and diffusion gauges suggested in Section 7.10 on dynamical simulations of a two-mode system will be considered (undamped, gainless, with two modes coupled by Rabi oscillations). This system is simple enough that there are only a few parameters, yet many of the changes caused by mode mixing are visible.

8.1 The model

The system consists of two orthogonal modes labeled 1 and 2 with interparticle interactions in each mode, and Rabi coupling between them. The coupling frequency will be restricted to be real for simplicity (complex values can always be considered in similar vein with no significant alterations to the following discussion). Time units are chosen so that the nonlinear interaction frequency is $\chi = 1$, and a transformation is made to an interaction picture in which the linear mode self-energies $\hbar\omega_{jj}\hat{n}_j$ are moved into the Heisenberg evolution of operators. The interaction picture

Hamiltonian then is

$$\hat{H} = \hbar\omega_{12} \left[\hat{a}_1^\dagger \hat{a}_2 + \hat{a}_2^\dagger \hat{a}_1 \right] + \hbar \sum_{j=1}^2 \hat{a}_j^\dagger{}^2 \hat{a}_j^2. \quad (8.1)$$

The Rabi frequency is ω_{12} (in scaled time units).

Physically this model can represent, for example, a simplified model of two internal boson states coupled by an EM field, or two trapped condensates spatially separated by a barrier. This approximation has been used by a multitude of authors to investigate the quantum behavior of BECs in two-state systems. For references to some of the work in this field see Leggett [30], Part VII.

Comparing to the lattice Hamiltonian (2.17) of Chapter 2, the modes 1 and 2 can be identified with some modes \mathbf{n} and \mathbf{m} , respectively. Then, $\omega_{12} = \omega_{\mathbf{nm}}/\chi$, while χ , $\omega_{\mathbf{nm}}$, and $\omega_{\mathbf{mm}}$ have been scaled out or moved into the Heisenberg evolution of the \hat{a}_j . It has been assumed that $\omega_{\mathbf{nm}}$ is real, and that all $\gamma_{\mathbf{n}}$ and $\varepsilon_{\mathbf{n}} = 0$.

From the gauge P equations (5.17) of Chapter 5, and not yet specifying the exact form of the noise matrices, one obtains gauged Ito stochastic equations

$$d\alpha_j = -i\omega_{12}\alpha_{-j} dt - 2i\alpha_j^2\beta_j dt + \sum_k B_{jk}^{(\alpha)}(dW_k - \mathcal{G}_k dt) \quad (8.2a)$$

$$d\beta_j = i\omega_{12}\beta_{-j} dt + 2i\alpha_j\beta_j^2 dt + \sum_k B_{jk}^{(\beta)}(dW_k - \mathcal{G}_k dt) \quad (8.2b)$$

$$d\Omega = \sum_k \mathcal{G}_k dW_k, \quad (8.2c)$$

where $-j = 1 + \delta_{1j}$, and dW_k are the usual Wiener increments. The noise matrices obey

$$\sum_l B_{jl}^{(\alpha)} B_{jk}^{(\alpha)} = -2i\delta_{jk}\alpha_j^2 \quad (8.3a)$$

$$\sum_l B_{jl}^{(\beta)} B_{jk}^{(\beta)} = 2i\delta_{jk}\beta_j^2 \quad (8.3b)$$

$$\sum_l B_{jl}^{(\alpha)} B_{jk}^{(\alpha)} = 0. \quad (8.3c)$$

Applying the drift and diffusion gauges (7.105) and (7.107) worked out in Chap-

ter 7 to each mode separately, these equations become

$$d\alpha_j = -i\omega_{12}\alpha_{-j} dt - 2i\alpha_j \text{Re}\{\check{n}_j\} dt + i\alpha_j\sqrt{2i} [\cosh g_j'' dW_{j1} + i \sinh g_j'' dW_{j2}] \quad (8.4a)$$

$$d\beta_j = i\omega_{12}\beta_{-j} dt + 2i\beta_j \text{Re}\{\check{n}_j\} dt + \beta_j\sqrt{2i} [-i \sinh g_j'' dW_{j1} + \cosh g_j'' dW_{j2}] \quad (8.4b)$$

$$d\Omega = -\sqrt{2i}\Omega \sum_j \text{Im}\{\check{n}_j\} e^{-g_j''} (dW_{j1} - idW_{j2}), \quad (8.4c)$$

where $\check{n}_j = \alpha_j\beta_j$, and

$$g_j'' = \frac{1}{6} \log \{8|\check{n}_j|^2 t_{\text{opt}} + (1 + 4 \text{Im}\{\check{n}_j\}^2)^{3/2}\}. \quad (8.5)$$

The four independent real Wiener increments dW_{jk} can be implemented by independent Gaussian random variables of variance dt at each time step.

Proceeding to check the no-moving-singularities condition (6.3) as for the single mode in Section 7.3, there is no super-exponential growth of any $|\alpha_j|$, $|\beta_j|$ or $|\Omega|$. Moving singularities or noise divergences will not occur.

Two kinds of initial conditions were numerically investigated in detail:

8.2 Case 1: Coupling to a vacuum mode

8.2.1 Description

The system starts initially with a coherent state of mean particle number n_0 in mode 1, and vacuum in mode 2. In all simulations of this case, the inter-mode coupling strength was taken to be $\omega_{12} = 5$, but the mean particle number n_0 was varied. This mean particle number $\bar{N} = n_0 = \sum_j \langle \hat{n}_j \rangle = \sum_j \langle \hat{a}_j^\dagger \hat{a}_j \rangle$ is conserved during each simulation.

At low particle number, the Rabi oscillations dominate the Hamiltonian, and particles oscillate between the modes, without much decoherence. At high particle number n_0 , on the other hand, phase decoherence dominates mode 1, suppressing also the coherent transfer of particles to mode 2.

The particular values chosen to simulate were

$$n_0 = \{1, 17, 200, 1500, 10^4\}. \quad (8.6)$$

The two observables considered were: The fraction of particles in the (initially empty) mode 2:

$$p_2 = \frac{\langle \hat{n}_2 \rangle}{N}, \quad (8.7)$$

and the local normalized second order correlation functions of modes j :

$$g_j^{(2)}(t, t) = \frac{\langle : \hat{n}_j(t) : \rangle}{\langle \hat{n}_j(t) \rangle^2} = \frac{\langle \hat{n}_j^2 \rangle}{\langle \hat{n}_j \rangle^2} - \frac{1}{\langle \hat{n}_j \rangle}. \quad (8.8)$$

The latter quantify the amount of (instantaneous) bunching/antibunching in the boson field, and for example, are unity for coherent states, two for thermal fields, and $1 - 1/n$ for Fock number states of n particles. Large values $g^{(2)} > 2$ occur e.g. for quantum superpositions of vacuum and Fock number states with two or more particles where the average particle number is small. For example in the state $|\psi\rangle = \sin\theta |0\rangle + \cos\theta |n\rangle$, $g^{(2)} = (1 - \frac{1}{n})/\cos^2\theta$.

8.2.2 Procedure

Three gauge choices were compared for this case.

1. *No gauge.* i.e. positive P.
2. *Drift and diffusion gauges* as in equations (8.4) with (8.5). The target time t_{opt} was varied to obtain the longest useful simulation time t_{sim} .
3. *Drift gauge* (7.99) of Carusotto *et al* [1]. As pointed out before, a subtle difference from Ref. [1] is that here particle number conservation is not explicitly hardwired into the simulation, but happens to arise because no environment interactions are included in this particular simulation. This freedom is left unconstrained to make conclusions reached on the basis of the two modes relevant to a larger open many-mode system.

In each run, $\mathcal{S} = 2 \times 10^5$ trajectories were used, and useful simulation precision was defined as in (7.88), and simulation time t_{sim} accordingly.

In Figure 8.1 actual simulation times are compared to physical timescales and expected simulation times based on the single-mode analysis of Chapter 7. Note that

fitting parameters to (7.89) for the Carusotto *et al*[1] gauge (7.99) were calculated from single-mode simulations to be

$$\{c_0, \dots, c_4\} = \{1, 3.3 \pm 0.5, 0.9^{+\infty}_{-0.2}, -0.3 \pm 0.4, 0.31 \pm 0.09\}, \quad (8.9)$$

and for (7.105) with (8.5) and $t_{\text{opt}} = 0$ were

$$\{c_0, \dots, c_4\} = \{1, 400 \pm 130, 3.6^{+\infty}_{-2.3}, 1.2 \pm 0.2, 0.42 \pm 0.03\}. \quad (8.10)$$

Example simulations are shown in Figures 8.2, 8.3, 8.4, and 8.5, and dependence of t_{sim} on the target time parameter t_{opt} are shown in Figure 8.6.

Any data from positive P simulations after the onset of spiking (numerical signature of possible boundary term errors) were discarded. The simulations up to such a point, showed no sign of any systematic biases¹.

8.2.3 Features seen

- With the single-mode gauge choices (7.105) and (8.5) applied directly to the twin-mode system, improvement over the positive P and Carusotto *et al* gauge (7.99)[1] is seen for all n_0
- This improvement becomes more and more significant at high mode occupations, like in the single-mode system. Generally speaking, the proportions between the simulation times obtained with the different gauge methods are qualitatively similar to the single-mode situation.
- At large mode occupations, one can still simulate well beyond coherence times, so most interesting phenomena should not be missed (apart from quantum revivals, of course). (See Figures 8.4–8.5)
- However, simulation times with all methods are significantly smaller than those predicted (by using expected mode occupation n_0) for single modes from Tables 7.1 or 7.2, or (8.9)–(8.10).

¹The onset of spiking naturally moves to slightly earlier times as larger ensembles of trajectories are simulated. This may cause a beneficial “bias-concealing” effect in the following sense: As precision increases due to larger \mathcal{S} , some small boundary term biases that were originally concealed by noise might emerge, but simultaneously some rare trajectories with earlier spikes become included in the larger ensemble. These then warn one to discard data from an earlier time than one would have with the original smaller ensemble.

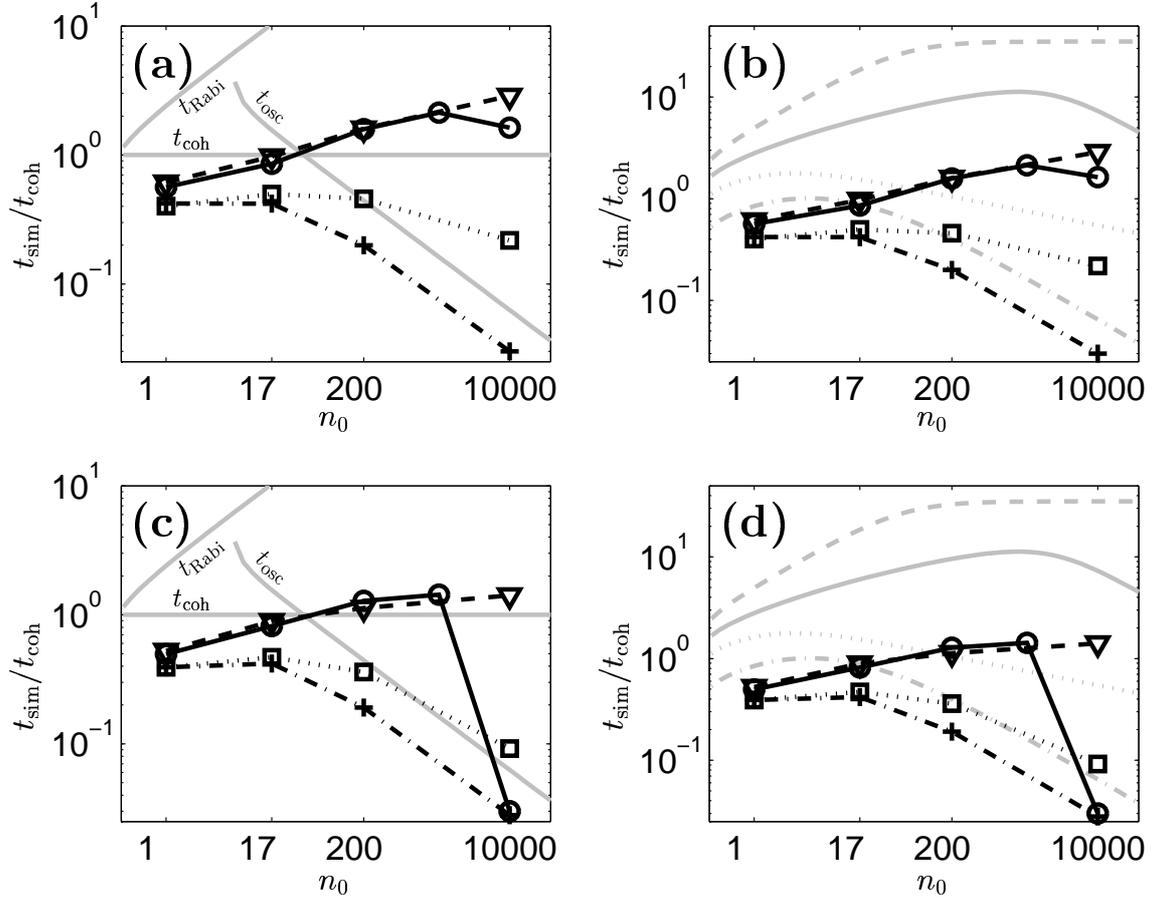


Figure 8.1: **Useful simulation times** t_{sim} for a mode coupled to vacuum as in Section 8.2. Calculated simulation times are shown as data points, with the symbols denoting gauge used. “ \square ”: positive P; “ \circ ”: drift gauge (7.105) and diffusion gauge (8.5) with $t_{\text{opt}} = 0$; “+” drift gauge (7.99) of Carusotto *et al*[1]; “ ∇ ”: drift and diffusion gauges (7.105) and (8.5) with best target time parameters: $\omega_{12}t_{\text{opt}} = \{2.5, 0.5, 0.5, 0.075\}$ for $n_0 = \{1, 17, 200, 10^4\}$, respectively. Subplots (a) and (b) show simulation times based on estimates of the observable p_2 , while (c) and (d) times based on $g_2^{(2)}$. Subplots (a) and (c) compare to physical time scales, including Rabi oscillation period $t_{\text{Rabi}} = 2\pi/\omega_{12}$, while subplots (b) and (d) compare to expected simulation times for a single mode using the empirical fits of Table 7.2 and (8.9)–(8.10). The expected t_{sim} are plotted as light lines: DOTTED: positive P; SOLID: drift gauge (7.105) and diffusion gauge (8.5) with $t_{\text{opt}} = 0$; DASHED: with optimum t_{opt} choice; DASH-DOTTED: drift gauge (7.99) of Carusotto *et al*[1].

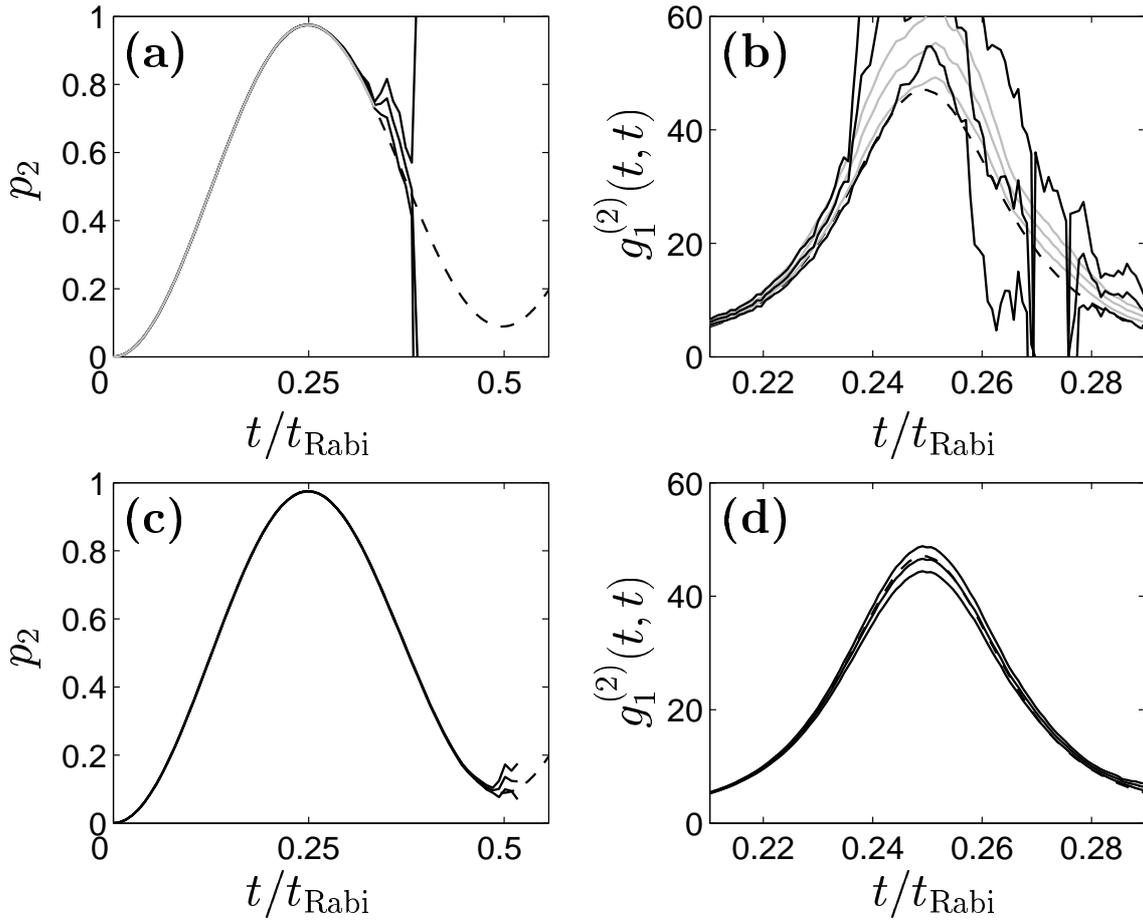


Figure 8.2: Coupling to vacuum mode: **System with few particles on average** ($n_0 = 1$). Time scaled to Rabi period $t_{\text{Rabi}} = 2\pi/\omega_{12}$. Subfigures (a) and (c) show mode 2 particle fraction, while (b) and (d) show instantaneous particle number correlations in mode 2. Subfigures (a) and (b) show results with positive P LIGHT SOLID and drift gauge (7.99) DARK SOLID, whereas (c) and (d) show results with combined drift gauge (7.105) and diffusion gauge (7.107) using target time $t_{\text{opt}} = 2.5/\omega_{12} \approx 0.4t_{\text{Rabi}}$. Exact results also shown (dashed). Triple lines indicate mean and error bars.

- No biases from the exact solution are seen for the cases ($n_0 = 1, 17$) where this was compared.
- The $n_0 = 17$ case was also simulated by Carusotto *et al*[1] with (solely) the drift gauge (7.99), and a simulation time of $\approx 0.55/\omega_{12}$ was seen. This is about 20% longer than with the same gauge here. This effect is most likely because there total particle number conservation was *a priori* imposed on each

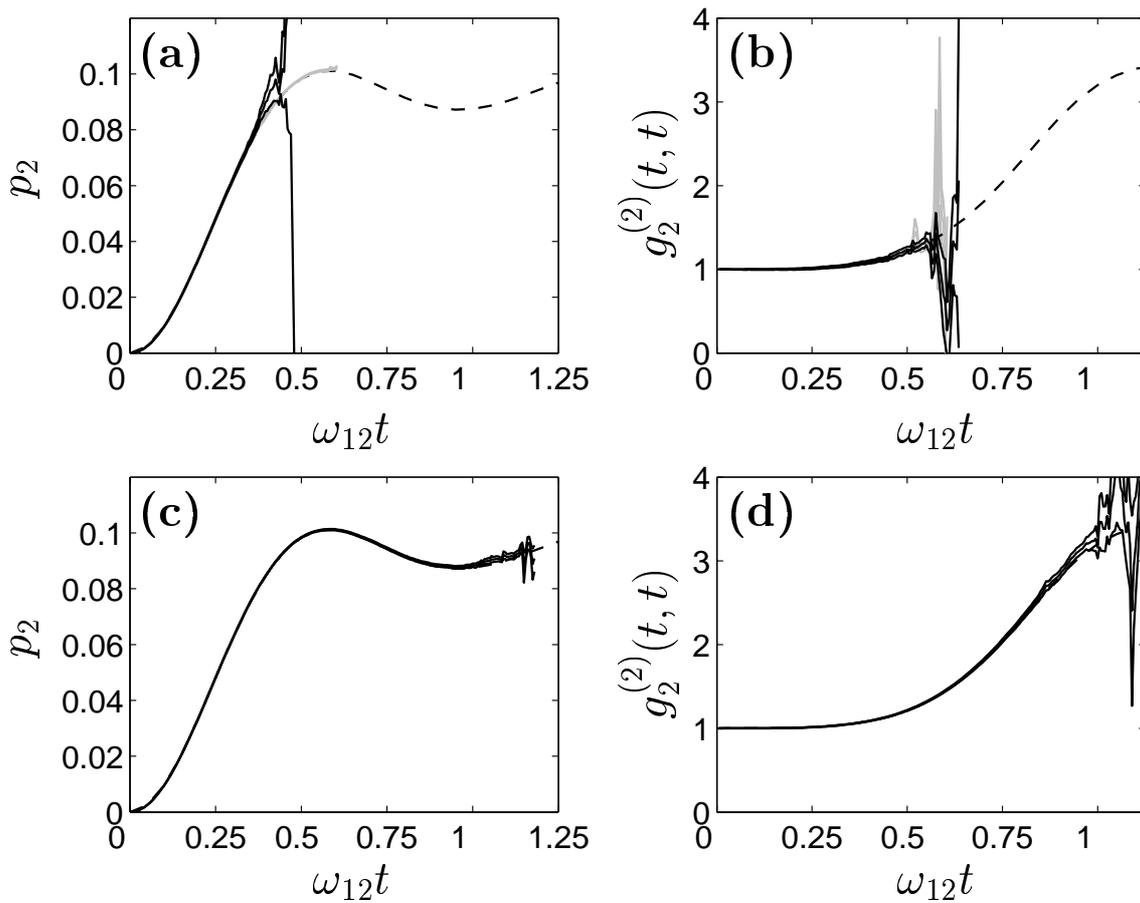


Figure 8.3: Coupling to vacuum mode: **System with intermediate number of particles** ($n_0 = 17$). Subfigures (a) and (c) show mode 2 particle fraction, while (b) and (d) show instantaneous particle number correlations in mode 2. Subfigures (a) and (b) show results with positive P LIGHT SOLID and drift gauge (7.99) DARK SOLID, whereas (c) and (d) show results with combined drift gauge (7.105) and diffusion gauge (7.107) using target time $\omega_{12}t_{\text{opt}} = 0.5$. Exact results also shown (dashed). Triple lines indicate mean and error bars.

trajectory rather than only in the ensemble mean as here. The extra freedom in the present simulations is necessary to allow straightforward extension to open systems, but the price paid appears to be some extra inefficiency.

- Precision in the highly-occupied mode appears largely insensitive to the choice of target time t_{opt} , as long as it is within reasonable values (See Figure 8.6). A choice of $t_{\text{opt}} = 0$ achieves very large improvement over positive P and drift-gauge only simulations for these observables. Note that $t_{\text{opt}} = 0$ is not the same as no diffusion gauge, because (8.5) becomes $g_j'' = \frac{1}{4} \log(1 + 4\text{Im}\{\tilde{n}_j\}^2)$,

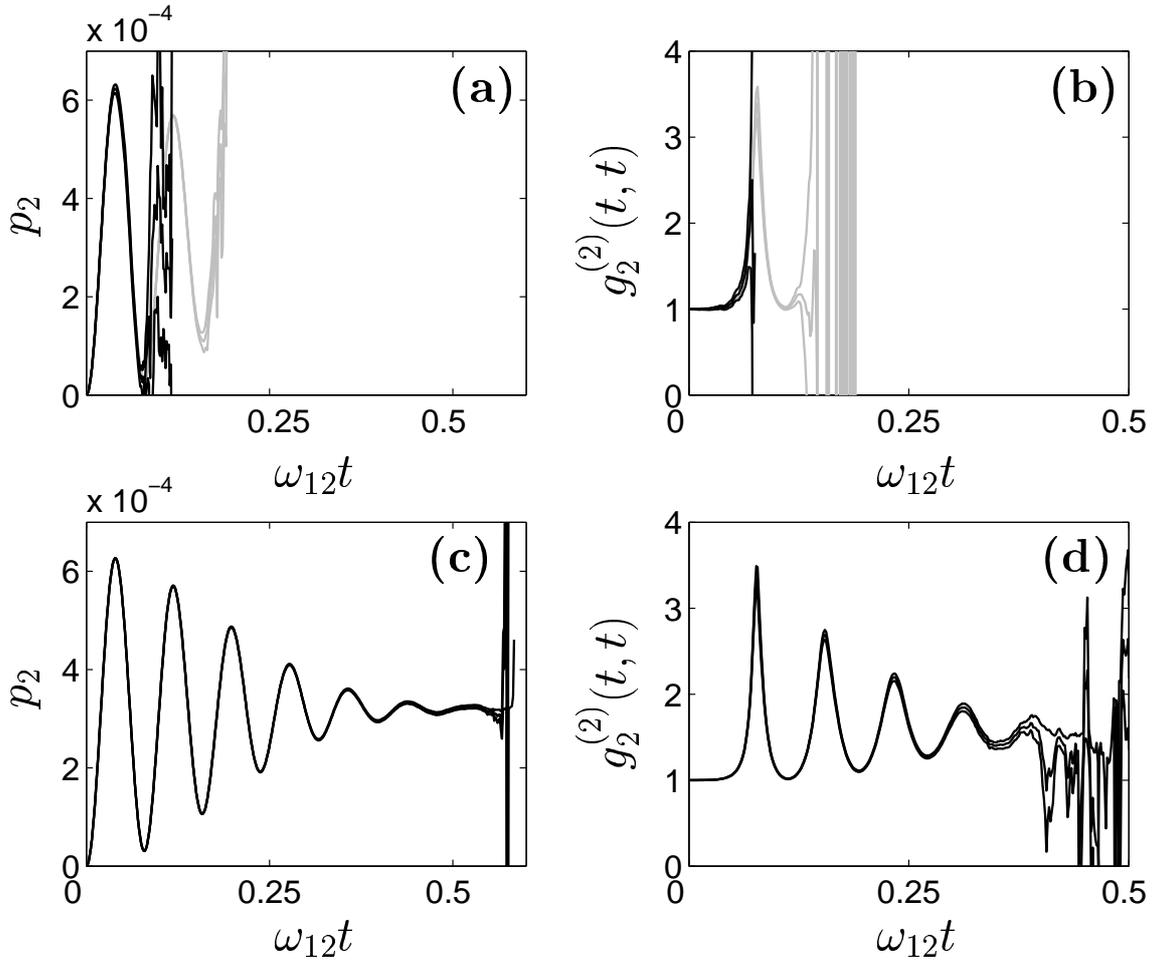


Figure 8.4: Coupling to vacuum mode: **System with large number of particles** ($n_0 = 200$). Subfigures (a) and (c) show mode 2 particle fraction, while (b) and (d) show instantaneous particle number correlations in mode 2. Subfigures (a) and (b) show results with positive P LIGHT SOLID and drift gauge (7.99) DARK SOLID, whereas (c) and (d) show results with combined drift gauge (7.105) and diffusion gauge (7.107) using target time $\omega_{12}t_{\text{opt}} = 0.25$. Triple lines indicate mean and error bars.

and this appears to have an important effect. The same applies to mode 2 when n_0 is not large.

- At high n_0 the observables for the nearly-vacuum mode 2 become sensitive to target time choice (See Figure 8.6). This is particularly evident for $g_2^{(2)}$ when $n_0 = 10^4$. The optimum target time t_{opt} is then sharply peaked, much as in the single-mode case, and $t_{\text{sim}} \approx \geq t_{\text{opt}}$ in the optimum region where t_{opt} is of the order of coherence time for the highly-occupied mode 1.

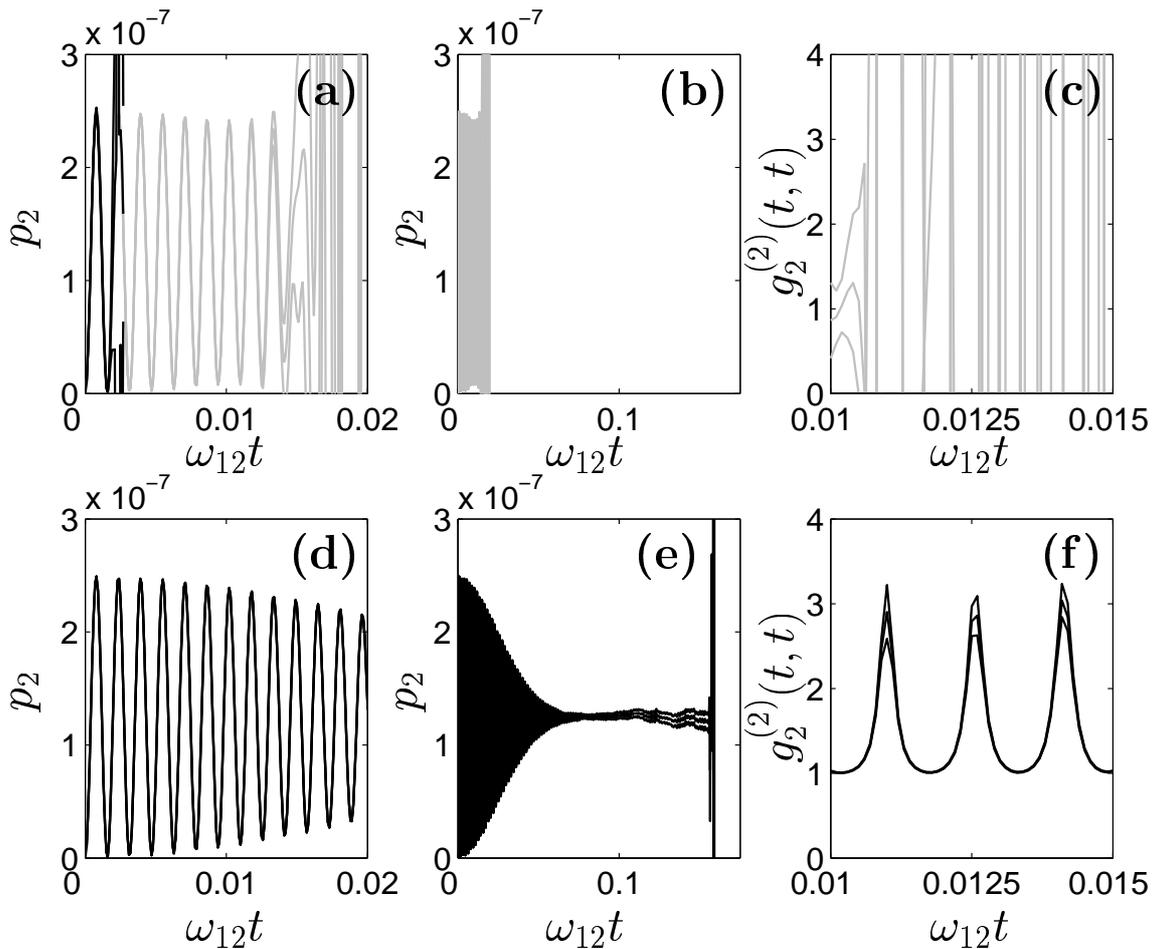


Figure 8.5: Coupling to vacuum mode: **System with very large number of particles** ($n_0 = 10^4$). Subfigures (a), (b), (d) and (e) show mode 2 particle fraction, while (c) and (f) show instantaneous particle number correlations in mode 2. Subfigures (a)–(c) show results with positive P LIGHT SOLID and drift gauge (7.99) DARK SOLID, whereas (d)–(f) show results with combined drift gauge (7.105) and diffusion gauge (7.107) using target time $\omega_{12}t_{\text{opt}} = 0.1$. Triple lines indicate mean and error bars.

- One point to note is that “useful simulation time” t_{sim} was based on the moment when relative error in a quantity was *first* found to be too large. In calculations of $g^{(2)}$, uncertainties are much greater when $g^{(2)}(0, t)$ peaks — see e.g. Figure 8.5(f). Good accuracy can often be obtained between peaks for much longer times than shown in Figure 8.6, up to about the same simulation time as worked out based on the occupation number. This is particularly evident in the $n_0 = 10^4$ case, where the sudden drop-off in t_{sim} is due to this

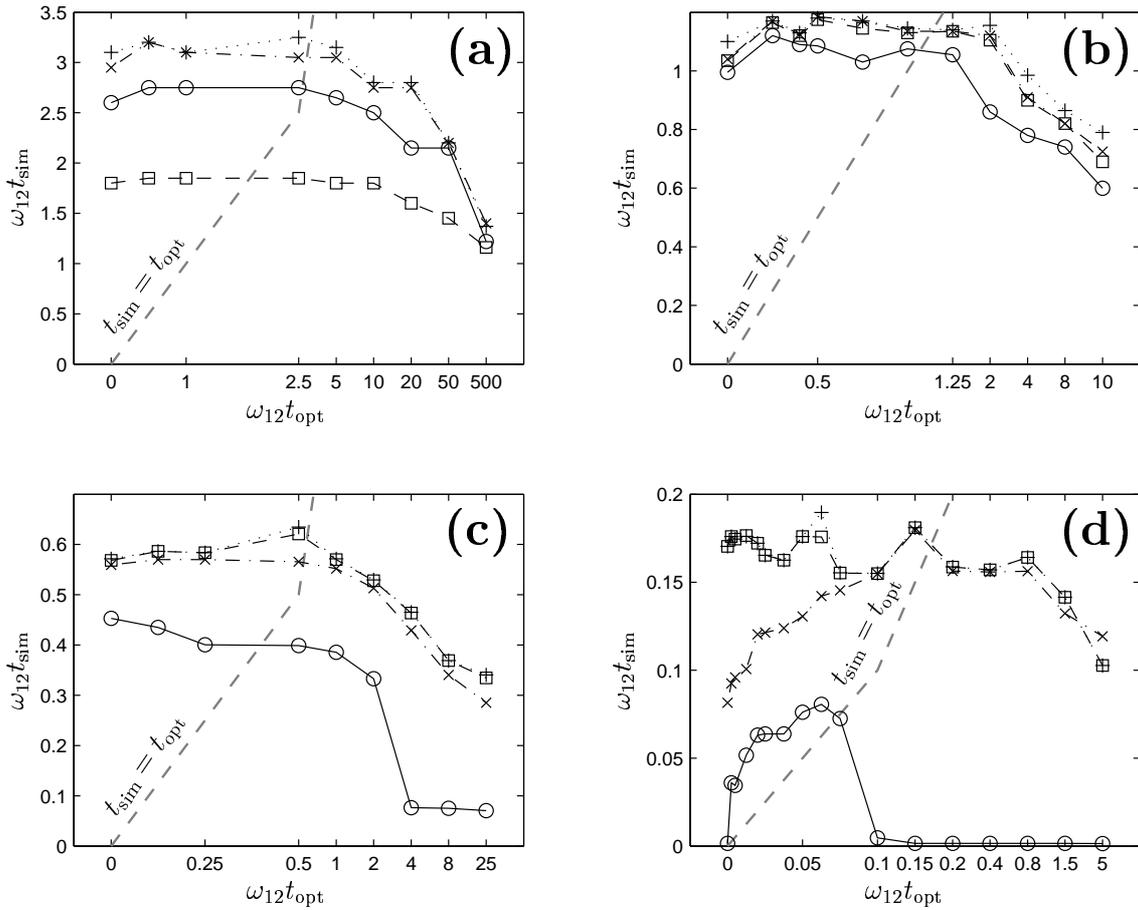


Figure 8.6: Dependence of **useful simulation time** t_{sim} on the target time parameter t_{opt} for simulations of Section 8.2 “+”: useful simulation of $\langle \hat{n}_1 \rangle$, “x”: of $\langle \hat{n}_2 \rangle$, “□”: of $g_1^{(2)}$, “○”: of $g_2^{(2)}$. Nonlinear but monotonic t_{opt} scale used to avoid overlap of data points, while showing trend. Subplot (a): $n_0 = 1$; (b): $n_0 = 17$; (c): $n_0 = 200$; (d): $n_0 = 10^4$. $t_{\text{opt}} = t_{\text{sim}}$ shown for reference as LIGHT DASHED line.

effect (see solid line of Figure 8.6(d)).

- Best simulation times at high n_0 occur for $t_{\text{opt}} \leq \approx t_{\text{sim}}$.
- Positive P simulations do better than those based on the drift gauge (7.99) for all n_0 .
- Best simulation times (based on $\langle \hat{n}_2 \rangle$) with both gauges appear to scale roughly as

$$\omega_{12} t_{\text{sim}} \approx 3/n_0^{1/3} \quad (8.11)$$

Collective conclusions are left till the end of the chapter.

8.2.4 Comparison to Fock state stochastic wavefunctions

The simulation of a seemingly similar 17-particle two-mode system was carried out by Carusotto *et al*[1] using the “simple Fock state scheme” with stochastic wavefunctions. This method gave a useful simulation time of $\approx 3/\omega_{12}$ with $\mathcal{S} = 2 \times 10^5$ trajectories. This is longer than seen with either the “simple coherent state scheme” therein ($\approx 0.55/\omega_{12}$), or the simulation in Figure 8.3 here ($\approx 1.18/\omega_{12}$). This was, however, for number-state initial conditions, so the two types of simulation cannot be easily directly compared, and the observable expectation values calculated are different.

Coherent states with mean occupation n_0 , used as starting points here, contain Fock components with much larger numbers of particles than n_0 . Since the spread in the norm, and hence the uncertainty of observables, has been found to rise exponentially depending on the quantity $\bar{N}t = n_0t$ (from equation (78) in [1]), one expects that useful simulation time of the simple Fock state scheme will be $\approx \propto 1/\bar{N} = 1/n_0$. This is a sharp decrease compared to the $1/n_0^{1/3}$ dependence with the combined gauges discussed here, and one expects that for $n_0 \gg 17$, the Fock state wavefunction method will give much shorter simulation times.

