INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND GENERAL

J. Phys. A: Math. Gen. **39** (2006) 2723–2755 doi:10.1088/0305-4470/39/11/011

First-principles quantum dynamics in interacting Bose gases II: stochastic gauges

P Deuar and P D Drummond[1]

Australian Centre for Quantum Atom Optics, The University of Queensland, Brisbane, Australia

E-mail: deuar@physics.uq.edu.au and drummond@physics.uq.edu.au

Received 21 June 2005, in final form 4 January 2006
Published 1 March 2006
Online at stacks.iop.org/JPhysA/39/2723

Abstract
First principles simulations of the quantum dynamics of interacting Bose gases using the stochastic gauge representation are analysed. In a companion paper, we showed how the positive-P representation can be applied to these problems using stochastic differential equations. That method, however, is limited by increased sampling error as time evolves. Here, we show how the sampling error can be greatly reduced and the simulation time significantly extended using stochastic gauges. In particular, local stochastic gauges (a subset) are investigated. Improvements are confirmed in numerical calculations of single-, double- and multi-mode systems in the weak-mode coupling regime. Convergence issues are investigated, including the recognition of two modes by which stochastic equations produced by phase-space methods in general can diverge: movable singularities and a noise-weight relationship. The example calculated here displays wave-like behaviour in spatial correlation functions propagating in a uniform 1D gas after a sudden change in the coupling constant. This could in principle be tested experimentally using Feshbach resonance methods.

PACS numbers: 03.75.Kk, 05.10.Gg, 03.75.Gg, 02.50.Ey

1. Introduction

Phase-space techniques [1–3] for the simulation of exact many-body quantum dynamics have been developing rapidly [4–8]. Perhaps the major driving force behind this is their ability to work around the exponential growth[2] of the Hilbert space that occurs in direct methods that work with an explicit quantum state [9]. This exponential scaling in the direct approach implies that it is virtually impossible to diagonalize general initial states in order to calculate

[1] www.physics.uq.edu.au/BEC.
[2] Exponential in the number of modes or orbitals involved.

dynamics, even in the exceptional case when the exact eigenstates are known. Path integral methods were developed to deal with the Hilbert space growth, but while they are very useful for calculating ground states [10, 11], they are not applicable to quantum dynamics due to the oscillatory phase terms which arise and swamp any observable predictions. In contrast, appropriate phase-space methods can scale linearly or polynomially with the system size (like path integrals) without being as affected by the oscillatory phases, and so can provide verifiable physical predictions.

In the companion paper [12], we analysed the performance of the nonclassical positive-P representation [4, 5] for many-body interacting bosonic systems. Here, we extend the analysis to treat the more flexible stochastic gauge method of weighted phase-space trajectories [7].

The aforementioned positive-P representation is a straightforward and successful method, which is a good 'baseline' on which more-involved methods can build. It is useful to give a brief overview of it and its limitations, at this point: the quantum state is written as a probability distribution of off-diagonal coherent states separable at each spatial lattice point. This corresponds to a nonclassical phase-space with twice the usual classical dimension. The number of variables needed to specify a single coherent state configuration grows only *linearly* with the lattice size because they are separable, while correlations between subsystems are contained in the details of the distribution (which is stochastically sampled). This very mild growth of the number of required variables is the reason that this method can lead to tractable many-body simulations.

Dynamics in the positive-P representation takes on the particularly simple form of the widely used Gross–Pitaevskii mean-field equations [13], but with appropriate Gaussian noise terms added. Several interacting Bose gas dynamics simulations with this method have made quantitative predictions, including: (1) the scattering dynamics during the collision of two BECs, and the evolution of momentum correlations between the scattered atoms [16]; (2) dynamics of evaporative cooling and incipient condensation of an interacting Bose gas [14, 15]; (3) dynamical behaviour of spatial correlations in one-dimensional interacting Bose gases [12, 16]; (4) dynamics of quantum soliton propagation in optical fibres [17–19]; and (5) dynamics and steady-state behaviour of open boson systems coupled to reservoirs [6]. Dynamics and thermodynamic spatial correlations of a similar but explicitly particle-conserving model have also been simulated using the related stochastic wavefunction method [8, 20].

While much can be simulated with the positive P, there is a limiting factor—the growth of the scatter of trajectories with time. This statistical error can be estimated, but eventually after some evolution time it becomes large enough to obscure the physical behaviour for any practical number of trajectories. Estimates of the useful simulation time with the positive-P method have been obtained [12], and indicate two major limitations.

(i) When occupation of individual lattice points is high ($\gg 1$), the simulation loses precision before phase-collapse can occur.
(ii) Useful simulation times decrease with reduced lattice spacing.

This loss of precision is often (but not always) associated with the development of power-law tails in the distribution function, which can also lead to systematic boundary term errors. A typical cause of power-law tails is the presence of movable singularities in the dimension-doubled phase-space equations (discussed in section 6). These problems are intimately related, and therefore the methods that improve precision can also solve boundary term issues.

One approach to this problem is to change the basis. For example, a different phase-space technique that can have lower sampling error when mode occupations are high is the truncated Wigner representation [1, 18, 21], sometimes called the classical field method [22].

The simulation can be quite accurate, particularly when all modes are strongly occupied. However, this method involves truncation of third-order differential terms in the evolution. This can lead to large systematic errors which are essentially intrinsic to the method itself, and tend to grow with time [23]. The conditions under which it is valid for interacting Bose gases have been investigated in significant detail in [24].

Here, however, we will investigate a different avenue of improvement. For a given basis, there is a wide range of freedom in the correspondence between quantum mechanics and the stochastic equations [7, 8]. The available freedoms can be constructively specified in a unified way by the stochastic gauge formalism [7, 16], which adds freely defined gauge functions and a corresponding global weight to the set of stochastic variables. Preliminary attempts at harnessing this freedom in single-mode systems have been successful [7, 25] with demonstrations of time-reversible quantum simulations with up to 10^{23} interacting particles [26]. In this paper we investigate these methods systematically, and also present the progress that has been achieved in exploiting stochastic gauges for large multi-mode systems.

We restrict ourselves to the case of gauge functions defined locally (i.e., separately for each lattice point), which is the simplest useful case. Sections 3 and 4 investigate two gauge approaches, each with its own merits. Subsequently, in sections 5, 7, and 8, the performance of the proposed improved methods is assessed for single-mode, double-mode and multi-mode cases, respectively. Section 6 investigates convergence issues and identifies two ways by which stochastic equations produced by phase-space methods in general can diverge: movable singularities and a noise-weight relationship. The last section simulates the uniform 1D gas as a non-trivial example. In particular, the dynamics of second-order spatial correlations $\overline{g}^{(2)}(x)$ after a sudden increase in the coupling constant at $t = 0$ are calculated as in the companion paper [12].

2. Model

2.1. The lattice model

We wish to solve for the quantum time-evolution of a dilute interacting Bose gas. Following the companion paper [12], we consider a lattice Hamiltonian that contains all the essential features of a continuum model, provided the lattice spacing is sufficiently small. For a rarefied gas of the kind occurring in contemporary BEC experiments, s-wave scattering dominates [27], and the s-wave scattering length a_s is much smaller than all other relevant length scales. If the lattice spacing is also much larger than a_s, then the two-body scattering is well described by an interaction local at each lattice point. Otherwise, a more careful renormalization procedure [28] than the one below is required.

Let us label the spatial dimensions by $d = 1, \ldots, \mathcal{D}$, and label lattice points by the vectors $\mathbf{n} = (n_1, \ldots, n_D)$. For lattice spacings $\Delta \mathrm{x}_d$, the spatial coordinates of the lattice points are $\mathbf{x_n} = (n_1 \Delta \mathrm{x}_1, \ldots, n_D \Delta \mathrm{x}_D)$. The volume per lattice point is $\Delta V = \prod_d \Delta \mathrm{x}_d$. This lattice implies effective momentum cutoffs [29] of $k_d^{\max} = \pi/\Delta x_d$. We also define the lattice annihilation operators $\widehat{a}_{\mathbf{n}}$ ($\approx \sqrt{\Delta V}\widehat{\Psi}(\mathbf{x_n})$ in the field notation of equation (1) of [12]), which obey the usual boson commutation relations of $\left[\widehat{a}_{\mathbf{n}}, \widehat{a}_{\mathbf{m}}^{\dagger}\right] = \delta_{\mathbf{nm}}$. With these definitions, one obtains

$$\widehat{H} = \hbar \left[\sum_{\mathbf{nm}} \omega_{\mathbf{nm}} \widehat{a}_{\mathbf{n}}^{\dagger} \widehat{a}_{\mathbf{m}} + \frac{\kappa}{2} \sum_{\mathbf{n}} \widehat{a}^{\dagger 2} \widehat{a}^{2} \right]. \tag{1}$$

In this Hamiltonian, the frequency terms $\omega_{\mathbf{nm}} = \omega_{\mathbf{mn}}^{*}$ 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 the kinetic processes. The nonlinearity due to local particle–particle interactions is of strength

$$\kappa = \frac{g}{\hbar \Delta V}, \tag{2}$$

with the standard coupling value [27] being $g = g_{3\mathrm{D}} = 4\pi a_s \hbar^2/m$ in 3D, and $g = g_{3\mathrm{D}}/\sigma$ in 2D and 1D, where σ is the effective thickness or cross-section of the collapsed dimensions.

When the interaction with the environment is a Markovian (i.e., no feedback), the evolution of the density matrix $\widehat{\rho}$ can be written as a master equation [30, 31] in the Lindblad [32] form

$$\frac{\partial \widehat{\rho}}{\partial t} = \frac{1}{\mathrm{i}\hbar}[\widehat{H}, \widehat{\rho}] + \frac{1}{2}\sum_j \left[2\widehat{L}_j \widehat{\rho} \widehat{L}_j^\dagger - \widehat{L}_j^\dagger \widehat{L}_j \widehat{\rho} - \widehat{\rho} \widehat{L}_j^\dagger \widehat{L}_j\right]. \tag{3}$$

For example, single-particle losses at rate $\gamma_{\mathbf{n}}$ (at $\mathbf{x}_{\mathbf{n}}$) to a $T = 0$ heat bath are described by $\widehat{L}_{\mathbf{n}} = \widehat{a}_{\mathbf{n}}\sqrt{\gamma_{\mathbf{n}}}$.

2.2. Gauge-P representation

The gauge-P representation was introduced in [6] and is described in more detail in [7, 16]. Here we summarize the issues relevant to the dynamics of the model (1), and present the stochastic equations to simulate.

For a lattice with M points, the density matrix is expanded as

$$\widehat{\rho} = \int G(\boldsymbol{\alpha}, \boldsymbol{\beta}, \Omega)\widehat{\Lambda}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \Omega) d^{2M}\boldsymbol{\alpha} d^{2M}\boldsymbol{\beta}\, \mathrm{d}^2\Omega, \tag{4}$$

where $\widehat{\Lambda}$ is an off-diagonal operator kernel, separable between the M lattice point subsystems. We use a coherent state basis so that

$$\widehat{\Lambda}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \Omega) = \Omega \|\boldsymbol{\alpha}\rangle\langle\boldsymbol{\beta}^*\| \exp[-\boldsymbol{\alpha}\cdot\boldsymbol{\beta}], \tag{5}$$

in terms of Bargmann coherent states with complex amplitudes $\boldsymbol{\alpha} = (\ldots, \alpha_{\mathbf{n}}, \ldots)$

$$\|\boldsymbol{\alpha}\rangle = \otimes_{\mathbf{n}} \exp\left[\alpha_{\mathbf{n}} \widehat{a}_{\mathbf{n}}^\dagger\right]|0\rangle. \tag{6}$$

Ω is a complex weight.

The gauge representation is a generalization of the positive-P representation, which has no weight term. With the choice

$$G(\boldsymbol{\alpha}, \boldsymbol{\beta}, \Omega) = P_+(\boldsymbol{\alpha}, \boldsymbol{\beta})\delta^2(\Omega - 1), \tag{7}$$

we recover the positive-P distribution P_+. It has been shown that all density matrices can be represented by a positive real P_+ [5], so the same is true for G. The expansion (4) then becomes a probability distribution of the $\widehat{\Lambda}$, or equivalently the variables $\boldsymbol{\alpha}, \boldsymbol{\beta}, \Omega$. A constructive expression for a distribution $P_+(\widehat{\rho})$ is given by expression (3.7) in [5], although more compact distributions may exist. In particular, a coherent state $|\boldsymbol{\alpha}_o\rangle$ will simply have

$$G = \delta^{2M}(\boldsymbol{\alpha} - \boldsymbol{\alpha}_o)\delta^{2M}(\boldsymbol{\beta} - \boldsymbol{\alpha}_o^*)\delta^2(\Omega - 1). \tag{8}$$

Using the identities

$$\widehat{a}_{\mathbf{n}}\widehat{\Lambda} = \alpha_{\mathbf{n}}\widehat{\Lambda}, \tag{9a}$$

$$\widehat{a}_{\mathbf{n}}^\dagger\widehat{\Lambda} = \left[\beta_{\mathbf{n}} + \frac{\partial}{\partial\alpha_{\mathbf{n}}}\right]\widehat{\Lambda}, \tag{9b}$$

$$0 = \left[1 - \Omega\frac{\partial}{\partial\Omega}\right]\widehat{\Lambda}, \tag{9c}$$

the master equation (3) in $\widehat{\rho}$ can be shown to be equivalent to a Fokker–Planck equation in G, and then to stochastic equations in the $\alpha_{\mathbf{n}}$, $\beta_{\mathbf{n}}$ and Ω, provided that the boundary terms vanish in carrying out the partial integration used to derive the Fokker–Planck equation. This implies a restriction on the phase-space distribution tails, which typically must vanish faster than any power law.

The standard method is described in [31]. The correspondence is in the sense that appropriate stochastic averages of these variables correspond to quantum expectation values in the limit when the number of trajectories $\mathcal{S} \to \infty$. In particular, one finds that [6]

$$\left\langle \prod_{jk} \widehat{a}^{\dagger}_{\mathbf{n}_j} \widehat{a}_{\mathbf{n}_k} \right\rangle = \lim_{\mathcal{S}\to\infty} \frac{\left\langle \prod_{jk} \beta_{\mathbf{n}_j} \alpha_{\mathbf{n}_k} \Omega + \prod_{jk} \alpha^*_{\mathbf{n}_j} \beta^*_{\mathbf{n}_k} \Omega^* \right\rangle_{\mathrm{s}}}{\langle \Omega + \Omega^* \rangle_{\mathrm{s}}}. \tag{10}$$

Any observable can be written as a linear combination of terms (10). We will use the notation $\langle\cdot\rangle_{\mathrm{s}}$ to distinguish averages of random variables from quantum expectations $\langle\cdot\rangle$.

Due to the fact that the basis vectors $\|\boldsymbol{\alpha}\rangle$ are not mutually orthogonal, many different distributions (and hence, sets of equations) correspond to the same master equation. It has been found [7, 16] that the family of stochastic equations corresponding to a given master equation is parameterized by *stochastic gauge functions* of several kinds. These are completely arbitrary functions that appear in the equations, but do not affect (10). However, the rate at which precision of the estimates (10) improves with more stochastic realizations can be strongly affected. It is shown below that judicious choices of the gauges can lead to large improvements in precision.

The $2M + 1$ Ito stochastic equations to simulate are found using methods described in [7, 31]. For the model (1), obeying the master equation (3) with coupling to a zero temperature heat bath, they are

$$\mathrm{d}\alpha_{\mathbf{n}} = \left[-\mathrm{i} \sum_{\mathbf{m}} \omega_{\mathbf{nm}} \alpha_{\mathbf{m}} - \mathrm{i}\kappa n_{\mathbf{n}} \alpha_{\mathbf{n}} - \gamma_{\mathbf{n}} \alpha_{\mathbf{n}}/2 \right] \mathrm{d}t + \sum_k B^{(\alpha)}_{\mathbf{n}k} (\mathrm{d}W_k - \mathcal{G}_k\, \mathrm{d}t), \tag{11a}$$

$$\mathrm{d}\beta_{\mathbf{n}} = \left[\mathrm{i} \sum_{\mathbf{m}} \omega^*_{\mathbf{nm}} \beta_{\mathbf{m}} + \mathrm{i}\kappa n_{\mathbf{n}} \beta_{\mathbf{n}} - \gamma_{\mathbf{n}} \beta_{\mathbf{n}}/2 \right] \mathrm{d}t + \sum_k B^{(\beta)}_{\mathbf{n}k} (\mathrm{d}W_k - \mathcal{G}_k\, \mathrm{d}t), \tag{11b}$$

$$\mathrm{d}\Omega = \Omega \left[\sum_k \mathcal{G}_k\, \mathrm{d}W_k \right]. \tag{11c}$$

Here, $n_{\mathbf{n}} = \alpha_{\mathbf{n}} \beta_{\mathbf{n}}$, and there are $M' \geqslant 2M$ labels k for the independent noise terms, to sum over. The $\mathrm{d}W_k$ are independent Wiener increments $\langle \mathrm{d}W_j(t)\, \mathrm{d}W_k(s)\rangle_{\mathrm{s}} = \delta_{jk}\delta(t-s)\, \mathrm{d}t^2$. In practice, these can be realized at each time step Δt by mutually- and time-independent real Gaussian noises of mean zero, and of variance Δt. The elements of the $M \times M'$ *noise matrices* B must satisfy

$$\sum_k B^{(\alpha)}_{\mathbf{n}k} B^{(\alpha)}_{\mathbf{m}k} = -\mathrm{i}\kappa \alpha^2_{\mathbf{n}} \delta_{\mathbf{nm}} \tag{12a}$$

$$\sum_k B^{(\beta)}_{\mathbf{n}k} B^{(\beta)}_{\mathbf{m}k} = \mathrm{i}\kappa \beta^2_{\mathbf{n}} \delta_{\mathbf{nm}}, \tag{12b}$$

$$\sum_k B^{(\alpha)}_{\mathbf{n}k} B^{(\beta)}_{\mathbf{m}k} = 0. \tag{12c}$$

The M complex quantities $\mathcal{G}_k$ are *arbitrary* complex gauge functions, referred to as *drift stochastic gauges*. There is also more gauge freedom here because (2.2) do not specify the noise matrices $B^{(\alpha)}$ and $B^{(\beta)}$ precisely. This freedom can be expressed as *diffusion stochastic gauges* [6, 7, 25].

The simulation strategy is (briefly) as follows.