One can attempt to compare schemes for $n_0 = 17$ in spite of the differences in the physical system simulated. It is not clear how best to sample a coherent state (which includes Fock state components with different particle numbers) with the Fock scheme, but one can estimate that for accuracy one should consider components with occupation at least three standard deviations away from the mean giving $\bar{N} \lesssim 30$ for such Fock state components. A trajectory corresponding to such a component might be expected to not diverge for times $\omega_{12}t \lesssim 1.7$.

This is about 30% longer than seen in Figure 8.3, a factor similar to that noticed between simulations using the gauge (7.99) in Ref. [1] using the “simple coherent scheme” and here. It is likely the price paid for having the formalism left open to simulating gains and losses.

8.3 Case 2: Coherent mixing of two identical modes

8.3.1 Description

The system starts initially with identical coherent states of mean particle number n_0 in each mode. The total mean particle number is $\bar{N} = 2n_0$. During time evolution, Rabi coupling of the modes occurs along with some decoherence from inter-atom collisions in each mode. This time simulations were carried out with constant particle number $n_0 = 100$, but the coupling frequency ω_{12} was varied.

At low frequency $\omega_{12} \ll n_0 = 100$, decoherence local to each mode dominates, and phase oscillations in each mode occur with period $t_{\text{osc}} = \pi/n_0$, while at high frequency $\omega_{12} \gg n_0 = 100$, the inter-mode coupling dominates and phase oscillations for each mode occur with period $2\pi/\omega_{12}$. One expects that for low coupling the two modes should behave largely as two independent single modes of Chapter 7.

The particular values chosen to simulate were

$$\omega_{12} = \{5000, 500, 50, 5, 0.5, 0.05, 0.005, 0.0005\}. \quad (8.12)$$

The observable of most interest here was $G^{(1)}(0, t)$ as in the single-mode case. Due to the symmetry of the system, all single-mode observables such as this have identical expectation values for both modes, so either mode can be considered.

The procedure was the same here as outlined in Section 8.2.2, with $\mathcal{S} = 10^4$ trajectories per simulation.

In Figure 8.7 actual simulation times are compared to physical timescales and expected simulation times based on the single-mode analysis of Chapter 7. Simulation times t_{sim} were based on the condition (7.88) for the observable $|G^{(1)}(0, t)|$.

Figure 8.8 shows data from a quick search for optimum target times t_{opt} , while some example simulations are shown in Figures 8.9, 8.10, and 8.11.

8.3.2 Features seen

- At weak coupling $\omega_{12} \ll n_0$, the simulation times with various gauge choices behave qualitatively similar to the single mode of Chapter 7, and vacuum

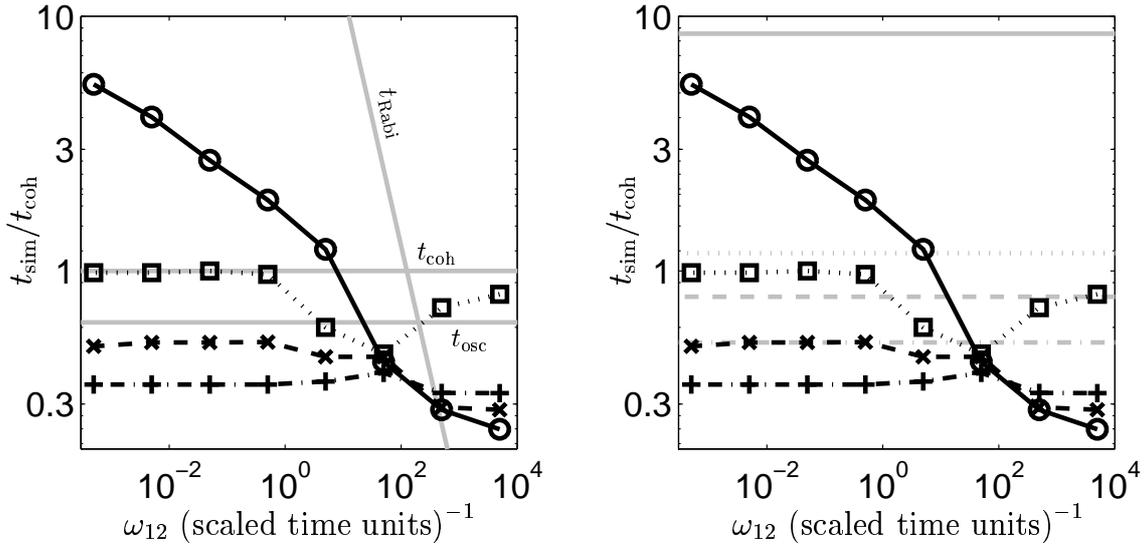


Figure 8.7: **Useful simulation times** t_{sim} for a two identical modes undergoing coherent mixing as in Section 8.3. Calculated simulation times are shown as data points, with the symbols denoting gauge used. “ \square ”: positive P; “ \circ ”: drift gauge (7.105) and diffusion gauge (8.5) with $t_{\text{opt}} = 0$; “+” drift gauge (7.99) of Carusotto *et al*[1]; “ \times ”: drift gauge (7.105) only. Subplot (a) compares to physical time scales, including Rabi oscillation period $t_{\text{Rabi}} = 2\pi/\omega_{12}$, while subplot (b) compares to expected simulation times for a single mode using the empirical fits of Table 7.2 and (8.9)–(8.10). The expected t_{sim} are plotted as light lines: DOTTED: positive P; SOLID: drift gauge (7.105) and diffusion gauge (8.5) with $t_{\text{opt}} = 0$; DASHED: with drift gauge (7.105) only; DASH-DOTTED: drift gauge (7.99) of Carusotto *et al*[1].

coupled modes of Section 8.2. That is, the dual gauge combination (7.105) and (8.5) allows simulation until the system has decohered, and give much longer simulation times than the positive P, drift gauge (7.105) only, or Carusotto *et al* gauge (7.99) (in that order). The drift-gauge-only methods simulate for less time than a single phase oscillation period.

- Simulation times are shorter for all methods than those given in Table 7.1 for the relevant single-mode case.
- At strong coupling, $\omega_{12} \gg n_0$, the plain positive P method appears to give the best results out of the gauges tried. The times achievable with the gauged methods are much shorter. Relatively longest with the Carusotto *et al* gauge

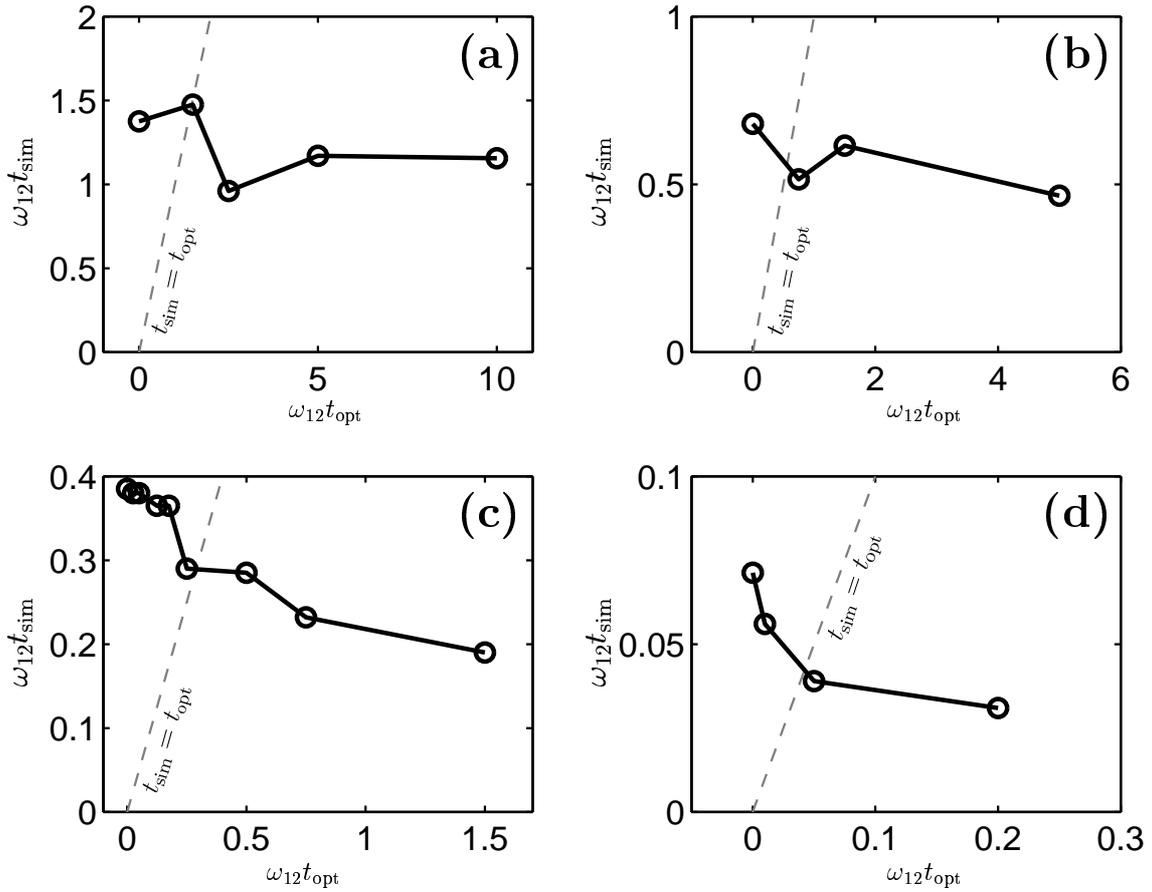


Figure 8.8: Mixing of two identical modes: Dependence of **useful simulation time** t_{sim} on the target time parameter t_{opt} for simulations of Section 8.3. Subplots (a) to (d) are for coupling strengths $\omega_{12} = \{0.0005, 0.05, 5, 500\}$, respectively. Data are from single simulations with $\mathcal{S} = 10^4$.

(7.99), then the drift gauge (7.105), and worst performance is obtained from the diffusion-gauged simulation.

- At strong coupling, none of the methods tried reach coherence time t_{coh} , but many Rabi oscillations can be simulated by all methods, since these are now much shorter than t_{osc} .
- Choosing various target times t_{opt} appears not to have no effect apart from a worsening of the simulation at large t_{opt} values. Note that this was also the situation for the vacuum-coupled simulations in Section 8.2 with the same total particle number $\bar{N} = 200$, and so this lack of improvement with t_{opt} may be typical of lower mode occupations when any significant coupling ω_{12}

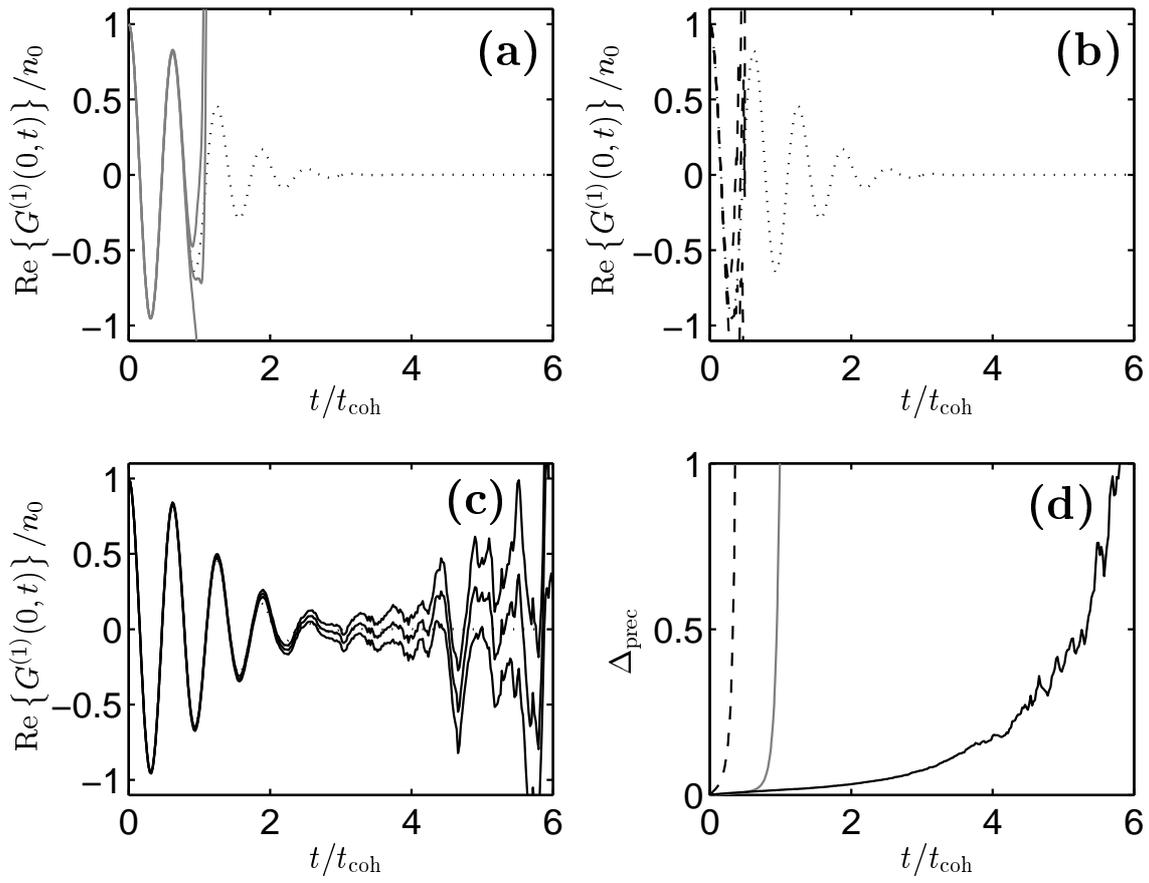


Figure 8.9: Mixing of two identical modes: **Weak coupling:** ($\omega_{12} = 0.0005$). Subplots (a)–(c) show, respectively, simulations with positive P (LIGHT LINE); with Carusotto *et al*[1] drift gauge (7.99) (DASHED LINE); and with both drift gauge (7.105) and diffusion gauge (8.5) (SOLID LINE). The exact solution (7.11b) for no coupling ($\omega_{12} = 0$) is shown as a DOTTED LINE. Subplot (d) shows relative sampling error in the estimate of $|G^{(1)}(0,t)|$, such that $\Delta_{\text{prec}} = 10\sqrt{\mathcal{S}/10^6}\Delta|G^{(1)}(0,t)|/|G^{(1)}(0,t)|$ is less than unity while “useful” precision is present, when defined as in (7.88).

is present.

- At strong coupling, the Carusotto *et al* gauge (7.99) allowed somewhat longer simulation times, while the signal to noise ratio in the estimate of $G^{(1)}(0,t)$ was lower with the gauge (7.105) at short and intermediate times — see Figure 8.11(d).

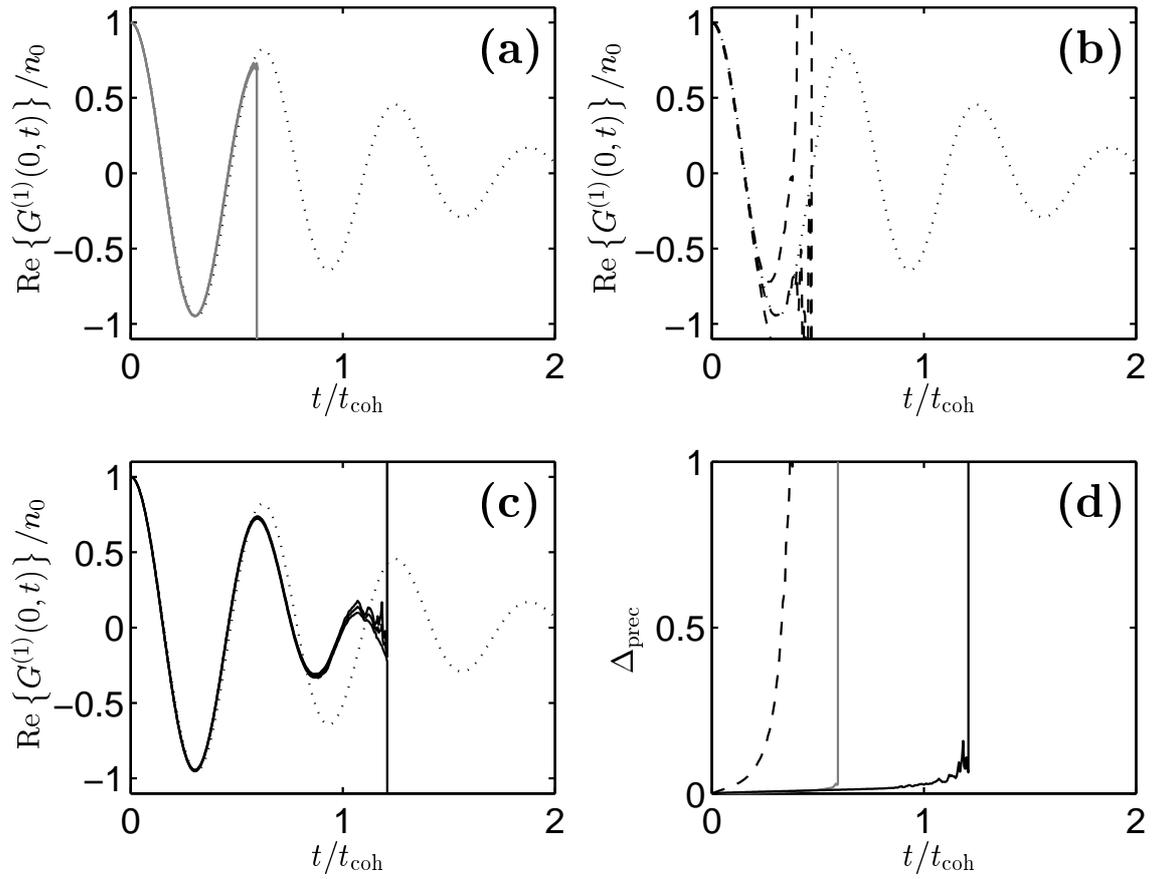


Figure 8.10: Mixing of two identical modes: **Intermediate coupling:** ($\omega_{12} = 5$). Subplots (a)–(c) show, respectively, simulations with positive P (LIGHT LINE); with Carusotto *et al*[1] drift gauge (7.99) (DASHED LINE); and with both drift gauge (7.105) and diffusion gauge (8.5) (SOLID LINE). The exact solution (7.11b) for no coupling ($\omega_{12} = 0$) is shown as a DOTTED LINE. Subplot (d) shows relative sampling error in the estimate of $|G^{(1)}(0,t)|$, such that $\Delta_{\text{prec}} = 10\sqrt{\mathcal{S}/10^6}\Delta|G^{(1)}(0,t)|/|G^{(1)}(0,t)|$ is less than unity while “useful” precision is present, when defined as in (7.88).

8.4 Analysis and Conclusions

The two kinds of initial conditions considered above represent the two kinds of situations generically occurring in all many-mode simulations. Coupling between modes of widely differing occupation will lead to behavior similar to that seen in Section 8.2 (Case 1). Coupling between modes of similar occupation will display features similar to those seen in Section 8.3 (Case 2). Adjacent spatial or momentum modes typically behave like Case 2, since if a field model is well resolved by the

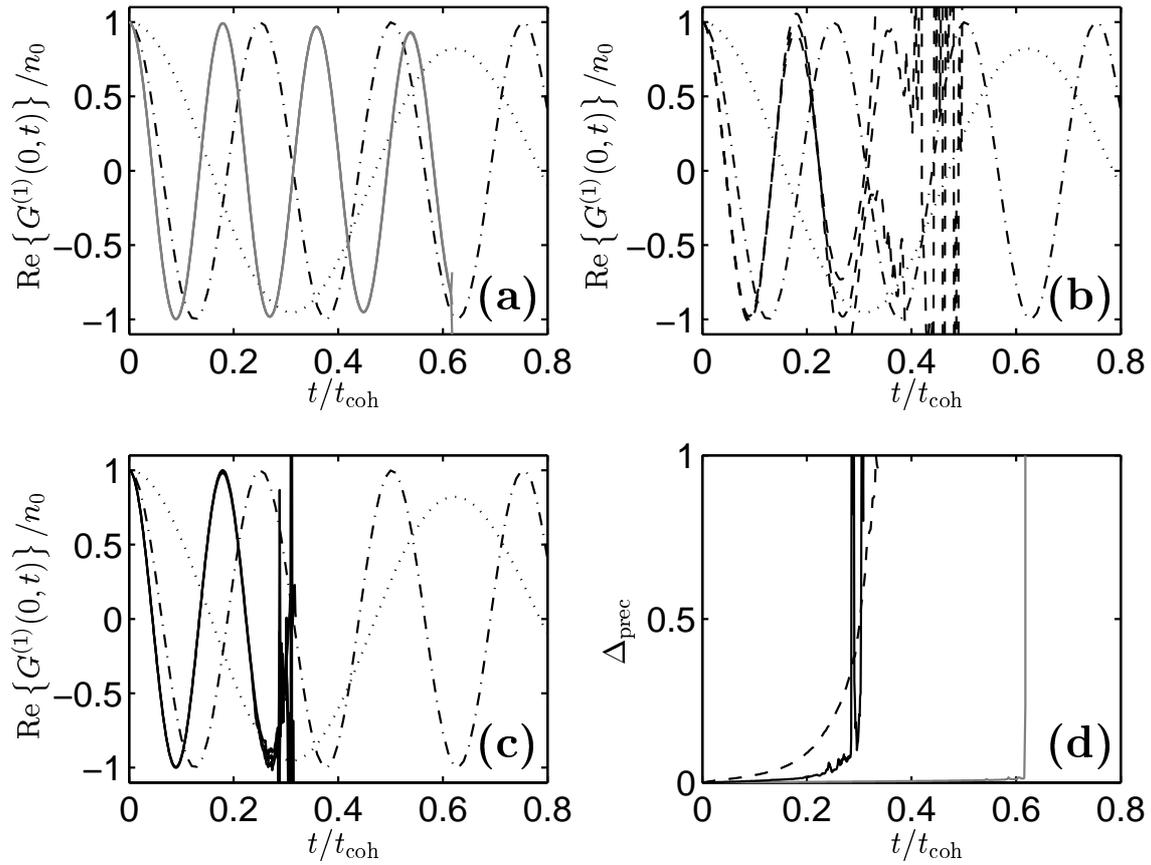


Figure 8.11: Mixing of two identical modes: **Strong coupling:** ($\omega_{12} = 500$). Subplots (a)–(c) show, respectively, simulations with positive P (LIGHT LINE); with Carusotto *et al*[1] drift gauge (7.99) (DASHED LINE); and with both drift gauge (7.105) and diffusion gauge (8.5) (SOLID LINE). The exact solution for (7.11b) no coupling ($\omega_{12} = 0$) is shown as a DOTTED LINE, and for no collisions ($\omega_{12} \rightarrow \infty$) as a DOT-DASHED LINE. Subplot (d) shows relative sampling error in the estimate of $|G^{(1)}(0, t)|$, such that $\Delta_{\text{prec}} = 10\sqrt{S/10^6}\Delta|G^{(1)}(0, t)|/|G^{(1)}(0, t)|$ is less than unity while “useful” precision is present, when defined as in (7.88).

lattice, then physical properties (e.g. density, and hence mode occupation) should not change much over the distance between neighboring lattice points.

Still, this chapter has not really been a comprehensive assessment of gauge performance for all coupled mode cases, since not all parameter regimes have been explored. For the most general two modes starting in (off-diagonal) coherent state initial conditions one could be looking at the 5 complex parameters \check{n}_j , α_j , ω_{12} — no mean feat to investigate. Rather, this chapter has been a study of several commonly occurring cases to get an indication of whether the simulation improvements found

for a single mode survive when inter-mode coupling is present, and to what degree.

Several aspects of the situation have emerged:

1. When scattering within a mode dominates over the coupling between modes, the response of the system to the gauges is similar to what was seen for a single mode. That is,
 - (a) When mode occupation is high $\check{n}_j \gg 1$, use of the combined drift (7.105) and diffusion (8.5) gauges developed in Chapter 7 gives dramatic improvement in useful simulation times and observable precision over the positive P method. In this regime, the drift-gauged only (e.g. (7.105) or (7.99)) simulations reduce simulation time from even the positive P case, although boundary terms are removed. Using both gauges (7.105) and (8.5) allows simulation to times by which full decoherence has occurred.
 - (b) When mode occupation is low $n_0 \lesssim \mathcal{O}(1)$, the improvement gained with the gauges is smaller although some improvement is always seen. The absolute simulation time is longer than for $n_0 \gg 1$, however.

Still, it turns out that in this strong scattering regime, one doesn't need much inter-mode coupling ω_{12} to reduce the simulation time in absolute terms by a factor $\mathcal{O}(2 - 10)$, although this reduction mainly affects the time after full decoherence has occurred, and not much is physically happening.

2. The beneficial effect of choosing $t_{\text{opt}} > 0$ appears to be strongly suppressed when the coupling ω_{12} is weak but not yet *very* weak, or the mode occupation is large but not yet *very* large. The $t_{\text{opt}} = 0$ diffusion gauge

$$g_j''(t_{\text{opt}} = 0) = \frac{1}{4} \log(1 + 4\text{Im}\{\check{n}_j\}^2), \quad (8.13)$$

does however have a marked beneficial effect whenever two body scattering is significant, and simulations using this gauge show very significant improvement over positive P or $g_j'' = 0$ simulations for a wide range of parameters. This parameter range appears to be $\omega_{12} \ll \check{n}_\chi$, and $\mathcal{O}(1) \lesssim n_0 \lesssim \mathcal{O}(10^3)$. Simulation times obtained are smaller by about a factor of $\mathcal{O}(2)$ than those

given for a single mode with the same gauge. At higher occupations, the benefit gained with (8.13) abates but nonzero t_{opt} values appear to become useful again, and continue to provide strong improvements over positive P simulations. (See, e.g. Figure 8.1(d).) The gauge (8.13) is convenient also because there is no *a priori* parameter t_{opt} .

3. When the inter-mode coupling dominates, the gauges developed in Chapter 7 or by Carusotto *et al*[1] do not appear to be very useful. They actually reduce simulation time as compared to the plain positive P method, although several Rabi oscillation periods can always be simulated.

The transition between the strong and weak coupling behavior appears to be at around

$$\omega_{12} \approx \chi \check{n}, \quad (8.14)$$

which is the point at which the expectation values of the coupling and two-body scattering energies in the Hamiltonian are approximately equal.

At the level of the stochastic equations (8.2), the evolution of (say) $d\alpha_j$ can gather noise from three sources

1. Directly from the local noise term $\propto \alpha_j \sqrt{\chi} dW_j$
2. Indirectly from the local nonlinear term $\propto \chi \alpha_j \check{n}_j$, which can amplify variation in the local noise term.
3. From the other mode through the coupling term $\propto \omega_{12} \alpha_{-j}$.

The drift gauges (7.105) developed here or (7.99) developed by Carusotto *et al*[1] neutralize source 2. The diffusion gauges (8.5) or (7.97d) suppresses the direct noise source 1. No local gauge can suppress the third source of randomness, however, because the randomness in α_{-j} is largely independent of any processes occurring in mode j . What happens is that even small randomness in one mode feeds into the other, can become amplified, and fed back again.

The part of the diffusion gauge that is retained at $t_{\text{opt}} = 0$ appears more robust to this mixing than the additional optimization obtained with $t_{\text{opt}} > 0$.

Based on this, and the behavior t_{sim} vs. t_{opt} shown e.g. in Figures 7.4 and 8.6, a way to proceed with a full many-mode simulation would be to

1. Run a simulation with local drift gauge (7.105), and $t_{\text{opt}} = 0$ diffusion gauge (8.13). Call the simulation time obtained $t_{\text{sim}}^{(0)}$.
2. Run a simulation with $t_{\text{opt}} = t_{\text{sim}}^{(0)}$, and obtain a simulation time $t_{\text{sim}}^{(1)}$.
3. If $t_{\text{sim}}^{(1)} > t_{\text{sim}}^{(0)}$ then nonzero t_{opt} is beneficial, one expects $t_{\text{sim}} \gtrsim t_{\text{opt}}$, and one can make a search for a reasonable value by now running a simulation with $t_{\text{opt}} = t_{\text{sim}}^{(1)}$, and iterating in this manner until no more significant improvement is gained.

If, on the other hand, $t_{\text{sim}}^{(1)} \lesssim t_{\text{sim}}^{(0)}$, then it appears that one is in a regime where the t_{opt} parameter does not affect the dominant noise source, and $t_{\text{sim}}^{(0)}$ is the best simulation time one can obtain with the (7.105)&(8.5) dual gauge method.

4. If simulation time is too short, it may be worth trying a plain positive P simulation, or even a diffusion-gauge-only simulation using (7.97d), while carefully monitoring for the presence of spiking, and discarding any data beyond the first spike time. (The simulations carried out in this chapter with this approach did not show any bias when compared with exact results — see e.g. Figures 8.2 and 8.3).

In conclusion, it has been found in this chapter that simulation time and precision improvement can still be gained with coupled modes by using the local drift and diffusion gauges (7.105) and (8.5), although marked improvement will mostly be seen only when two-particle scattering dominates the inter-mode coupling. Other non-local gauges may improve simulations in a regime where inter-mode coupling dominates but development of such is for future research.

Chapter 9

Single-mode interacting Bose gas thermodynamics

In Section 5.6 it was shown that the gauge P representation can also be used to calculate properties of the grand canonical ensemble of an interacting Bose gas described by the lattice Hamiltonian (2.12). A convenient feature of calculations made with this method (rather than, say path integral Monte Carlo, or variational Monte Carlo methods) is that a single simulation can estimate any observables, and gives results for a wide range of temperatures $T \geq 1/k_B \max[\tau]$.

Before attempting full many-mode simulations (see Chapter 11), it is pertinent to make sure that the simplest special case of a single-mode is correctly simulated. This is a good test case because it can easily be solved exactly for a unambiguous comparison, and because it is simple enough that a broad analysis of the statistical behavior of the simulation is possible.

The single-mode model is particularly relevant to a locally-interacting many-mode Hamiltonian (2.17). Here the stochastic equations to simulate are given by (5.50), and it can be seen that the nonlinearity and noise in the evolution of local amplitudes $\alpha_{\mathbf{n}}$ for each mode \mathbf{n} depend only on the local variables $\alpha_{\mathbf{n}}$ and $\beta_{\mathbf{n}}$. In such a situation the nonlinear and stochastic features of the many-mode behavior should appear in their entirety in the single-mode toy model.

9.1 The single-mode model

9.1.1 Quantum description

Isolating a single mode from the description of Section 2.6, the master equation for the evolution of the (un-normalized) density matrix is

$$\frac{\partial \hat{\rho}_u}{\partial \tau} = \left[\mu_e(\tau) \hat{n} - \hat{H} \right] \hat{\rho}_u \quad (9.1)$$

with $\hat{n} = \hat{a}^\dagger \hat{a}$ (the number operator) defined in the usual way in terms of boson creation and annihilation operators. The “imaginary time” parameter is $\tau = 1/k_B T$, and the “effective chemical potential” μ_e is defined in (2.30). That expression can be integrated to give

$$\mu(\tau) = \frac{1}{\tau} \left[-\lambda_n + \int_0^\tau \mu_e(t) dt \right] \quad (9.2)$$

with constant λ_n , and it is then seen that at low temperatures $\mu(T) \rightarrow \mu_e(T)$. The single-mode analogue of the interacting Bose gas Hamiltonian is

$$\hat{H} = \hbar \chi \hat{n} (\hat{n} - 1), \quad (9.3)$$

where any linear contribution $\propto \hat{n}$ has been amalgamated into the chemical potential. The initial state is (as per (2.31))

$$\hat{\rho}_u(0) = \exp[-\lambda_n \hat{n}]. \quad (9.4)$$

Physically, this single-mode model is an approximation of e.g. a boson orbital in a heat bath at temperature T , and chemical potential μ , where this chemical potential includes any kinetic and external potential effects.

9.1.2 Gauge P stochastic equations

Defining

$$\check{n} = \alpha\beta = n' + in'' \quad (9.5)$$

as in previous chapters, the Ito stochastic equations in a gauge P representation are obtained from (5.50):

$$d\alpha = (\mu_e - 2\hbar\chi\check{n})\alpha d\tau + i\alpha\sqrt{2\hbar\chi}(dW - \mathcal{G}d\tau) \quad (9.6a)$$

$$d\beta = 0 \quad (9.6b)$$

$$d\Omega = \Omega[(\mu_e - \check{n})n d\tau + \mathcal{G}dW]. \quad (9.6c)$$

The Wiener increments dW can be implemented as independent real Gaussian noises of mean zero, and variance $d\tau$ at each time step of the simulation. The single complex drift gauge is \mathcal{G} .

Because only a single non-weight variable α experiences diffusion, then the ungauged diffusion matrix is $\underline{D} = -2\hbar\chi\alpha^2$, a single complex function. There are then no standard diffusion gauges (as described in Section 4.4.2) because the orthogonal 1×1 matrix (4.68) is just $O = 1$, apart from a phase factor.

Initial conditions on the variables α and β are (from (5.59))

$$P_G(\alpha, \beta, \Omega) = \delta^2(\Omega - 1)\delta^2(\beta - \alpha^*)\frac{1}{\pi\bar{n}_0}\exp\left(\frac{-|\alpha|^2}{\bar{n}_0}\right), \quad (9.7)$$

where

$$\bar{n}_0 = \frac{1}{e^{\lambda_n} - 1}, \quad (9.8)$$

is the mean particle number at high temperature. α is easily sampled using Gaussian random variables.

The mean number of particles is $\langle\hat{n}\rangle = \text{Tr}[\hat{\rho}_u\hat{n}]/\text{Tr}[\hat{\rho}_u]$, and is estimated by

$$\bar{n} = \frac{\langle\text{Re}\{\check{n}\Omega\}\rangle_{\text{stoch}}}{\langle\text{Re}\{\Omega\}\rangle_{\text{stoch}}}. \quad (9.9)$$

These stochastic equations exhibit completely different behavior to the dynamics of this system (7.3). Rather they are more similar to the two-boson absorption of Section 6.4.

9.1.3 Exact solution

The quantum evolution of this model is easily evaluated exactly in a Fock number state basis $|n\rangle$. This is useful to make a definitive check of the correctness of the

stochastic simulations. Using $\widehat{\rho}_u = \exp[\mu\tau\widehat{n} - \widehat{H}\tau]$, the density matrix elements in this basis $\rho_{n\tilde{n}} = \langle n | \widehat{\rho}_u | \tilde{n} \rangle$ are found to be:

$$\rho_{n\tilde{n}} = \delta_{n\tilde{n}} \exp\{[\mu - \hbar\chi(n-1)]n\tau\}. \quad (9.10)$$

A long times $\tau \rightarrow \infty$, i.e. low temperatures, the density matrix will be dominated by the populations for which the exponent is largest. This maximum of the exponent occurs when $n_\infty = [\lim_{\tau \rightarrow \infty} \mu(\tau) + \hbar\chi]/2\hbar\chi$. Since n takes on only integer values, then the dominant mode occupation will be at the nearest integer to n_∞ (or the two nearest components if n_∞ takes on half-integer values). Using (9.2),

$$n_\infty = \frac{1}{2} + \frac{1}{2\hbar\chi} \lim_{\tau \rightarrow \infty} \left[\frac{1}{\tau} \int_0^\tau \mu_e(t) dt \right], \quad (9.11)$$

and the $T \rightarrow 0$ state $\widehat{\rho}(\tau \rightarrow \infty)$ is:

- A vacuum if $n_\infty < 0.5$.
- Otherwise, if n_∞ is a half integer, then $\widehat{\rho}(\tau \rightarrow \infty)$ is an equal mixture of the $|n_\infty \pm \frac{1}{2}\rangle$ states.
- Otherwise, $\widehat{\rho}(\tau \rightarrow \infty)$ is the Fock number state with the nearest whole occupation number to n_∞ .

9.2 Moving singularities and removal with gauges

9.2.1 Moving singularity

Consider, for now, the case of constant effective chemical potential μ_e . From (9.6), and using (B.7), the Stratonovich stochastic equation for the complex occupation variable \check{n} is

$$d\check{n} = \check{n} \left[\mu_e d\tau - \hbar\chi(2\check{n} - 1) d\tau + i\sqrt{2\hbar\chi} (dW - \mathcal{G} d\tau) \right]. \quad (9.12)$$

In a positive-P-like simulation¹, $\mathcal{G} = 0$, and the deterministic part of the $d\check{n}$ equation has the phase-space structure shown in Figure 9.1(a). There are stationary points

¹But with a weighting gauge to allow for complex weight evolution $d\Omega$.

at vacuum and at

$$\check{n} = n_\infty = \frac{1}{2} + \frac{\mu_e}{2\hbar\chi}, \quad (9.13)$$

with the more positive stationary point being an attractor, and the more negative a repellor. The deterministic evolution is easily solved, and with initial condition $\check{n}(0) = n_0$ gives

$$\check{n}(\tau) = \frac{n_\infty n_0}{n_0 + (n_\infty - n_0)e^{-2\hbar\chi n_\infty \tau}}. \quad (9.14)$$

A moving singularity appears along the negative real axis, and if one starts with a negative real $n_0 = -|n'_0|$, then $\check{n} \rightarrow \infty$ at time

$$\tau_{\text{sing}} = \frac{1}{2\hbar\chi n_\infty} \log \left(1 - \frac{n_\infty}{n_0} \right). \quad (9.15)$$

Since initial conditions (9.7) lead to samples with $n_0 \in (0, \infty)$, then some low- n_0 trajectories can be expected to rapidly diffuse into the negative real part of phase space where the super-exponential growth occurs. This will tend to cause either misleading systematic boundary term errors or uncontrolled spiking, which renders the simulation useless after a short time. As is seen in Figure 9.5 in this case it is spiking.

9.2.2 Minimal drift gauge

To correct the problem one has to change the phase-space topology in some way to prevent the occurrence of such moving singularities. The radial evolution $d|\check{n}| = -2\hbar\chi|\check{n}|n' d\tau + \dots$ is at fault, and the offending term can be removed with the gauge

$$\mathcal{G} = i\sqrt{2\hbar\chi} (\text{Re}\{n\} - |\check{n}|). \quad (9.16)$$

This gauge has the effect of replacing the $-2\hbar\chi|\check{n}|\text{Re}\{n\} d\tau$ term in $d|\check{n}|$ with $-2\hbar\chi|\check{n}|^2 d\tau$, which always attracts trajectories into the phase space in the vicinity of the origin. Polar phase evolution is unchanged, and the resulting deterministic phase space is shown in Figure 9.1(c). If $n_\infty > 0$ the attractor and repellor remain, while for $n_\infty \leq 0$, there is just an attractor at vacuum. For near-classical trajectories having real \check{n} , phase-space evolution is unchanged, and the gauge is zero.

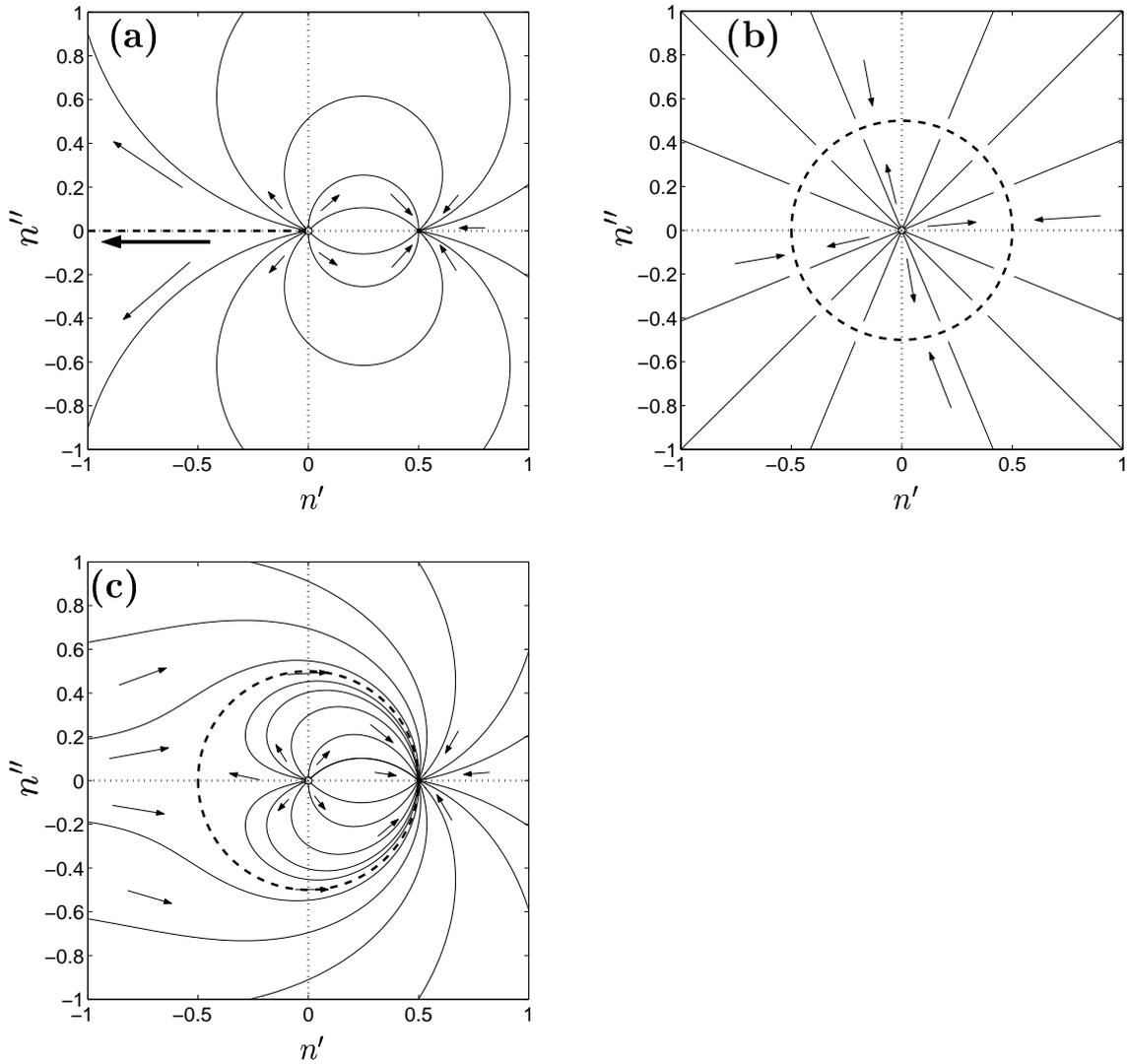


Figure 9.1: **Deterministic phase space** for Stratonovich form of the $d\check{n}$ equation. The case of $n_\infty = 0.5$ is shown. (a): positive-P-like simulation $\mathcal{G} = 0$; (b): radial gauge (9.17); (c): minimal gauge (9.16). The moving singularity in (a) is shown with large arrow, the attractors at $|\check{n}''| = n_\infty$ with a thick dashed line.

How do the modified equations measure up to the gauge choice criteria of Section 6.3.2? In order:

1. The moving singularity in $d\check{n}$ has been removed (at high $|\check{n}''|$ the deterministic behavior is restorative towards the origin), and no other moving singularities are present. Note that if \check{n} remains finite, then so does Ω , which just experiences exponential growth with an exponent that is always finite, since it depends only on \check{n} .

2. No new moving singularities have been introduced.
3. No noise divergences (new or old) are present since all noise terms satisfy (6.3).
4. No discontinuities in the drift or diffusion coefficients are present provided the time-dependence of $\mu_e(\tau)$ is not singular.
5. Is the weight spread minimized?
 - (a) Gauge is zero when \check{n} is positive real, and small in its neighborhood, where gauge corrections are unnecessary.
 - (b) Explicit variational minimization of \mathcal{G} has not been carried out, but rather criteria 6 have been applied.
 - (c) Gauge is zero at deterministic attractor $\check{n} = n_\infty$.
6. No unwanted gauge behaviors are present:
 - (a) Gauge is nonzero in a large part of phase space, so most trajectories contribute to removal of bias.
 - (b) Gauge changes smoothly in phase space.
 - (c) Gauge is autonomous.
 - (d) Gauge breaks the analyticity of the equations.

9.2.3 Radial drift gauge

The gauge (9.16) looks good, but it was found that a more severe phase-space modification usually gives smaller statistical uncertainties. This better gauge form is

$$\mathcal{G} = i\sqrt{2\hbar\chi}(\check{n} - |\check{n}|), \quad (9.17)$$

which will be called here the “radial” gauge, due to elimination of any deterministic polar-angle evolution as shown in Figure 9.1(b). The deterministic attractor is the entire $|\check{n}| = n_\infty$ circle (or the vacuum if $n_\infty \leq 0$), with some phase diffusion in the phase of \check{n} .

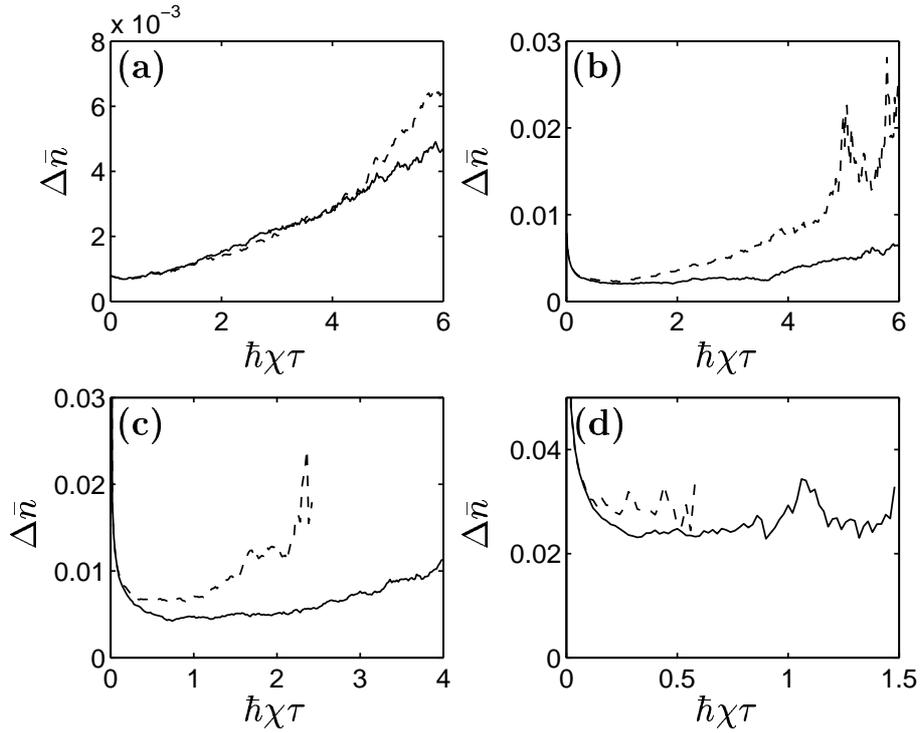


Figure 9.2: **Uncertainty in \bar{n}** , the estimate of $\langle \hat{n} \rangle$ for the $\mu_e = 0$ system. (a): $\bar{n}_0 = 0.1$; (b): $\bar{n}_0 = 1$; (c): $\bar{n}_0 = 10$; (d): $\bar{n}_0 = 100$; SOLID LINE: radial gauge (9.17); DASHED LINE: gauge (9.16). Simulations were carried out with $\mathcal{S} = 2 \times 10^4$ trajectories.

The efficiency of the two gauges is compared in Figure 9.2 for $\mu_e = 0$ and a variety of initial \bar{n}_0 . It is not immediately clear why the radial gauge is better, but the improvement is significant whenever $\bar{n}_0 \geq \mathcal{O}(1)$. Similar behavior was seen for other μ_e values. In particular also for $\mu_e = \hbar\chi$, which corresponds to a different low temperature state (i.e. a pure Fock number state — see Figure 9.3).

Comparison to the heuristic gauge choice criteria of Section 6.3.2 follows through identically as for the previous gauge (9.16), apart from the gauge not being zero on the entire $|\check{n}\rangle = n_\infty$ attractor but only at $\check{n} = n_\infty$.

9.3 Numerical simulations

The results of some example simulations using the radial gauge (9.17) are shown in Figures 9.3 and 9.4. These are compared to exact results obtained from (9.10) and also to mean field semiclassical calculations. As outlined in Section 5.4, mean

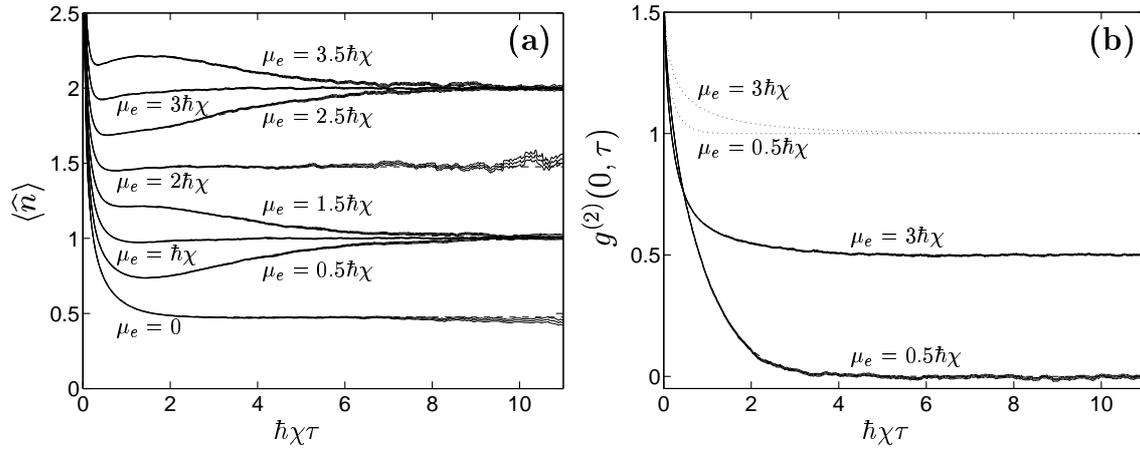


Figure 9.3: **Temperature dependent observables** given constant μ_e and $\bar{n}_0 = 10$ at high temperature. SOLID LINES: simulation with $\mathcal{S} = 2 \times 10^5$ trajectories. Triple lines indicate error bars (often not resolved at this scale). DASHED LINES: exact values (mostly obscured by simulation results) DOTTED LINES: mean field calculation (subplot (b)). The second order correlation function $g^{(2)}(0, \tau)$ is given by (8.8).

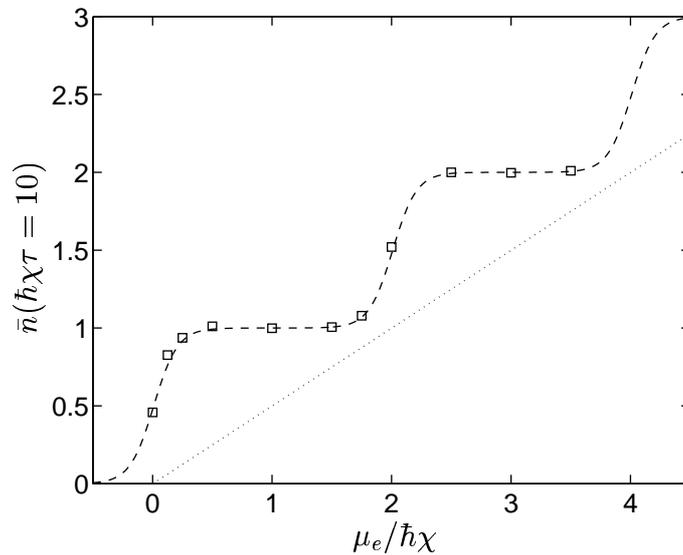


Figure 9.4: **Mean mode occupations at low temperature** $T = \hbar\chi/10k_B$ as a function of constant μ_e . DASHED: exact result — note the quantization, DOTTED: mean field result, SQUARES: simulation results with radial gauge (9.17). Uncertainty was of symbol size or smaller.

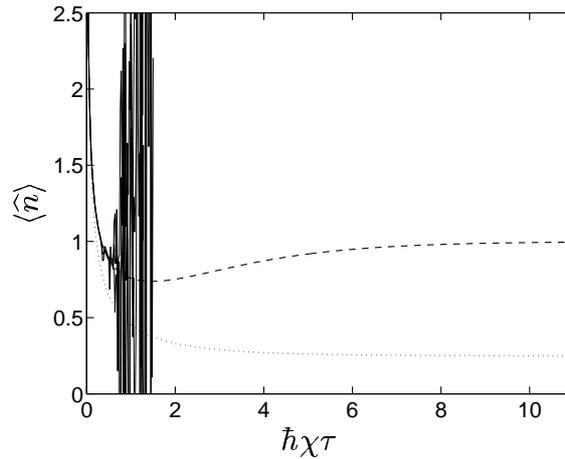


Figure 9.5: Performance of the mean field and un-gauged simulations for comparison with Figure 9.3(a). **Mean mode occupation** with $\mu_e = 0.5\hbar\chi$ and $\bar{n}_0 = 10$. DASHED: exact result, DOTTED: mean field result, SOLID: un-gauged simulation.

field calculations such as solution of Gross-Pitaevskii equations are equivalent to simulation of only the deterministic part of the Ito gauge P equations.

Additionally, a positive-P-like simulation with no gauge ($\mathcal{G} = 0$) is shown in Figure 9.5 for comparison with the gauged technique.

A subtlety to keep in mind in these thermodynamics simulations is that (in contrast to dynamics) the normalization $\langle \text{Re} \{ \Omega \} \rangle$, which appears in the denominator of observable estimates (3.14), is not unity in the $\mathcal{S} \rightarrow \infty$ limit. This then requires a sufficient number of trajectories per subensemble so that the denominator is always positive and not too close to zero for any of these, as discussed in detail in Appendix C.

Features seen in these figures include:

- Convergence of the gauged simulation is excellent, and precisely reproduces the exact quantum behavior.
- At low temperatures $\tau \gtrsim 1/\hbar\chi$ the semiclassical approximation gives completely wrong results for $g^{(2)}$ and is out by $\mathcal{O}(1)$ in $\langle \hat{n} \rangle$, which is also completely wrong if mode occupation is $\leq \mathcal{O}(1)$.
- The gauge P method, on the other hand, reproduces the exact quantum behavior precisely despite using a semiclassical coherent-state basis.

- The un-gauged simulation breaks down while still at high temperatures due to boundary term induced spiking.
- Low temperature observable estimates are consistent with the Fock state low temperature ground state described in Section 9.1.3. ($\langle \hat{n} \rangle$ approaches integer values² and the two-particle correlation $g^{(2)}$ approaches $1 - 1/\langle \hat{n} \rangle$.)
- Mid-temperature behavior is also simulated precisely.
- At low temperature (long times) uncertainty in the observable estimates is greatest for μ_e even integer multiples of $\hbar\chi$ (and hence even integer multiple chemical potential μ at low temperatures), and lowest for μ_e odd integer multiples of $\hbar\chi$. These cases correspond to the low temperature state being a mixture of two Fock number states, or just a single pure Fock state, respectively.

9.4 Chemical potential as a free gauge parameter

Suppose one is interested in obtaining properties at a given temperature T and chemical potential μ . This determines the final simulation target time τ_T and target chemical potential $\mu(\tau_T)$, but the chemical potential at intermediate times $\tau < \tau_T$ is not specified. The only conditions are 1) that the quantity $\mu(\tau)\tau$ be (piecewise) differentiable so that $\mu_e(\tau)$ can be calculated, and 2) that $\lambda_n = -\lim_{\tau \rightarrow 0}(\mu\tau)$ be positive so that the initial distribution (9.7) is a normalizable probability. Apart from these, $\mu(\tau)$ is formally an arbitrary function (and λ_n an arbitrary positive number).

This is all in the limit of many trajectories, however. The form of $\mu(\tau)$ at intermediate τ has no effect on observable estimate means in the limit $\mathcal{S} \rightarrow \infty$, but can have a strong effect on the broadness of the trajectory distribution. And, hence — on the precision of a finite sample estimate. In this sense then, $\mu_e(\tau)$ (given by (2.30)), acts as effectively an additional gauge function.

²Apart from when μ_e is an even integer multiple of $\hbar\chi$, since then $\langle \hat{n} \rangle$ approaches half-integer values.

An indication of the characteristics of an efficient form of $\mu_e(\tau)$ can be gained by comparing calculated observable uncertainties for several basic cases. Let us look at the situation when $\tau_T = 1/\hbar\chi$, and $\mu(\tau_T) = \hbar\chi$. (This gives mean occupation $\bar{n}(\tau_T) = 0.5628$).

- **Case 1:** Variation in $\bar{n}_0(\lambda_n)$, i.e. in the high-temperature starting mode occupation, while $\mu_e(\tau)$ is a constant chosen appropriately using (9.2). Calculated uncertainties in $\langle \hat{n} \rangle$ and $g^{(2)}(0, \tau)$ are compared in Figure 9.6. It is seen that initial occupation $\mathcal{O}(1) - \mathcal{O}(10)$ gives the best performance, while too large or too small initial occupation lead to excessive uncertainty in the observables.
- **Case 2a:** Time-varying μ_e , while $\bar{n}_0 = 1$ is held constant. μ_e is chosen to be a nonzero constant up to a time τ_c , and zero for times $\tau > \tau_c \leq \tau_T$. Note that as $\tau_c \rightarrow 0$, the size of μ_e grows to make up for the shorter time over which it acts. Uncertainties are compared in Figure 9.7(a). It is seen that if μ_e is varied too strongly, an inefficient simulation results, although some relatively small variation may be beneficial for the $g^{(2)}$ calculation.
- **Case 2b:** Again time-varying μ_e , while $\bar{n}_0 = 1$ is held constant, but this time μ_e is held zero at high temperatures (low times $\tau < \tau_c$), and makes up for this by becoming a nonzero constant for $\tau_c < \tau < \tau_T$. As $\tau_c \rightarrow \tau_T$, the size of the required μ_e grows. Uncertainties are compared in Figure 9.7(b). Again, it is seen that strong variation in μ_e leads to an inefficient simulation.

From the simulation data it appears that, other things being equal,

1. $\bar{n}_0 \approx \mathcal{O}(\bar{n}(\tau_T))$ or slightly greater appears to give the best performance
2. A constant (or nearly so) $\mu_e(\tau)$, which keeps the phase-space behavior time independent appears to give better performance than simulations for which μ_e strongly varies with τ .

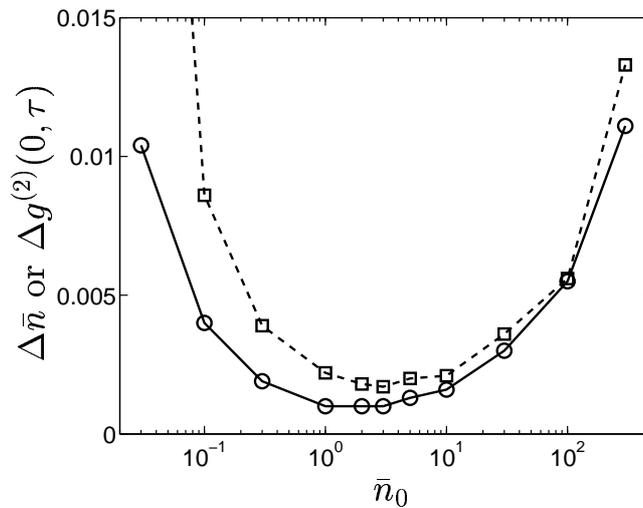


Figure 9.6: **Dependence of observable estimate uncertainties on the starting occupation \bar{n}_0 .** Shown are calculated uncertainties in CIRCLES: $\bar{n} = \langle \hat{n} \rangle$ SQUARES: $g^{(2)}(0, \tau)$, for $\hbar\chi\tau = 1$ and $\mu = \hbar\chi$. Calculations were made with the radial gauge (9.17) varying the starting occupation \bar{n}_0 and choosing a constant μ_e according to (9.2). Each calculation was with 2×10^4 trajectories.

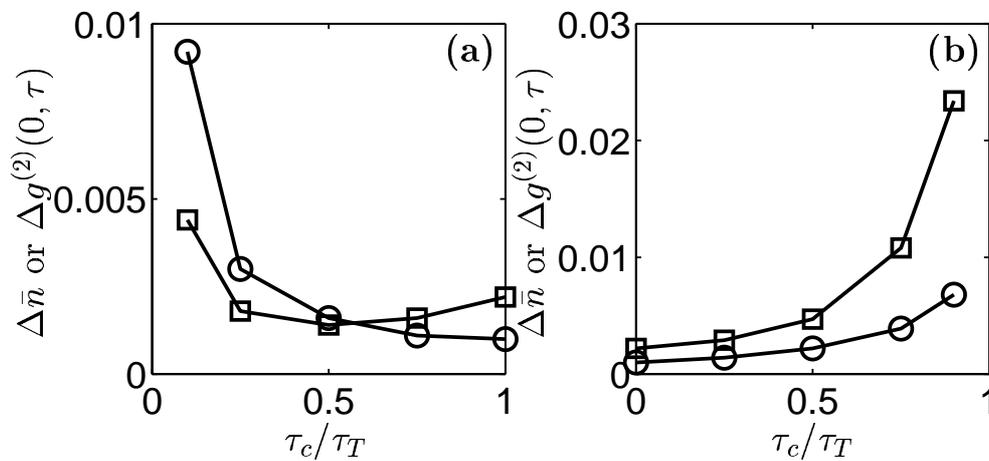


Figure 9.7: **Dependence of observable estimate uncertainties on the form of μ_e .** Shown are calculated uncertainties in CIRCLES: $\bar{n} = \langle \hat{n} \rangle$ SQUARES: $g^{(2)}(0, \tau)$, for $\hbar\chi\tau = 1$ and $\mu = \hbar\chi$. Calculations were made with the radial gauge (9.17), starting with $\bar{n}_0 = 1$ initially, but varying the form of μ_e . In (a), this was taken to be a nonzero constant for $\tau \leq \tau_c$ and zero for $\tau > \tau_c$. In (b), μ_e was zero for $\tau < \tau_c$, and nonzero constant for $\tau \geq \tau_c$. Each calculation was with 2×10^4 trajectories.

9.5 Summary

The drift gauges (9.16) and (9.17) have been developed to overcome shortcomings of the un-gauged positive-P-like technique. Using these gauges it has been shown that a stochastic simulation can precisely simulate the full quantum grand canonical thermodynamics of a two-body interacting Bose gas mode or orbital. Features include

1. Simulation with good precision down to the lowest temperatures $\hbar\chi\tau \gg 1$ where discretization of mode occupation takes place (See e.g. Figures 9.3 and 9.4).
2. Simulation of a range of temperatures down to a minimum $1/k_B\tau_T$ in one run.
3. Calculation of any desired set of observables from one run.

The uncontrollable spiking (which was present in an un-gauged simulation) is removed.

Numerical simulations indicate that the radial gauge (9.17) gives superior efficiency for a wide range of parameters. Also, it has been found that if one is aiming for a particular target temperature and chemical potential, it appears advantageous to choose the intermediate-time chemical potentials in the simulation such that: 1) μ_e is slowly varying with τ (or constant), and 2) The starting occupation $\bar{n}_0(\lambda_n)$ is of the same order as the mean occupation at the target time.

The gauges developed should carry over in a straightforward manner to multi-mode simulations when the interparticle scattering is local to the lattice sites as described by the Hamiltonian (2.17). This is implemented in Chapter 11.

Part C

Examples of many-mode simulations

Chapter 10

Many-mode dynamics

Part C of this thesis investigates the performance of the gauge P representation (including the special case of the positive P with $\mathcal{G}_n = 0$) for many-mode calculations. Specifically, the interacting Bose gas. Chapter 10 will consider simulations of many-mode dynamics, while Chapter 11 concerns calculations of grand canonical ensembles.

Regarding dynamics, it will be seen that an un-gauged simulation (equivalent to positive P) can already give many useful results, while under some conditions the diffusion gauges of Chapter 7 can be applied directly (i.e. without any additional gauge improvements for kinetic coupling) to lengthen the simulation time and improve precision.

The particular example systems simulated in this chapter do not have particle losses, which is a “worst case” in terms of simulation time and stability, as has been discussed in Section 7.2.4. This is also the case in which nonclassical phenomena are most pronounced. It is worth pointing out, however, that the gauge P representation method allows the addition of damping and/or losses in a completely straightforward way by adding some linear terms of the forms (5.18), or (5.21) into the stochastic equations. Chapter 11, on the other hand, considers simulations where the particle number strongly varies during the simulation.

The stochastic wavefunction method of Carusotto *et al*[1] and gauge P representation methods share many common features, but the former is applicable only to

explicitly particle-conserving systems. That method has been used to calculate the evolution of the width of an atomic cloud with extended interactions in a similar breathing trap arrangement to the example calculation of Section 10.6 here.

10.1 Simulation procedure

To simulate dynamics with the gauge P representation, one starts from the equations (5.17) to (5.21). For an M -mode system, there are then $2M + 1$ complex stochastic differential equations. The Wiener stochastic increments were implemented using Gaussian noises of variance Δt , independent at each time step of length Δt .

For efficiency reasons, it is convenient to use a split-step algorithm to evolve the kinetic energy part of the terms $\omega_{\mathbf{nm}}$ in Fourier space, while doing the rest of the evolution in position space. For notational convenience let us define the dual coherent state amplitudes (of the bra vector in the kernel (5.4)) as

$$\alpha'_{\mathbf{n}} = \beta_{\mathbf{n}}^*. \quad (10.1)$$

If the label z is allowed to stand for either α or α' , then the (discrete) Fourier space variables are

$$\tilde{z}_{\mathbf{n}} = \frac{(2\pi)^{D/2}}{V} \sum_{\mathbf{n}} e^{-i\mathbf{k}_{\mathbf{n}} \cdot \mathbf{x}_{\mathbf{n}}} z_{\mathbf{n}}, \quad (10.2)$$

where V is the volume spanned by the spatial lattice. Expanding $\omega_{\mathbf{nm}}$ as per (2.13) and (2.14), the kinetic evolution in momentum space takes the form

$$d\tilde{\alpha}_{\mathbf{n}} = -\frac{i}{2} |\mathbf{k}_{\mathbf{n}}|^2 \tilde{\alpha}_{\mathbf{n}} dt + \dots \quad (10.3a)$$

$$d\tilde{\alpha}'_{\mathbf{n}} = -\frac{i}{2} |\mathbf{k}_{\mathbf{n}}|^2 \tilde{\alpha}'_{\mathbf{n}} dt + \dots \quad (10.3b)$$

after some algebra. (The remainder of the terms denoted by “...” will be evaluated in position space.) The split-step simulation algorithm is then, for each trajectory out of \mathcal{S} :

1. Initialize variables choosing randomly either from explicitly known initial conditions, or by passing variables from a trajectory in a previous calculation (e.g. a thermal grand canonical ensemble as will be calculated in Chapter 11).

2. For each time step Δt ,
 - (a) Transform variables to $\tilde{\alpha}_{\mathbf{n}}, \tilde{\alpha}'_{\mathbf{n}}$ in Fourier space $\{\tilde{\mathbf{n}}\}$.
 - (b) Evolve variables forward by a time step Δt by applying the change (10.3) due to kinetic terms.
 - (c) Accumulate any moments of variables in momentum space required for observable estimates.
 - (d) Transform variables to $\alpha_{\mathbf{n}}, \alpha'_{\mathbf{n}}$ in normal space $\{\mathbf{n}\}$.
 - (e) Evolve variables (including weight Ω) forward by a time step Δt by applying the rest of the evolution due to interparticle collisions, particle gains or losses, and external potential. This last appears as diagonal linear frequency terms, which now have the form $\omega_{\mathbf{nn}} = V_{\mathbf{n}}^{\text{ext}}/\hbar$ after kinetic processes have been moved to Fourier space.
 - (f) Accumulate any moments of variables in normal space required for observable estimates.

In the case when there is no particle gain from the environment (but losses to a zero temperature heat bath can be tolerated), the differential equations for all the variables $\alpha_{\mathbf{n}}, \alpha'_{\mathbf{n}}, \Omega$, take on an exponential and local form $dz_j \propto z_j$ (j is here a generic variable label). It is then convenient to change to logarithmic variables $\log(\alpha_{\mathbf{n}})$, $\log(\alpha'_{\mathbf{n}})$, and $z_0 = \log \Omega$. The radial and tangential evolution are then simulated separately, which is found to lead to superior numerical stability. This allows one to use larger time steps. Efficiency is improved despite the need to change back to the non-logarithmic amplitude variables $\alpha_{\mathbf{n}}, \alpha'_{\mathbf{n}}$ to make the Fourier transform.

The integration algorithm used was a semi-implicit half-step iterative method, ($\kappa = \frac{1}{2}$ in the notation of (B.10) in Appendix B). The variable increments used ($d\alpha_{\mathbf{n}}, dz_0$ etc.) are in the Stratonovich calculus. This algorithm has been found by Drummond and Mortimer[82] to have much superior stability to a plain first order method (although it is still first order in Δt).

During a simulation, several indicators were monitored to make sure no sampling biases occur. These indicators were

1. **Excessive variance in logarithmic variables.** All or most of the evolution for the interacting Bose gas occurs in an exponential fashion as per $dz_j \propto z_j$, when written in terms of variables that are averaged to obtain observable estimates. As described in Appendix A, bias may occur if the variance of $\log |z_j|$ exceeds $\mathcal{O}(10)$.
2. **Sudden spiking**, which may be an indicator of boundary term errors in the simulations with no drift gauge. This is described by Gilchrist *et al*[64]. Some care had to be taken not to confuse this spiking with the far more benign kind caused by too-small ensemble sizes \mathcal{S} , which occurs when evaluating quotients of variable moments (as described in Appendix C). A characteristic distinguishing feature between these is what happens to the time of first spiking when the ensemble size is changed.
 - Boundary term related spiking becomes more severe and/or first occurs at slightly earlier times as the number of trajectories \mathcal{S} is increased. This is because as \mathcal{S} increases, trajectories closer to the divergent moving singularity (and hence spiking sooner) become included in the ensemble.
 - Subensemble-size related spiking abates and/or first occurs at later times as the number of trajectories in an ensemble \mathcal{S} is increased, due to better precision in evaluating the denominator in the quotient.
3. **Change in mean with ensemble size.** A robust and widely applicable indicator of possible bias is when means of variables or of their functions change in a systematic manner (with statistical significance) as the ensemble size \mathcal{S} is increased.

On the whole, reasonably precise observable estimates were not brought into question by any of the above indicators, apart from some unusual situations. (An example of such a pathological simulation were thermodynamic calculations of ideal gases, where no noise occurs during the evolution. See Section 11.2.3.) Typically,

worrying symptoms were seen only at times when noise had already obscured any observable estimates.

10.2 Lattice size and simulation performance

A lattice simulation will be equivalent to the continuum system provided that 1) the lattice encompasses the entire system (i.e. the volume spanned by the lattice is large enough) and 2) all relevant features are resolved (i.e. the lattice is fine enough). Once in this limit, all simulations will converge to the same physical predictions in the many trajectories limit $\mathcal{S} \rightarrow \infty$ (provided there are no boundary term errors). However, the *rate* of convergence turns out to be affected by the size of the lattice spacing. This affects the useful simulation time.

Before considering some example simulations in Sections 10.4 to 10.7, the effect of lattice size on the simulation time is investigated.

10.2.1 Scaling due to interparticle scattering

To get a rough feel for the issue, let us see what happens in an un-gauged (positive P) simulation when the kinetic coupling is negligible, and the scattering process dominates. Consider the case of local rather than extended ($U_{\mathbf{n}}$) interactions. Each lattice point represents an effective volume $\Delta V = \prod_d \Delta x_d$, and the mean occupation is $\bar{n}_{\mathbf{n}} \approx \rho(\mathbf{x}_{\mathbf{n}})\Delta V$ for a system with density $\rho(\mathbf{x})$. The self-interaction strength is $\chi = g/2\hbar\Delta V$ from (2.19).

At large mean lattice point occupations \bar{n} , the simulation time for a single mode using the positive P is given from Table 7.1 by $t_{\text{sim}} \approx 1.27/\chi\bar{n}^{2/3}$, based on numerical calculations. This is also in agreement with the analytic result (7.85). At low occupations the empirical fits of Table 7.2 can be used instead. Since single-mode simulation time decreases in absolute terms with mode occupation, then the expected simulation time is limited by the most highly occupied mode:

$$t_{\text{sim}} \lesssim \frac{2.5\hbar(\Delta V)^{1/3}}{g(\max[\rho(\mathbf{x})])^{2/3}}. \quad (10.4)$$

Kinetic processes and external potentials have been ignored for the time being,

although at least the kinetic mode coupling will reduce the useful simulation time below this level, as was seen in the calculations of Chapter 8. The expression (10.4) (or its extensions to include the empirical fits (7.89) at low mode occupation) is qualitatively borne out in the simulations presented later in this chapter.

Summarizing, **finer lattices lead to shorter simulations**, scaling as $(\Delta V)^{1/3}$ in the strong interaction limit.

10.2.2 Effect of kinetic interactions

Let us now consider the effect of the kinetic interactions. The simplest case is a uniform coherent gas with density ρ , scattering strength g , and lattice point volume ΔV . Also, no environment interactions, no external potential, and a constant standard diffusion gauge g'' applied to each mode separately.

When this is implemented as described in Chapter 7 and Section 4.4.2 (e.g. as in (8.4)), the evolution of the amplitudes of a single spatial mode \mathbf{n} is given (via (5.17), (2.13) and $\check{n}_{\mathbf{n}} = \alpha_{\mathbf{n}}\alpha'_{\mathbf{n}}^*$) by

$$d\alpha_{\mathbf{n}} = -i\mathcal{E}_{\mathbf{n}}^{(\alpha)} dt - iV_{\mathbf{n}}^{\text{ext}}\alpha_{\mathbf{n}} dt/\hbar - 2i\chi\alpha_{\mathbf{n}}\check{n}_{\mathbf{n}} dt + dX_{\mathbf{n}}^{(\alpha)} \quad (10.5a)$$

$$d\alpha'_{\mathbf{n}} = -i\mathcal{E}_{\mathbf{n}}^{(\alpha')} dt - iV_{\mathbf{n}}^{\text{ext}}\alpha'_{\mathbf{n}} dt/\hbar - 2i\chi\alpha'_{\mathbf{n}}\check{n}_{\mathbf{n}}^* dt + dX_{\mathbf{n}}^{(\alpha')}. \quad (10.5b)$$

The quantities $\mathcal{E}_{\mathbf{n}}^{(z)}$ (with z representing either α or α') contain the kinetic interactions, and from (5.17), (2.13), and (2.14),

$$\mathcal{E}_{\mathbf{n}}^{(z)} = \frac{\hbar\Delta V}{2mV} \sum_{\mathbf{m}, \tilde{\mathbf{n}}} z_{\mathbf{m}} |\mathbf{k}_{\tilde{\mathbf{n}}}|^2 e^{i\mathbf{k}_{\tilde{\mathbf{n}}}\cdot(\mathbf{x}_{\mathbf{n}}-\mathbf{x}_{\mathbf{m}})}, \quad (10.6)$$

where V is the entire lattice volume. The direct noise terms are explicitly

$$dX_{\mathbf{n}}^{(\alpha)} = i\sqrt{2i\chi}\alpha_{\mathbf{n}} \left[\cosh g'' dW_{\mathbf{n}} + i \sinh g'' d\widetilde{W}_{\mathbf{n}} \right] \quad (10.7a)$$

$$dX_{\mathbf{n}}^{(\alpha')} = i\sqrt{2i\chi}\alpha'_{\mathbf{n}} \left[-i \sinh g'' dW_{\mathbf{n}} - \cosh g'' d\widetilde{W}_{\mathbf{n}} \right]. \quad (10.7b)$$

Let us write the $\mathcal{E}_{\mathbf{n}}^{(z)}$ in terms of a mean $\overline{\mathcal{E}}^{(z)}$ and a fluctuating part $\delta\mathcal{E}_{\mathbf{n}}^{(z)}$ of mean zero:

$$\mathcal{E}_{\mathbf{n}}^{(z)}(t) = \overline{\mathcal{E}}^{(z)}(t) + \delta\mathcal{E}_{\mathbf{n}}^{(z)}(t). \quad (10.8)$$

Similarly the solution of (10.5) will have a constant and fluctuating part:

$$z_{\mathbf{n}}(t) = \bar{z}(t) + \delta z_{\mathbf{n}}(t). \quad (10.9)$$

Since the initial conditions are uniform and coherent then both \bar{z} and $\bar{\mathcal{E}}^{(z)}$ are the same for all modes and trajectories. Also, because there are no external forces on this system,

$$\bar{z}(t) = \sqrt{\bar{n}} e^{i\theta(t)} = \sqrt{\rho \Delta V} e^{i\theta(t)} \quad (10.10)$$

with some real phase $\theta(t)$.