(i) Sample a trajectory according to the known initial condition $G(0) = G(\widehat{\rho}(0))$.
(ii) Evolve according to the stochastic equations (2.2), calculating moments of interest and recording.
(iii) Repeat for $\mathcal{S} \gg 1$ independent trajectories and average.

2.3. Single-mode model

The single-mode model is a good test bed for gauge choices in quantum many-body theory. It is exactly solvable, yet it already contains all the essential features that lead to the rapid growth of fluctuations which limits simulation time with the positive-P method [12]. It has also been experimentally realized in recent optical lattice experiments [33] with Bose–Einstein condensates. We proceed as in the companion paper [12].

To simplify the notation, the mode labels $\mathbf{n} = (0, \ldots, 0)$ will be omitted when referring to the single-mode system. Furthermore, we move to an interaction picture where the harmonic oscillator evolution due to the $\omega\widehat{a}^\dagger\widehat{a}$ term in the Hamiltonian is implicitly contained within the Heisenberg evolution of the operators. Then, this 'anharmonic oscillator' model simply has

$$\widehat{H} = \frac{\hbar\kappa}{2}\widehat{a}^{\dagger 2}\widehat{a}^2. \tag{13}$$

When dealing with this system, it will be convenient to consider the evolution starting with an off-diagonal coherent-state kernel

$$\widehat{\Lambda}_0 = \widehat{\Lambda}(\alpha_0, \beta_0, 1), \tag{14}$$

with 'particle number' $n_0 = \alpha_0\beta_0$ and initial unit weight $\Omega(0) = 1$. This is because for any general initial state, each sampled trajectory will start as an $\widehat{\Lambda}_0$ with some coherent amplitudes.

With initial condition $\widehat{\Lambda}_0$, analytic expressions for observables can be readily obtained. In particular, the first-order time-correlation function

$$G^{(1)}(0, t) = \beta_0\langle\widehat{a}\rangle = \alpha_0^*\langle\widehat{a}^\dagger\rangle^*, \tag{15}$$

which contains phase coherence information. Normalizing by $\mathrm{Tr}[\widehat{\Lambda}_0]$,

$$G^{(1)}(0, t) = n_0\,\mathrm{e}^{-\gamma t/2}\exp\left\{\frac{n_0}{1 - \mathrm{i}\gamma/\kappa}(\mathrm{e}^{-\mathrm{i}\kappa t - \gamma t} - 1)\right\}. \tag{16}$$

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 is

$$t_{\mathrm{osc}} = \frac{1}{\kappa}\sin^{-1}\left(\frac{2\pi}{n_0}\right) \approx \frac{2\pi}{\kappa n_0}, \tag{17}$$

and the phase collapse time [21] over which $|G^{(1)}(0, t)|$ decays[3] is

$$t_{\mathrm{coh}} = \frac{1}{\kappa}\cos^{-1}\left(1 - \frac{1}{2n_0}\right) \approx \frac{1}{\kappa\sqrt{n_0}}. \tag{18}$$

When there is no damping, the first quantum revival occurs—as observed experimentally [33]—at

$$t_{\mathrm{revival}} = \frac{2\pi}{\kappa}. \tag{19}$$

[3] We have kept the notation t_{coh} from the companion paper [12], where 'coh' indicates that coherence is maintained.

Table 1. Some properties of simulation times $t_{\rm sim}$ for various gauges when applied to the undamped one-mode anharmonic oscillator. Initial coherent state mean occupation $n_0 = n_0'$.

Drift gauge $\mathcal{G}_k$	Diffusion gauge g''	Useful simulation time $t_{\rm sim}$ maximized over $t_{\rm opt}$, when $n_0 = n_0' \gg 1$	Maximum n_0 for which $\chi t_{\rm sim} \geqslant 1$
(67)	0	$(1.06 \pm 0.16)\ t_{\rm osc}$	$0.014\ ^{+0.016}_{-0.008}$
(59)	0	$(1.7 \pm 0.4)\ t_{\rm osc}$	$0.08\ ^{+0.07}_{-0.05}$
0	0	$(2.5 \pm 0.2)\ t_{\rm coh}/{n_0'}^{1/6}$	0.11 ± 0.06
0	(68) or (36)	$(8.2 \pm 0.4)\ t_{\rm coh}$	12 ± 3
0	(42)	$(10.4 \pm 0.7)\ t_{\rm coh}$	19 ± 4
(59)	(56)	$(30 \pm 3)\ t_{\rm coh}$	150 ± 40
(59)	(58)	$(35 \pm 4)\ t_{\rm coh}$	240 ± 70

2.4. Positive-P noise behaviour

The noise behaviour of this model and simulation times has been investigated in detail for the positive-P case in [12]. Some useful conclusions include.

(i) If ξ is a Gaussian random variable, and $v = \mathrm{e}^{\sigma\xi}$, then finite-sample estimates of $\langle v\rangle_{\rm s}$ using $\mathcal{S}$ realizations of the random variable have relative fluctuations that scale as $\propto\sqrt{\left(\mathrm{e}^{\sigma^2}-1\right)/\mathcal{S}}$. Due to the rapid rise after $\sigma \gtrsim 1$, reasonable precision can only be obtained with practical sample sizes ($\mathcal{S} \lesssim 10^6$, say) while

$$\sigma^2 \lesssim 10. \tag{20}$$

(ii) In a single-mode system, the variable $n(t)$ is the exponential of a Gaussian random variable, as is $\alpha(t)$ at short times. Since observable estimates (10) involve means of polynomial functions f of these variables $\langle f\rangle_{\rm s}$, then one needs

$$\mathrm{var}[\log|f|] \lesssim \mathrm{O}(10) \tag{21}$$

for reasonable precision.

(iii) The mechanism responsible for limiting simulation time is growth of fluctuations in $\log|\alpha|$ due to the real part of the $\mathrm{d}\log\alpha = -\mathrm{i}\kappa n\,\mathrm{d}t$ term. When $|n| \gg 1$, the fluctuations in $\mathrm{Im}[n]$ are sizeable, and $\mathrm{var}[\log|\alpha|]$ rapidly reaches large values that exceed (21).

(iv) In the multi-mode system, this nonlinear (in $\alpha_{\mathbf{n}}$) term becomes $\mathrm{d}\alpha_{\mathbf{n}} = -\mathrm{i}\kappa n_{\mathbf{n}}\alpha_{\mathbf{n}}\,\mathrm{d}t$, which depends only on the variables in the local mode $\mathbf{n}$. For this reason, the single-mode analysis continues to give a good qualitative description of the noise behaviour in multi-mode systems, especially while the inter-mode coupling is weak in comparison with the local scattering.

2.5. Aims

In what follows, an estimate $t_{\rm sim}$ of the useful simulation time will be evaluated using condition (21) for the amplitude variables α and β. For the positive-P method applied to the system (13), $t_{\rm sim}$ was found to be relatively shortest at high-mode occupation and weak damping, scaling as $\propto n^{-2/3}$. This is shown in figure 3 and table 1. For $n \gg 1$, we see that the positive-P method does not reach even the phase collapse times $t_{\rm coh}$ [21]. This is caused by the phase-diffusion inherent in any initial state with a range of particle numbers, where the self-interactions cause the phase to evolve at a different rate for different total particle number.

However, indications that simulation improvements can be obtained using nonzero gauge choices were seen by Deuar and Drummond [34] and Plimak *et al* [25], using drift and diffusion modifications, respectively. Carusotto *et al* [8] also found improvements in a

different (number-conserving) model, using an approach that effectively uses a drift gauge (see section 5.4 for more on this).

The aim here is to develop gauges that improve simulation times at high-mode occupation, while being applicable to the many-mode situation where conditions for a single mode are dynamically changing. To do this we wish to find the dependence of advantageous gauges on mode parameters such as $n(t)$, constants κ, γ, and time t itself. We are especially interested in the case of low or absent damping $\gamma \ll \kappa$, in highly occupied modes $n \gg 1$. This is the regime where the simulation time is most limited with the positive-P method.

3. Local diffusion gauges

In the positive-P simulation of modes with $n \gg 1$ occupations, t_{sim} is limited by the following process: in order to allow for an exact simulation of quantum fluctuations, the nonclassical phase space requires the effective particle number $n = n' + \mathrm{i}n''$ to develop complex values. This in turn gives rise to exponential growth in amplitude fluctuations in either $|\alpha|$ or $|\beta|$, since the drift equations have the structure

$$\frac{\mathrm{d}|\alpha|}{\mathrm{d}t} = \kappa n''|\alpha| + \cdots \tag{22a}$$

$$\frac{\mathrm{d}|\beta|}{\mathrm{d}t} = -\kappa n''|\beta| + \cdots. \tag{22b}$$

These instabilities lead to growth in sampling error, and possibly ultimately to systematic boundary term errors as well. In this section we consider the freedoms present in the choice of noise matrix B, to limit the stochastic growth in the effective 'gain' n'', without introducing any drift gauges. Such techniques can extend the useful time-scale of a simulation, although they are typically unable to remove boundary terms at long times if caused by movable singularities.

In this case $\mathcal{G}_k = 0$, $\mathrm{d}\Omega = 0$, so there is no weight evolution and we can assume $\Omega = 1$ for all practical purposes. Analytic expressions for an optimized diffusion gauge choice will be found and discussed. Their performance in actual numerical simulations is reported in sections 5–8.