If (temporarily) the influence of the kinetic terms was ignored, then the analysis of the logarithmic variances of a single mode from Sections 7.5.3 or 7.8.2 would apply, and the variance of $\log |z_{\mathbf{n}}|$ would be given by (7.65) or (7.92). (Compare $z_{\mathbf{n}}$ to the definition of G_L in (7.51)). In particular, at short times the dominant fluctuating contribution would be due directly to the noise terms rather than amplification of $\check{n}_{\mathbf{n}}$ fluctuations by the nonlinear drift term. From (7.53) and (7.15), then

$$\begin{bmatrix} \log \alpha_{\mathbf{n}}(t) \\ \log \alpha'_{\mathbf{n}}(t) \end{bmatrix} \approx \frac{1}{2} \log \bar{n} - 2i\chi \bar{n}t + \frac{\sqrt{\chi}}{2} \left\{ e^{-g''} [i\zeta_{\mathbf{n}}^+(t) - \zeta_{\mathbf{n}}^-(t)] \begin{bmatrix} + \\ - \end{bmatrix} e^{g''} [i\zeta_{\mathbf{n}}^-(t) - \zeta_{\mathbf{n}}^+(t)] \right\}, \quad (10.11)$$

where the $\zeta_{\mathbf{n}}^{\pm}(t)$ are time-correlated Gaussian random variables defined as in (7.56) and (7.57), and are independent for each mode \mathbf{n} . At short enough times that $\text{var} [\text{Re} \{\log z_{\mathbf{n}}(t)\}] \lesssim 1$, one would have

$$\begin{aligned} z_{\mathbf{n}}(t) &\approx \bar{z}(t) [1 + \log z_{\mathbf{n}}(t) - i\theta(t)] \\ &= \bar{z} + \delta^{\text{direct}} z_{\mathbf{n}}(t), \end{aligned} \quad (10.12)$$

where $\theta(t) = -2\chi \bar{n}t$, and $z_{\mathbf{n}}$ represents either $\alpha_{\mathbf{n}}$ or $\alpha'_{\mathbf{n}}$.

In reality there is also, a fluctuating contribution from the other modes mediated by the kinetic interactions so that the total short time fluctuations are

$$\delta z_{\mathbf{n}}(t) \approx \delta^{\text{direct}} z_{\mathbf{n}}(t) + \delta^{\text{kinetic}} z_{\mathbf{n}}(t). \quad (10.13)$$

In the Ito calculus, where Wiener increments are uncorrelated with variables at the same time step, the direct and kinetic fluctuations at t are un-correlated, and so the

variance of fluctuations will be

$$\text{var} [|\delta z_{\mathbf{n}}(t)|] = \text{var} [|\delta^{\text{direct}} z_{\mathbf{n}}(t)|] + \text{var} [|\delta^{\text{kinetic}} z_{\mathbf{n}}(t)|]. \quad (10.14)$$

One expects that as long as the first term (due to the direct noise terms $dX_{\mathbf{n}}^{(z)}$) dominates the fluctuations, then the single-mode analysis of Chapter 7 is accurate. In such a regime:

- The diffusion gauge (7.107) (or (7.94) for the case of $\mathcal{G}_{\mathbf{n}} = 0$) will be well optimized.
- The simulation times of Table 7.1 and 7.2 are accurate (at least for $\mathcal{G}_{\mathbf{n}} = 0$ simulations — see Section 10.2.4).
- The diffusion gauges (7.107) or (7.94) will give improvements of simulation time while the mean mode occupation is high (i.e. $\bar{n} = \rho\Delta V \gg 1$).

Let us estimate the relative size of these two fluctuation contributions at short times. Using (10.12) and (10.11), the properties (7.57) of the $\zeta_{\mathbf{n}}^{\pm}$, one finds

$$\text{var} [|\delta^{\text{direct}} z_{\mathbf{n}}(t)|] \approx \frac{\rho g t \cosh 2g''}{\hbar}. \quad (10.15)$$

Also, directly from the equations (10.5) and (10.8)

$$\delta^{\text{kinetic}} z_{\mathbf{n}}(t) = -i \int_0^t \delta \mathcal{E}_{\mathbf{n}}^{(z)}(t') dt'. \quad (10.16)$$

Using (10.6), and substituting in for $z_{\mathbf{m}}$ with the approximate (direct noise only) short time expression $z_{\mathbf{m}} \approx \bar{z} + \delta^{\text{direct}} z_{\mathbf{m}}$, one obtains

$$\begin{aligned} \text{var} [|\delta^{\text{kinetic}} z_{\mathbf{n}}(t)|] &= \left(\frac{\hbar \Delta V}{2mV} \right)^2 \int_0^t dt' \int_0^{t'} dt'' \sum_{\mathbf{m}, \mathbf{m}', \tilde{\mathbf{n}}, \tilde{\mathbf{n}'}} \langle \delta^{\text{direct}} z_{\mathbf{m}'}^*(t'') \delta^{\text{direct}} z_{\mathbf{m}}(t') \rangle_{\text{stoch}} \\ &\quad \times |\mathbf{k}_{\tilde{\mathbf{n}}}|^2 |\mathbf{k}_{\tilde{\mathbf{n}'}}|^2 e^{-i\mathbf{k}_{\tilde{\mathbf{n}'}} \cdot (\mathbf{x}_{\mathbf{n}} - \mathbf{x}_{\mathbf{m}'})} e^{i\mathbf{k}_{\tilde{\mathbf{n}}} \cdot (\mathbf{x}_{\mathbf{n}} - \mathbf{x}_{\mathbf{m}})}. \end{aligned} \quad (10.17)$$

Since the direct noise at each lattice point in the locally-interacting model is independent, then

$$\langle \delta^{\text{direct}} z_{\mathbf{m}'}^*(t'') \delta^{\text{direct}} z_{\mathbf{m}}(t') \rangle_{\text{stoch}} = \delta_{\mathbf{m}, \mathbf{m}'} \frac{\rho g \cosh 2g'' \min[t'', t']}{\hbar}, \quad (10.18)$$

similarly to (10.15), and the same for all modes. After performing the integrations over t' and t'' , and simplifying the Fourier transforms, one obtains

$$\text{var} [|\delta^{\text{kinetic}} z_{\mathbf{n}}(t)|] = \text{var} [|\delta^{\text{direct}} z_{\mathbf{n}}(t)|] \frac{\hbar^2 t^2 \Delta V}{12m^2 V} \sum_{\tilde{\mathbf{n}}} |\mathbf{k}_{\tilde{\mathbf{n}}}|^4. \quad (10.19)$$

For a \mathcal{D} -dimensional system with many modes, one can approximate

$$\begin{aligned} \sum_{\tilde{\mathbf{n}}} |\mathbf{k}_{\tilde{\mathbf{n}}}|^4 &\approx \frac{V}{(2\pi)^{\mathcal{D}}} \int |\mathbf{k}|^4 d^{\mathcal{D}}\mathbf{k} \\ &= \frac{V}{(2\pi)^{\mathcal{D}}} c_1(\mathbf{k}^{\text{max}}) \left(\frac{(2\pi)^{\mathcal{D}}}{5\Delta V} \right) \left(\frac{\pi}{(\Delta V)^{1/\mathcal{D}}} \right)^4 \\ &= c_1(\mathbf{k}^{\text{max}}) \frac{V\pi^4}{5(\Delta V)^{1+4/\mathcal{D}}}, \end{aligned} \quad (10.20)$$

where $c_1(\mathbf{k}^{\text{max}})$ is a shape factor $\mathcal{O}(1)$ that depends on the ratios between momentum cutoffs k_d^{max} for the various lattice dimensions. For example in $1D$, $c_1 = 1$, while for $2D$ and $3D$ when the momentum cutoffs in each dimension are equal, one has $c_1 = 3\frac{1}{9}$, and $c_1 = 6\frac{1}{3}$. Using (10.20), one obtains the final estimate

$$\text{var} [|\delta^{\text{kinetic}} z_{\mathbf{n}}(t)|] \approx \text{var} [|\delta^{\text{direct}} z_{\mathbf{n}}(t)|] \frac{\hbar^2 t^2 \pi^4 c_1}{60m^2 (\Delta V)^{4/\mathcal{D}}}. \quad (10.21)$$

It can be seen that: the fluctuations due to kinetic terms

- become relatively more important with time, and
- are more dominant for fine lattices (i.e. when ΔV is small).

10.2.3 Lattice size and diffusion gauge usefulness

If the kinetic coupling remains weak (in relative terms) for the duration of a ungauged or diffusion-gauged simulation, then one expects that the local diffusion gauges (7.94) are fairly well optimized, and will give simulation time improvements. Let us investigate under what conditions this is expected to occur based on (10.21).

Firstly, simulation times were found in Chapter 7 to be significantly improved when the mean particle occupation per mode is $\bar{n} \gtrsim 1$. The expected single-mode simulation time was found to be $t_{\text{sim}} \approx \mathcal{O}(10) / \chi \sqrt{\bar{n}}$ — see Table 7.1. It will be

convenient to write the lattice interaction strength $\chi = g/2\hbar\Delta V$ in terms of the **healing length**

$$\xi^{\text{heal}} = \frac{\hbar}{\sqrt{2m\rho g}}. \quad (10.22)$$

This is the minimum length scale over which a local density inhomogeneity in a Bose condensate wavefunction can be in balance with the quantum pressure due to kinetic effects (e.g. it is the typical size of a quantized vortex in a BEC[42, 44], or the typical size of a density correlation disturbance, as will be seen here). The healing length is discussed in more detail e.g. in Dalfovo *et al*[31], p. 481.

In terms of ξ^{heal} , the diffusion gauged simulation time is $t_{\text{sim}} \approx \mathcal{O}(40) (\xi^{\text{heal}})^2 m\sqrt{\bar{n}}/\hbar$. From (10.21), the kinetic fluctuations are expected to be weak ($\text{var} [|\delta^{\text{kinetic}} z_{\mathbf{n}}(t)|] \ll \text{var} [|\delta^{\text{direct}} z_{\mathbf{n}}(t)|]$) up to this time provided that

$$\frac{80}{3} c_1 \bar{n} \left(\frac{\pi \xi^{\text{heal}}}{(\Delta V)^{1/\mathcal{D}}} \right)^4 \ll 1. \quad (10.23)$$

The quantity $(\Delta V)^{1/\mathcal{D}} = \overline{\Delta x}$ is the geometric mean of the lattice spacing. One sees that simulation time improvement using a local gauge diffusion gauge occurs only when

$$\begin{aligned} \overline{\Delta x} &\gtrsim \xi^{\text{heal}} \bar{n}^{1/4} c_3 \\ &\gtrsim \mathcal{O}(\xi^{\text{heal}}) \end{aligned} \quad (10.24)$$

with $c_3 = \pi(3/80c_1)^{1/4} \approx \mathcal{O}(1)$ a weakly lattice shape dependent constant. (The second line follows since improvements occur only for $\bar{n} \gtrsim 1$.)

That is, **the local diffusion gauges (7.107) or (7.94) can be expected to give simulation time improvements only if the lattice spacing is of the order of the healing length or greater.**

10.2.4 Drift gauges and many-mode dynamics

As has been noted in Section 7.8.1, drift gauged simulations additionally encounter a scaling problem in many-mode systems because the single log-weight variable z_0 accumulates fluctuations from all modes.

One has the Ito equation for the log-weight:

$$dz_0 = \sum_{\mathbf{n}} \left\{ \mathcal{G}_{\mathbf{n}} \left[dW_{\mathbf{n}} - \frac{1}{2} \mathcal{G}_{\mathbf{n}} dt \right] + \tilde{\mathcal{G}}_{\mathbf{n}} \left[d\tilde{W}_{\mathbf{n}} - \frac{1}{2} \tilde{\mathcal{G}}_{\mathbf{n}} dt \right] \right\}, \quad (10.25)$$

with independent contributions from each mode. In a uniform gas on $M = V/\Delta V$ modes each with occupation \bar{n} , the contribution from each mode is identical on average, and the variance of real and imaginary parts of dz_0 will scale as $\text{var}[z_0] \propto M$. Because z_0 appears as $\Omega = e^{z_0}$ in observable estimates, there is a limit (7.43) to how large the variance of $\text{Re}\{z_0\}$ can be if any precision is to be retained in the simulation (see Section 7.4). At short times $\text{var}[\text{Re}\{z_0\}] \propto tM$, which leads to a reduction of simulation time $t_{\text{sim}} \propto 1/M$ in this case.

In the simulations with drift gauge (7.105) the diffusion gauge g'' acts to trade-off fluctuations in the amplitudes $\propto e^{g''}$ with fluctuations in the log-weight $\propto \sqrt{M}e^{-g''}$. This indicates that when many modes are present, a more advantageous trade-off might be achieved by increasing g'' relative to the single-mode expression than (7.107).

Let us investigate this. With M identical modes in a uniform gas, the changes in the optimization of Section 7.5.3 can be easily tracked. The last term of (7.60) for each mode is M times greater, and so the $\propto |n_0|^2$ term in (7.65) also acquires a factor M . This factor also appears in the third-order (in V_g) term of the cubics (7.67) and (7.72). In the end, the new optimization gives

$$g_{\mathbf{n}}'' = \frac{1}{6} \log \left\{ 8M |\check{n}_{\mathbf{n}}(t)|^2 \chi t_{\text{rem}} + a_2^{3/2} (\check{n}_{\mathbf{n}}(t), \gamma_{\mathbf{n}} t_{\text{rem}}) \right\} \quad (10.26)$$

instead of the single-mode expression (7.107). Simulation time can be estimated as in Section 7.6. At short times

$$\text{var}[G_L] = \frac{\chi t}{2} \left(V_g + \frac{1}{V_g} \right) + 2M (\chi V_g t n_0')^2 \quad (10.27)$$

(compare to (7.80)), while for large particle number $V_g \approx 1/2(n_0'^2 \chi t M)^{1/3}$. One obtains

$$t_{\text{sim}} \approx \frac{(20/3)^{3/2}}{M^{1/4} \chi \sqrt{n_0'}} \approx \mathcal{O} \left(\frac{40 t_{\text{coh}}}{M^{1/4}} \right). \quad (10.28)$$

so simulation appears to be reduced by a factor of $M^{-1/4}$ as compared to the single-mode case (Much better than a $1/M$ reduction with the single-mode expression (7.107)).

Unfortunately there is some further noise process that limits simulation times with drift gauged dynamics simulations on large lattices. A preliminary trial with a uniform 1D gas on a 50 point lattice with $\Delta x_1 = \Delta V = 10 \xi^{\text{heal}}$, and $\bar{n} = 1000$ bosons per lattice point was tried, but numerical simulations with the gauge (10.26) did not show improvement to simulation times over the $\mathcal{G}_{\mathbf{n}} = 0$ case. (Nor did a combination of drift gauge (7.30) with (7.107) or $g''_{\mathbf{n}} = 0$.) Only some improvement of precision at very small times $\ll t_{\text{sim}}$ was seen. Presumably kinetic effects lead to a large increase in z_0 fluctuations in this regime in comparison with what is expected for M uncoupled modes, and neither the original optimization nor the above re-optimization of $g''_{\mathbf{n}}$ gives improvement.

A combination of local drift and diffusion gauges *was* seen to give improvement under some conditions in Chapter 8. There it was concluded that the kinetic coupling needs to be relatively weak for this to occur, but details remain to be investigated. Simulations shown in the remainder of this chapter were made with zero drift gauge, and diffusion gauges only.

10.3 Correlation functions

In most experimentally realized Bose gas systems, low order local observables such as density or energy are often well described by approximate theories such as Gross-Pitaevskii (GP) equations (for a BEC), or statistical approaches in a high temperature gas. Multi-particle correlations are not well described by these theories, however, and so are of more interest for first-principles calculations. In terms of local lattice annihilation/creation operators those correlation functions that will be considered here are:

1. First order:

$$g^{(1)}(\mathbf{x}_{\mathbf{n}}, \mathbf{x}_{\mathbf{m}}) = \frac{\langle \hat{a}_{\mathbf{n}}^\dagger \hat{a}_{\mathbf{m}} \rangle}{\sqrt{\langle \hat{a}_{\mathbf{n}}^\dagger \hat{a}_{\mathbf{n}} \rangle \langle \hat{a}_{\mathbf{m}}^\dagger \hat{a}_{\mathbf{m}} \rangle}}. \quad (10.29)$$

This is a phase-dependent correlation function. Its magnitude $|g^{(1)}|$ tells one the degree of first-order spatial coherence, while its phase gives the relative phase of the wavefunction at spacing $\mathbf{x}_{\mathbf{n}} - \mathbf{x}_{\mathbf{m}}$.

2. Second order:

$$g^{(2)}(\mathbf{x}_n, \mathbf{x}_m) = \frac{\langle \hat{a}_n^\dagger \hat{a}_m^\dagger \hat{a}_n \hat{a}_m \rangle}{\langle \hat{a}_n^\dagger \hat{a}_n \rangle \langle \hat{a}_m^\dagger \hat{a}_m \rangle}. \quad (10.30)$$

This two-particle number correlation function is not phase dependent, and always real positive. It describes the likelihood of finding two particles at a spacing $\mathbf{x}_m - \mathbf{x}_n$ from each other, relative to what is expected of a coherent field. For a bunched field, $g^{(2)}(\mathbf{x}, \mathbf{x}) > 1$ (e.g. a thermal state has $g^{(2)}(\mathbf{x}, \mathbf{x}) = 2$), while antibunching is evidenced by $g^{(2)}(\mathbf{x}, \mathbf{x}) < 1$.

3. Third order:

$$g^{(3)}(\mathbf{x}_n, \mathbf{x}_m, \mathbf{x}_{m'}) = \frac{\langle \hat{a}_n^\dagger \hat{a}_m^\dagger \hat{a}_{m'}^\dagger \hat{a}_n \hat{a}_m \hat{a}_{m'} \rangle}{\langle \hat{a}_n^\dagger \hat{a}_n \rangle \langle \hat{a}_m^\dagger \hat{a}_m \rangle \langle \hat{a}_{m'}^\dagger \hat{a}_{m'} \rangle}. \quad (10.31)$$

This three-particle correlation function describes the likelihood of particles at \mathbf{x}_n , \mathbf{x}_m , and $\mathbf{x}_{m'}$ (relative to a coherent field). The rate of three-body processes is proportional to $g^{(3)}(\mathbf{x}, \mathbf{x}, \mathbf{x})$ (or, if the effective radius of particles is r , three body processes will depend on values of $g^{(3)}(\mathbf{x}, \mathbf{x} + \epsilon_1, \mathbf{x} + \epsilon_2)$, where $|\epsilon_j| \lesssim r$.) For example in a BEC, the rate of three-body recombination, which can limit the condensate's lifetime, is proportional to $g^{(3)}(\mathbf{x}, \mathbf{x}, \mathbf{x})$ [48].

For the uniform gas, average values of these (over all space of M lattice points) are of relevance:

$$\bar{g}^{(j)}(\mathbf{x}) = \frac{1}{M} \sum_{\mathbf{m}} g^{(j)}(\mathbf{x}_m, \mathbf{x}_m + \mathbf{x}) \quad (10.32a)$$

$$\bar{g}^{(3)}(\mathbf{x}, \mathbf{y}) = \frac{1}{M} \sum_{\mathbf{m}} g^{(3)}(\mathbf{x}_m, \mathbf{x}_m + \mathbf{x}, \mathbf{x}_m + \mathbf{y}). \quad (10.32b)$$

10.4 Example 1:

Correlation waves in a uniform gas

10.4.1 The system

This system consists of a uniform gas of bosons with density ρ and interparticle s -wave scattering length a_s . The lattice is chosen with a spacing $\Delta x_d \gg a_s$ so that

interparticle interactions are effectively local at each lattice point with effective field interaction strength g given by (2.16) or (2.19), depending on the dimensionality of the system. Coupling strength between lattice points, χ , is then given by (2.18). Periodic boundary conditions are assumed. The results obtained from a simulation will be invariant of lattice size provided the lattice spacing is fine enough to resolve all occupied momenta, and the lattice volume V is large enough to encompass all phenomena.

The initial state is taken to be a coherent wavefunction, which is a stationary state of the ideal gas with no interparticle interactions (i.e. $g = 0$). Subsequent evolution is with $g > 0$, so that there is a disturbance at $t = 0$ when interparticle interactions are rapidly turned on. Physically, this kind of disturbance can be created in a BEC by rapidly increasing the scattering length at $t \approx 0$ by e.g. tuning the external magnetic field near a Feshbach resonance.

The behavior of a uniform gas can tell us a lot about what goes on in more complicated inhomogeneous systems. If the density of a system is slowly varying on length scales of L_ρ or less, then any uniform gas phenomena of spatial extent L_ρ or smaller will also be present in the inhomogeneous system. In a trapped gas, for example, uniform gas phenomena that occur on length scales of the order of a fraction of the trap size or smaller.

10.4.2 Correlation waves in a one-dimensional gas

The evolution of some of the quantum correlations that appear in a one-dimensional gas with density $\rho = 100/\xi^{\text{heal}}$ is shown in Figures 10.1, 10.2, and 10.4, based on $\mathcal{G}_n = 0$ simulations. Figure 10.3 also shows the momentum distribution of particles at times after being excited by the disturbance at $t = 0$.

Time is given in units of the timescale

$$t_\xi = \frac{m(\xi^{\text{heal}})^2}{\hbar} = \frac{\hbar}{2\rho g} \quad (10.33)$$

(the “healing time”), which is approximately the time needed for the short-distance ($\mathcal{O}(\xi^{\text{heal}})$) inter-atomic correlations to equilibrate after the disturbance (see Figure 10.4(a)).

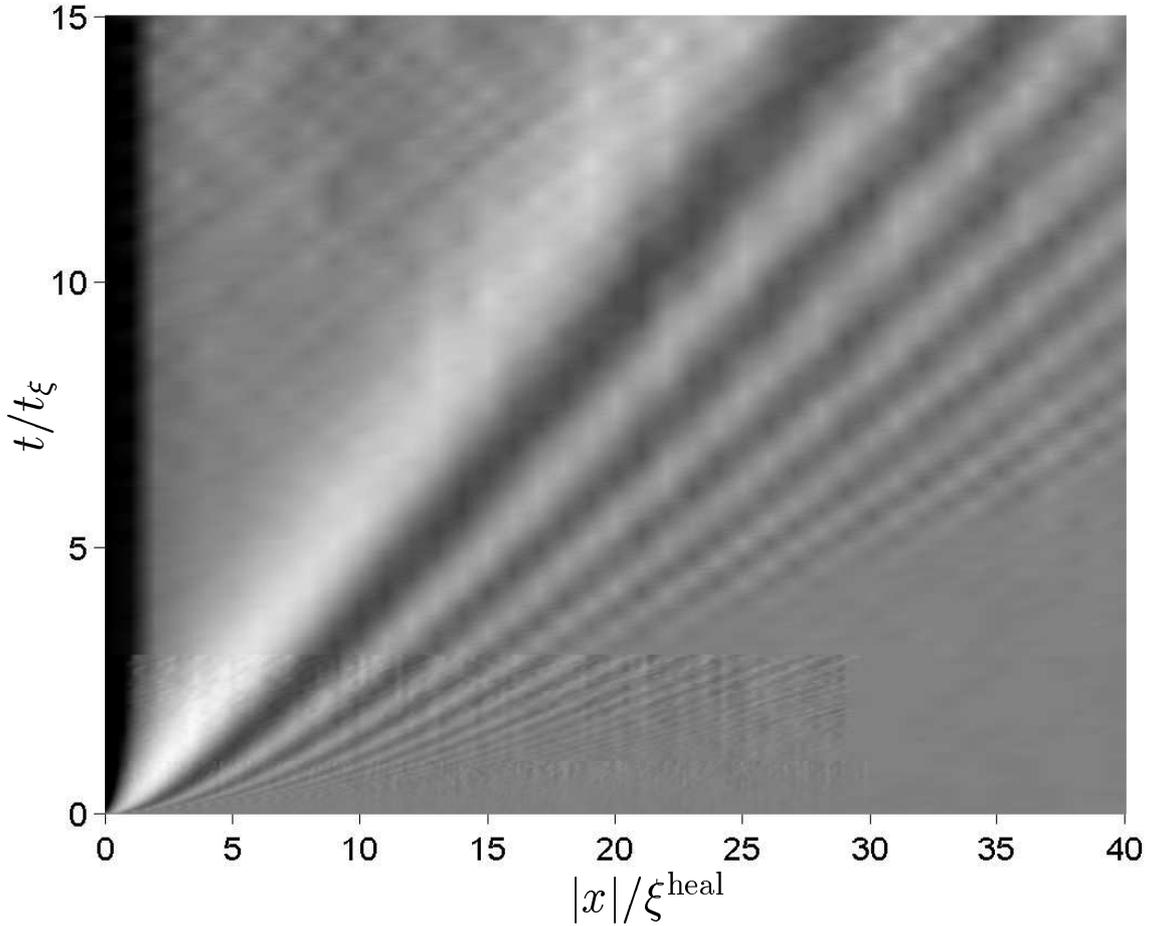


Figure 10.1: **Evolution of $\bar{g}^{(2)}(x)$ in a uniform 1D gas** with density $\rho = 100/\xi^{\text{heal}}$. Shading indicates relative magnitude of $\bar{g}^{(2)}(x)$, with light regions indicating high values > 1 , and dark indicating low values < 1 . The diagram is a composite of data from three simulations with different lattices: $\Delta x = 0.12\xi^{\text{heal}}$, 500 lattice points and $\mathcal{S} = 3000$ for $t \in [0, t_\xi]$ and $|x| \in [0, 29\xi^{\text{heal}}]$. $\Delta x = 0.24\xi^{\text{heal}}$, 250 points and $\mathcal{S} = 10^4$ for $t \in [t_\xi, 3t_\xi]$. $|x| \in [0, 29\xi^{\text{heal}}]$, and $\Delta x = \xi^{\text{heal}}$, 200 points and $\mathcal{S} = 10^4$ for the rest of the data. These changes in resolution are responsible for the apparent slight discontinuities in the data. The apparent weak cross-hashing superimposed on the plot at long times and/or large distances is due to finite sample size and/or lattice discretization effects and is not statistically significant. Magnitudes of $\bar{g}^{(2)}$ are shown in Figure 10.4.

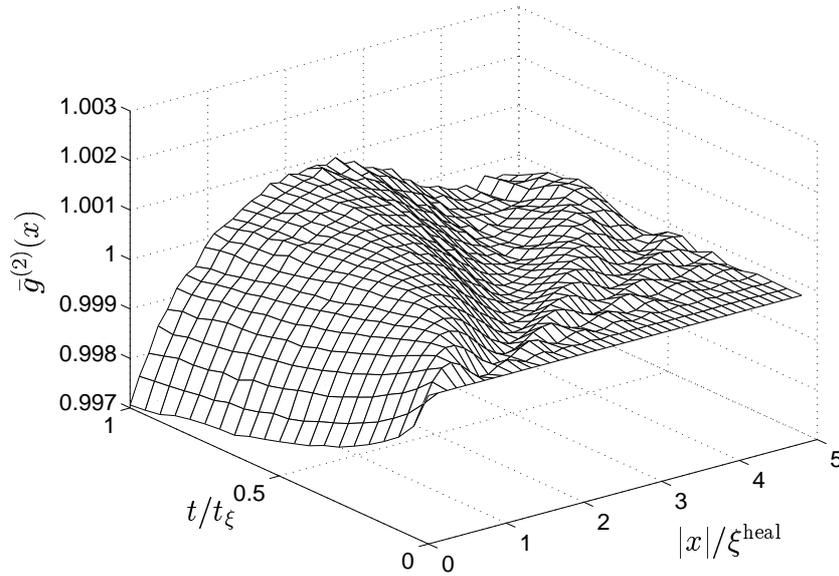


Figure 10.2: **Short time evolution of $\bar{g}^{(2)}(x)$ in a 1D gas with density $\rho = 100/\xi^{\text{heal}}$.** From a simulation with $\mathcal{S} = 3000$ trajectories, 500 points and $\Delta x = 0.12\xi^{\text{heal}}$.

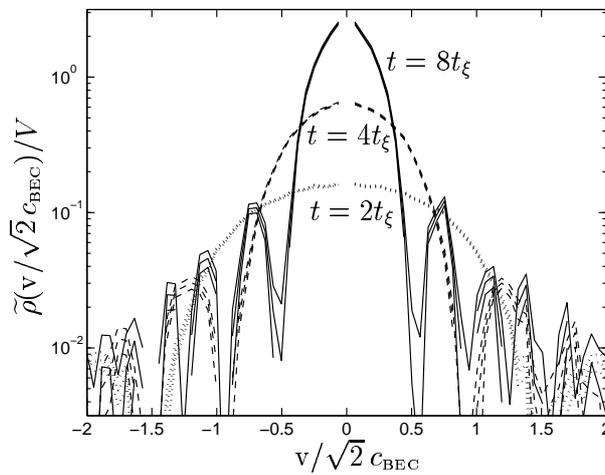


Figure 10.3: **Distribution of velocity in the uniform 1D gas of spatial density $\rho = 100/\xi^{\text{heal}}$** at several times t after the initial disturbance. Plotted is the velocity density $\tilde{\rho}$ per unit volume of the gas. Velocities v are given in units of $\sqrt{2}c_{\text{BEC}}$, which is the speed of the strongest correlation wave. The background density of stationary atoms has been omitted. Triple lines indicate error bars at one standard deviation significance. $\mathcal{S} = 10^4$ trajectories, $\Delta x = \xi^{\text{heal}}$, 100 lattice points.

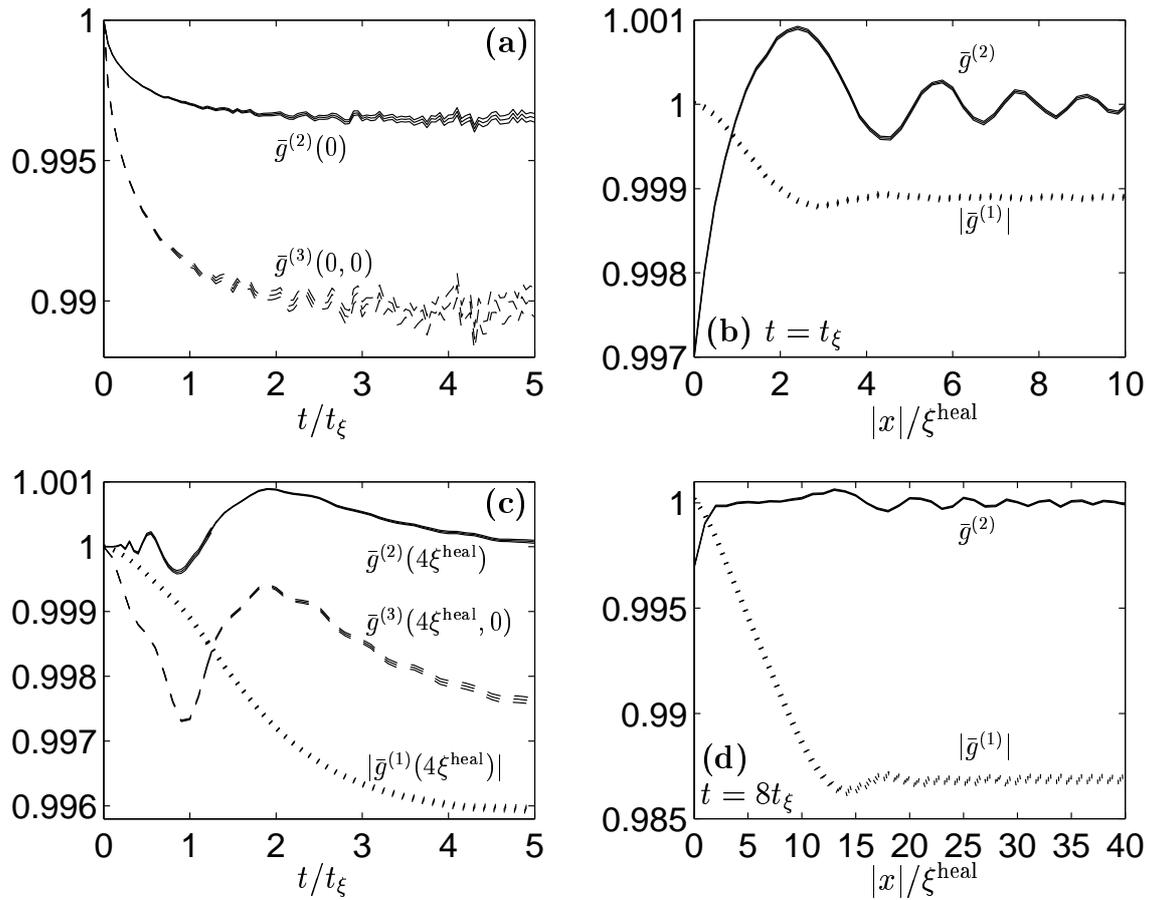


Figure 10.4: **Correlation functions in a uniform 1D gas** with density $\rho = 100/\xi^{\text{heal}}$. SOLID lines show $\bar{g}^{(2)}$, DASHED lines show $\bar{g}^{(3)}$, DOTTED lines show $|\bar{g}^{(1)}|$. Triple lines indicate error bars at one standard deviation significance. Subplot (a): Time dependence of local correlations. (b): Correlations at time $t = t_\xi$. (c): Time dependence of correlations at distance $x = 4\xi^{\text{heal}}$. (d): Long time correlations at $t = 8t_\xi$. Data in (a), (b), and $t < 1.25t_\xi$ in (c) are from a simulation with $\Delta x = 0.24\xi^{\text{heal}}$ and 250 points, while the rest is from a $\Delta x = \xi^{\text{heal}}$ simulation with 200 points. $S = 10^4$ trajectories in both cases.

Perhaps the most interesting feature noted in the simulations is the propagating wave train that appears in the *two-particle correlations*. These waves are seen only in the two- (or more) particle correlations, while the density always remains uniform (by symmetry, since there are no gains or losses in this model), and shows no sign of any wave behavior. This “correlation wave” is best seen in the time-dependent behavior of $\bar{g}^{(2)}(x)$ in Figure 10.1.

Once some initial transient effects at times $t \lesssim \mathcal{O}(t_\xi)$ after the disturbance die

out (these are shown in Figure 10.2), the long time behavior of first and second order correlations appears as follows:

- At distances $\mathcal{O}(\xi^{\text{heal}})$, antibunching occurs — See e.g. Figures 10.1, 10.4(b), 10.4(d).
- Enhanced two-particle correlations occur at a quite well-defined interparticle spacing, which increases at a constant rate: $\approx \sqrt{2}\xi^{\text{heal}}$ per t_ξ . This is a factor $\sqrt{2}$ faster than the low-momentum sound velocity in a BEC, obtained from the Bogoliubov dispersion law:

$$c_{\text{BEC}} = \sqrt{\rho g/m}. \quad (10.34)$$

See e.g. Dalfovo *et al*[31] p. 481.

- There are also many weaker correlations (and anti-correlations) at larger distances that move apart at increasingly faster rates, as seen in Figure 10.1. These leading correlation wavelets are particularly well visible in Figure 10.4(b). The leading disturbances appear to move at $\approx 20c_{\text{BEC}}$ for the $\rho = 100/\xi^{\text{heal}}$ system, although this is only clearly seen in this simulation for times $\lesssim 1.5t_\xi$ when transient effects may still be significant.
- Looking at the momentum spectrum of the particles (Figure 10.3), there are peaks, but their position changes with time (towards lower momenta), and in particular there is no peak corresponding to any particles travelling at the speed of the main (trailing) correlation wave.
- The peak to peak width of these correlation waves is of the order of several healing lengths for the times simulated, although some slow spreading is seen.
- At long times $\gtrsim \mathcal{O}(t_\xi)$ at distances shorter than the strongest trailing correlation wave with velocity $\sqrt{2}c_{\text{BEC}}$, but longer than the healing length, long range second order coherence ($\overline{g}^{(2)}(x) = 1$) between particles reappears.
- Long range phase coherence $|\overline{g}^{(1)}(x)|$ decays with time due to the scattering processes. (approximately linearly, at least while $|\overline{g}^{(1)}| \approx \mathcal{O}(1)$). See e.g. the

shorter time behavior of Figure 10.4(c), and the long distance behavior of Figure 10.4(b) and (d).

- However, in the (2nd order) coherent neighborhood that appears after the last correlation wave has passed, phase coherence decays only up to a certain value (see Figure 10.4(c) after the last $\bar{g}^{(2)}$ correlation wave has passed. This steady state phase coherence value drops off approximately linearly with distance, as seen especially in Figure 10.4(d).
- The third-order correlation $\bar{g}^{(3)}$ function also displays some wave behavior at similar times and distances as $g^{(2)}(x)$. This is visible e.g. in Figure 10.4(c).