3.1. Diffusion gauge mechanism

The noise matrices appearing in the stochastic equations must obey (2.2) which can be written as a matrix equation $BB^T = D$, where

$$B = \begin{bmatrix} B^{(\alpha)} \\ B^{(\beta)} \end{bmatrix}, \tag{23}$$

and D is completely determined by the master equation. Otherwise the B are free, including the number of noises M', i.e. columns in the matrix. These freedoms in B choices are investigated in [16]. A broad class of suitable B which are square complex orthogonal matrices have been described in [7]. These are given by

$$B(g_{jk}) = \sqrt{D}\,O(g_{jk}), \tag{24}$$

where O are arbitrary complex $2M \times 2M$ orthogonal matrices such that $OO^T = I$. In the general case, these can be constructed using $(2M-1)M$ complex *diffusion gauge* functions g_{jk}

$$O = \exp\left(\sum_{j<k} g_{jk}\sigma^{(jk)}\right), \tag{25}$$

where the antisymmetric matrix basis ($j \neq k = 1, \ldots, 2M$) has elements

$$\sigma_{lp}^{(jk)} = \delta_{jl}\delta_{kp} - \delta_{jp}\delta_{kl}. \tag{26}$$

3.2. Single mode

We proceed similarly to the companion paper [12] for the positive-P case. For the model (13) with damping to a zero temperature heat bath, the equations to simulate are

$$\begin{aligned} \mathrm{d}\alpha &= -\alpha[\mathrm{i}\kappa n + \gamma/2]\,\mathrm{d}t + \mathrm{i}\sqrt{\mathrm{i}\kappa}\alpha[\cos(g)\,\mathrm{d}W_1 + \sin(g)\,\mathrm{d}W_2] \\ \mathrm{d}\beta &= \beta[\mathrm{i}\kappa n - \gamma/2]\,\mathrm{d}t + \sqrt{\mathrm{i}\kappa}\beta[-\sin(g)\,\mathrm{d}W_1 + \cos(g)\,\mathrm{d}W_2]. \end{aligned} \tag{27}$$

with $n = \alpha\beta$. Here we include only one complex diffusion gauge $g = g_{12} = g' + \mathrm{i}g''$.

3.2.1. Real gauges. Note that

$$B = B_o O(g) = B_o S(g'')R(g'). \tag{28}$$

in terms of a rotation R and a transformation S. The rotation serves only to mix the noises $\mathrm{d}W_j$ together

$$R(g')\,\mathrm{d}W = \begin{bmatrix} \cos g' & \sin g' \\ -\sin g' & \cos g' \end{bmatrix} \begin{bmatrix} \mathrm{d}W_1 \\ \mathrm{d}W_2 \end{bmatrix} = \begin{bmatrix} \mathrm{d}\widetilde{W}_1 \\ \mathrm{d}\widetilde{W}_2 \end{bmatrix}, \tag{29}$$

where the new noises $\mathrm{d}\widetilde{W}_j$ have the same statistical properties as the original noises $\mathrm{d}W_j$. Because of this, g' has no impact on the statistical properties of the simulation. Henceforth we will consider only $g = \mathrm{i}g''$, where $\cos g = \cosh g''$, and $\sin g = \mathrm{i}\sinh g''$.

3.2.2. Logarithmic variance. As in [12] for the standard positive-P case, we consider the mean phase variable variance

$$\mathcal{V} = \tfrac{1}{2}\{\mathrm{var}[\log|\alpha(t)|] + \mathrm{var}[\log|\beta(t)|]\}, \tag{30}$$

which was found to be the limiting factor for simulation time when it exceeds $\mathcal{V} \gtrsim \mathrm{O}(10)$ as per (21).

Taking $g = \mathrm{i}g''$, the equations (27) can be formally solved using the rules of Ito calculus as

$$\log n(t) = \log n(0) - \gamma t + \sqrt{\mathrm{i}\kappa}\,\mathrm{e}^{-g''}[\zeta_2(t) + \mathrm{i}\zeta_1(t)], \tag{31a}$$

$$\log\alpha(t) = \log\alpha(0) + (\mathrm{i}\kappa - \gamma)t/2 + \mathrm{i}\sqrt{\mathrm{i}\kappa}\zeta_1(t)\cosh g'' - \sqrt{\mathrm{i}\kappa}\zeta_2(t)\sinh g'' - \mathrm{i}\kappa\int_0^t n(s)\,\mathrm{d}s, \tag{31b}$$

where

$$\zeta_j(t) = \int_{s=0}^{t} \mathrm{d}W(s) \tag{31c}$$

are Gaussian-distributed random variables of mean zero, and

$$\langle\zeta_j(t)\zeta_k(s)\rangle_{\mathrm{s}} = \delta_{jk}\min[t, s]. \tag{32}$$

It is convenient to define

$$n_0 = \alpha_0\beta_0 = n_0' + \mathrm{i}n_0''. \tag{33}$$

Note that $n(t)$ is the exponential of a Gaussian random variable, and that if a generic variable ξ is Gaussian then

$$\lim_{\mathcal{S}\to\infty} \langle \xi(t)^m \rangle_{\mathrm{s}} = \frac{m!}{2^{m/2}(m/2)!} \tag{34}$$

for even m, and zero for odd m. From this, $\lim_{\mathcal{S}\to\infty}\langle \mathrm{e}^{\sigma\xi}\rangle_s = \exp(\sigma^2/2)$ and $\lim_{\mathcal{S}\to\infty}\langle \cos(\sigma\xi)\rangle_s = \exp(-\sigma^2/2)$. Using these we obtain that the limiting variance $\mathcal{V}$ in the limit of a large number of samples $\mathcal{S}\to\infty$ is given by

$$\mathcal{V} = \kappa t \cosh(2g'')/2 - \kappa^2 n_0' \left\{ \frac{1-\mathrm{e}^{-\gamma t}(1+\gamma t)}{\gamma^2} \right\} + \frac{\kappa^2 |n_0|^2}{q\gamma} \left\{ 1 - \frac{q\,\mathrm{e}^{-\gamma t} - \gamma\,\mathrm{e}^{-qt}}{(q-\gamma)} - \frac{q}{2\gamma}(1-\mathrm{e}^{-\gamma t})^2 \right\}, \tag{35a}$$

where

$$q = 2(\gamma - \kappa\,\mathrm{e}^{-2g''}) \tag{35b}$$

is a 'damping strength' parameter.

At short times the first term (direct fluctuations from noise in $\mathrm{d}\log|\alpha|$ etc.) certainly dominates. For long time-scales, the indirect noise mediated by a growing spread in n'' (remaining terms) has the most serious consequences, since it causes an exponential growth in the sampled fluctuations.

3.2.3. Optimum gauge. The first term in (35*a*) (due to direct noise in $\log\alpha$ and $\log\beta$) grows with g'', while the later terms decrease. This indicates that there is a trade-off parameterized by g''. The optimum choice is given by a balance between the direct noise in the amplitudes, and the indirect noise in the gain term n''. The lowest fluctuations $\mathcal{V}$ at a given time t will occur when $\partial\mathcal{V}(t)/\partial g'' = 0$. To choose g'', we must decide upon a 'target time' $t = t_{\mathrm{opt}}$ at which to minimize $\mathcal{V}$.

When we aim for relatively short target times t_{opt} such that $|q|t_{\mathrm{opt}} \ll 1$ and $2\kappa t_{\mathrm{opt}}\,\mathrm{e}^{-2g''} \ll 1$, the optimized gauge is approximated by

$$g''_{\mathrm{opt}} \approx g''_{\mathrm{approx}} = \frac{1}{4}\log\left[\frac{(2\kappa t_{\mathrm{opt}}|n_0|)^2}{3} a_2(\gamma t_{\mathrm{opt}}) + 1\right], \tag{36a}$$

where, on defining $v = \gamma t_{\mathrm{opt}}$, the coefficient $a_2(v)$ is

$$a_2(v) = \frac{3}{2}\left(\frac{\mathrm{e}^{-2v}(3+2v) + 1 - 4\,\mathrm{e}^{-v}}{v^3}\right). \tag{36b}$$

This reduces to $a_2(0) = 1$ in the undamped case. The discrepancy between (36) and the exact optimization g''_{opt} found using $\partial\mathcal{V}/\partial g'' = 0$ is shown for real $n_0 = n_0'$ in figure 1. It can be seen that for $n_0' \gtrsim \mathrm{O}(10)$ and/or for times shorter than singly-occupied coherence time $1/\kappa$, the approximate expression for the optimized gauge choice is still useful.

3.2.4. Useful simulation times. By inspection of (31), the behaviour of $\log|\alpha|$ and $\log|\beta|$ is Gaussian-like due to the ζ_j terms. Since $G^{(1)} \propto \langle\alpha\rangle_{\mathrm{s}}$, condition (21) implies that useful precision in $G^{(1)}$ is obtainable only while $\mathcal{V} \lesssim 10$.

In the simplest case of no damping and coherent state initial conditions at large occupation $n_0 = n_0' \gg 1$, for target times $t_{\mathrm{opt}} \gg 3t_{\mathrm{osc}}/4\pi$ one has $\mathrm{e}^{-2g''_{\mathrm{approx}}} \approx \sqrt{3}/2\kappa t_{\mathrm{opt}} n_0$. These are times longer than an oscillation period, but shorter than coherence time t_{coh}. Taking then the

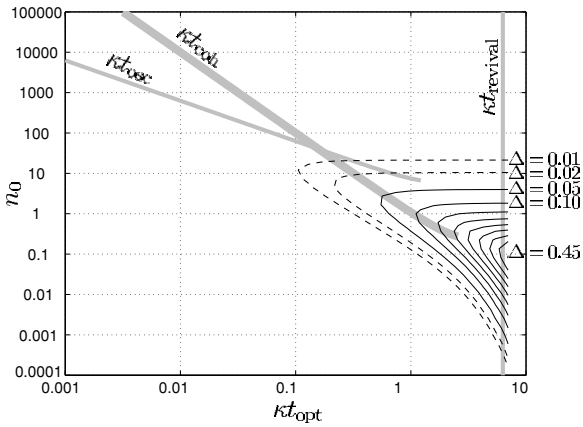

Figure 1. Discrepancy $\Delta = g''_{\rm opt} - g''_{\rm approx}$ between $g''_{\rm opt}$ (the exact optimization of g'' for the diffusion-gauge-only case) and the approximate expression (36) for an undamped mode with diagonal coherent initial occupation $n_0 = n'_0$. Discrepancy values Δ are shown as contours.

terms of $\mathcal{V}(t, g'')$ from (36) at $t = t_{\rm opt}$ of highest order in n_0, and imposing the precision limit $\mathcal{V}(t_{\rm sim}, g''_{\rm approx}) \lesssim 10$, one obtains the expected simulation time for an undamped system as

$$t_{\rm sim} \approx {\rm O}(10) t_{\rm coh}. \tag{37}$$

This is a large improvement at high mode occupations—compare to the positive-P ($g'' = 0$) results in table 1.

Checking back, we see that when the target time is given by $t_{\rm opt} = t_{\rm sim}$ from (37), it is indeed much longer than an oscillation period, and also that the terms of $\mathcal{V}$ highest order in n_0 remain so after taking $\kappa t \lesssim \propto n_0^{-1/2}$ into account. The clearest evidence for the validity of the above reasoning—which involves a few approximations—is that the numerical simulations of section 5 agree with (37).

At low occupations and with $n_0 = n'_0$, on the other hand, $g''_{\rm approx} \to 0$, and one again expects the same simulation time as with the standard positive-P equations.

3.2.5. Strong damping. If damping is present, then with a large enough gauge $g'' > \frac{1}{2}\log(\kappa/\gamma)$ such that $q > 0$ the regime of linear increase in sampling variance can always be reached for times $qt \gg 1$. Here, fluctuations grow slowly as

$$\mathcal{V} = \kappa t \cosh(2g'')/2 - b. \tag{38}$$

The constant b is

$$b = \frac{\epsilon^2}{2}[2n'_0 - |n_0|^2 \epsilon/({\rm e}^{2g''} - \epsilon)], \tag{39}$$

where $\epsilon = \kappa/\gamma < {\rm e}^{2g''}$. The required $q > 0$ implies $\gamma > \kappa\, {\rm e}^{-2g''}$, i.e. either that damping rates are strong compared to the nonlinear detuning at the two-particle level, or the diffusion gauge is large.

3.3. Extension to many-modes

Expression (36) was worked under the assumption that the mean occupation of the mode is conserved. In coupled-mode simulations, this is no longer the case, and can be adapted for by replacing n_0 in expression (36) by $n_{\mathbf{n}}(t)$.

Next, consider the situation when we aim to simulate until a target time $t_{\rm opt}$. At some intermediate time $0 < t < t_{\rm opt}$ it may be advantageous to optimize g'' only for the *remaining* time to target

$$t_{\rm rem} = \max[t_{\rm opt} - t, 0], \tag{40}$$

rather than for a constant target time $t_{\rm opt}$ ahead of t. Indeed, this is demonstrated to be an improvement in section 5.3.

There will be a separate local gauge at each lattice point, such that for an M-mode system (with $j = 1, \ldots, M$ labelling the modes) the only nonzero noise matrix elements are

$$\begin{aligned} B^{(\alpha)}_{\mathbf{n}_j,j} &= \mathrm{i}\sqrt{\mathrm{i}\kappa}\alpha_{\mathbf{n}_j}\cosh g''_{\mathbf{n}_j} & B^{(\alpha)}_{\mathbf{n}_j,(j+M)} &= -\sqrt{\mathrm{i}\kappa}\alpha_{\mathbf{n}_j}\sinh g''_{\mathbf{n}_j} \\ B^{(\beta)}_{\mathbf{n}_j,j} &= -\mathrm{i}\sqrt{\mathrm{i}\kappa}\beta_{\mathbf{n}_j}\sinh g''_{\mathbf{n}_j} & B^{(\beta)}_{\mathbf{n}_j,(j+M)} &= \sqrt{\mathrm{i}\kappa}\beta_{\mathbf{n}_j}\cosh g''_{\mathbf{n}_j}. \end{aligned} \tag{41}$$

The suggested diffusion gauges are

$$g''_{\mathbf{n}} = \frac{1}{4}\log\left[\frac{(2\kappa t_{\rm rem}|n_{\mathbf{n}}(t)|)^2}{3}a_2(\gamma_{\mathbf{n}}t_{\rm rem}) + 1\right]. \tag{42}$$

3.4. Effect of lattice spacing

Labelling α or β as z, it was found [12] that for a uniform gas of density ρ with volume ΔV per lattice point in a $\mathcal{D}$-dimensional lattice one has

$$\frac{\mathrm{var}[|\delta^{\rm kinetic}z_{\mathbf{n}}(t)|]}{\mathrm{var}[|\delta^{\rm direct}z_{\mathbf{n}}(t)|]} = \mathcal{R} \approx \mathrm{O}\left[\frac{\hbar^2 t^2 \pi^4}{60 m^2 (\Delta V)^{4/\mathcal{D}}}\right] \tag{43}$$

at short times (such that the first term in (35*a*) dominates) when using the standard positive-P equations. Here $\delta^{\rm direct}z_{\mathbf{n}}$ are the fluctuations in $z_{\mathbf{n}}$ directly due to the noise in the equations, and present with zero coupling $\omega_{\mathbf{nm}} = 0$, whereas $\delta^{\rm kinetic}z_{\mathbf{n}}$ are the remaining fluctuations induced by coupling between modes. The short time assumption was $\mathrm{var}[\log|z|] \ll 1$, which is approximately $\kappa t \cosh 2g'' \ll 2$, from the formal single-mode solutions (31). A calculation that takes into account nonzero g'' is found to give the same expression (43).

The single-mode analysis is expected to be accurate only when this ratio $\mathcal{R} \ll 1$. For this reason, the simulation time improvements due to the local diffusion gauge (42) are only to be expected while $\mathcal{R} \lesssim 1$, or perhaps even $\mathcal{R} \ll 1$. Let us see what this translates to for the case for a uniform gas.

For a lattice derived from a continuum model, it will be convenient to write the lattice interaction strength $\kappa = g/\hbar\Delta V$ in terms of the healing length

$$\xi^{\rm heal} = \frac{\hbar}{\sqrt{2m\rho g}}. \tag{44}$$

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 [35].

In terms of $\xi^{\rm heal}$, the expected simulation time with diffusion gauges (42) is $t_{\rm sim} \approx \mathrm{O}(10)m\sqrt{\overline{n}}(\xi^{\rm heal})^2/\hbar$ from (37), where $\overline{n}$ is the mean-mode occupation. Expression (43) can then be evaluated at $t_{\rm sim}$. Imposing $\mathcal{R} \ll 1$ for times $t \leqslant t_{\rm sim}$ leads to the condition $\overline{\Delta x} \gtrsim \xi^{\rm heal}\overline{n}^{1/4}\mathrm{O}(1)$ where the quantity $\overline{\Delta x} = (\Delta V)^{1/\mathcal{D}}$ is the geometric mean of the lattice spacing. Local gauges only lead to significant simulation time improvements in the single mode when $\overline{n} \gg 1$.

We conclude that *local diffusion gauges will lead to simulation time improvements when*

$$\overline{\Delta \mathrm{x}} \gtrsim \mathrm{O}(\xi^{\mathrm{heal}}). \tag{45}$$

The calculations of section 8 will be seen to be consistent with this. This is a strong limitation on local diffusion gauges, since interesting dynamical effects can readily occur over shorter length scales than this. It suggests that nonlocal diffusion gauges may be more useful for general spatially-varying problems, but these are outside the scope of the present paper.

4. Local drift gauges

A complementary approach to the diffusion gauge is to remove the source of the drift instability in equation (22*b*) by using drift gauges that directly alter the unstable drift equations. Since the nonlinearity takes the local form $\mathrm{d}\alpha_{\mathbf{n}} = -\mathrm{i}\kappa n_{\mathbf{n}}\alpha_{\mathbf{n}}\,\mathrm{d}t + \cdots$ etc, this can be done using drift gauges $\mathcal{G}_k$ dependent only on local parameters. These methods are certainly capable of eliminating movable singularities, which are a common cause of systematic boundary term errors. However, this is accompanied by a corresponding growth in sampling error, due to the introduction of a weight term in the stochastic equations.