The uniform one-dimensional gas is peculiar in some respects, one of these being that as density increases, the system becomes more like an ideal gas[86, 87]. From (10.4) and (10.33) the expected simulation time in units of the healing time t_ξ is

$$t_{\text{sim}} \approx 5\bar{n}^{1/3}t_\xi, \quad (10.35)$$

at high lattice point occupation \bar{n} . To resolve the correlation waves (with width $\approx \xi^{\text{heal}}$), one requires $\Delta x \approx \xi^{\text{heal}}$, which gives the scaling

$$\frac{t_{\text{sim}}}{t_\xi} \propto (\rho \xi^{\text{heal}})^{1/3}. \quad (10.36)$$

That is, at lower densities, it is more difficult to observe the long time correlation wave behavior that evolves on the t_ξ timescale. This is borne out in Figure 10.6(b), and is the reason why the relatively weak correlations in the $\rho = 100/\xi^{\text{heal}}$ have been investigated here in most detail.

Lower density simulations were also carried out for $\rho = 1/\xi^{\text{heal}}$ and $\rho = 10/\xi^{\text{heal}}$. Qualitatively the same kind of correlation wave behavior was seen for these densities up to simulation time, but the correlations are much stronger. Some examples for $\rho = 1/\xi^{\text{heal}}$ are shown in Figure 10.5.

When the length and time scales are given in density-dependent dimensionless units x/ξ^{heal} and t/t_ξ , the scaling of correlations in the simulated parameter regimes was closely approximated by

$$\bar{g}^{(j)} \propto \frac{\xi^{\text{heal}}}{\rho}. \quad (10.37)$$

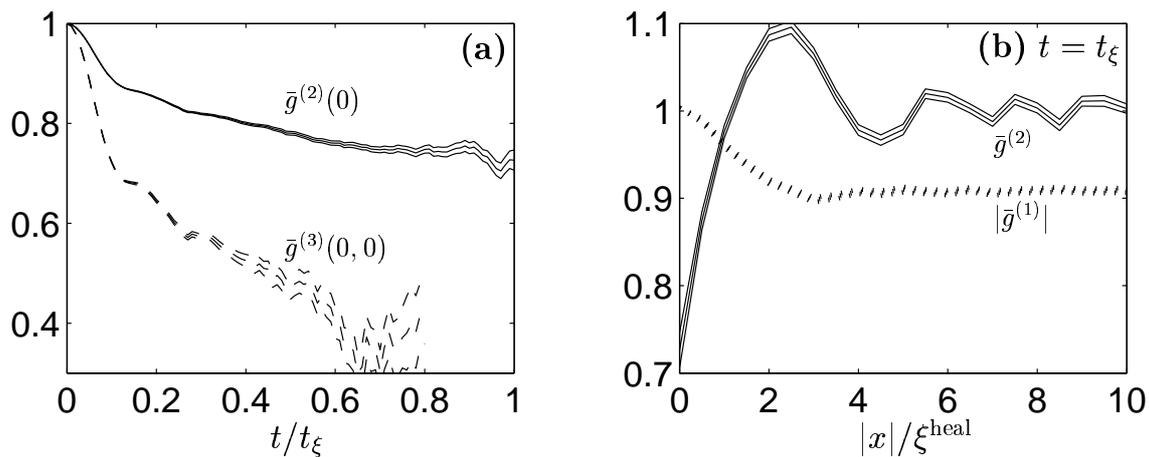


Figure 10.5: **Correlation functions in a uniform 1D gas** with density $\rho = 1/\xi^{\text{heal}}$. SOLID lines show $\bar{g}^{(2)}$, DASHED lines show $\bar{g}^{(3)}$, DOTTED lines show $|\bar{g}^{(1)}|$. Triple lines indicate error bars at one standard deviation significance. Subplot (a): Time dependence of local correlations. (b): Correlations at time $t = t_\xi$. Simulations with $\Delta x = 0.5\xi^{\text{heal}}$, 50 lattice points, and $\mathcal{S} = 10^4$ trajectories.

10.4.3 Lattice dependence of simulation time

The observed scaling of simulation time with lattice spacing Δx and lattice occupation $\bar{n} = \rho\Delta x$ is shown in Figure 10.6, and compared to the expected un-gauged times when kinetic effects are ignored. The estimated times shown in gray are based on the empirical relationship (7.89) with the positive P (i.e. un-gauged) fitting parameters given in Table 7.2. It can be seen that for the physical regimes simulated, this empirical estimate of simulation time is very good also for many-mode simulations. (Most of the residual variation in t_{sim} is due to statistical effects of finite sample size).

10.4.4 Diffusion gauge dependence of simulation time

From the analysis of Section 10.2.3, it is expected that diffusion gauges will give significant improvements once $\Delta x \gtrsim \xi^{\text{heal}}$. This is confirmed in Figure 10.7, which shows the dependence of simulation time on t_{opt} for some simulations using the gauge (7.94). At large lattice spacings, significant gains can be made with the local diffusion gauge, while for spacing of the order of the healing length the improvement

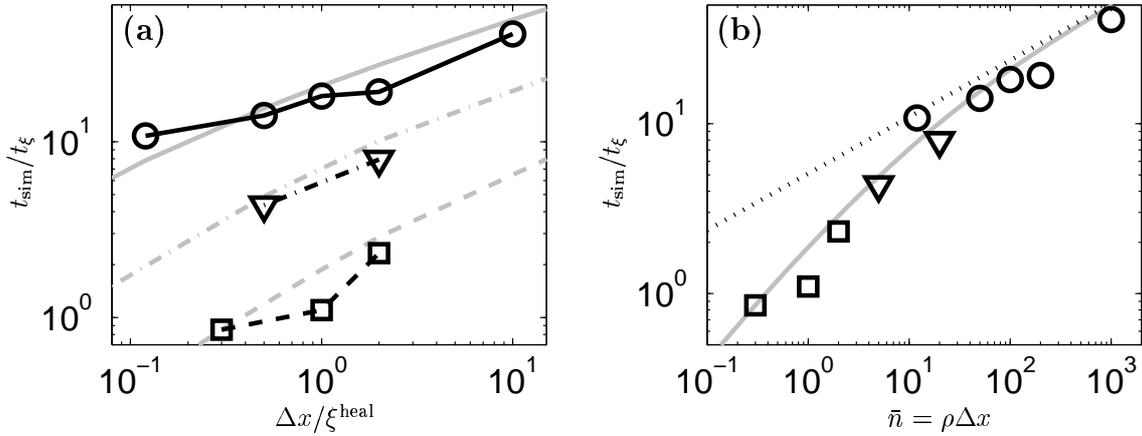


Figure 10.6: **Scaling of useful simulation time** with lattice spacing Δx and mean lattice point occupation $\bar{n} = \rho\Delta x$. Data points for simulations of 1D gases of densities $100/\xi^{\text{heal}}$, $10/\xi^{\text{heal}}$, and $1/\xi^{\text{heal}}$, are shown with CIRCLES, TRIANGLES, and SQUARES, respectively. $\mathcal{S} = 10^4$ trajectories. Also shown for comparison as LIGHT LINES are expected simulation times based on the empirical relationship (7.89) for the single-mode positive P. The DOTTED LINE in subplot (b) gives the high-occupation estimate (10.35).

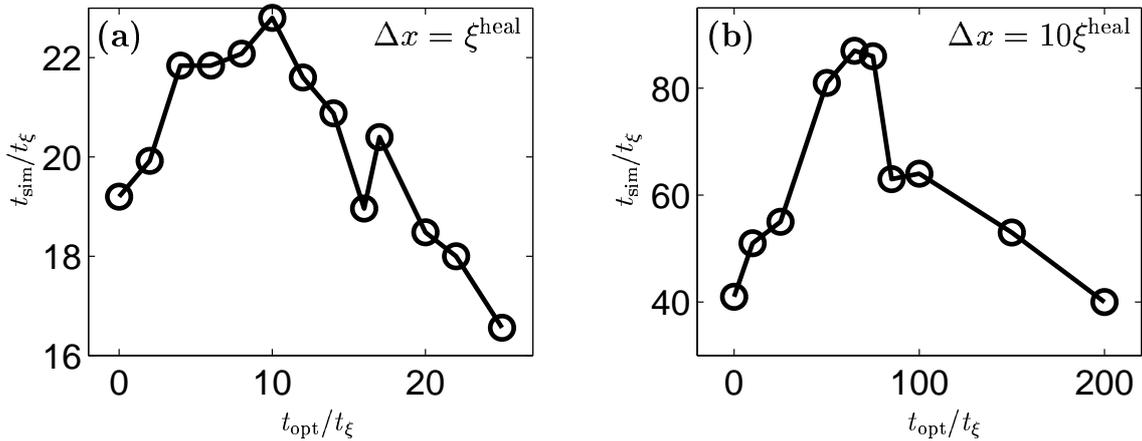


Figure 10.7: **Improvement of simulation time with diffusion gauges** Simulation times as a function of the target time t_{opt} when using the diffusion (only) gauge (7.94). All simulations are of a $\rho = 100/\xi^{\text{heal}}$ gas. In subplot (a): $\Delta x = \xi^{\text{heal}}$ with $\mathcal{S} = 10^4$ trajectories, and in (b), $\Delta x = 10\xi^{\text{heal}}$ with $\mathcal{O}(500)$ trajectories.

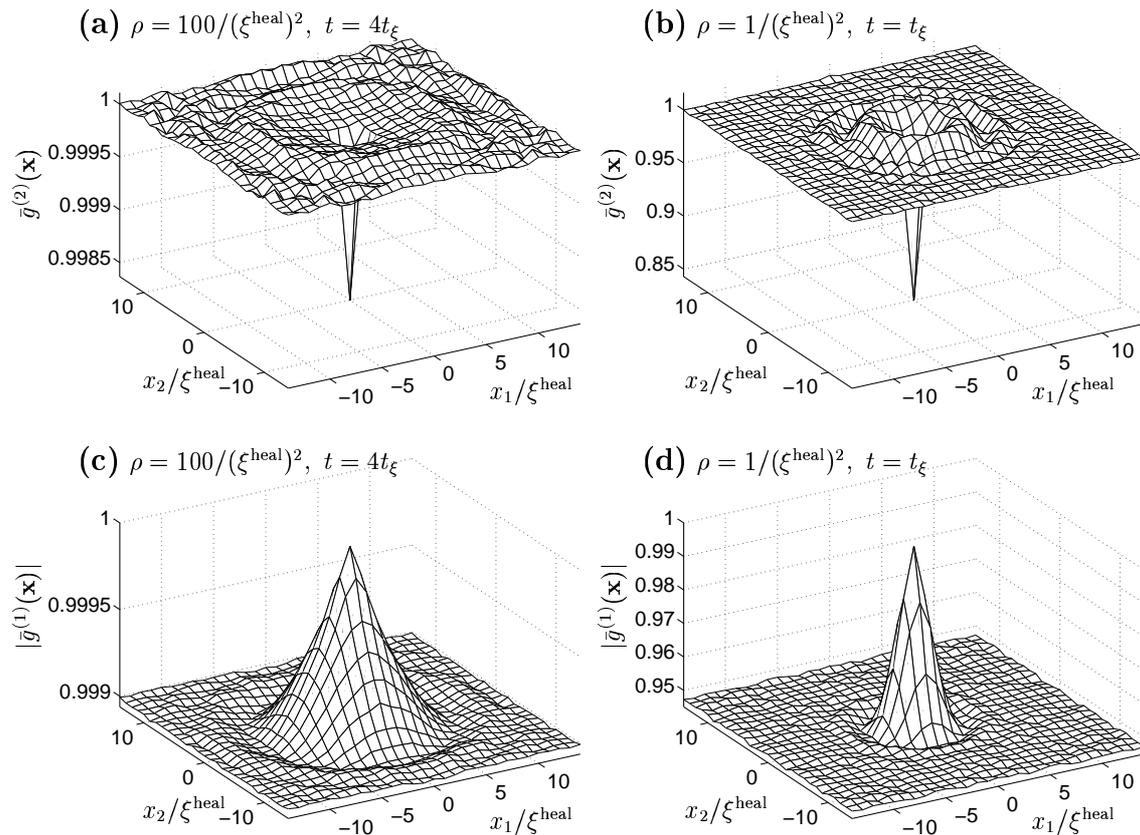


Figure 10.8: **Correlation functions in 2D uniform gases** at times $4t_\xi$ and t_ξ for gases of density $100/(\xi^{\text{heal}})^2$ and $1/(\xi^{\text{heal}})^2$, respectively. Coordinates in dimension d are denoted x_d . Simulations used a lattice spacing of $\Delta x_d = \xi^{\text{heal}}$, 30×30 lattice points, and $\mathcal{S} = 10^4$ trajectories.

is small. Note the similarity of Figure 10.7(b) to the single-mode case Figure 7.8(a). In both cases simulation time increases steadily while $t_{\text{opt}} \leq \max[t_{\text{sim}}]$, but drops off sharply once the critical t_{opt} is reached ($t_{\text{opt}} \approx 80t_\xi$ in this case).

10.4.5 Two-dimensional gas

Simulation of a two-dimensional gas is also straightforward using the equations (5.17). Some examples of calculated correlation functions for 2D uniform gases with densities $\rho = 100/(\xi^{\text{heal}})^2$ and $\rho = 1/(\xi^{\text{heal}})^2$ are shown in Figure 10.8. Correlation wave behavior similar to the 1D case is seen.

10.5 Example 2:

Extended interactions in a uniform gas

The simulation of a gas with extended interparticle interactions is also possible with the gauge P representation, as described in Section 5.5. One uses the equations (5.17) but with scattering effects modified to become the nonlinear terms (5.47) and noise terms given by (5.46) and (5.45).

A Gaussian interparticle potential

$$U(x) = \frac{g}{\sigma_U \sqrt{2\pi}} \exp \left[\frac{-x^2}{2\sigma_U^2} \right]. \quad (10.38)$$

is assumed for the interparticle potential in this example. The bulk interaction energy density in a system with density varying much slower in x than $U(x)$ is $\bar{u} = \frac{1}{2V} \int U(x-y) \hat{\Psi}^\dagger(x) \hat{\Psi}^\dagger(y) \hat{\Psi}(x) \hat{\Psi}(y) dx dy \approx \frac{\rho^2}{2} \int U(x) dx$. The potential (10.38) has been normalized so that \bar{u} is the same as for a local-only “delta-function” interaction with strength g .

The simulated second order correlation function is shown in Figures 10.9 and 10.10 for a potential of standard deviation $\sigma_U = 3\xi^{\text{heal}}$. Density was again chosen $\rho = 100/\xi^{\text{heal}}$, so that at large length scales $\gg \sigma_U$, the behavior of the gas approaches the behavior of the locally-interacting model shown in Section 10.4.2 and Figures 10.1, 10.2, and 10.4.

Correlation wave phenomena are seen in this system as well, but with significant differences to the locally interacting gas of Section 10.4.2:

- The antibunching at short distance scales occurs out to a range $\mathcal{O}(\sigma_U)$ rather than ξ^{heal} , and takes on approximately a Gaussian form. See Figure 10.10(a) and (b), and compare to the linear (in x) growth of $\bar{g}^{(2)}(x)$ in Figure 10.4(a).
- The strongest positive correlation wave moves at approximately the speed of sound c_{BEC} , not at the faster rates seen in the locally-interacting gas.
- A train of weaker correlation waves is created (at least two more wavefronts are seen in Figure 10.9), but these form *behind* the strongest leading wavefront

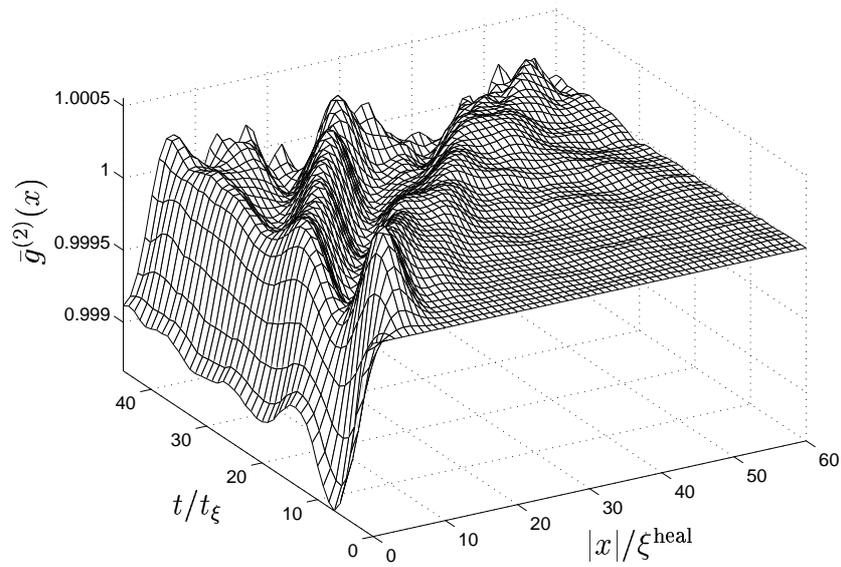


Figure 10.9: **Evolution of $\bar{g}^{(2)}(x)$** in a uniform 1D gas with extended interparticle interactions (10.38): $\sigma_U = 3\xi^{\text{heal}}$. Gas of density $\rho = 100/\xi^{\text{heal}}$. Simulation was on a lattice with spacing $\Delta x = \xi^{\text{heal}}$, 200 lattice points (thus with 20 000 total particles on average), and $\mathcal{S} = 10^4$ trajectories.

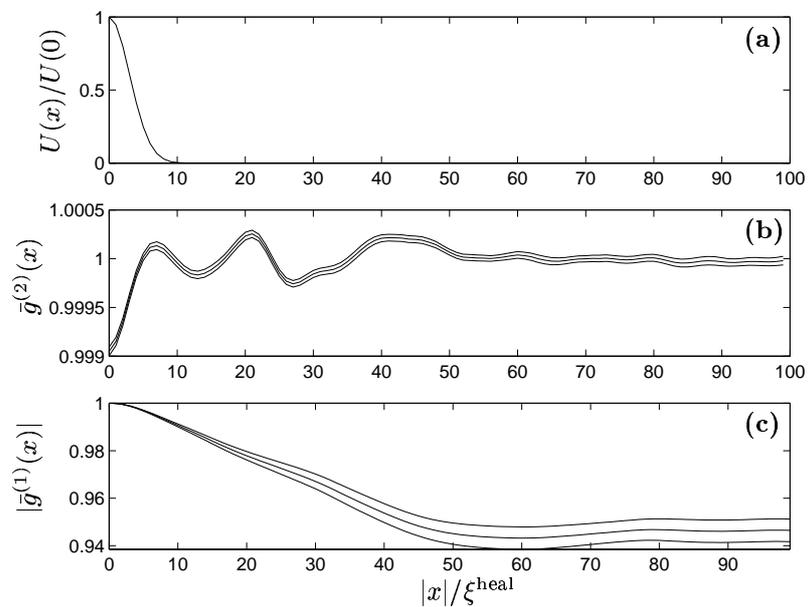


Figure 10.10: **Correlations at $t = 35t_\xi$** for the uniform 1D gas with extended interparticle interactions. Gas and simulation parameters as in Figure 10.9. Triple lines indicate error bars at one standard deviation. The form of the interparticle potential is shown to scale in (a).

rather than in front of it, and these secondary waves appear to also move at the sound velocity.

- Some decaying oscillations of $\bar{g}^{(2)}(0)$ occur.

Phase coherence (Figure 10.10(c)) decays in similar fashion to the locally-interacting gas.

The differences in the long time behavior between this system and the locally interacting gas may be reconciled despite the potential having the same coarse-grained interaction strength g . An assumption made to arrive at an effectively locally-interacting gas was that the s-wave scattering length $a_s = mg/4\pi\hbar^2$ is much smaller than all relevant length scales – in this case e.g. ξ^{heal} . This is satisfied for the simulations in Section 10.4.2, but the equivalent quantity here (which is σ_U) is *not* smaller than ξ^{heal} .

As in the gas with only local interactions on the lattice, long time behavior is best observed with high density but small correlations as in Figure 10.9. Much stronger correlations can be simulated, but for shorter times (scaling will again be $t_{\text{sim}} \propto \bar{n}^{1/3}$).

Lastly, a convenient feature of these extended interaction simulations is that the useful simulation time is found to be significantly longer than for a system with local interactions of the same coarse-grained strength g . For example in the $\rho = 100/\xi^{\text{heal}}$ system, $t_{\text{sim}} \approx 46t_\xi$ for the $\sigma_U = 3\xi^{\text{heal}}$ gas, while the locally-interacting gas had $t_{\text{sim}} \approx 17t_\xi$ (Note that $g = \hbar/2\rho t_\xi$).

10.6 Example 3: Correlations in a trap

Simulations in an external trapping potential V^{ext} pose no particular problem. For example, the following system was simulated:

- Bosons are prepared in a harmonic trap with trapping potential

$$V^{\text{ext}}(x) = \frac{1}{2}m\omega_{\text{ho}}^2x^2, \quad (10.39)$$

which has a harmonic oscillator length $a_{\text{ho}} = \sqrt{\hbar/m\omega_{\text{ho}}}$. Initially they are in the coherent zero temperature ground state obtained by solving the Gross-Pitaevskii mean field equations¹. The mean number of atoms in the trap in this example is $\overline{N} = 10$.

- The bosons experience two-body interactions with an effective Gaussian interparticle potential (10.38) with radius $\sigma_U = a_{\text{ho}}$, and a strength $g = 0.4\hbar a_{\text{ho}}\omega_{\text{ho}}$.
- At $t = 0$, breathing of the atomic cloud is induced by switching to a more confined harmonic potential with double the trapping frequency: i.e. $V^{\text{ext}} \rightarrow \frac{1}{2}m(2\omega_{\text{ho}})^2x^2$.

Some data from the simulation are shown in Figures 10.11 and 10.12.

This model is in a regime where qualitative results are hard to achieve using approximate methods because several length scales from different processes are of the same order: Trap width is a_{ho} , initial cloud width is $\approx 2a_{\text{ho}}$, interparticle scattering range is also a_{ho} .

Phenomena seen in this the simulation include:

- The two-particle density correlations $g^{(2)}(0, x)$ display different behaviors in the contracting and expanding phase. See Figure 10.11(a).
 - When the particle cloud is contracting, *antibunching* appears at the center of the trap, while during the later part of the contraction there is an enhanced likelihood of pairs of atoms with one in the outer region of the cloud and one in the center.
 - During expansion, on the other hand, the particles tend to bunch in the center of the trap (i.e. there is increased likelihood of two particles at small separation), while pairs of particles with one in the tails, one in the center are suppressed.

¹i.e. the stationary state of (5.17) with noise terms removed. — see Section 5.4 for more on the correspondence between the deterministic part of the gauge P equations and the Gross-Pitaevskii equations.

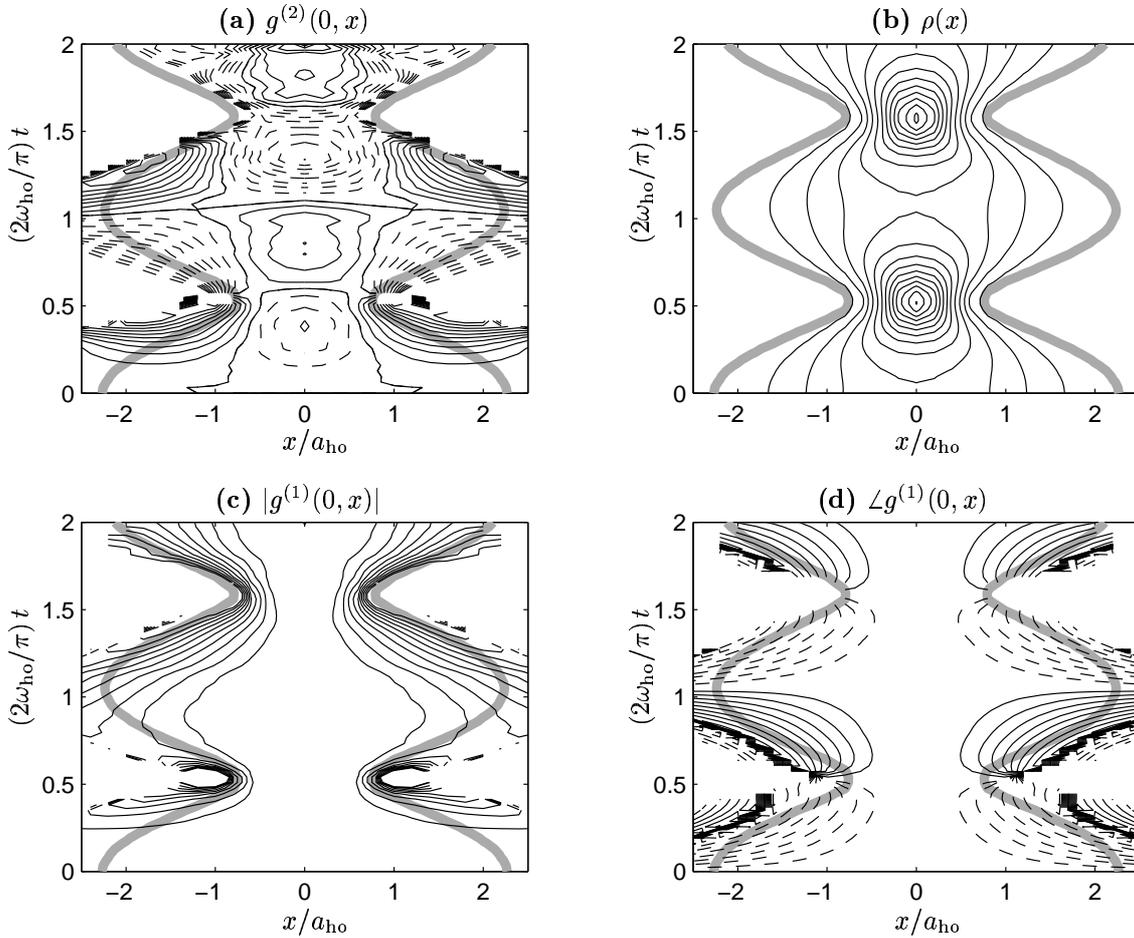


Figure 10.11: **Contour plots of the evolution of correlations and boson density** in the trapped Bose gas described in Section 10.6. Correlations shown are between bosons in the center of the trap, and those a distance x from the center. **(a)**: Contours of density correlations $g^{(2)}(0, x)$ with a spacing of 0.01. SOLID lines indicate $g^{(2)} > 1$, DASHED $g^{(2)} < 1$. **(b)**: contours of mean density $\rho(x)$ with spacing of $1/a_{\text{ho}}$. **(c)**: contours of phase coherence $|g^{(1)}(0, x)| \leq 1$ with spacing 0.01 (note: $g^{(1)}(0, 0) = 1$). **(d)**: contours of the relative condensate phase $\angle g^{(1)}(0, x)$ with spacing $\pi/10$. SOLID lines indicate $0 < \angle g^{(1)} < \pi$, DASHED $-\pi < \angle g^{(1)} < 0$. $\angle g^{(1)} = 0$ contour omitted. LIGHT BROAD lines are contours of density at 5% of the central value, and indicate the approximate extent of the boson cloud. Simulation with $S = 10^4$ trajectories on a 60 point lattice with $\Delta x = 0.2a_{\text{ho}}$. Note that a_{ho} is the harmonic oscillator length of the initial cloud, while $a_{\text{ho}}/2$ is the width of the narrower trap at $t > 0$ with frequency $2\omega_{\text{ho}}$.

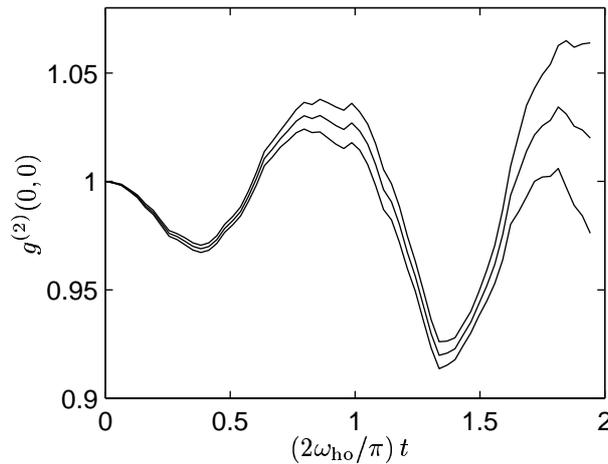


Figure 10.12: **Local two-particle correlations in the center of the trap** for the breathing condensate described in Section 10.6. Triple lines indicate error bars at one standard deviation. Parameters as in Figure 10.11.

-
- This bunching during the expansion and antibunching during the contraction appear counter-intuitive. The reason for this appears to be a time lag before the dominant effect makes itself felt in the correlations. This results in the bunching trailing the contraction by a significant part of the breathing period. Initially, the cloud is stationary, and during the contraction phase, there is a long period of time when the atoms are largely in free fall, and interparticle repulsion dominates the correlations causing antibunching despite the contraction. Eventually, the atoms become squashed together leading to bunching, but this occurs only at the end of the contraction phase. Bunching now remains for a large part of the expansion phase before it is finally overcome by the interparticle repulsion.
 - The oscillations of $g^{(2)}(0,0)$ (i.e. bunching at the center of the trap) due to the breathing of the atomic cloud become more pronounced with time — see Figure 10.12. This may indicate a resonance between the breathing and the repulsion, although it is also possible that this is a transient initial effect.
 - Coherence between the center of the trap and outlying regions of the cloud deteriorates as time proceeds – compare the first and second contraction phase in Figure 10.11(c).

10.7 Example 4:

Bosonic enhancement of atoms scattered from colliding condensates

10.7.1 The Vogels *et al* four wave-mixing experiment

In a recent experiment of Vogels *et al* [41] at MIT, strong coherent four-wave mixing between components of a Bose-Einstein condensate at different velocities was observed. The matter wave components were created by applying Bragg pulses to an initially stationary trapped condensate so as to impart a velocity $2\mathbf{v}^{\text{cm}}$ to approximately half of the atoms. (\mathbf{v}^{cm} is their velocity with respect to the center of mass). A small seed population ($\mathcal{O}(1\%)$ of all the atoms) moving at velocity \mathbf{v}_s was also created. Bose-enhanced scattering of the atoms during the half-collision of the two main wavepackets led to $\mathcal{O}(10\times)$ coherent amplification of the of the seed population and the creation of a fourth coherent population at velocity $\mathbf{v}_4 = 2\mathbf{v}^{\text{cm}} - \mathbf{v}_s$.

As the wavepackets move through each other, scattering of atom pairs occurs into velocities \mathbf{v} and $2\mathbf{v}^{\text{cm}} - \mathbf{v}$, with energy conservation favoring $|\mathbf{v}| \approx \mathbf{v}^{\text{cm}}$. If the seed wavepacket is present at some $|\mathbf{v}_s| \approx \mathbf{v}^{\text{cm}}$ then scattering of atoms into its modes is preferred due to Bose enhancement, analogously to stimulated emission in photonic systems.

In the experiment the initially empty momentum modes also acquired sufficient density after some time, so that Bose enhancement of scattering occurred into these non-seed modes as well. These momentum modes (which are much more numerous than the seed wave) eventually competed with the seed wavepacket and limited its growth.

A first-principles method could be desirable to quantitatively describe the effect of the initially empty modes in this system. Such a calculation is difficult with approximate methods as both single-particle effects (to occupy the empty modes in the first place), and subsequent amplification of the coherent many-particle wavefunction are involved. The first process can be estimated with perturbative methods, and the second with mean field GP equations, but combining the two has been

difficult.

Experimental parameters (relevant to the subsequent simulation described in the next subsection) were : $\approx 30\,000\,000$ atoms of ^{23}Na ($a_s = 2.75\text{nm}$, $m = 3.82 \times 10^{-26}\text{kg}$) in the initial wavepacket. This was created in an axially symmetric longitudinal trap with frequencies $\omega/2\pi$ of 20Hz in the longitudinal and 80Hz in the axial directions. This trap is subsequently turned off, and the two main wavepackets created with the Bragg pulse move at a velocity $\pm \mathbf{v}^{\text{cm}}$ relative to the center of mass, with $\mathbf{v}^{\text{cm}} \approx 10\text{mm/s}$.

10.7.2 Simulating Bose-enhanced scattering into initially empty modes

To study the enhanced scattering into the initially empty modes, an un-gauged positive P representation simulation was made with similar parameters as the Vogels *et al* experiment. There were two differences:

- In the simulation there were 1.5×10^5 atoms (on average), while there were $\approx 3 \times 10^7$ atoms in the experiment. This smaller atom number was needed to achieve a long enough simulation time to see significant Bose enhancement of the scattered modes.
- In the simulation no seed wave was placed.

The simulation was carried out in the center-of-mass frame, and the initial coherent wavefunction was taken to be

$$\psi(\mathbf{x}) = \psi_{\text{GP}}(\mathbf{x})\sqrt{2} \cos\left(\frac{m\mathbf{v}^{\text{cm}} \cdot \mathbf{x}}{\hbar}\right), \quad (10.40)$$

where $\psi_{\text{GP}}(\mathbf{x})$ is the coherent $T = 0$ ground state solution of the Gross-Pitaevskii (GP) equation in a trap of the same dimensions as in the experiment, and the full number of atoms. The Bragg pulse used to impart velocity to the moving wavepacket was of short duration $\approx 40\mu\text{s}$, and so to a good approximation no significant evolution of the condensate took place during the pulse, hence (10.40) is a good approximation to the state of the system just after the pulse. The initial

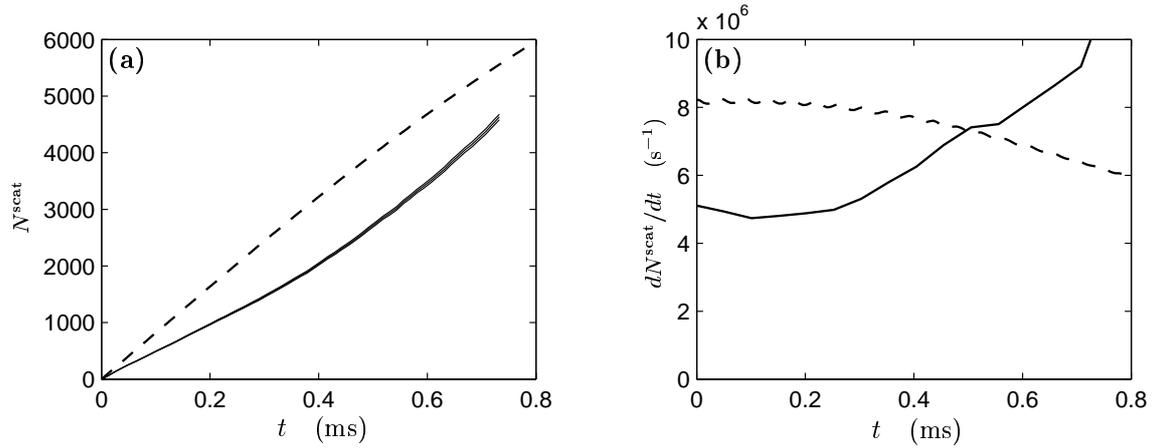


Figure 10.13: (a): Total number of scattered atoms N^{scat} , and (b): rate of scattering in the colliding BEC system of Section 10.7. SOLID LINE: scattered atoms calculated as described in Section 10.7.2, from a positive P simulation. Triple lines indicate uncertainty at one standard deviation. DASHED LINE: Estimate obtained with the imaginary scattering length method[88].

wavepackets move in the longitudinal direction $\pm \mathbf{v}^{\text{cm}} = [\pm v_1^{\text{cm}}, 0, 0]$. The simulation was on a $438 \times 48 \times 48$ lattice with $\mathcal{S} = 200$ trajectories, and lattice spacing of $\Delta x_1 = 0.2209 \mu\text{m}$ in the longitudinal, and $\Delta x_{2,3} = 0.4921 \mu\text{m}$ in the axial directions. Simulation time is $\mathcal{O}(6\text{days})$ on a PC of 2002 vintage.

Figure 10.13 shows the change of scattering rate with time, and the total number of scattered atoms for this simulation. Since the overlap between the two wavepackets travelling at $\pm \mathbf{v}^{\text{cm}}$ decreases as they move apart, and the rate of scattering into empty modes is proportional to this overlap, then the increasing scattering rate in the simulation is **clear evidence of the beginning of Bose enhancement of the initially empty modes**.

An interesting technical difficulty occurs when estimating the number of scattered atoms. Just as in a real experiment, the atoms in the first-principles calculation are not labeled as “scattered” or “non-scattered”. Some of them are scattered back into momentum modes already occupied by the initial wavepackets, and cannot be separated from the unscattered atoms by counting. To nevertheless make an estimate of scattering rate, a similar procedure was used to what would be needed for experimental data. It was aimed to count only those atoms with momenta beyond the coherent wavepackets. Explicitly, momentum modes with longitudinal veloci-

ties differing by less than 1mm/s from $\pm v_1^{\text{cm}}$, and with radial velocities of less than 2.3mm/s in both orthogonal directions are excluded from the count. (For comparison, the rms velocity deviation of atoms from $\pm \mathbf{v}^{\text{cm}} = [\pm v_1^{\text{cm}}, 0, 0]$ in the condensates is initially about 0.13mm/s and 0.5mm/s in the longitudinal and radial directions, although some later spreading occurs). This gives a scattered atom count N_1^{scat} .

Furthermore, in a mean field GP equation calculation, there appear some atoms at momenta that would be counted as “scattered” with the above counting method. (These occur mainly in the far tails of the wavepacket momentum distribution due to spreading with time, and also at velocities $\pm 3\mathbf{v}^{\text{cm}}$ due to the stimulation of a weak scattering process $\mathbf{v} \& \mathbf{v} \rightarrow 3\mathbf{v} \& -\mathbf{v}$ when $\mathbf{v} \approx \pm \mathbf{v}^{\text{cm}}$ to a state with short lifetime (due to lack of energy conservation). Initially their number is $\mathcal{O}(200)$ initially, growing to $\mathcal{O}(3000)$ at the end of the simulation). These GP background atoms, which are not due to the spontaneous scattering process of interest or its Bose enhancement, are subtracted away from the N_1^{scat} count obtained previously to arrive at the final scattered atoms estimate N^{scat} shown in Figure 10.13.

To check how well the expression (10.4) assesses simulation time *a priori*, let us see what estimate it gives in this more complex system. The peak density in the middle of the initial cloud in the Thomas-Fermi approximation is $\rho_0^{\text{TF}} = \hbar\omega_{\text{ho}}(15\bar{N}a_s/a_{\text{ho}})^{2/5}/2g$. The maximum density is well approximated by twice this Thomas-Fermi peak density (the factor of two arises because the initial condition (10.40) has local density peaks with twice the local average density). Using $\max[\rho(\mathbf{x})] = 2\rho_0^{\text{TF}}$ in (10.4), one obtains an expected simulation time of 0.6ms, which agrees fairly well with the observed simulation time of 0.78ms. The extra simulation time as compared to the estimate may be due to the decrease of peak density with time as the wavepackets move apart.

Lastly, the local diffusion gauges (7.94) and (7.107) of Chapter 7 were not expected (nor found) to give significant simulation time improvements for this calculation, because the healing length is larger than the lattice spacing. (Using ρ_0^{TF} again, one obtains $\xi^{\text{heal}} \approx 0.6\mu\text{m}$. The small lattice size used is needed to resolve the phase oscillations in the moving condensate.

10.7.3 Comparison to imaginary scattering length estimate

The first-principles scattering rate is also compared to that obtained using the imaginary scattering length technique[88] (see Figure 10.13). This approximate method applies to colliding BECs when (among other conditions) their relative velocity is much larger than the momentum spread in a single wavepacket. (Here, the packet velocity relative to the center of mass is $\approx 10\text{mm/s}$ in the longitudinal direction, while the initial velocity spread in this direction is $\approx 0.13\text{mm/s}$ rms). Scattering losses from the condensate wavefunction to empty modes are estimated from the GP equations by making the replacement

$$a_s \rightarrow a'_s = a_s(1 - i|\mathbf{k}|a_s), \quad (10.41)$$

where $\mathbf{k} = m\mathbf{v}^{\text{cm}}/\hbar$. a'_s enters the GP equations as

$$g = \frac{4\pi\hbar^2 a'_s}{m} \quad (10.42)$$

in the (now complex) scattering strength, which leads to particle loss from the wavefunction.

In Figure 10.13, one immediately sees that the scattering rate is significantly overestimated with this method. This difference is due to suppression of scattering to momenta lying close to the primary direction of motion $\pm\mathbf{v}^{\text{cm}}$, which is not taken into account in the imaginary scattering length approximation.

In Figure 10.14, marginals of the (positive P) simulated momentum distribution are plotted, and in particular, density of the velocity component v_1 in the longitudinal direction is shown in subplot (c). One sees that the density of scattered atoms decreases (slowly) as $|v_1| \rightarrow v_1^{\text{cm}}$, whereas if the scattering was to an isotropic spherical shell of momenta, this linear density would rise towards v_1^{cm} .

This suppression of scattering has been predicted by Bach *et al*[89] when considering the scattering from two plane waves under the Bogoliubov approximation. They found that the suppression of scattering was dependent on the ratio r_E of kinetic single-particle energy due to the plane wave motion with respect to the center of mass, and the interaction energy per particle. This ratio was

$$r_E = m|\mathbf{v}^{\text{cm}}|^2/2g\rho, \quad (10.43)$$

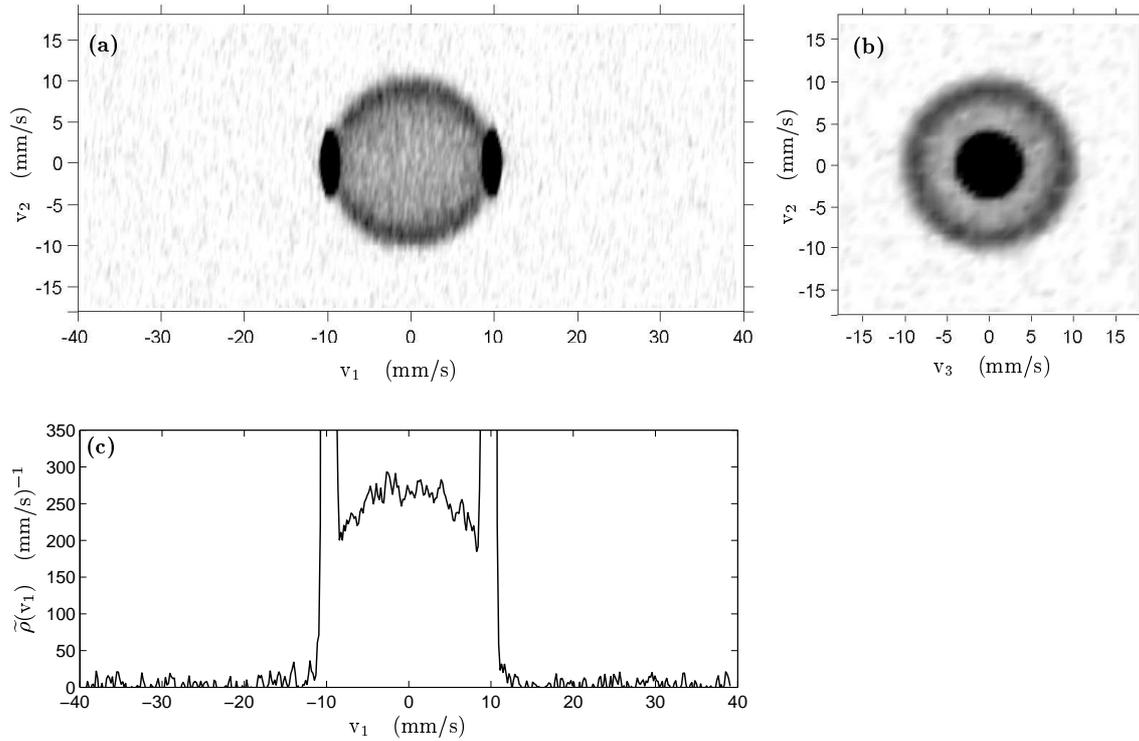


Figure 10.14: **Velocity distribution** of atoms at a time $t = 0.783\text{ms}$ after the end of the Bragg pulse that separates the two condensate wavepackets. Subplot (a) shows the combined distribution of longitudinal velocity v_1 and velocity in one axial direction v_2 , after summing over all v_3 values. Density of atoms grows as shading darkens, with no shading corresponding to vacuum. The high spatial frequency noise is due finite sample ($S = 200$) uncertainty rather than actual atoms. Subplot (b) shows the atom distribution in the axial directions, with summing over all longitudinal velocity values v_1 . Subplot (c) shows the distribution of longitudinal atom velocity $\tilde{\rho}(v_1)$. The noise is due to finite sample effects, and shows the degree of uncertainty in the calculated values. Data are from the positive P simulation described in Section 10.7.2.

where ρ was the spatial density of each plane wave component. The suppression of scattering in the $\pm\mathbf{v}^{\text{cm}}$ directions becomes less significant at large r_E , but is still strong when $r_E \approx \mathcal{O}(10)$. For the system simulated here, r_E can be estimated by using ρ_0^{TF} as an estimate for ρ in (10.43), and gives $r_E \approx 5$, well in the regime where scattering along the direction $\pm\mathbf{v}^{\text{cm}}$ direction is suppressed.

A reduction of the magnitude of scattered particle momenta from the condensate values was also predicted in Ref. [89] and is also seen in the simulation here (Note how the circle of scattered atoms in Figure 10.14(c) lies at a slightly smaller radius in momentum space than the condensates.)

10.8 Summary

The above examples demonstrate that qualitative predictions of dynamics can be made with the gauge P representation (or the special case of the positive P representation) for a wide range of many-mode interacting Bose gas systems. This includes predictions for spatial correlation functions (including non-local and/or high order) as well as local observables such as densities. Processes simulated included interaction with external trap potentials and two-body scattering under the influence of an extended interparticle potential, or of one acting only locally at each lattice point as in a Bose-Hubbard model. Spatial and momentum densities and their fluctuations have also been previously calculated for a lossy system with local interactions by Corney and Drummond [46, 15] using the un-gauged method. There, the onset of condensation in an evaporative cooling simulation was seen.

The examples in this chapter show that such first-principles simulations can be tractable even with very large numbers of modes or particles in the system given the right conditions. This is evidenced by example 4 above, where there were 1009 152 lattice points and on average 150 000 atoms.

Situations where there are several length scales of similar order, or processes of similar strength are of particular interest for first-principles simulations because it is difficult to make accurate quantitative predictions otherwise. The examples of Sections 10.6 and 10.7 shown that predictions for systems in such regimes can be made with the present method.

Improvement of simulation times by use of local diffusion gauges was seen in Section 10.4. This occurs when the lattice spacing is of the order of the (local) healing length or greater. This confirms what was expected from the analysis of Section 10.2.3.

In Section 10.2.1, an analytic estimate of useful simulation time with the ungauged positive P method was derived. This was subsequently seen to be quite accurate both in the uniform gas with local interactions, and in the more complex colliding condensates system of Section 10.7. The simulation time is seen to scale as $(\Delta V)^{1/3}$ with the effective volume of a lattice point ΔV , so that coarse-grained simulations last longer. There is a tradeoff between a fine enough lattice to include all relevant processes, and one coarse enough for the simulation to last long enough to see the desired phenomena. The analysis of Section 10.2.2 also shows that kinetic coupling between spatial modes (which can also reduce simulation time) is more dominant for fine lattices. Simulations with Gaussian extended interparticle interactions (10.38) were seen in Sections 10.5 and 10.6 to last at least several times longer than those with the same bulk interaction energy density.

Some of the phenomena predicted by the example simulations include:

- The propagation of correlation disturbances (“correlation waves”) in a coherent gas without any corresponding density waves, and at a velocity $\approx \sqrt{2}$ times faster than the speed of sound. (Section 10.4)
- The complex interaction between the breathing motion of a condensate in a trap and the interparticle correlations in the condensate. (Section 10.6)
- Bosonic enhancement of initially empty momentum modes during the collision of two condensates. (Section 10.7)
- Suppression of spontaneous scattering processes in these colliding condensates in the direction of motion. (Section 10.7.3)

Chapter 11

Thermodynamics of a one-dimensional Bose gas

The one-dimensional (1D) regime has been recently observed in trapped Bose gases[38, 39, 40], giving strong impetus to theoretical studies of such systems. This chapter describes exact predictions of spatial correlation functions and momentum densities, which have been obtained for 1D uniform Bose gases in a grand canonical ensemble by using gauge P simulations.

A part of these results have been recently published[61].

11.1 One-dimensional uniform Bose gas at finite temperature

11.1.1 Exact solutions

The interacting uniform 1D Bose gas model described by the Hamiltonian

$$\hat{H} = \frac{\hbar^2}{2m} \int \frac{\partial \hat{\Psi}^\dagger(x)}{\partial x} \frac{\partial \hat{\Psi}(x)}{\partial x} dx + \frac{g}{2} \int \hat{\Psi}^{\dagger 2}(x) \hat{\Psi}^2(x) dx \quad (11.1)$$

was found to be exactly solvable several decades ago, and is one of the few exactly solvable nontrivial many-body problems[90]. The solutions were found in the pioneering works of Girardeau[91, 92] at $g \rightarrow \infty$, Lieb and Liniger[86, 93] at $T = 0$,

and by Yang and Yang[87] for the grand canonical ensemble at finite temperature $T > 0$. See also [94, 95, 96] for reviews.

Explicitly, these solutions provide a numerical algorithm to calculate a density of states ρ_k and holes ρ_h . These can then be used to obtain some intensive physical quantities of the system such as density, energy density, and pressure. Recently there have been some further advances, especially in the zero temperature and strong interactions limit. These include $T = 0$ expressions for the tails of the momentum distribution and short range $\bar{g}^{(1)}(x)$ [97]; $T = 0$ expressions for $\bar{g}^{(2)}(0)$, $\bar{g}^{(3)}(0, 0)$, and $\bar{g}^{(1)}(x)$ [98]; and finite T values of $\bar{g}^{(2)}(0)$ [99]. These last finite temperature $\bar{g}^{(2)}(0)$ solutions indicate that there is a rich variety of correlation phenomena in the $T > 0$ regime.

Still, exact results for the vast majority of observables remain hitherto inaccessible because they cannot be obtained from the density of states and holes (or perhaps, a way to obtain them has not yet been found). This includes such basic observables as spatial correlations and even momentum distributions for $T > 0$.

The Hamiltonian (11.1) is a special case of (2.1), and as such, its grand canonical thermodynamics can be simulated using the gauge P equations (5.50) in a wide range of physical regimes. This has allowed the calculation of exact spatial correlation functions and momentum distributions for this model here — see Section 11.4.

Noteworthy in this context are also the stochastic wavefunction simulations of Carusotto and Castin[65], who calculated spatial correlation functions $\bar{g}^{(2)}(x)$ and $\bar{g}^{(1)}(x)$ of a different but related model: The canonical ensemble (i.e. with set particle number, as with all stochastic wavefunction calculations) of a uniform interacting gas in a system of small spatial extent where finite size effects were important.

11.1.2 Correspondence with trapped gases

Under present day experimental conditions, the Bose gas confinement is well approximated by a harmonic trap of transverse angular frequency ω_{\perp} and longitudinal frequency ω_0 . The correspondence between the uniform gas model, and a trapped gas experiment can be summarized in the following points:

- The atoms behave effectively as a 1D gas provided that the transverse zero point oscillations

$$l_{\perp} = \sqrt{\frac{\hbar}{m\omega_{\perp}}} \quad (11.2)$$

are much smaller than other characteristic length scales of the system such as the healing length (11.9) and the thermal de Broglie wavelength of excitations (11.8). This regime has been reached in recent experiments[38, 39, 40].

- On length scales smaller than the trap size

$$x \ll l_0 = \sqrt{\frac{\hbar}{m\omega_0}} \quad (11.3)$$

the density of the gas is slowly varying. On this and smaller scales the 1D trapped gas and uniform gas will show the same physical behavior, and so the results obtained in Section 11.4 will be directly applicable to a trapped gas while $x \ll l_0$. It is worth noting that if, instead, details of the trap are important, a gauge P simulation with the external potential V^{ext} explicitly specified can be used instead. A dynamical example of this was given in Section 10.6.

- The field interaction strength g is related to the 3D scattering length a_s by the expression (2.19), where the effective trap cross-section in the transverse direction is[100]

$$\lambda_0 = 2\pi l_{\perp}^2 = \frac{2\pi\hbar}{m\omega_{\perp}}. \quad (11.4)$$

11.1.3 Uniform gas regimes

The properties of a the finite temperature uniform 1D gas with the Hamiltonian (11.1) depend on two dimensionless parameters: The first is the **coupling parameter**¹

$$\gamma = \frac{mg}{\hbar^2\rho}, \quad (11.5)$$

¹Not to be confused with the unrelated single-boson loss rate γ of Chapters 6 and 7. γ is used here in Chapter 11 for the quantity (11.5) as this is the standard notation in the literature.

dependent on interaction strength g and density ρ . The ideal Bose gas is reached as $\gamma \rightarrow 0$, while the Tonks-Girardeau (TG) hard sphere gas occurs in the limit $\gamma \rightarrow \infty$. In this limit the particles undergo an effective “fermionization” [91, 92, 86, 93].

The second dimensionless parameter is the **reduced temperature**

$$\tilde{T} = \frac{T}{T_d} = \frac{mk_B T}{2\pi\hbar^2\rho^2}, \quad (11.6)$$

where T_d is the quantum degeneracy temperature. For $T < T_d$, the gas is denser than the quantum concentration $n_Q = \sqrt{mk_B T/2\pi\hbar^2}$ and the average spacing between particles $1/\rho$ is less than the thermal de Broglie wavelength (11.8).

Several characteristic length scales in the uniform gas are:

- Mean interparticle spacing

$$l_\rho = \frac{1}{\rho}. \quad (11.7)$$

- Thermal de Broglie wavelength

$$\lambda_T = \sqrt{\frac{2\pi\hbar^2}{mk_B T}} = \frac{l_\rho}{\sqrt{\tilde{T}}}. \quad (11.8)$$

Atoms with typical energy behave as particles on length scales $\gg \lambda_T$, as waves on $\ll \lambda_T$

- Healing length

$$\xi^{\text{heal}} = \frac{\hbar}{\sqrt{2m\rho g}} = \frac{l_\rho}{\sqrt{2\gamma}}, \quad (11.9)$$

which is the length scale of typical spatial density fluctuations.

- Degenerate phase coherence length

$$l_\phi = \frac{\rho\lambda_T^2}{2\pi} = \frac{l_\rho}{2\pi\tilde{T}}, \quad (11.10)$$

which applies when $\tilde{T} \ll 1$ and $\gamma \ll 1$. When $\tilde{T} \gg 1$, the phase coherence length is $l_\phi = \lambda_T/\sqrt{2\pi}$.

- 1D scattering length[100]

$$a_{1D} = \frac{\hbar^2}{mg} = \frac{l_\rho}{\gamma} \quad (11.11)$$

is the characteristic distance over which the influence of a single particle is felt. In the fermionized large γ regime, this is the effective size of the hard sphere due to a single particle.

The 1D gas behaves rather differently than the 2D and 3D cases, in that, for example, in a strong coupling regime the density is *low*. Also, a true BEC does not form, because long range order on length scales $\gg l_\phi$ is destroyed. This follows from the Bogoliubov k^{-2} theorem, as explained in [101, 102]. Nevertheless, at degenerate temperatures $\tilde{T} \lesssim 1$ and relatively weak couplings $l_\phi \ll \xi^{\text{heal}}$ and $\gamma \ll 1$, a finite-range quasicondensate with phase coherence exists.

11.2 Simulation details

11.2.1 Overview

Integration of the gauge P stochastic differential equations (5.50) proceeds in a similar fashion to the dynamics calculations of Chapter 10. These are based on the master equation formulation of the grand canonical ensemble, as described in Section 2.6. A split-step algorithm is used, as described in Section 10.1 with integration using the semi-implicit method described in Appendix B. To exclude potential finite-sample bias of the kind described in Appendix A, the variance of the logarithms of quantities to be averaged for observable estimates was also tracked.

To simulate the *uniform* gas, care needed to be taken to avoid finite size effects. This can usually be achieved for all practical purposes by choosing a lattice spacing several times smaller than the smallest length scale of Section 11.1.3, and a number of lattice points large enough so that the total spatial extent is several times larger than the largest length scale. Operationally, the lattice resolution was increased in normal and/or momentum space until no changes with lattice resolution were seen in the observables of interest.

There are several further issues that arise only in the thermodynamics calculations and require more detailed discussion:

11.2.2 Drift Gauge

As was discussed in more detail in Section 9.2.1, the un-gauged positive P equations are not suitable for simulation because of severe moving singularities as $\text{Re}\{\alpha_{\mathbf{n}}\beta_{\mathbf{n}}\}$ takes on negative values. This behavior appears in the many-mode case as well, and leads to rapid onset of spiking, as well as possible systematic errors after only a short simulation time.

The radial drift gauge (9.17) acting locally on each mode was used to remove the moving singularities in the same manner as for the single-mode system of Chapter 9. Explicitly, when adapted to the many-mode situation this gauge is

$$\mathcal{G}_{\mathbf{n}} = i\sqrt{2\hbar\chi}(\alpha_{\mathbf{n}}\beta_{\mathbf{n}} - |\alpha_{\mathbf{n}}\beta_{\mathbf{n}}|), \quad (11.12)$$

where χ is given by (2.18), as usual.

Spiking abates, and boundary term errors are not seen.

11.2.3 Importance sampling

The simulated equations (5.50) include evolution of both amplitudes $\alpha_{\mathbf{n}}$ and weight $\Omega = e^{z_0}$. Deterministic evolution of the weight is a new feature compared to dynamics and can cause sampling problems at nonzero times $\tau > 0$.

It is convenient to initialize in Fourier space. Here the variables $\tilde{\alpha}_{\tilde{\mathbf{n}}}(\tau)$ and $\tilde{\alpha}'_{\tilde{\mathbf{n}}}(\tau)$ are given by (10.2), and bold quantities indicate vectors with M elements — one per mode. For an M -mode system of length L , the standard initial distribution (5.59) becomes in Fourier space

$$P_G(\tilde{\boldsymbol{\alpha}}, \tilde{\boldsymbol{\alpha}}', z_0) = \delta^2(z_0)\delta^{2M}(\tilde{\boldsymbol{\alpha}}' - \tilde{\boldsymbol{\alpha}}) \prod_{\tilde{\mathbf{n}}} \frac{1}{\pi\tilde{n}_0} \exp\left(\frac{-|\tilde{\alpha}_{\tilde{\mathbf{n}}}|^2}{\tilde{n}_0}\right), \quad (11.13)$$

where $\tilde{n}_0 = L^2\bar{n}_0/2\pi M$ is the mean of all $|\tilde{\alpha}_{\tilde{\mathbf{n}}}|^2$ at $\tau = 0$. The lattice factors arise because the number of particles per Fourier mode is $\propto \frac{L^2}{2\pi M}$. A sample is generated

by the simple procedure

$$\tilde{\alpha}_{\tilde{\mathbf{n}}} = \sqrt{\tilde{n}_0} \eta_{\tilde{\mathbf{n}}} = \tilde{\alpha}_{\tilde{\mathbf{n}}}^{(0)} \quad (11.14a)$$

$$\tilde{\alpha}'_{\tilde{\mathbf{n}}} = \tilde{\alpha}_{\tilde{\mathbf{n}}}^{(0)} \quad (11.14b)$$

$$z_0 = 0, \quad (11.14c)$$

where the $\eta_{\tilde{\mathbf{n}}}$ are independent complex Gaussian noises of variance unity.

The problem is that this is a good sample of the $\tau = 0$ distribution, but not necessarily so at later times. This is so even for an ideal gas when two-body interactions are absent. In the simplest case of the ideal gas with no external potential $V^{\text{ext}} = 0$, the equations of motion (5.50) can be solved exactly in Fourier space, which gives

$$\tilde{\alpha}_{\tilde{\mathbf{n}}}(\tau) = \tilde{\alpha}_{\tilde{\mathbf{n}}}^{(0)} e^{h_{\tilde{\mathbf{n}}}(\tau)\tau} \quad (11.15a)$$

$$\tilde{\alpha}'_{\tilde{\mathbf{n}}}(\tau) = \tilde{\alpha}_{\tilde{\mathbf{n}}}^{(0)} \quad (11.15b)$$

$$z_0(\tau) = \frac{L^2}{2\pi M} \sum_{\tilde{\mathbf{n}}} \left\{ |\tilde{\alpha}_{\tilde{\mathbf{n}}}(\tau)|^2 - |\tilde{\alpha}_{\tilde{\mathbf{n}}}^{(0)}|^2 \right\}, \quad (11.15c)$$

where

$$h_{\tilde{\mathbf{n}}}(\tau) = \mu(\tau) - \frac{\hbar^2 k_{\tilde{\mathbf{n}}}^2}{2m} \quad (11.16)$$

is the Gibbs factor exponent.

For the ideal gas, the exact finite temperature solution consists of an independent thermal state with mean occupation $\bar{n}_{\tilde{\mathbf{n}}}$ for each Fourier mode. An unbiased sample of this is

$$\tilde{\alpha}_{\tilde{\mathbf{n}}}(\tau) = L \sqrt{\frac{\bar{n}_{\tilde{\mathbf{n}}}(\tau)}{2\pi M}} \eta_{\tilde{\mathbf{n}}} \quad (11.17a)$$

$$\tilde{\alpha}'_{\tilde{\mathbf{n}}}(\tau) = \tilde{\alpha}_{\tilde{\mathbf{n}}}(\tau) \quad (11.17b)$$

$$z_0 = 0, \quad (11.17c)$$

where for the ideal Bose gas, the mode occupations obey the Bose-Einstein distribution

$$\bar{n}_{\tilde{\mathbf{n}}}(\tau) = \bar{n}_{\tilde{\mathbf{n}}}^{\text{BE}}(\tau) = \frac{1}{e^{-h_{\tilde{\mathbf{n}}}(\tau)\tau} - 1}. \quad (11.18)$$

The actual simulated $\tilde{\alpha}_{\tilde{\mathbf{n}}}(\tau)$ in (11.15a) are not necessarily anywhere near (11.17a) with Bose-Einstein occupations (11.18), so the nonzero variable weight (11.15c) is needed to compensate for this.

This is fine in the limit of infinite samples $\mathcal{S} \rightarrow \infty$, but in practice the distribution of the weight $\Omega = e^{z_0}$ can be badly sampled once the variance in z_0 becomes too large. Since $|\tilde{\alpha}_{\mathbf{n}}^{(0)}|^2$ is a Gaussian, then z_0 also will have a similar nature and the usual limit (7.43) for exponentials of a Gaussian applies: The weights (and hence, observable estimators, which all average quantities $\propto e^{z_0}$) are well sampled only while

$$\text{var} [\text{Re} \{z_0\}] \lesssim \mathcal{O}(10). \quad (11.19)$$

A method known as **importance sampling** offers a way to alleviate this weighting problem if one is primarily interested in the gas around a particular target time τ_T (i.e. temperature $1/k_B\tau$ and chemical potential $\mu(\tau_T)$). The importance sampling approach is widely used for weighted stochastic integration, and a comprehensive discussion of this for general situations can be found in [103], Section 7.8. In the case considered here, the idea is as follows:

In the gauge P representation, the density matrix is written as

$$\hat{\rho}_u(\tau) = \int P_G(\tau, \mathbf{z}, z_0) \hat{\Lambda}(\mathbf{z}, z_0) d^{4M} \mathbf{z} d^2 z_0, \quad (11.20)$$

where \mathbf{z} stands for all the phase-space variables $\tilde{\alpha}, \tilde{\alpha}'$. The simulation here uses standard drift gauges, so that the evolution of the weight z_0 is actually deterministically dependent on the evolution of the \mathbf{z} , as given in (4.43). In the initial distribution (11.13), $z_0 = 0$ is independent of \mathbf{z} . Together, these properties allow the \mathbf{z} and z_0 evolutions to be separated. Explicitly,

$$z_0(\tau, \mathbf{z}) = \int_0^\tau \frac{dz_0(\tau', \mathbf{z}(\tau'))}{d\tau'} d\tau', \quad (11.21)$$

where dz_0 is given by (4.43). After defining $\hat{\Lambda}(\mathbf{z}, z_0) = e^{z_0} \hat{\underline{\Lambda}}(\mathbf{z})$, this allows us to rewrite (11.20) as

$$\hat{\rho}_u(\tau) = \int \underline{P}_G(\tau, \mathbf{z}) \hat{\underline{\Lambda}}(\mathbf{z}) e^{z_0(\tau, \mathbf{z})} d^{4M} \mathbf{z}, \quad (11.22)$$

where the underlined quantities depend only on the amplitude variables \mathbf{z} .

The multiplying factor of the operator kernel $\hat{\underline{\Lambda}}$ in (11.22) is $\underline{P}_G e^{z_0}$, and in fact there is a whole range of possible initial distributions provided that this factor

remains the same. In particular we could start with a different initial distribution of the amplitude variables $\underline{P}_{G_0}(0, \mathbf{z})$ provided that this is compensated for in the initial weights by

$$z_0 = \int_0^\tau \frac{dz_0(\tau', \mathbf{z}(\tau'))}{d\tau'} d\tau' + \log \underline{P}_G(0, \mathbf{z}) - \log \underline{P}_{G_0}(0, \mathbf{z}). \quad (11.23)$$

Then (and this is the crucial point), if we could choose some initial distribution \underline{P}_{G_0} that leads to only a small spread in z_0 at target time τ_T , then the simulation would be well sampled there. (It would also be sampled much worse at $\tau = 0$, but this doesn't bother us). This then is the sampling according to importance (i.e. according to the situation at τ_T).

For the purposes of the simulations reported in this chapter, a fairly crude yet effective importance sampling was applied: At relatively weak coupling, a very rough but useful estimate of the state is that the Fourier modes are uncoupled, and thermally distributed with mean occupations $\bar{n}_{\tilde{\mathbf{n}}}(\tau_T)$. The equal-weight samples (11.17) correspond to the distribution

$$\underline{P}_G^{\text{est}}(\tau_T, \mathbf{z}) \propto \delta^{2M}(\tilde{\boldsymbol{\alpha}}' - \tilde{\boldsymbol{\alpha}}) \exp \left[-\frac{2\pi M}{L^2} \sum_{\tilde{\mathbf{n}}} \frac{|\tilde{\alpha}_{\tilde{\mathbf{n}}}|^2}{\bar{n}_{\tilde{\mathbf{n}}}(\tau_T)} \right] \quad (11.24)$$

with appropriate normalization. This is not yet quite what is wanted because the desired sampling distribution \underline{P}_{G_0} is to be at $\tau = 0$. An estimate of the initial distribution that leads to $\underline{P}_G^{\text{est}}(\tau, \mathbf{z})$ can be obtained by evolving (11.24) back in time with only kinetic interactions. This is again rather rough, since deterministic interaction terms $\propto g$ have been omitted (not to mention noise), but is simple to carry out, and was found to be sufficient for the purposes of the preliminary calculations presented here. One obtains the sampling distribution

$$\underline{P}_{G_0}(0, \mathbf{z}) \propto \delta^{2M}(\tilde{\boldsymbol{\alpha}}' - \tilde{\boldsymbol{\alpha}}) \exp \left[-\frac{2\pi M}{L^2} \sum_{\tilde{\mathbf{n}}} \frac{|\tilde{\alpha}_{\tilde{\mathbf{n}}}|^2}{\bar{n}_{\tilde{\mathbf{n}}}^{(0)}} \right], \quad (11.25)$$

where

$$\bar{n}_{\tilde{\mathbf{n}}}^{(0)} = \bar{n}_{\tilde{\mathbf{n}}}(\tau_T) \exp[-\lambda_n - h_{\tilde{\mathbf{n}}}(\tau_T)\tau_T]. \quad (11.26)$$

(Note that $\lim_{\tau \rightarrow 0} [h_{\tilde{\mathbf{n}}}(\tau)\tau] = -\lambda_n$). Using (11.23), the initial variables are then

sampled according to

$$\tilde{\alpha}_{\tilde{\mathbf{n}}}(0) = L \sqrt{\frac{\bar{n}_{\tilde{\mathbf{n}}}^{(0)}}{2\pi M}} \eta_{\tilde{\mathbf{n}}} \quad (11.27a)$$

$$\tilde{\alpha}'_{\tilde{\mathbf{n}}}(0) = \tilde{\alpha}_{\tilde{\mathbf{n}}}(0) \quad (11.27b)$$

$$z_0(0) = \frac{L^2}{2\pi M} \sum_{\tilde{\mathbf{n}}} |\tilde{\alpha}_{\tilde{\mathbf{n}}}(0)|^2 \left(\frac{1}{\sqrt{\bar{n}_{\tilde{\mathbf{n}}}^{(0)}}} - \frac{1}{\bar{n}_0} \right). \quad (11.27c)$$

For most of the simulations of Section 11.4, the target Fourier space distribution $\bar{n}_{\tilde{\mathbf{n}}}(\tau_T)$ was just taken to be the plain ideal gas Bose-Einstein distribution (11.18). However, once the interactions become significant, this is not satisfactory because the actual interacting gas momentum distribution is far from the Bose-Einstein form (See e.g. Figure 11.8). A better choice of $\bar{n}_{k_{\tilde{\mathbf{n}}}}(\tau_T)$ is the density of states function ρ_k of the exact Yang & Yang solution[87]. In practice, to simplify the calculation, an estimate of ρ_k was used instead. This was obtained by making a least squares (unweighted) fit of parameters C_{est} and σ_{est} in the expression

$$\bar{n}_{\tilde{\mathbf{n}}} = \bar{n}_0 e^{h_{\tilde{\mathbf{n}}}(\tau_T)\tau_T + \lambda_{\tilde{\mathbf{n}}}} + C_{\text{est}} e^{-k_{\tilde{\mathbf{n}}}^2 \tau_T / 2\sigma_{\text{est}}^2} \quad (11.28)$$

to ρ_k . This form was chosen because in the tails of the distribution in Fourier space when $k_{\tilde{\mathbf{n}}}^2 \tau_T \gg 1$, $\bar{n}_{\tilde{\mathbf{n}}}^{(0)} \approx \bar{n}_0$, and so the initial weights $z_0(0)$ in (11.27c) are not dependent on the momentum cutoff.

11.2.4 Momentum cutoff

In a dynamical simulation, one requires that all occupied momentum modes are contained within the momentum cutoff $k_d^{\text{max}} = \pi/\Delta x_d$. Strictly speaking, this cannot ever be satisfied in these thermodynamic simulations because the $T \rightarrow \infty$ ideal gas grand canonical ensemble is used as a starting condition. This has all momentum modes occupied up to infinite energy. This turns out not to be a practical problem because the precision of calculated results is limited anyway by the finite sample size \mathcal{S} . It is then sufficient to increase the k_d^{max} until all observable estimates calculated are invariant with k_d^{max} to the precision achievable with a given ensemble size. This was carried out.

11.2.5 Chemical potential at intermediate temperatures

As was noted in Section 9.4, if one is primarily interested in the behavior of the system at (or around) a given temperature τ_T , and chemical potential μ , then the values of μ at intermediate times $\tau < \tau_T$ can in principle be chosen at will.

In practice, some choices lead to better precision per ensemble size than others. The analysis in Section 9.4 indicated that the single-mode simulations led to the best precision when the “effective” chemical potential $\mu_e = \frac{\partial \mu \tau}{\partial \tau}$ was roughly constant during the simulation. It was assumed that a similar dependence occurs in the many-mode system, and so, accordingly a constant μ_e was chosen so that at τ_T the chemical potential becomes $\mu(\tau_T)$. Using (9.2), the constant μ_e assumption leads to

$$\mu_e = \mu(\tau_T) + \frac{\lambda_n}{\tau_T}. \quad (11.29)$$

It remains to choose the initial density $\rho_0 = M\bar{n}_0(\lambda_n)/L$ (with initial occupation \bar{n}_0 given by (2.33)). It was found that for target densities $\rho(\tau_T)$ a choice of initial density

$$\rho_0 = \mathcal{O}\left(\frac{\rho(\tau_T)}{10}\right) \quad (11.30)$$

gave the best precision at τ_T in most cases. Details varied depending on system parameters, although a clear tradeoff between two noise generating processes was seen:

1. When ρ_0 was too large, a lot of added randomness is introduced into z_0 at early times, and precision is lost very rapidly once $\text{var}[\text{Re}\{z_0\}] \gtrsim \mathcal{O}(10)$, as discussed in Section 7.4 and Appendix A. This randomness arises largely independently of gauges or noise since the spread of $\alpha_{\mathbf{n}}$ in the initial thermal state is proportional to $\sqrt{\rho_0}$, and this then leads directly to a spread in the deterministic evolution terms of dz_0 .
2. When ρ_0 is too small, excessive spread in the amplitude variables arises as well (which also then feeds into the log-weight z_0).

11.2.6 Scaling of weight variance

For uniform gas calculations it was found to be desirable to use the smallest system size that does not introduce finite size effects. The reason is that (apart from gauge-dependent terms), the weight $\Omega(\tau) = \exp \int_0^\tau \frac{dz_0}{d\tau'} d\tau'$ has the form of a Gibbs factor $\exp((\mu\bar{N} - E)/k_B T)$. This can be verified by inspection of the equations (5.50). The magnitude but also the *spread* in log-weights z_0 will thus grow with energy E and particle number \bar{N} . Since $\text{var}[\text{Re}\{z_0\}]$ must be $\lesssim \mathcal{O}(10)$ for good sampling of the distribution, this sets an upper limit on how large a system can be simulated.

11.3 Physical regimes simulated

11.3.1 Parameter targeting

The procedure to calculate properties at given target gas parameters γ and \tilde{T} was as follows:

1. Three physical quantities specify units. For example, m , \hbar , and $k_B T = 1/\tau_T$ can be chosen unity. For the purposes of investigating gases at a given γ and \tilde{T} , these can be chosen arbitrarily and properties of the whole family of physical gases with those γ and \tilde{T} follow by scaling.
2. The target time τ_T is determined by the choice of units above, and the remaining essential parameters required to simulate (5.50) are g and $\mu(\tau_T)$. The coupling strength is determined by

$$g = \hbar\gamma \sqrt{\frac{k_B T}{2\pi m \tilde{T}}}, \quad (11.31)$$

while the required target density is

$$\rho(\tau_T) = \sqrt{\frac{m k_B T}{2\pi \hbar^2 \tilde{T}}}. \quad (11.32)$$

3. The target density is not a direct input parameter into the simulation, but it has a one-to-one correspondence with $\mu(\tau_T)$ (while in dimensionless units).

The Yang & Yang exact solutions[87] provide an algorithm to calculate $\rho(T, \mu, g)$, and the chemical potential needed to obtain the desired density (11.32) was calculated by numerically inverting this relation.

11.3.2 Regimes attained

A wide range of physical regimes were simulated directly with the procedure described in the previous sections. In terms of the two characteristic dimensionless gas parameters γ and \tilde{T} , the physical regions that have been explored are shown in Figure 11.1.

In terms of the physical regimes classification of Kheruntsyan *et al*[99] based on $\bar{g}^{(2)}(0)$ behavior (These are described in more detail in Section 11.4.1), the accessible regions are both the **quantum degenerate decoherent** and **classical decoherent** regimes, as well as the **nondegenerate fermionized** (strong coupling) regime. Simulations also access **transition regions** between these. The interesting transition region $\gamma \approx \tilde{T} \approx \mathcal{O}(1)$ where several length scales are of the same order is also accessible in part. The lower limit of accessible \tilde{T} does not appear to be a hard limit, and more sophisticated importance sampling techniques and/or choices of $\mu_e(\tau)$ may be able to make some further inroads into lower temperature regions.