4.1. Single mode

For a single mode with imaginary diffusion gauge as in section 3, the Ito equations are

$$\begin{aligned}
\mathrm{d}\alpha &= \alpha[-\mathrm{i}\kappa n - \gamma/2 - \mathrm{i}\sqrt{\mathrm{i}\kappa}(\mathcal{G}_1\cosh g'' + \mathrm{i}\mathcal{G}_2\sinh g'')]\,\mathrm{d}t \\
&\quad + \mathrm{i}\sqrt{\mathrm{i}\kappa}\alpha[\mathrm{d}W_1\cosh g'' + \mathrm{i}\,\mathrm{d}W_2\sinh g''] \\
\mathrm{d}\beta &= \beta[\mathrm{i}\kappa n - \gamma/2 - \sqrt{\mathrm{i}\kappa}(-\mathrm{i}\mathcal{G}_1\sinh g'' + \mathcal{G}_2\cosh g'')]\,\mathrm{d}t \\
&\quad + \sqrt{\mathrm{i}\kappa}\beta[-\mathrm{i}\,\mathrm{d}W_1\sinh g'' + \mathrm{d}W_2\cosh g''] \\
\mathrm{d}\Omega &= \Omega\sum_{k=1}^{2}\mathcal{G}_k\,\mathrm{d}W_k.
\end{aligned} \tag{46}$$

4.1.1. Drift gauge. When $\mathcal{G}_j = 0$, the rapid increase in phase variable variance is due to the process outlined in point 3 of section 2.4. A drift gauge that interrupts this process by removing the offending drift terms $\mathrm{d}\log|\alpha| = \kappa n''\,\mathrm{d}t$ and $\mathrm{d}\log|\beta| = -\kappa n''\,\mathrm{d}t$ is

$$\mathcal{G}_1 = \mathrm{i}\mathcal{G}_2 = -\sqrt{\mathrm{i}\kappa}\,\mathrm{e}^{-g''}n''. \tag{47}$$

4.1.2. Logarithmic variances. There is a price paid, of course, and this is fluctuations in the global weight Ω, such that $\mathrm{var}[\log|\Omega|]$ now scales $\propto \mathrm{e}^{-2g''}$. With a fluctuating weight, quantum phase correlations such as $G^{(1)}$ are now given by

$$G^{(1)}(0,t) = \beta_0\langle\Omega\alpha(t)\rangle_{\mathrm{s}} = \alpha_0^*\langle\Omega\beta(t)\rangle_{\mathrm{s}}^*, \tag{48}$$

using (10). Note that since from (11*c*) by the properties of Ito calculus $\mathrm{d}\langle\Omega\rangle_{\mathrm{s}} = 0$. Thus the denominator $\langle\Omega+\Omega^*\rangle_{\mathrm{s}}$ appearing in (10) can be ignored if the initial distribution is normalized so that $\langle\Omega\rangle_{\mathrm{s}} = 1$. From (48), the relevant random variables in phase-dependent calculations are $(\Omega\alpha)$ and $(\Omega\beta)$, so that the log-phase quantity analogous to $\mathcal{V}$ in the calculations of section 3 is now

$$\mathcal{V}_\Omega = \tfrac{1}{2}(\mathrm{var}[\log|\Omega(t)\alpha(t)|] + \mathrm{var}[\log|\Omega(t)\beta(t)|]) \tag{49}$$

(compare to (30)). Proceeding as before by formally solving (46), we find that for off-diagonal coherent state initial conditions,

$$\mathcal{V}_\Omega = \frac{\kappa}{2}\left\{t\cosh(2g'') - \mathrm{e}^{-2g''}|n_0|^2\left(\frac{\mathrm{e}^{-qt}-1}{q}\right) - \mathrm{e}^{-2g''}[(n_0')^2-(n_0'')^2]\left(\frac{1-\mathrm{e}^{-2\gamma t}}{2\gamma}\right)\right\}. \tag{50}$$

Just as in the pure diffusion gauge case, there is a trade-off between fluctuations due to direct phase variable noise $\propto\cosh(2g'')$ (first term), and fluctuations in the global weight Ω, which are dependent on $\mathrm{e}^{-2g''}$ (following terms). The optimum $g''=g''_{\mathrm{opt}}$ can be calculated by solving $\partial\mathcal{V}_\Omega/\partial g''=0$.

4.1.3. Approximate optimum diffusion gauge. Since the aim is a gauge adaptable to changing mode occupation like (42), it is desirable to obtain an approximate expression for g''_{opt} that can be rapidly evaluated at each time step in a calculation. To do this, we first consider some special cases.

At *short times* such that $|q|t_{\mathrm{opt}} \ll 1$ and $2\kappa t_{\mathrm{opt}}\,\mathrm{e}^{-2g''} \ll 1$ the optimum is given by the roots of the cubic in $V_g = \mathrm{e}^{-2g''_{\mathrm{opt}}}$

$$4\kappa t_{\mathrm{opt}}|n_0|^2V_g^3 + a_3(n_0,\gamma t_{\mathrm{opt}})V_g^2 - 1 = 0 \tag{51}$$

where

$$a_3(n_0,v) = 1 + 4(n_0'')^2\left(\frac{1-\mathrm{e}^{-2v}}{2v}\right) - 2|n_0|^2\left(\frac{1-2v+2v^2-\mathrm{e}^{-2v}}{v}\right). \tag{52}$$

For zero damping $a_3(n_0,0) = 1+4(n_0'')^2$.

In the case of simulations with sizable occupation of modes, and times up to $\mathrm{O}(t_{\mathrm{coh}})$ the short time conditions are satisfied and this cubic applies. When the V_g^2 term is negligible, we have

$$g''_{\mathrm{opt}} \approx \tfrac{1}{3}\log(|n_0|\sqrt{4\kappa t_{\mathrm{opt}}}). \tag{53}$$

This occurs when n_0 is large enough and mostly real, and t is big enough: i.e. when $a_3(n_0,\gamma t) \ll (4\kappa t_{\mathrm{opt}}|n_0|^2)^{2/3}$. For example, when undamped, κt_{opt} must be at least $\gg 1/4|n_0|^2$ with classical initial conditions ($n_0''=0$). 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

$$g''_{\mathrm{opt}} \approx \tfrac{1}{4}\log a_3(n_0,\gamma t_{\mathrm{opt}}). \tag{54}$$

For strong damping $q>0$, we again have linear growth $\mathcal{V}_\Omega = \kappa t\cosh(2g'')/2 - b_2$, where the constant is now

$$b_2 = \frac{\epsilon}{4}\{[(n_0')^2-(n_0'')^2]\,\mathrm{e}^{-2g''} - |n_0|^2/(\mathrm{e}^{2g''}-\epsilon)\}. \tag{55}$$

An approximation for the diffusion gauge that is found to work well in practice (see section 5) is

$$g''_{\mathrm{approx}} = \tfrac{1}{6}\log\{4|n_0|^2\kappa t_{\mathrm{opt}} + a_3(n_0,\gamma t_{\mathrm{opt}})^{3/2}\}, \tag{56}$$

which reduces to (53) and (54) in their limits of applicability. The discrepancy Δ between (56) and the exact optimization obtained by solving $\partial\mathcal{V}_\Omega/\partial g''=0$ is shown for real n_0 in figure 2.

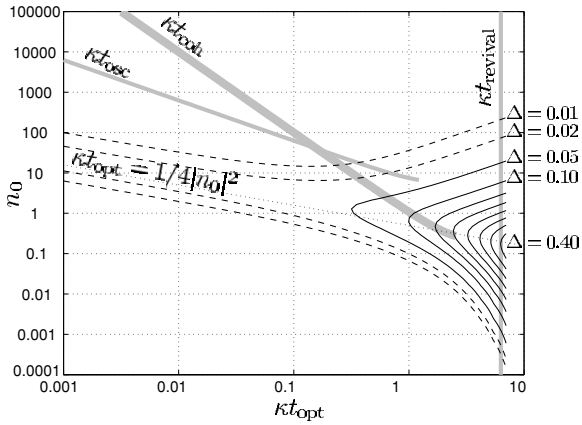

Figure 2. Discrepancy $\Delta = g''_{\rm opt} - g''_{\rm approx}$ between $g''_{\rm opt}$ (the exact optimization of g'' from (50)) and the approximate expression (56) for an undamped mode with diagonal coherent initial occupation $n_0 = n'_0$. Discrepancy values Δ are shown as contours. The dotted line approximates region of greatest discrepancy $\kappa t_{\rm opt} \approx 1/4|n'_0|^2$.

4.1.4. Useful simulation time. Consider the undamped high occupation regime with coherent state initial conditions $n_0 = n'_0 \gg 1$. Using (53), at target times $t_{\rm opt} \gg 1/4\kappa n_0^2$, one has $\mathrm{e}^{-2g''_{\rm approx}} \approx 1\big/\big[4n_0^2\kappa t_{\rm opt}\big)^{1/3} \ll 1$. Then the terms in (50) of highest order in n_0 give $\mathcal{V}_\Omega(t = t_{\rm opt}) \approx \frac{3}{2}[(\kappa t_{\rm opt})^2 n_0/4]^{2/3}$. (Target times of interest will almost always satisfy the prior condition.) We can use the condition $\mathcal{V}_\Omega \lesssim 10$ from (21) to estimate the useful simulation time in this regime as

$$t_{\rm sim} \approx \mathrm{O}(10) t_{\rm coh}. \tag{57}$$

This is again a large improvement compared to the positive-P results in table 1.

4.2. Extension to many modes

We now wish to consider how the drift gauge approach can be used to treat a multi-mode situation.