The basic limiting factor is growth of the variance of the real and/or imaginary parts of the log-weight z_0 . As explained Section 11.2.6, there is a tradeoff between the variance of z_0 and the system size (e.g. lattice length L if the density is set), but there is a limit on how small the system can be made if it is to correspond to the uniform gas model. L must be significantly larger than the longest relevant length scale. However, this limit is only applicable if an unconstrained uniform gas is required. The behavior of a gas in a finite box, torus, or trap is an easier simulation and can reach temperatures not accessible for uniform gas simulations if L can be made small enough to control the growth of $\text{var}[z_0]$. The previously mentioned calculations of by Carusotto and Castin[65] for the different but related closed system model were made in such a “low temperature with finite-size effects” regime, with $\gamma \approx \mathcal{O}(0.001) - \mathcal{O}(0.01)$, $\tilde{T} \approx \mathcal{O}(0.001) - \mathcal{O}(0.005)$.

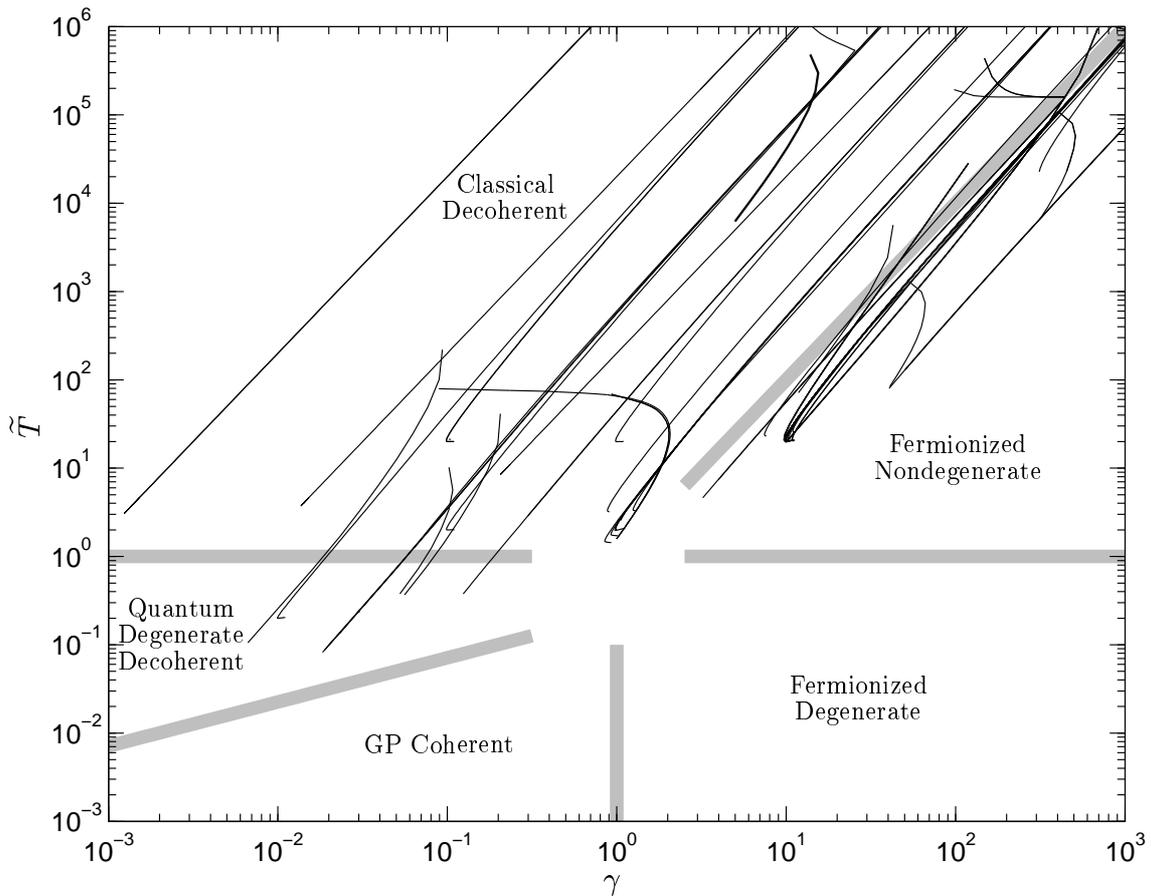


Figure 11.1: **Physical regimes attained** for the uniform interacting 1D Bose gas. DARK SOLID lines show the paths in γ , \tilde{T} parameter space taken by simulations for the τ range where useful precision in $\bar{g}^{(2)}(x)$ was obtained. Paths shown include simulations with various forms of μ_e : either given by (11.29), or by $2e^{\mu\tau} = e^{-\lambda_n} + e^{\mu(\tau_T)\tau_T} + (e^{\mu(\tau_T)\tau_T} - e^{-\lambda_n}) \cos[4\pi(\tau - \tau_T)/3\tau_T]$ described in [61]. The constant form (11.29) was found to be as good or superior at reaching low temperatures. THICK SHADED bars indicate approximate transition regions between physical regimes based on the classification by Kheruntsyan *et al*[99] on the basis of $\bar{g}^{(2)}(0)$.

11.4 Observable predictions

11.4.1 Spatial correlation functions

A classification of uniform gas behaviors based on the local second order correlation function $\bar{g}^{(2)}(0)$ has been determined from the Yang & Yang solution by Kheruntsyan *et al*[99]. In this chapter, the finite temperature behavior of $\bar{g}^{(1)}(x)$, $\bar{g}^{(2)}(x)$, and $\bar{g}^{(3)}(x, y)$ has been calculated in several of these regimes and is shown in Figures 11.2

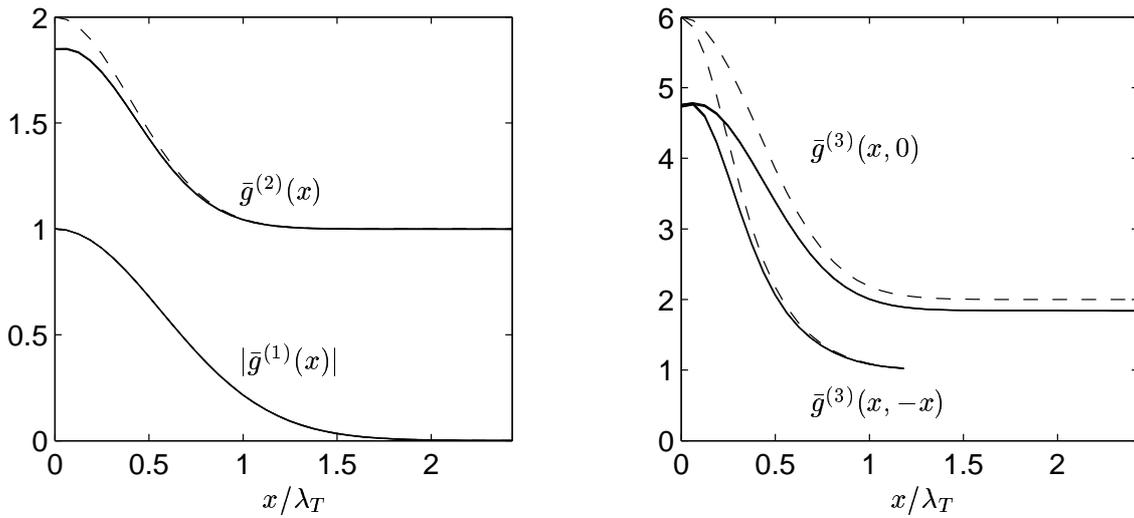


Figure 11.2: **Correlation functions in $\gamma = 5$, $\tilde{T} = 1000$ decoherent classical gases.** Exact results are shown as SOLID lines, with triple lines indicating uncertainty (mostly not visible at this scale). DASHED lines show the ideal Bose gas values for comparison ($g^{(1)}(x)$ is indistinguishable). Here, $l_\rho = 10\pi\lambda_T$. $\mathcal{S} = 10^5$, $M = 80$, $L = 5\lambda_T$.

to 11.7. Exact results are also compared to ideal gas values. The simulated physical regimes were:

Classical decoherent regime ($\tilde{T} \gg \max[1, \gamma^2]$)

This regime occurs while the thermal wavelength dominates the small scale behavior of the particles. In particular, while it is smaller than the interparticle spacing l_ρ and the effective scattering length a_{1D} . Correlation functions in a fairly strongly interacting part of this regime are shown in Figure 11.2. Correlations decay on the length scale of λ_T , with only a small reduction in close range correlation with respect to the ideal gas.

Quantum degenerate decoherent regime ($\sqrt{\gamma/2\pi^2} \ll \tilde{T} \ll 1$)

Here the gas is degenerate, but phase coherence exists only on length scales l_ϕ shorter or of the same order as the healing length ξ^{heal} . This is not enough for any significant quasicondensate to form. Example correlations shown in Figure 11.3. The short-range correlations are still largely thermal $g^{(2)} > 1$, as for the classical gas, but a

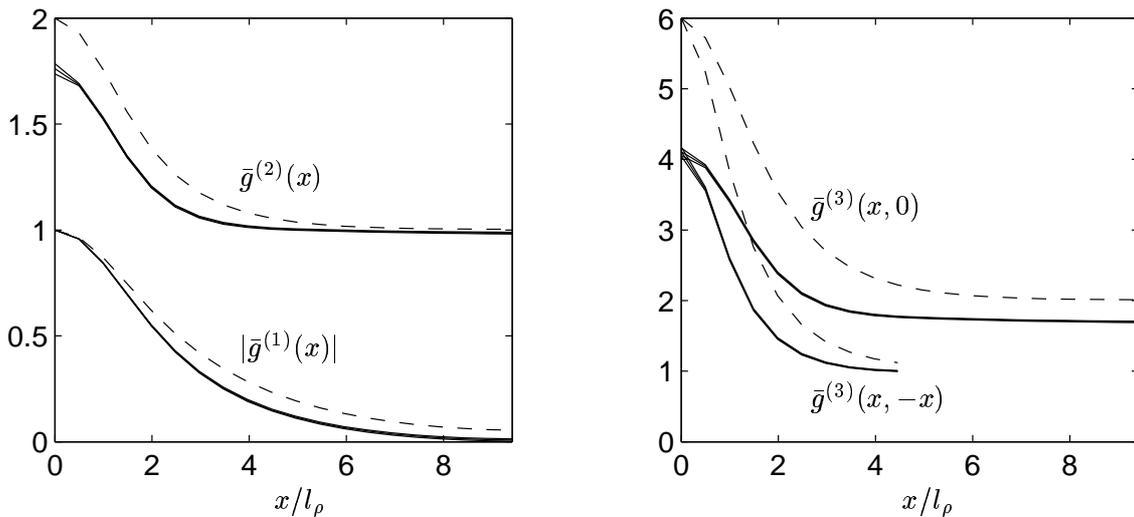


Figure 11.3: **Correlation functions in $\gamma = 0.03$, $\tilde{T} = 0.236$ decoherent quantum degenerate gases.** Exact results are shown as SOLID lines, with triple lines indicating uncertainty (mostly not visible at this scale). DASHED lines show the ideal Bose gas values (for comparison). Here, $\lambda_T = 2.06l_\rho$, $S = 10^5$, $M = 40$, $L = 5.14l_\rho$.

systematic reduction of the range of all multi-particle correlations in comparison to the ideal gas is seen.

Fermionized nondegenerate regime ($1 \gg \tilde{T} \ll \gamma^2$)

Here the fermion-like nature of strongly repulsive atoms dominates the small-scale behavior of the gas. The effective scattering length a_{1D} is much smaller than both the thermal wavelength and the interparticle spacing l_ρ leading to behavior similar to a hard-sphere (Tonks-Girardeau) model. In Figure 11.4, one sees strong antibunching on length scales a_{1D} , several times smaller than the phase coherence length. Three-particle correlations are also very strongly reduced at small distances $\lesssim a_{1D}$, with $|\bar{g}^{(3)}(0, 0)| \lesssim 0.2$ in the example at $\gamma = 300$, $\tilde{T} = 6000$ shown in Figure 11.4. This will lead to a strong reduction in three-particle inelastic scattering losses despite the very high temperature.

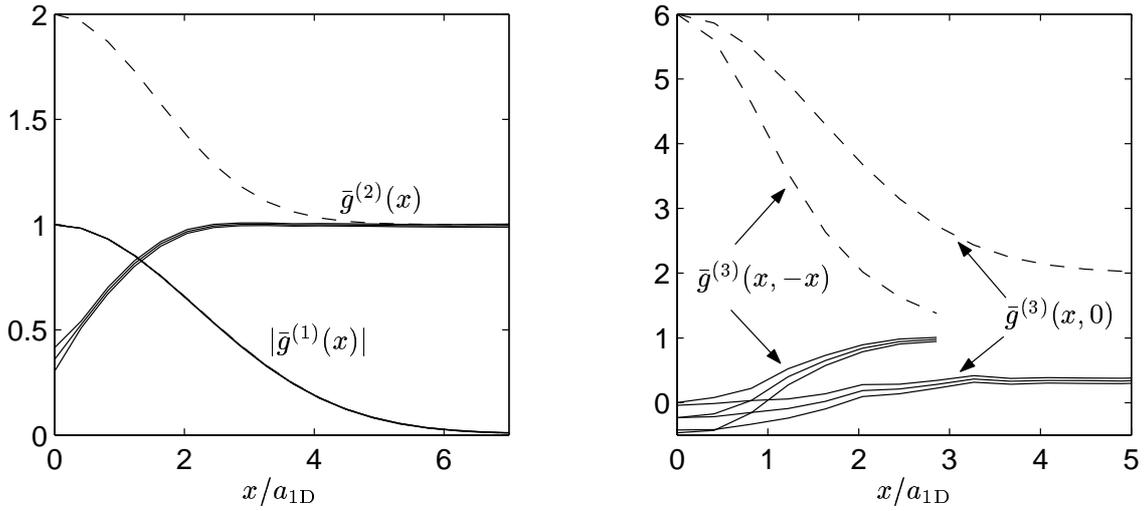


Figure 11.4: **Correlation functions in $\gamma = 300$, $\tilde{T} = 6000$ strongly fermionized nondegenerate gases.** Exact results are shown as SOLID lines, with triple lines indicating uncertainty. DASHED lines show the ideal Bose gas values for comparison ($g^{(1)}(x)$ is indistinguishable). Here, $l_p = 300a_{1D}$, $\lambda_T = 3.873a_{1D}$, and $\xi^{\text{heal}} = 12.24a_{1D}$. $S = 10^5$, $M = 36$, $L = 14.7a_{1D}$.

Transition regime with atom pairing ($\lambda_T \approx a_{1D} \approx \mathcal{O}(1)$)

Interesting new physical phenomena were seen in simulations of the transition regime where $\gamma \approx \tilde{T} \approx \mathcal{O}(1)$. Here all the length scales of the system noted in Section 11.1.3 are of similar order, the exact relationship between them depending fairly sensitively on γ and \tilde{T} . Because of the presence several competing processes, approximate methods do not usually give accurate quantitative predictions.

The transition between the nondegenerate fermionized gas and a decoherent quantum degenerate gas is shown in the sequence of Figures 11.4, 11.5, 11.6, and 11.3. Details of the three-particle correlations in the still-fermionized part of the transition are shown in Figure 11.7.

There is a parameter regime where a peak in correlations is seen at finite separations, indicating the appearance of **pairing between atoms**. The pairing arises at interparticle distances of

$$l_{\text{pair}} \approx \frac{\lambda_T}{2} \quad (11.33)$$

when λ_T and a_{1D} are of the same order. It appears to be a consequence of competition between repulsion, which promotes relative antibunching at short lengths, and

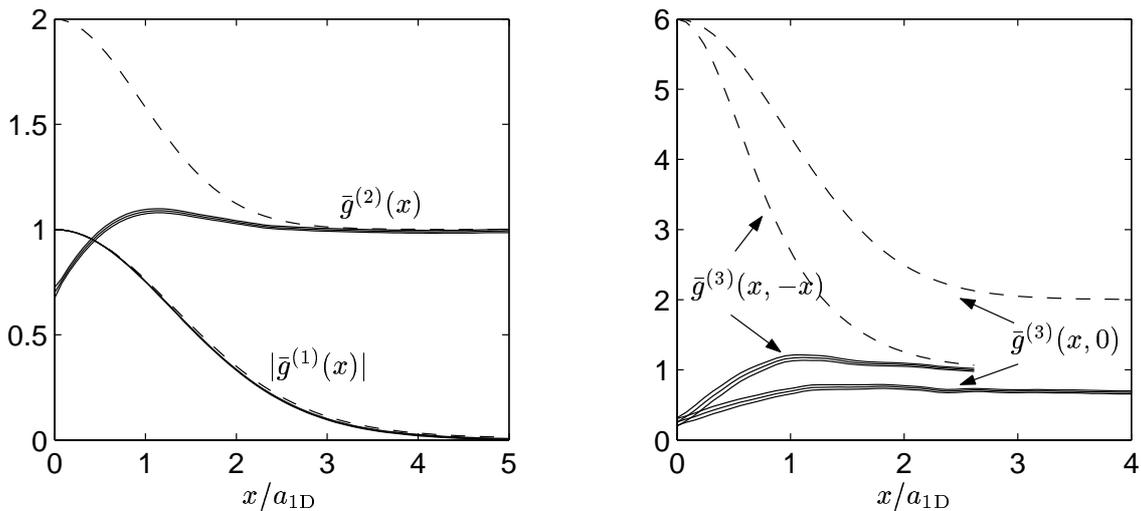


Figure 11.5: **Correlation functions in $\gamma = 10$, $\tilde{T} = 20$ fermionized nondegenerate gases.** Exact results are shown as SOLID lines, with triple lines indicating uncertainty. DASHED lines show the ideal Bose gas values for comparison. Here, $l_p = 10a_{1D}$, $\lambda_T = \xi^{\text{heal}} = 2.24a_{1D}$. $S = 10^5$, $M = 135$, $L = 10.7a_{1D}$.

the inherent bunching in a thermal field on length scales $\lesssim \lambda_T$. Enhanced pairing at distances l_{pair} is seen to occur both in situations where the point density correlations $\bar{g}^{(0)}$ indicate local bunching (> 1 — e.g. Figure 11.6) or antibunching (< 1 — e.g. Figure 11.5). Enhanced three-particle correlations at distances $\approx l_{\text{pair}}$ were also seen.

11.4.2 Momentum distributions

An example of a calculated momentum distribution in the transition region is shown in Figure 11.8. At low momenta $k \lesssim k_T = \lambda_T/\sqrt{2\pi}$, the distribution is seen to be intermediate between the ideal gas Bose-Einstein distribution and the Fermi-Dirac distribution for the same μ and T . One also sees that the density of states function ρ_k from the exact Yang & Yang solution[87] differs significantly from the actual distribution of momenta.

At intermediate momenta $k_T \lesssim k \lesssim 2k_T$, a depletion in comparison to the ideal gas is seen, while very high momentum particles $k \gtrsim 2k_T$ are more common than in the ideal gas. This regime is shown in Figure 11.9, and one sees that the high

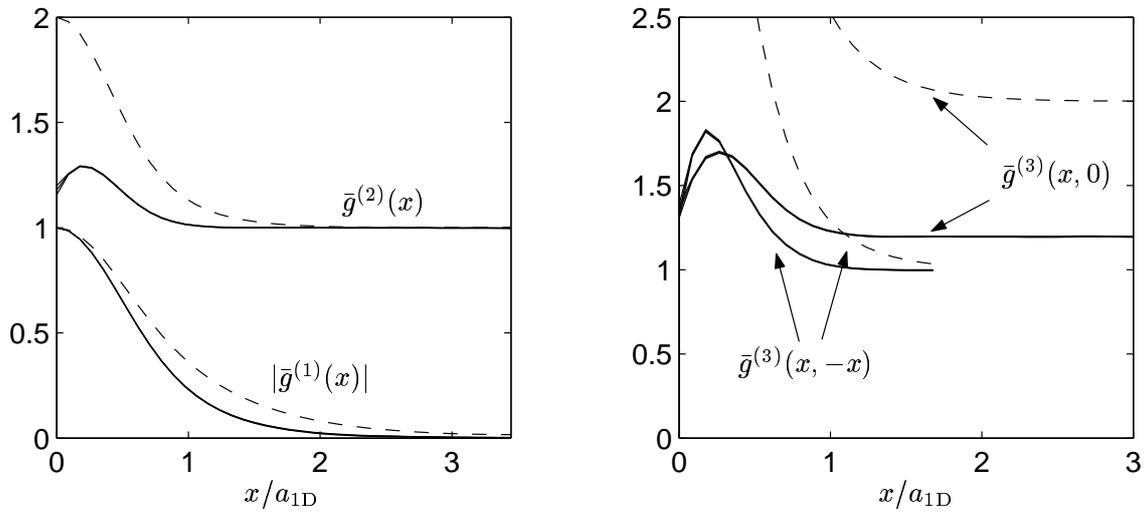


Figure 11.6: **Correlation functions in $\gamma = 1$, $\tilde{T} = 1.6$ gases.** Exact results are shown as SOLID lines, with triple lines indicating uncertainty (mostly not visible at this scale). DASHED lines show the ideal Bose gas values for comparison. Here, $l_p = a_{1D}$, $\lambda_T = 0.793a_{1D}$, and $\xi^{\text{heal}} = 0.708a_{1D}$. $\mathcal{S} = 10^5$, $M = 80$, $L = 7.07a_{1D}$.

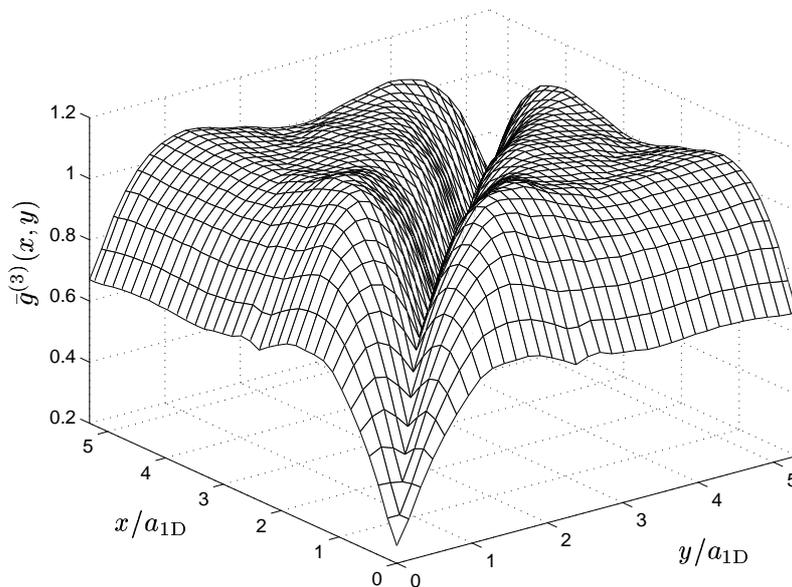


Figure 11.7: **General three-particle correlation function in $\gamma = 10$, $\tilde{T} = 20$ fermionized nondegenerate gases.** Simulation details as in Figure 11.5.

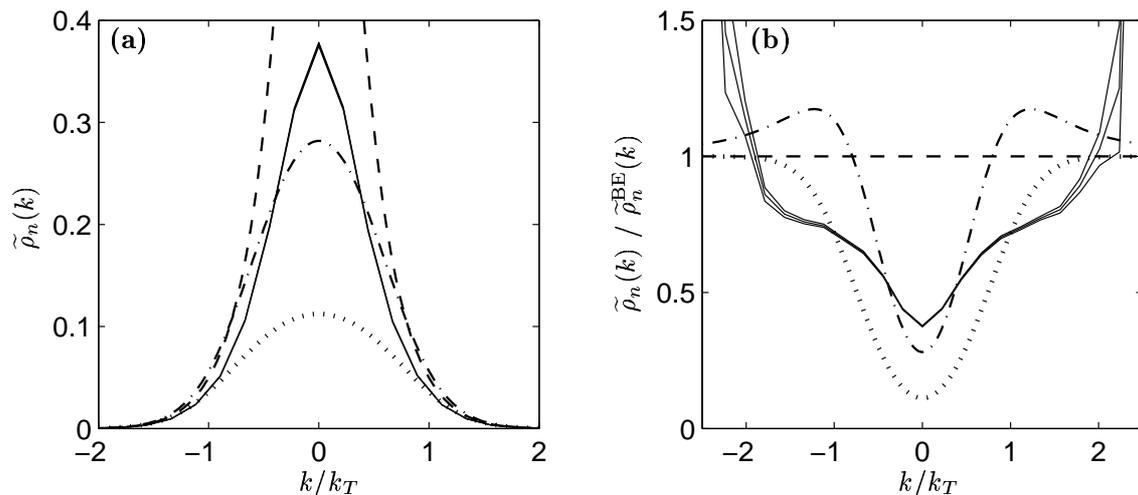


Figure 11.8: **Momentum distribution in the $\gamma = 1$, $\tilde{T} = 1.6$ gases.** SOLID LINES indicate the results (with error bars) of a calculation with parameters as in Figure 11.6. Wave vectors k are scaled with respect to the thermal momentum $k_T = \lambda_T/\sqrt{2\pi}$, while the momentum distribution is normalized so that $\int \tilde{\rho}_n(k)dk = 1$. Also shown for comparison are the Bose-Einstein ideal gas distribution (DASHED), Fermi-Dirac distribution (DOTTED), and density of states ρ_k from the Yang & Yang exact solution[87] (DASH-DOTTED), all calculated at the same T , μ , and g . These are normalized with the same factor as $\tilde{\rho}_n$, rather than to unity. Subplot (b) shows the ratio with the ideal gas distribution $\tilde{\rho}_n^{\text{BE}}$ to resolve the high k behavior.

momentum distribution approaches the $T = 0$ power-law decay $\tilde{\rho}(k) \propto k^{-4}$ found by Olshanii and Dunjko[97] from the exact Lieb and Liniger solution[86].

11.4.3 Comparison to exact Yang & Yang solution

As a test, these calculations have been compared with the density and energy per particle calculated from the Yang & Yang exact solution[87]. This is shown in Figure 11.10. At and around the importance sampling target time τ_T , the agreement is excellent. As expected, at $\tau \ll \tau_T$ the true distribution is badly sampled due to excessive spread in z_0 . The indicator used to catch this problem is $\text{var}[\text{Re}\{z_0\}]$ (described in Section 10.1), which should be $\lesssim \mathcal{O}(10)$ to be confident of accuracy. For the case shown in Figure 11.10, this indicates accurate calculations at $\tau \geq 0.89\tau_T$, which is quite conservative. A similar comparison with $\bar{g}^{(2)}(0)$ values calculated[99] from the exact solution also shows agreement.

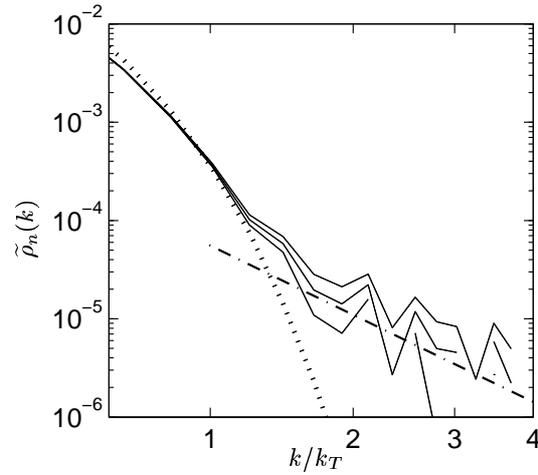


Figure 11.9: **Far tails of the momentum distribution in the $\gamma = 1$, $\tilde{T} = 1.6$ gases.** Parameters as in Figure 11.8. SOLID triple lines indicate calculated normalized momentum distribution with error bars, DOTTED the Bose-Einstein ideal gas momentum distribution. The DOT-DASHED line shows the $T = 0$ high k asymptotic behavior[97].

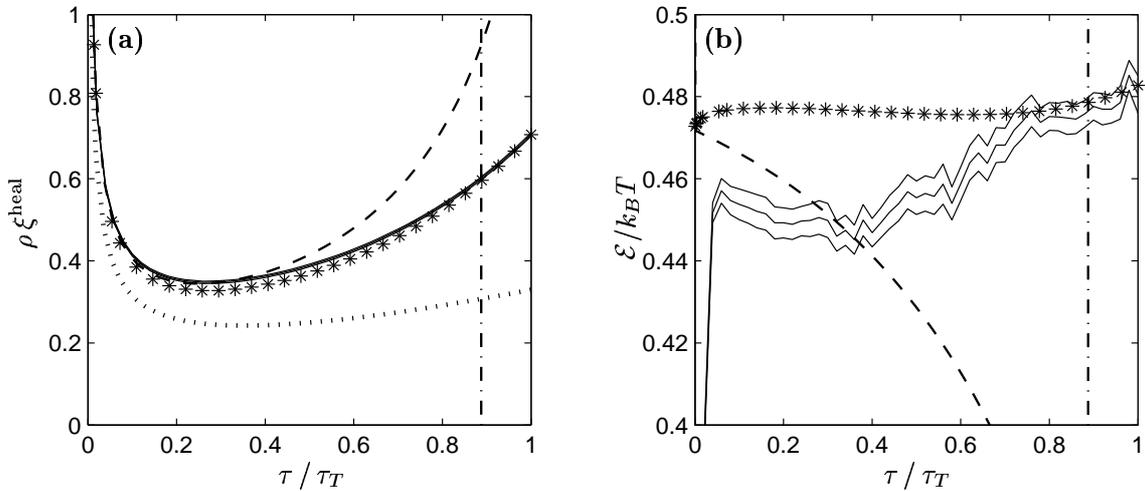


Figure 11.10: **Comparison of numerical calculation (SOLID LINES with error bars) to exact solution (ASTERIXES) for the $\gamma = 1$, $\tilde{T} = 1.6$ calculation with parameters given in Figure 11.6.** (a): density ρ , (b): energy per particle \mathcal{E} . Ideal Bose gas (DASHED) and Fermi gas values also shown for comparison. The DASH-DOTTED vertical line indicates the lower limit on times τ for which accuracy is expected with a finite sample on the basis of the $\text{var}[\text{Re}\{z_0\}] < 10$ indicator (See Appendix A).

Chapter 12

Conclusions

12.1 Overview of main results

As with the thesis, the main results obtained can be divided into three parts:

1. **Theoretical background work regarding general phase-space distributions.** The first aim here has been to specify what freedoms are available to the broadly conceived phase-space distribution approach when it is to be used for first-principles simulations of mesoscopic quantum mechanics. The considerations of Chapter 3 led to a set of necessary requirements for any useful simulation, summarized in Section 3.6. Some fundamental expressions (e.g. observable estimators) have been given for general representations (i.e. choices of kernel). The usefulness of a given representation (which is closely related to basis choice) is highly system dependent, and so it is hoped that the results and considerations of Chapter 3 will aid in matching representations to problems.

With a given representation, often there still remains a wide range of stochastic equations that simulate the same physical problem, but strongly differ in efficiency. The stochastic gauge formalism developed in Chapter 4 describes the freedoms available in a systematic way. At the coarsest level there are two main kinds: Those that arise when making the correspondences from master to Fokker-Planck equations (kernel gauges), and from Fokker-Planck to

stochastic equations (diffusion gauges). Some analysis of these freedoms for general gauge choices has also been made. Uses have been found to include

- (a) Improvement of efficiency by tailoring the shape of the distribution of random variables to improve sampling.
- (b) Removal of biases due to a broad class of so-called boundary term errors, which are due to instabilities or divergences in the stochastic equations.
- (c) Allowance for calculations of grand canonical ensemble properties by the inclusion of a dynamically varying (with simulation steps) weight.

Part of the freedoms (e.g. real diffusion gauges) have been determined to not be useful. The stochastic gauge formalism includes in a systematic way several recent developments[2, 1, 65, 3, 66, 56, 61] as special cases. A relatively restricted set of gauges that still allows the uses listed above has been identified in Section 4.5 as the “standard gauges”.

The causes and symptoms of boundary term errors (sampling biases that do not abate in the infinite sample number limit), have been considered in considerable detail in Chapter 6. These have been major obstacles for many phase-space simulations in the past. Two kinds of processes have been identified as causing these — either arising when making the correspondence between master and Fokker-Planck equations (first kind), or when calculating observable estimates using the simulated random variables (second kind). Several warning symptoms of these errors, which can be identified by inspection of the stochastic equations prior to simulation, have been identified.

It has been found that appropriate choices of drift stochastic gauges (a class of Kernel gauge) can be used to remove the biases and instabilities in many (or possibly all) cases of boundary term errors of the first kind. This is demonstrated in several simple examples where boundary term errors have been known to occur, and also used in some subsequent many-mode calculations where required. Heuristic guidelines for appropriate gauge choice have been found and given in Section 6.3.2.

2. Investigation of gauge P representation and development of gauges for use with interacting Bose gases. This representation is an extension of the successful positive P representation based on coherent states. The added complex weight factor is shown to allow the use of the standard gauges to improve simulations, remove possible boundary term bias, and extend the range of physical models that can be simulated (thermodynamic calculations).

The emphasis here has been almost exclusively on simulations of open Bose gases with binary interactions. Equations for dynamics and thermodynamics calculations have been obtained (including non-delta-function interparticle potentials). Part B searched for useful choices of gauges for dynamics and thermodynamics calculations, given the constraint that they be local (i.e. dependent only on variables at a single lattice point). This constraint is not optimal, but is a basic starting point for possible further investigation. Both diffusion and drift gauges that improve simulations under appropriate conditions were found. Diffusion gauges were able to significantly improve sampling and useful simulation times in dynamics calculations, while drift gauges are essential for any accurate thermodynamic calculation, and were found to also give sampling improvement in dynamics under some conditions (See Chapter 8).

Analyses were made of the expected behavior of multi-mode simulations, and subsequently confirmed in the actual calculations. In dynamics calculations, simulation time is found to be limited by properties of the most highly occupied lattice point, and a robust estimate of this time for positive P simulations is given by (10.4). Relative to this positive P simulation time, it was found that simulation time improvement with diffusion gauges occurs when lattice spacing is greater than the healing length in the gas. Local drift gauges were found to give even more improvement in simulation time but for a much narrower range of systems in which two-body collision effects dominate kinetic processes, as in some simulations in Chapter 8. Thermodynamic simulations were made possible with judicious use of a drift gauge. It was also found that

simulation precision at a target temperature and chemical potential was dependent (sometimes strongly) on the (otherwise free) choice of chemical potential at higher temperatures.

3. Simulations of many-mode interacting Bose gases Several nontrivial mesoscopic systems of open interacting Bose gases have been simulated. These were dynamics of:

- Uniform 1D and 2D gas with effectively local scattering. Wave behavior occurring in the two- or three-particle correlations (but not density) was found, and displays some interesting properties (e.g. movement of main wavefront at $\sqrt{2}$ the speed of sound).
- Uniform 1D gas with extended nonlocal interparticle interactions. Correlation wave behavior was also seen, but with different properties than for the locally-interacting gas. Simulation time was found to increase significantly with respect to a locally-interacting gas with the same interaction energy density.
- Trapped 1D gas with extended interactions on the length scale of the trap. Strong interplay between the scattering and breathing of the atom cloud in the trap was seen.
- Scattering of atoms from colliding BECs in 3D. Bosonic enhancement of initially empty modes was seen in a physical situation similar to the Vogels *et al* four wave mixing experiment[41] (the difference was that there were less atoms in the simulation). Additionally, a significant suppression of the spontaneous scattering process compared to the imaginary scattering length estimate is predicted.

Also:

- Calculations of grand canonical ensembles were made for the uniform 1D gas at temperature T and chemical potential μ . Physical regimes reached included the nondegenerate strongly interacting fermionized regime, and the quantum degenerate decoherent regime. Spatial correlation functions

of first $\bar{g}^{(1)}(x)$, second $\bar{g}^{(s)}(x)$ and third order $\bar{g}^{(3)}(x, y)$ have been calculated as well as momentum distributions. A transition regime where enhanced atom pairing occurs at a preferred interparticle distance has been found. This occurs when the de Broglie thermal wavelength λ_T is of the same order as the 1D scattering length a_{1D} (effectively the “size” of the atom for scattering processes). The pairing was found to occur at distances $\approx \lambda_T/2$ for the parameters simulated.

These simulations can be tractable even with very large numbers of modes or particles in the system given the right conditions, as evidenced by Section 10.7, where there were 1009 152 lattice points and 150 000 atoms on average.

Situations where there are several length scales of similar order, or processes of similar strength are of particular interest for first-principles simulations because it is difficult to make accurate quantitative predictions otherwise. The calculations in Sections 10.6 and 10.7 and the thermodynamics calculations were of this type, and were seen to be amenable to simulation using the gauge P method. In fact, even in a perturbative regime (e.g. the correlation phenomena at low intensity can probably be treated by Bogoliubov approximation approach), the equations are convenient to use: One simply repeatedly simulates a GP equation with noise and takes appropriate averages. All quantum effects are included to within statistical precision. The caveat is that simulation time is limited. Nevertheless, many phenomena can still be seen, and simulation time can be extended with appropriate gauges under some conditions.

12.2 Future directions of improvement

The local diffusion gauges developed here give significant improvements only when the lattice is relatively sparse, and/or scattering interactions are dominant. A way to overcome these limitations might be found by using nonlocal diffusion gauges that would depend on neighboring, or preferably all, lattice amplitudes $\alpha_{\mathbf{n}}$, $\beta_{\mathbf{n}}$. A possible starting point would be to consider optimizing diffusion gauges in the

opposite regime of strong kinetic interactions and weak scattering. The equations for such a model are probably best considered in Fourier space where kinetic processes take on a simpler form.