4.2.1. Adaptive gauge. Proceeding to as in section 3.3, the suggested gauges with the present approach are

$$g''_{\mathbf{n}} = \tfrac{1}{6}\log\{4|n_{\mathbf{n}}(t)|^2\kappa t_{\rm rem} + a_3(n_{\mathbf{n}}(t), \gamma t_{\rm rem})^{3/2}\}, \tag{58}$$

appearing in noise matrices of the form (41), and

$$\mathcal{G}_j = \mathrm{i}\mathcal{G}_{j+M} = -\sqrt{\mathrm{i}\kappa}\,\mathrm{Im}\big[n_{\mathbf{n}_j}(t)\big]\exp\big(-g''_{\mathbf{n}_j}(t)\big). \tag{59}$$

for $j = 1, \ldots, M$.

4.2.2. Drift gauges and weight spread. Drift gauged simulations using (59) encounter a scaling problem in many-mode systems because the single weight variable Ω accumulates fluctuations from all modes (see (11*c*)). There are precisely two independent noises and two drift gauges for each mode. Consider for example, a uniform gas of density ρ, and volume

V on M modes, so that the mean mode occupation is $\overline{n} = \rho\Delta V$. Writing $\mathcal{G}_k = \mathcal{G}'_k + \mathrm{i}\mathcal{G}''_k$ and $\log\Omega = \theta' + \mathrm{i}\theta''$, from (11*c*) one can show that

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\mathrm{var}[\theta'] = \sum_k \langle (\mathcal{G}'_k)^2 + \mathrm{covar}[\theta', (\mathcal{G}''_k)^2] - \mathrm{covar}[\theta', (\mathcal{G}'_k)^2]\rangle_{\mathrm{s}}. \tag{60}$$

For the uniform gas, the contribution from each mode is identical on average, and from (59), $\mathcal{G}'_k$ and $\mathcal{G}''_k$ are of the form $\pm\sqrt{\kappa/2}\,\mathrm{e}^{-g''}\,\mathrm{Im}[n]$. Hence, approximately,

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{var}[\log|\Omega|] \propto M\kappa\,\langle\,\mathrm{e}^{-2g''}\,\mathrm{Im}[n]^2\rangle_{\mathrm{s}}. \tag{61}$$

From the formal solution (31*a*) (which also applies with drift gauges (59), one has $\mathrm{var}[\mathrm{Im}[n]] = \langle(|n|\sin\mathrm{Im}[\log n])^2\rangle_{\mathrm{s}} \approx \overline{n}^2\langle\sin^2\mathrm{Im}[\log n(t)]\rangle_{\mathrm{s}}$, which (at short times $\kappa t\,\mathrm{e}^{-2g''} \ll 1$) is $\approx \overline{n}^2\kappa t\,\mathrm{e}^{-2g''}$. Thus at these short times

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\mathrm{var}[\log|\Omega|] \approx\propto M\kappa^2\overline{n}^2 t\,\mathrm{e}^{-4g''}, \tag{62}$$

and at high lattice occupations $\overline{n} \gg 1$ when using (58) $\mathrm{e}^{-2g''} \approx 1/(4\overline{n}\kappa t)^{1/3}$, so, substituting

$$\mathrm{var}[\log|\Omega|] \approx\ \propto M(g\rho t)^{4/3}. \tag{63}$$

Imposing the log variance limit (21) on $|\Omega|$ (because the factor $|\Omega|$ appears in all observable estimates (10), one obtains the estimate that

$$t_{\mathrm{sim}} \approx\ \propto\frac{1}{g\rho M^{3/4}}. \tag{64}$$

For this reason we expect that simulations using the local diffusion and local drift gauges (59) and (58) will only give significant simulation time improvements when the number of highly occupied modes is relatively small. Indeed, the two mode cases of section 7 show strong simulation time improvement with this gauge, while we have found that the many-mode uniform gases of section 8 do not simulate well with the gauge form (59).

Note, however, that since single-mode t_{sim} decreases rapidly with occupation n, it is the most highly occupied modes that limit the simulation time. So, even a very large M system may still experience improvements in simulation time with the present method if there are only a few modes with the highest occupations.

As in the pure diffusion gauge case, this does not preclude that better scaling may be obtainable with appropriately tailored *nonlocal* drift and diffusion gauges. In particular, the local drift gauge employed here does not take into account the fact that neighbouring lattice points with spacings of less than a healing length become strongly correlated due to the kinetic energy terms. These cause particle exchange, and hence effectively average out local fluctuations in n''.

5. Single mode: numerical results

Simulations of an undamped single-mode anharmonic oscillator (see section 2.3) were performed for a wide range of initial coherent states $n_0 = n'_0$ from 10^{-5} to 10^{10}. Results for the standard positive-P method were reported in [12]. Here the gauged methods described in sections 3 and 4 are tested.

5.1. Procedure

In what follows, the term *useful precision* for an observable $\widehat{O}$ has been taken to indicate the situation where the estimate $\overline{O} = \langle f\rangle_{\mathrm{s}}$ of $\langle\widehat{O}\rangle$ using $\mathcal{S} = 10^6$ trajectories has a relative

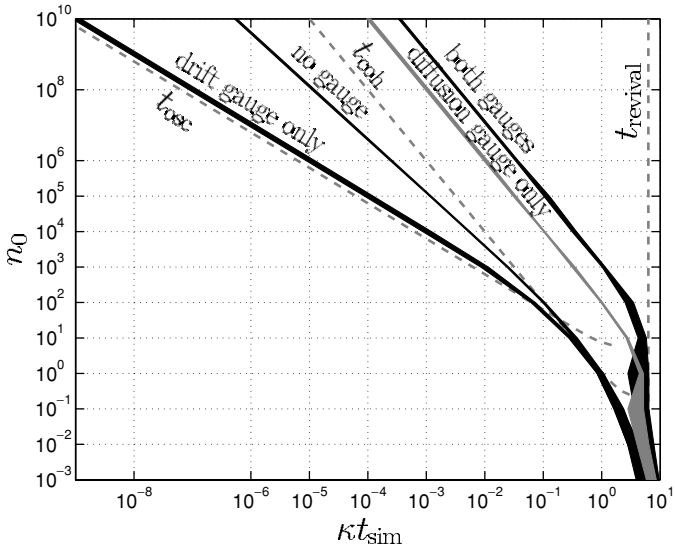

Figure 3. Maximum useful simulation time $t_{\rm sim}$, of the one-mode undamped anharmonic oscillator with various gauge choices. Initial coherent state mean-mode occupations $n_0 = n_0'$. Width of plotted lines shows range, when using ten runs of $\mathcal{S} = 10^4$ trajectories each. The drift gauge is (59), while the diffusion gauge is (42) when on its own, or (58) when combined with the drift gauge.

precision of at least 0.1 at the one-sigma confidence level. This is assuming the CLT holds so that

$$\Delta\overline{O} \approx \sqrt{\frac{\mathrm{var}[f]}{\mathcal{S}}} \tag{65}$$

is used to assess uncertainty in $\overline{O}$. For the model here, we consider useful precision in the magnitude of phase-dependent correlations $|G^{(1)}(0,t)|$, which is the low-order observable most sensitive to the numerical instabilities in the equations.

Uncertainties in the calculated useful times $t_{\rm sim}$ arise because the $\Delta|G^{(1)}|$ were themselves estimated from finite ensembles of $\mathcal{S} = 10^4$ trajectories. The range of $t_{\rm sim}$ indicated in figure 3 was obtained from ten independent runs with identical parameters.

Taking the analytic scalings (37), (57) at high n_0, and from [12] at low n_0 into account, parameters in an approximate curve

$$t_{\rm est} = \frac{1}{\kappa}\left\{[c_1(n_0')^{-c_0}]^{-c_2} + \left[2\log\left(\frac{\mathrm{e}^{c_3}}{{n_0'}^{c_4}} + 1\right)\right]^{-c_2}\right\}^{-1/c_2} \tag{66}$$

have been fitted to the empirical data, as was done for the standard positive-P method in [12]. Best values of c_j are given in table 2. Expression (66) reduces to $c_1 {n_0'}^{-c_0}/\kappa$ and $(c_3 - c_4 \log n_0')/(\kappa/2)$, when $n_0' \gg 1$ and $n_0' \ll 1$, respectively. c_0 determines the high n_0' scaling (here, assumed from analysis of $\mathcal{V}$, or $\mathcal{V}_\Omega$), c_1 characterizes the pre-factor for high n_0', c_3 is proportional to a constant residual $t_{\rm sim}$ at near vacuum, c_4 characterizes the curvature at small n_0', while c_2 is related to the stiffness of the transition between the two regimes. Uncertainty Δc_j in parameters c_j was worked out by requiring $\sum_{n_0}\{[t_{\rm est}(c_j \pm \Delta c_j, n_0) - t_{\rm sim}(n_0)]/\Delta t_{\rm sim}\}^2 = \sum_{n_0}\{1 + ([t_{\rm est}(c_j, n_0) - t_{\rm sim}(n_0)]/\Delta t_{\rm sim})^2\}$.