Another major issue is the reduction of simulation time with system size due to increasing weight or z_0 spread when drift gauges are used. (As was discussed in Section 10.2.4). A starting point here may be to determine why the lattice-dependent diffusion gauge (10.26) did not give the expected simulation time improvements (10.28). If the expected $M^{-1/4}$ scaling could be achieved, many useful mesoscopic simulations should become accessible with a drift gauge. Such a drift-gauged simulation would be free of moving singularities, and possibly also extend simulation time as was seen for the drift gauged one- and two-mode systems.

A completely different but potentially very promising approach is to use non-coherent state kernels for the representation. Certainly in many low temperature regimes this would be a more efficient approach if viable representations could be found because the low temperature ground and low excited states are often far from coherent states. Some preliminary attempts with squeezed state kernels give mixed results[104].

For thermodynamics calculations of properties at and around a target temperature and chemical potential, an important outstanding issue is to characterize the influence of the choice of $\mu(T)$ at intermediate temperature values. It is conceivable that a good choice of $\mu(T)$ may give dramatic improvement in precision and significantly extend the physical region that can be simulated. Some precision and simulation time improvement should also be obtained by using a more sophisticated importance sampling technique than the rough method used in Chapter 11. Some preliminary results indicate that using a Metropolis sampling algorithm[105] extends the reachable γ, \tilde{T} regime[106].

Bibliography

- [1] I. Carusotto, Y. Castin, and J. Dalibard. N-boson time-dependent problem: A reformulation with stochastic wave functions. *Physical Review A* **63**, 023606 (Jan. 2001).
- [2] L. I. Plimak, M. K. Olsen, and M. J. Collett. Optimization of the positive-P representation for the anharmonic oscillator. *Physical Review A* **64**, 025801 (July 2001).
- [3] P. Deuar and P. D. Drummond. Stochastic gauges in quantum dynamics for many-body simulations. *Computer Physics Communications* **142**, 442–445 (Dec. 2001).
- [4] R. P. Feynman. Simulating physics with computers. *International Journal of Theoretical Physics* **21**, 467–488 (June 1982).
- [5] D. M. Ceperley. Microscopic simulations in physics. *Reviews of Modern Physics* **71**, S438–S443 (Mar. 1999).
- [6] M. H. Anderson, J. R. Ensher, M. R. Matthews, and C. E. Wieman. Observation of Bose-Einstein condensation in a dilute atomic vapor. *Science* **269**, 198–201 (July 1995).
- [7] K. B. Davis, M.-O. Mewes, M. R. Andrews, N. J. van Druten, D. S. Durfee, D. M. Kurn, and W. Ketterle. Bose-Einstein condensation in a gas of sodium atoms. *Physical Review Letters* **75**, 3969–3973 (Nov. 1995).

-
- [8] C. C. Bradley, C. A. Sackett, J. J. Tollett, and R. G. Hulet. Evidence of Bose-Einstein condensation in an atomic gas with attractive interactions. *Physical Review Letters* **75**, 1687–1690 (Aug. 1995).
- [9] D. G. Fried, T. C. Killian, L. Willmann, D. Landhuis, S. C. Moss, D. Kleppner, and T. J. Greytak. Bose-Einstein condensation of atomic hydrogen. *Physical Review Letters* **81**, 3811–3814 (Nov. 1998).
- [10] S. Chaturvedi, P. D. Drummond, and D. F. Walls. 2 photon absorption with coherent and partially coherent driving fields. *Journal of Physics A: Mathematical and General* **10**, L187–L192 (Nov. 1977).
- [11] P. D. Drummond and C. W. Gardiner. Generalised P-representations in quantum optics. *Journal of Physics A: Mathematical and General* **13**, 2353–2368 (July 1980).
- [12] S. J. Carter, P. D. Drummond, M. D. Reid, and R. M. Shelby. Squeezing of quantum solitons. *Physical Review Letters* **58**, 1841–1844 (May 1987).
- [13] P. D. Drummond, R. M. Shelby, S. R. Friberg, and Y. Yamamoto. Quantum solitons in optical fibers. *Nature* **365**, 307–313 (Sep. 1993).
- [14] M. J. Steel, M. K. Olsen, L. I. Plimak, P. D. Drummond, S. M. Tan, M. J. Collett, D. F. Walls, and R. Graham. Dynamical quantum noise in trapped Bose-Einstein condensates. *Physical Review A* **58**, 4824–4835 (Dec. 1998).
- [15] P. D. Drummond and J. F. Corney. Quantum dynamics of evaporatively cooled Bose-Einstein condensates. *Physical Review A* **60**, R2661–R2664 (Oct. 1999).
- [16] K. Southwell. Ultracold matter. *Nature* **416**, 205 (Mar. 2002).
- [17] S. Chu. Cold atoms and quantum control. *Nature* **416**, 206–210 (Mar. 2002).
- [18] J. R. Anglin and W. Ketterle. Bose-Einstein condensation of atomic gases. *Nature* **416**, 211–218 (Mar. 2002).

-
- [19] K. Burnett, P. S. Julienne, P. D. Lett, E. Tiesinga, and C. J. Williams. Quantum encounters of the cold kind. *Nature* **416**, 225–232 (Mar. 2002).
- [20] C. Monroe. Quantum information processing with atoms and photons. *Nature* **416**, 238–246 (Mar. 2002).
- [21] P. Zoller and C. W. Gardiner. Quantum noise in quantum optics: the stochastic Schrödinger equation. In *Quantum fluctuations: Les Houches Summer School LXIII*, edited by S. Reynaud, E. Giacobino, and J. Zinn-Justin (Elsevier, Amsterdam, 1997).
- [22] M. B. Plenio and P. L. Knight. The quantum-jump approach to dissipative dynamics in quantum optics. *Reviews of Modern Physics* **70**, 101–144 (Jan. 1998).
- [23] H. M. Wiseman and G. J. Milburn. Quantum theory of field-quadrature measurements. *Physical Review A* **47**, 642–662 (Jan. 1993).
- [24] G. G. Carlo, G. Benenti, and G. Casati. Teleportation in a noisy environment: A quantum trajectories approach. *Physical Review Letters* **91**, 257903 (Dec. 2003).
- [25] R. P. Feynman. The lambda-transition in liquid helium. *Physical Review* **90**, 1116–1117 (June 1953).
- [26] H. F. Trotter. On the product of semigroups operators. *Proceedings of the American Mathematical Society* **10**, 545–551 (1959).
- [27] D. M. Ceperley. Path integrals in the theory of condensed helium. *Reviews of Modern Physics* **67**, 279–355 (Apr. 1995).
- [28] D. M. Ceperley. Lectures on quantum monte carlo (May 1996). <http://archive.ncsa.uiuc.edu/Science/CMP/papers/cep96b/lnotes.ps>
- [29] K. E. Schmidt and M. H. Kalos. In *Monte Carlo Methods in Statistical Physics II*, edited by K. Binder (Springer, Berlin, 1984).

- [30] A. J. Leggett. Bose-Einstein condensation in the alkali gases: Some fundamental concepts. *Reviews of Modern Physics* **73**, 307–356 (Apr. 2001).
- [31] F. Dalfovo, S. Giorgini, L. P. Pitaevskii, and S. Stringari. Theory of Bose-Einstein condensation in trapped gases. *Reviews of Modern Physics* **71**, 463–512 (Apr. 1999).
- [32] A. S. Parkins and D. F. Walls. The physics of trapped dilute-gas Bose-Einstein condensates. *Physics Reports* **303**, 1–80 (Sep. 1998).
- [33] S. Chu. Nobel lecture: The manipulation of neutral particles. *Reviews of Modern Physics* **70**, 685–706 (July 1998).
- [34] C. Cohen-Tannoudji. Nobel lecture: Manipulating atoms with photons. *Reviews of Modern Physics* **70**, 707–719 (July 1998).
- [35] W. D. Phillips. Nobel lecture: Laser cooling and trapping of neutral atoms. *Reviews of Modern Physics* **70**, 721–741 (July 1998).
- [36] W. Ketterle and N. J. van Druten. Evaporative cooling of atoms. In *Advances in Atomic, Molecular, and Optical Physics*, edited by B. Bederson and H. Walther, volume **37**, pages 181–236 (Academic, San Diego, 1996).
- [37] J. T. M. Walraven. Atomic hydrogen in magnetostatic traps. In *Quantum Dynamics of Simple Systems*, edited by G. L. Oppo, S. M. Barnett, E. Riis, and M. Wilkinson, page 315 (Institute of Physics, Bristol, 1996).
- [38] A. Görlitz, J. M. Vogels, A. E. Leanhardt, C. Raman, T. L. Gustavson, J. R. Abo-Shaeer, A. P. Chikkatur, S. Gupta, S. Inouye, T. Rosenband, and W. Ketterle. Realization of Bose-Einstein condensates in lower dimensions. *Physical Review Letters* **87**, 130402 (Sep. 2001).
- [39] F. Schreck, L. Khaykovich, K. L. Corwin, G. Ferrari, T. Bourdel, J. Cubizolles, and C. Salomon. Quasipure Bose-Einstein condensate immersed in a Fermi sea. *Physical Review Letters* **87**, 080403 (Aug. 2001).

- [40] M. Greiner, I. Bloch, O. Mandel, T. W. Hansch, and T. Esslinger. Exploring phase coherence in a 2D lattice of Bose-Einstein condensates. *Physical Review Letters* **87**, 160405 (Oct. 2001).
- [41] J. M. Vogels, K. Xu, and W. Ketterle. Generation of macroscopic pair-correlated atomic beams by four-wave mixing in Bose-Einstein condensates. *Physical Review Letters* **89**, 020401 (June 2002).
- [42] E. P. Gross. Structure of quantized vortex. *Il Nuovo Cimento* **20**, 454 (1961).
- [43] E. P. Gross. Hydrodynamics of a superfluid condensate. *Journal of Mathematical Physics* **4**, 195 (1963).
- [44] L. P. Pitaevskii. *Soviet Physics JETP* **13**, 451 (1961). [Original: *Zhurnal Eksperimental'noi i Teoreticheskoi Fiziki* **40**, 646 (1961)].
- [45] N. N. Bogoliubov. On the theory of superfluidity. *Journal of Physics* **11**, 23–32 (1947).
- [46] J. F. Corney. *Open Quantum Dynamics of Mesoscopic Bose-Einstein Condensates*. PhD thesis, The University of Queensland (1999).
- [47] Yu Kagan, B. V. Svistunov, and G. V. Shlyapnikov. Effect of Bose condensation on inelastic processes in gases. *JETP Letters* **42**, 209–212 (Aug. 1985). [Original: *Pis'ma v Zhurnal Eksperimental'noi i Teoreticheskoi Fiziki* **42**, 169–172 (Aug. 1985)].
- [48] Yu Kagan, B. V. Svistunov, and G. V. Shlyapnikov. Quantum correlations in the optical characteristics of low-temperature H down arrow and D down arrow gases. *JETP Letters* **48**, 56–59 (July 1988). [Original: *Pis'ma v Zhurnal Eksperimental'noi i Teoreticheskoi Fiziki* **48**, 54–57 (July 1988)].
- [49] J. Dalibard. Bose-Einstein condensation in atomic gases. In *International School of Physics “Enrico Fermi” Course 140*, edited by M. Inguscio, S. Stringari, and C. Wieman (IOS Press, Amsterdam, 1999).

-
- [50] J. Wiener, V. S. Bagnato, S. Zilio, and P. S. Julienne. Experiments and theory in cold and ultracold collisions. *Reviews of Modern Physics* **71**, 1–85 (Jan. 1999).
- [51] L. D. Landau and E. M. Lifshitz. *Quantum Mechanics* (Pergamon, London, 1959).
- [52] H. M. J. M. Boesten, C. C. Tsai, J. R. Gardner, D. J. Heinzen, and B. J. Verhaar. Observation of a shape resonance in the collision of two cold Rb-87 atoms. *Physical Review A* **55**, 636–640 (Jan. 1997).
- [53] C. W. Gardiner. Handbook of stochastic methods for physics, chemistry, and the natural sciences. In *Springer series in synergetics*, volume **13** (Springer-Verlag, Berlin, New York, 1983).
- [54] W. Louisell. *Quantum Optics* (Academic Press, New York, 1970).
- [55] W. Louisell. *Quantum Statistical Properties of Radiation* (Wiley, New York, 1974).
- [56] P. D. Drummond and P. Deuar. Quantum dynamics with stochastic gauge simulations. *Journal of Optics B – Quantum and Semiclassical Optics* **5**, S281–S289 (June 2003).
- [57] E. P. Wigner. On the quantum correction for thermodynamic equilibrium. *Physical Review* **40**, 749–759 (June 1932).
- [58] R. J. Glauber. Coherent and incoherent states of the radiation field. *Physical Review* **131**, 2766–2788 (Sep. 1963).
- [59] E. C. G. Sudarshan. Equivalence of semiclassical and quantum mechanical descriptions of statistical light beams. *Physical Review Letters* **10**, 277–279 (Apr. 1963).
- [60] J. R. Klauder and E. C. G. Sudarshan. *Fundamentals of Quantum Optics* (Benjamin, New York, 1970).

-
- [61] P. D. Drummond, P. Deuar, and Kheruntsyan K. V. Canonical Bose gas simulations with stochastic gauges. *Physical Review Letters* **92**, 040405 (Jan. 2004).
- [62] C. W. Gardiner. *Quantum Noise* (Springer-Verlag, Berlin, Heidelberg, 1991).
- [63] M. K. Olsen, Plimak L. I., and M. Fleischhauer. Quantum-theoretical treatments of three-photon processes. *Physical Review A* **65**, 053806 (May 2002).
- [64] A. Gilchrist, C. W. Gardiner, and P. D. Drummond. Positive P representation: Application and validity. *Physical Review A* **55**, 3014–3032 (Apr. 1997).
- [65] I. Carusotto and Y. Castin. An exact stochastic field method for the interacting Bose gas at thermal equilibrium. *Journal of Physics B: Atomic and Molecular Optical Physics*, **34**, 4589–4608 (Nov. 2001).
- [66] P. Deuar and P. D. Drummond. Gauge P-representations for quantum-dynamical problems: Removal of boundary terms. *Physical Review A* **66**, 033812 (Sep. 2002).
- [67] J. F. Corney and P. D Drummond. Gaussian quantum operator representation for bosons. *Physical Review A* **68**, 063822 (Dec. 2003).
- [68] P. D Drummond. To be published.
- [69] M. Dowling. *The Metropolis Algorithm for Real Time Quantum Dynamics*. Honours thesis, The University of Queensland (Nov. 2003).
- [70] A. M. Smith and C. W. Gardiner. Simulations of nonlinear quantum damping using the positive P representation. *Physical Review A* **39**, 3511–3524 (Apr. 1989).
- [71] R. Schack and A. Schenzle. Positive P-representation. *Physical Review A* **44**, 682–687 (July 1991).
- [72] I. S. Gradshteyn and I. M. Ryzhik. *Table of Integrals, Series, and Products*, page 337 (Academic Press, New York, 4th edition, 1965).

-
- [73] K. Husimi. *Proc. Phys. Math. Soc. Japan* **22**, 264 (1940).
- [74] K. E. Cahill and R. J. Glauber. Density operators and quasiprobability distributions. *Physical Review* **177**, 1882–1902 (Jan. 1969).
- [75] C. W. Gardiner and S. Chaturvedi. The Poisson representation I. A new technique for chemical master equations. *Journal of Statistical Physics* **17**, 429–468 (Dec. 1977).
- [76] S. Chaturvedi and C. W. Gardiner. The Poisson representation II. Two-time correlation functions. *Journal of Statistical Physics* **18**, 501–522 (May 1978).
- [77] S. Mancini, V. I. Man’ko, and P. Tombesi. Symplectic tomography as classical approach to quantum systems. *Physics Letters A* **213**, 1–6 (Apr. 1996).
- [78] S. Mancini, V. I. Man’ko, and P. Tombesi. Classical-like description of quantum dynamics by means of symplectic tomography. *Foundations of Physics* **27**, 801–824 (June 1997).
- [79] M. J. Ablowitz, A. Ramani, and H. Segur. Non-linear evolution equations and ordinary differential-equations of Painleve type. *Lettere al Nuovo Cimento* **23**, 333–338 (Oct. 1978).
- [80] T. Bountis, H. Segur, and F. Vivaldi. Integrable Hamiltonian-systems and the Painleve property. *Physical Review A* **25**, 1257–1264 (Mar. 1982).
- [81] J. Weiss, M. Tabor, and G. Carnevale. The Painleve property for partial differential equations. *Journal of Mathematical Physics* **24**, 522–526 (Mar. 1983).
- [82] P. D. Drummond and Mortimer I. K. Computer simulations of multiplicative stochastic differential equations. *Journal of Computational Physics* **93**, 144–170 (Mar. 1991).
- [83] I. J. D. Craig and K. J. McNeil. Quantum simulations of nonlinear optical damping: An exact solution for the stochastic differential equations and an interpretation of “spiking”. *Physical Review A* **39**, 6267–6270 (June 1989).

-
- [84] K. J. McNeil and I. J. D. Craig. Positive-P representation for second-harmonic generation: Analytic and computational results. *Physical Review A* **41**, 4009–4018 (Apr. 1990).
- [85] C. W. Gardiner, A. Gilchrist, and P. D. Drummond. Using the positive P-representation. In *Lecture Notes in Physics*, edited by F. Ehlotzky, volume **420**, page 239 (Springer-Verlag, Berlin, 1993).
- [86] E. H. Lieb and Liniger W. Analysis of an interacting Bose gas I. The general solution and the ground state. *Physical Review* **130**, 1605–1616 (May 1963).
- [87] C. N. Yang and Yang C. P. Thermodynamics of a one-dimensional system of bosons with repulsive delta-function interaction. *Journal of Mathematical Physics* **10**, 1115–1122 (July 1969).
- [88] Y. B. Band, M. Trippenbach, J. P. Jr. Burke, and P. S. Julienne. Elastic scattering loss of atoms from colliding Bose-Einstein condensate wavepackets. *Physical Review Letters* **84**, 5462–5465 (June 2000).
- [89] R. Bach, M Trippenbach, and Rzażewski. Spontaneous emission of atoms via collisions of Bose-Einstein condensates. *Physical Review A* **65**, 063605 (May 2002).
- [90] D. C. Mattis. *The Many-Body problem: An Encyclopedia of Exactly Solved Models in 1D* (World Scientific, Singapore, 1994).
- [91] M. D. Girardeau. Relationship between systems of impenetrable bosons and fermions on one dimension. *Journal of Mathematical Physics* **1**, 516–523 (1960).
- [92] M. D. Girardeau. Permutation symmetry of many-particle wave functions. *Physical Review* **139**, B500–B508 (July 1965).
- [93] E. H. Lieb. Analysis of an interacting Bose gas II. The excitation spectrum. *Physical Review* **130**, 1616–24 (May 1963).
- [94] V. N. Popov. *Functional Integrals in Quantum Field Theory and Statistical Physics* (Reidel, Dordrecht, 1983).

-
- [95] H. B. Thacker. Exact integrability in quantum field theory and statistical systems. *Reviews of Modern Physics* **53**, 253–285 (Apr. 1981).
- [96] V. E. Korepin, N. M. Bogoliubov, and A. G. Izergin. *Quantum Inverse Scattering Method and Correlation Functions* (Cambridge University Press, Cambridge, 1993).
- [97] M. Olshanii and V. Dunjko. Short-distance correlation properties of the Lieb-Liniger system and momentum distributions of trapped one-dimensional atomic gases. *Physical Review Letters* **91**, 090401 (Aug. 2003).
- [98] D. M. Gangardt and G. V. Shlyapnikov. Stability and phase coherence of trapped 1D Bose gases. *Physical Review Letters* **90**, 010401 (Jan. 2003).
- [99] K. V. Kheruntsyan, D. M. Gangardt, P. D. Drummond, and G. V. Shlyapnikov. Pair correlations in a finite-temperature 1D Bose gas. *Physical Review Letters* **91**, 040403 (July 2003).
- [100] M. Olshanii. Atomic scattering in the presence of external confinement in a gas of impenetrable bosons. *Physical Review Letters* **81**, 938–941 (Aug. 1998).
- [101] N. D. Mermin and H. Wagner. Absence of ferromagnetism or antiferromagnetism in one- and two-dimensional isotropic Heisenberg models. *Physical Review Letters* **17**, 1133–1136 (Nov. 1966).
- [102] P. C. Hohenberg. Existence of long-range order in one and two dimensions. *Physical Review* **158**, 383–386 (June 1967).
- [103] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical recipes* (Cambridge University Press, Cambridge, 1988).
- [104] P. Deuar and P. D Drummond. To be published.
- [105] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. Equation of state calculations by fast computing machines. *Journal of Chemical Physics* **21**, 1087–1092 (June 1953).
- [106] P. Deuar and P. D Drummond. To be published.

Appendix A

Exponentials of Gaussian random variables

Gauge P simulations of interacting Bose gases (whether dynamic or thermodynamic) involve multiplicative noise terms of the form

$$dz = czdW(t) + \dots, \quad (\text{A.1})$$

where z is a complex variable $(\alpha_{\mathbf{n}}, \beta_{\mathbf{n}}, z_0)$, and c a constant. This leads to $\log z$ being distributed approximately as a Gaussian. This is exact at short times, but some modification of the distribution occur also as a consequence of the other “...” drift (or noise) terms in the equations. Observable estimates are usually obtained through averages of non-logarithmic combinations of z , however, and this requires some care. Let us consider the idealized situation of $z \approx v_\sigma = v_0 e^{\sigma\xi} = e^{vL}$, where ξ is a Gaussian random variable of mean zero, variance unity. The notation of Section 7.4 will be used.

Using the distribution of ξ , (7.36), one obtains

$$\Pr(v_\sigma) = \frac{1}{\sigma v_\sigma \sqrt{2\pi}} \exp \left\{ -\frac{(\log[v_\sigma/v_0])^2}{2\sigma^2} \right\}. \quad (\text{A.2})$$

This distribution falls off slowly as $v_\sigma \rightarrow \infty$, and for a finite sample may not be sampled correctly if ξ is chosen according to its Gaussian distribution. With \mathcal{S} samples, the greatest value of ξ obtained can be expected to be ξ_{\max} , the solution of $\frac{1}{2}[1 - \text{erf}(\xi_{\max}/\sqrt{2})] \approx 2/\mathcal{S}$, where $\text{erf}(x) = 2 \int_0^x e^{-t^2} dt$ is the error function. For

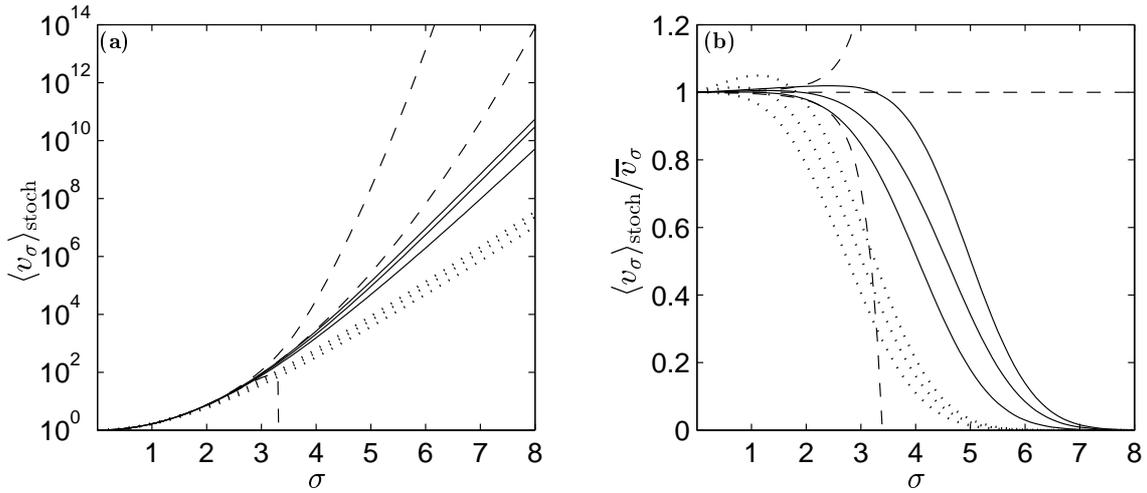


Figure A.1: **Finite sample estimates** of means of exponentials of gaussian random variables $v_\sigma = e^{\sigma\xi}$. SOLID lines: $\mathcal{S} = 10^5$ sample estimate of $\langle v_\sigma \rangle_{\text{stoch}}$ and error bars, where ξ was sampled from a Gaussian distribution. DOTTED lines: Same but using $\mathcal{S} = 1000$ samples. DASHED lines: Exact ($\mathcal{S} \rightarrow \infty$) value of the mean $\langle v_\sigma \rangle \rightarrow \bar{v}_\sigma$, along with error estimates for the $\mathcal{S} = 10^5$ calculation based on the exact value of $\text{var}[v_\sigma]$, from (7.39) and (7.40). This corresponds to what would be expected from an unbiased sample of v_σ . All error bars were calculated according to (7.41). Subplot (b) shows the same values, but scaled with respect to the $\mathcal{S} \rightarrow \infty$ mean \bar{v}_σ . Values at all σ were made with the same sample of ξ to show the direct dependence on σ only.

$\mathcal{S} \approx \mathcal{O}(1000)$ to $\mathcal{O}(10^5)$, this gives $\xi_{\max} \approx 3$ to 4. Some bias in $\langle v_\sigma \rangle_{\text{stoch}}$ (underestimation) may be expected due to the lack of sampling of the far tails once

$$\int_{v_0 \exp(\sigma\xi_{\max})}^{\infty} \text{Pr}(v_\sigma) dv_\sigma \gg 1/\mathcal{S}. \quad (\text{A.3})$$

For typical sample sizes $\mathcal{S} \approx \mathcal{O}(1000)$ to $\mathcal{O}(10^5)$ this will usually occur around $\sigma^2 \gtrsim 10$. This is actually the same large σ region given by (7.43) where all precision is lost anyway because of excessive distribution spread (as shown in Section 7.4).

However, the problem is that *both* the mean $\langle v_\sigma \rangle_{\text{stoch}}$ and the CLT precision estimate $\sqrt{\text{var}[v_\sigma]/\mathcal{S}}$ are underestimated when σ is large, but the error estimate is underestimated *by an even greater amount* than the mean. As a result, badly sampled data may still appear to be significant for large σ values. The situation is shown in Figure A.1.

The simplest practical solution is to simply discard any calculated means $\langle z \rangle_{\text{stoch}}$

when

$$\text{var} [\log |z|] \gtrsim 10. \quad (\text{A.4})$$

If an unbiased mean $\langle z \rangle_{\text{stoch}}$ and precision estimate Δz was obtained by some more sophisticated method, then it would not be significant anyway since one expects $\Delta z \gg \langle z \rangle_{\text{stoch}}$ at these large σ values. The main point here is that quantities (A.4) should be monitored when dealing with multiplicative noise of the form (A.1) in the stochastic equations.

Some subtleties can arise because in realistic simulations $\log z$ is not exactly Gaussian. If the large $\log z$ distribution tails fall off more rapidly than Gaussian, then an unbiased simulation can be obtained for larger $\text{var} [\log |z|]$, while the converse is true if these tails fall off less rapidly. A **more robust** indicator of possible bias than (A.4) is to compare the observable estimate obtained with two sample sizes \mathcal{S} , differing by at least an order of magnitude. If sampling bias is present, the two estimates of the average $\langle z \rangle_{\text{stoch}}$ will usually differ by a statistically significant amount. An example of this for the idealized v_σ model is shown in Figure A.1. This kind of sample-size-dependent behavior is always a strong warning sign.

Appendix B

Some details of Stochastic Calculus

Some results regarding random terms in differential equations relevant to the thesis are gathered here. Proofs of these can be found in Gardiner[53], with details on computer algorithms in [82].

In many cases, a set of stochastic differential equations (i.e. differential equations with random terms) can be written in the Langevin form

$$dx_j(t) = A_j(\mathbf{x}, t) dt + \sum_k B_{jk}(\mathbf{x}, t) dW_k(t), \quad (\text{B.1})$$

where \mathbf{x} contains all variables x_j . The $dW_j(t)$ are Wiener increments and (B.1) is to be interpreted according to the Ito calculus (see below).

A Wiener increment is defined in terms of the Wiener process $W(t)$, which is the special case of (B.1) with $j = 1$, $A_1 = 0$ and $B_{1k} = \delta_{k1}$. The probability distribution of W is governed by the Fokker-Planck equation (FPE)

$$\frac{\partial P_W(W, t)}{\partial t} = \frac{\partial^2 P_W(W, t)}{\partial W^2}, \quad (\text{B.2})$$

a special case of Brownian motion. The individual realizations of W are continuous but not differentiable. The (infinitesimal) Wiener increment can, however, be defined, and is

$$dW(t) = W(t + dt) - W(t), \quad (\text{B.3})$$

with dt infinitesimal. This random quantity has the expectation values

$$\langle dW(t) \rangle_{\text{stoch}} = 0 \quad (\text{B.4a})$$

$$\langle dW(t)^2 \rangle_{\text{stoch}} = dt \quad (\text{B.4b})$$

$$\langle dW(t)dW(t') \rangle_{\text{stoch}} = 0 \quad \text{if } t \neq t' \quad (\text{B.4c})$$

$$\left\langle \prod_{j=1}^{\max[j]>2} dW(t_j) \right\rangle_{\text{stoch}} = 0 \quad (\text{B.4d})$$

The Wiener increment is related to processes with white noise correlations such that if $\langle \xi(t)\xi(t') \rangle_{\text{stoch}} = \delta(t - t')$ then one can write $\xi(t) = dW(t)/dt$. In (B.1), the Wiener increments are independent: $\langle dW_j dW_{k \neq j} \rangle_{\text{stoch}} = \langle dW_j \rangle_{\text{stoch}} \langle dW_k \rangle_{\text{stoch}}$ etc. In a numerical calculation, the Wiener increment is usually implemented as independent Gaussian random variables ΔW_j at each time step of length Δt with mean zero and variance

$$\langle \Delta W_j \Delta W_k \rangle_{\text{stoch}} = \Delta t \delta_{jk}, \quad (\text{B.5})$$

although other choices of the distribution of ΔW_j are possible provided only that the discrete step analogues of (B.4) are satisfied, as in (B.5).

An equation (B.1) in the Ito calculus is equivalent to the Stratonovich calculus equation

$$\begin{aligned} dx_j(t) &= A_j(\mathbf{x}, t) dt + \sum_k B_{jk}(\mathbf{x}, t) dW_k(t) + S_j(\mathbf{x}, t) \\ &= dx_j^{\text{Ito}}(t) + S_j(\mathbf{x}, t), \end{aligned} \quad (\text{B.6})$$

where the *Stratonovich correction* is

$$S_j(\mathbf{x}, t) = -\frac{1}{2} \sum_{kl} B_{lk}(\mathbf{x}, t) \frac{\partial B_{jk}(\mathbf{x}, t)}{\partial x_l}. \quad (\text{B.7})$$

These two forms arise from different ways of defining the integral of the differential equations, both useful. For practical purposes, the main differences are that:

- In the Ito calculus, the time-dependent variables $x_j(t)$ are independent of the same-time Wiener increments $W_k(t)$, while this is not the case in the Stratonovich calculus.

- In the Stratonovich calculus, the x_j obey the usual rules of deterministic calculus, in particular the chain rule. In the Ito calculus one instead has

$$dy(\mathbf{x}) = \sum_j dx_j(\mathbf{x}, t) \frac{\partial y(\mathbf{x})}{\partial x_j} + \frac{dt}{2} \sum_{jkl} B_{jk}(\mathbf{x}, t) B_{lk}(\mathbf{x}, t) \frac{\partial^2 y(\mathbf{x})}{\partial x_j \partial x_l}. \quad (\text{B.8})$$

Because of the first point above, the Ito calculus corresponds to a discrete step integration algorithm with the derivative approximated using values at the beginning of the time interval:

$$x_j(t + \Delta t) = x_j(t) + A_j(\mathbf{x}(t), t) \Delta t + \sum_k B_{jk}(\mathbf{x}(t), t) \Delta W_k(t). \quad (\text{B.9})$$

Implicit integration algorithms estimate the derivative using quantities evaluated at later times¹

$$t_\kappa = t + \kappa \Delta t \quad (\text{B.10a})$$

(where $\kappa \in (0, 1]$) during the timestep. That is,

$$x_j(t + \Delta t) = x_j(t) + A_j(\mathbf{x}(t_\kappa), t_\kappa) \Delta t + \sum_k B_{jk}(\mathbf{x}(t_\kappa), t_\kappa) \Delta W_k(t). \quad (\text{B.10b})$$

In this case, one must use a Stratonovich (or Stratonovich-like) form of the drift:

$$dx_j(t) = dx_j^{\text{Ito}}(t) + 2\kappa S_j(\mathbf{x}, t) \quad (\text{B.11})$$

to ensure that (B.10b) is the same as (B.9) up to lowest deterministic order Δt . This follows by use of (B.5) on (B.10b). For multiplicative noise it has been found that using a semi-implicit method ($\kappa = \frac{1}{2}$) integration method gives superior numerical stability[82].

¹A common algorithm for implicitly estimating the $\mathbf{x}(t_\kappa)$ is to iterate (B.10b) several times using the appropriate smaller timestep $\Delta t \rightarrow \kappa \Delta t$.

Appendix C

Quotients of means or random variables

In the gauge P representation, estimates of the expectation value of an observable \hat{O} are made by evaluating the expression (3.14) for a finite ensemble of \mathcal{S} samples. This has the form of a quotient

$$\langle \hat{O} \rangle \approx \bar{O} = \frac{\langle f \rangle_{\text{stoch}}}{\langle \text{Re} \{ \Omega \} \rangle_{\text{stoch}}} = \frac{\langle f_N \rangle_{\text{stoch}}}{\langle f_D \rangle_{\text{stoch}}}. \quad (\text{C.1})$$

Error estimates are made by subensemble averaging as explained in Section 3.3.2. While for dynamics calculations, one is assured of $\langle \text{Re} \{ \Omega \} \rangle_{\text{stoch}} = 1$ in the limit $\mathcal{S} \rightarrow \infty$ at all times, so that the denominator can be ignored, thermodynamics calculations require the evaluation of both denominator $\langle f_D \rangle_{\text{stoch}}$ and numerator in (C.1). In particular, the error estimate (3.18) using subensembles requires some care.

The problem is that if there are not enough samples s in (say) the j th subensemble, then the j th denominator average may have the wrong sign or be close to zero. The result is that this j th subensemble average $\bar{O}^{(j)}$ takes on very large or non-realistic values, which are nowhere near $\langle \hat{O} \rangle$. Since their absolute values can be very large if the denominator was close to zero, these outlier $\bar{O}^{(j)}$ strongly influence the final error estimate. An example of such excessive error estimates is shown in Figure C.1, and typically goes in hand with spiking in the uncertainty $\Delta \bar{O}$.

Roughly, one expects that if there are $\mathcal{S}_E \approx 1000$ subensembles, then the far-

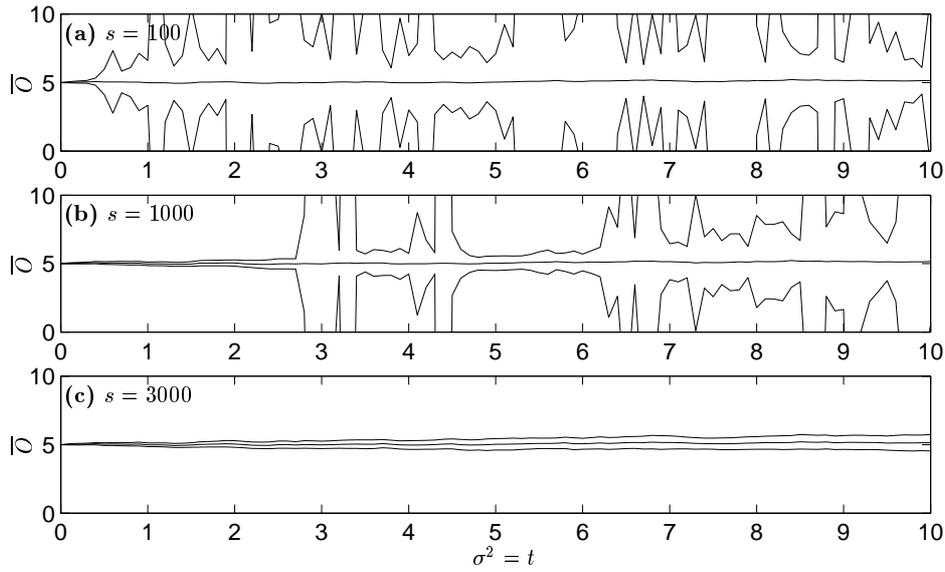


Figure C.1: **Error estimates with different subensemble sizes s .** SOLID lines show the means and error estimates obtained using (3.18) for the quantity $\bar{O} = 1/(\langle v \rangle_{\text{stoch}} + 0.2)$, where $v(t)$ obeys the Brownian motion stochastic equation $dv = dW(t)$ ($dW(t)$ is a Wiener increment). The standard deviation of v is $\sigma = \sqrt{t}$. In each simulation there were $\mathcal{S} = 10^5$ trajectories in all. Expression (C.2) suggests $s \gtrsim 2250$ is required at $t = 10$.

thermost outlier will be at about 3σ from the mean. Thus, excessive uncertainty is likely to occur once $3\sqrt{\text{var}[\bar{f}_D^{(j)}]} \gtrsim \langle f_D \rangle_{\text{stoch}}$, where $\bar{f}_D^{(j)}$ is the denominator average for the j th subensemble. For subensembles with s elements, this gives the rough limit

$$s \gtrsim 9 \frac{\text{var}[f_D]}{\langle f_D \rangle_{\text{stoch}}^2}. \quad (\text{C.2})$$

Often the variation in the weight $\Omega = e^{z_0}$ is mainly due to a close-to-gaussian distributed $\text{Re}\{z_0\}$. In this case, using (7.39) and (7.40), the requirement (C.2) becomes

$$s \gtrsim 9 (e^{\text{var}[\text{Re}\{z_0\}]} - 1). \quad (\text{C.3})$$