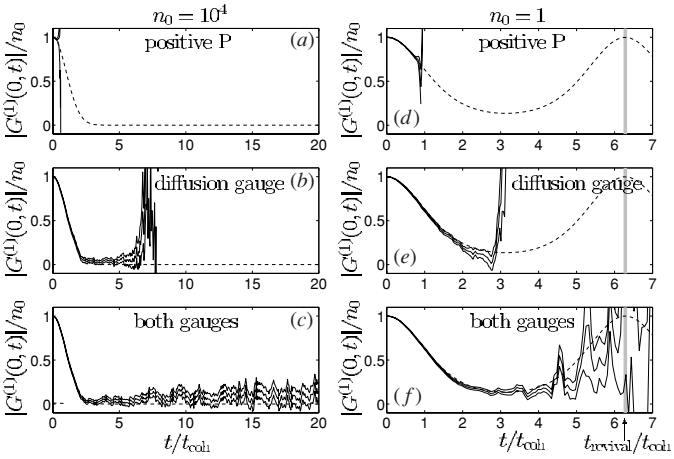

Figure 4. Modulus of the phase correlation function $G^{(1)}(0,t)$. Comparison of calculations with different gauges. Subplots (*b*) and (*e*): use diffusion gauge (68) of Plimak *et al* [25] with $t_{\rm opt} = 3t_{\rm coh}$, while subplots (*c*) and (*f*) use the combined drift and adaptive diffusion gauges (59) and (58) with the choice $t_{\rm opt} = 20t_{\rm coh}$. The initial conditions: coherent state with $\langle\hat{n}\rangle = n_0$. The triple solid lines indicate $G^{(1)}$ estimate with error bars, the dashed lines are exact values. $\mathcal{S} = 10^4$ trajectories in all cases.

Table 2. Empirical fitting parameters for maximum useful simulation time $t_{\rm sim}$ with several different gauge choices when applied to the undamped one-mode anharmonic oscillator. The fit is to expression (66).

	Positive P	Drift gauge	Diffusion gauge	Both gauges
G_k	0	(59)	0	(59)
g''	0	0	(42)	(58)
c_0	2/3	1	1/2	1/2
c_1	2.5 ± 0.2	11 ± 3	10.4 ± 0.7	35 ± 4
c_2	$3.2 \pm {}^{\infty}_{1.2}$	$1.4 \pm {}^{\infty}_{0.4}$	$2.7 \pm {}^{\infty}_{1.0}$	$03.6 \pm {}^{\infty}_{2.3}$
c_3	-0.5 ± 0.3	-0.5 ± 0.3	02.4 ± 0.6	2.8 ± 0.9
c_4	0.45 ± 0.07	0.49 ± 0.08	0.23 ± 0.13	0.23 ± 0.13

For calculations involving a diffusion gauge dependent on the target time $t_{\rm opt}$, a wide variety of target times were tried to investigate the dependence between $t_{\rm sim}$ and $t_{\rm opt}$, and ascertain what are the longest simulation times achievable.

5.2. *Simulation times*

Figure 3 compares $t_{\rm sim}$, as defined via 'useful precision' in $|G^{(1)}(0,t)|$, for several gauge choices. Note that a logarithmic scale is used. Results at high occupation are tabulated in table 1, which includes data for a larger set of gauge choices. Figure 4 gives examples of calculated values $|G^{(1)}(0,t)|$ along with error estimates. Table 2 gives empirical fitting parameters to expression (66).

We see that

- 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

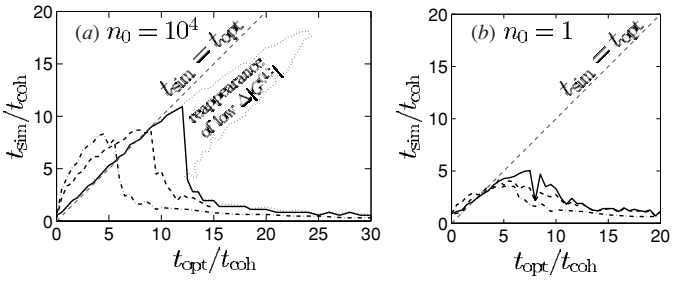

Figure 5. Dependence of actual useful simulation time t_{sim} on the *a priori* target time t_{opt} for a variety of diffusion gauges in the $\mathcal{G}_k = 0$ schemes. Solid line: full adaptive gauge (42); dashed line: number-adaptive-only gauge of form (36), but with the modification $n_0 \to n(t)$; dash-dotted line: gauge (68) of Plimak *et al* [25]. For the gauge (36), the whole region where useful precision occurs is shown by the dotted line. The undamped anharmonic oscillator system (13), with coherent state initial conditions $n_0 = n_0'$.

away for highly-occupied modes—potentially up to about 35 collapse times t_{coh} in the cases treated here.

- Diffusion-gauge-only simulations also give quite good statistical behaviour (although useful simulation times are about four 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. Such simulations are restricted in time to about one-phase oscillation. This indicates that the drift gauge choice made here is not an optimum choice when used in isolation.
- At low occupation, i.e. of the order of one boson or less, combined gauge methods still give the best results, but the advantage is marginal.
- The simulation times with nonzero diffusion gauges (whether accompanied by drift gauge (59) or not) not only have better scaling with n_0 when n_0 is large, but also this power-law decay of simulation time starts much later, as seen in figure 3 and the right-hand column of table 1.

5.3. Target time dependence

Figures 5 and 6 show the dependence of simulation times on the target time parameter t_{opt} for a variety of gauges. Comparison is made between the full adaptive forms $g''(n_{\mathbf{n}}(t), t_{rem})$ (solid lines), number-adaptive-only forms $g''(n_{\mathbf{n}}(t), t_{opt})$ that do not use the explicit time-dependence t_{rem} (dashed lines), and the Plimak*et al* [25] gauge (68) (dash-dotted lines).

Some comments

- The diffusion gauge forms that optimize for the 'remaining time to target' t_{rem} give the longest simulation times, and these times are well controlled. 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 1.
- Diffusion gauges that instead use a constant t_{opt} parameter give.
  - (i) Somewhat shorter simulation times.
  - (ii) A complicated relationship between target and useful simulation times. Broadly speaking $t_{opt} \ll t_{sim}$ for the optimum cases.

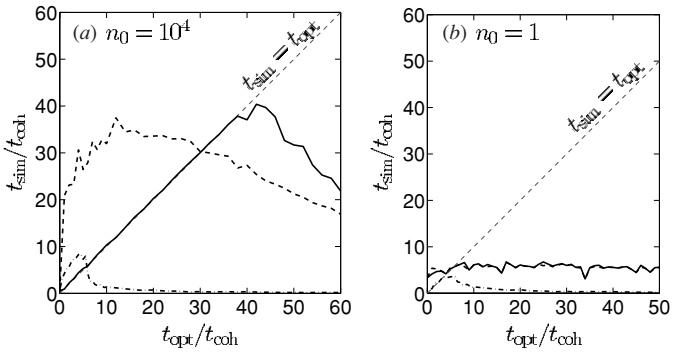

Figure 6. Dependence of t_{sim} on t_{opt}, as in figure 5, with drift gauge (59) as well as diffusion gauges. Details as in figure 5.

These forms of diffusion gauges require tedious parameter searching to find the best t_{opt} choice for given initial conditions. The probable reason why $t_{sim} \neq t_{opt}$ here is that g'' has been optimized to minimize the variance of *logarithmic* variables. This is not the same as the variance of the non-logarithmic variables α, or β that actually appear in the observable calculations. Hence, different g'' forms may extend simulation time by a further amount.

- When drift gauges (59) are used, an adaptive diffusion gauge $g''(n(t))$ rather than a constant $g''(n_0)$ can give much longer simulation times. (Compare to the Plimak *et al* gauge in figure 6(*a*) which effectively scans through a range of constant g'' values).
- At low n_0', t_{sim} is only weakly dependent on local diffusion gauge choice.
- When there is no drift gauge, for $n_0' \gg 1$, the time-adaptive gauge forms $g''(t_{rem})$ lead to a peculiar effect if the optimization time t_{opt} is chosen larger than the usual maximum t_{sim} given in table 1. The statistical error in the $G^{(1)}$ estimate first rises rapidly, then falls again, and finally grows definitively. The parameter region in which this occurs is shown in figure 5(*a*). In effect, one has two time intervals when the simulation gives useful results: at short times, and later in a time interval around $t \approx \mathrm{O}(10t_{coh})$.

More detail on these numerical investigations can be found in [16], chapter 7.

5.4. 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 [7, 8, 25, 34]. Here we compare these with the stochastic gauge formalism, and make some comparison to the results and analysis in the present chapter.

5.4.1. The work [8] of Carusotto, Castin and Dalibard. An isolated (i.e., particle-conserving) system of exactly $\overline{N}$ interacting bosons was considered (on a 1D lattice). The 'coherent state simple scheme' for stochastic wavefunctions described in section 3.2.2 therein can be identified as using drift gauges of the form

$$\mathcal{G}_1 = \mathrm{i}\sqrt{\mathrm{i}\kappa}(n - |\alpha|^2) \qquad \mathcal{G}_2 = \sqrt{\mathrm{i}\kappa}(n - |\beta|^2) \tag{67}$$

when re-written our notation for the single-mode system (13). This gauge causes a full decoupling of the complementary α and β equations by making the replacements $n \to |\alpha|^2$ or $n \to |\beta|^2$ in the nonlinear terms. Like (59) it is also successful in removing movable singularities, since the nonlinear terms in the radial equations for $\mathrm{d}|\alpha|$ and $\mathrm{d}|\beta|$ are removed.

The stochastic wavefunction in the form presented in [8] is only applicable to closed systems with a definite and explicit number of particles, so e.g. evaporative cooling or coherent out-coupling from a system cannot be treated.

5.4.2. The work [25] of Plimak, Olsen and Collett. A single-mode undamped, gainless system (13) at high Bose occupation with coherent state initial conditions was considered. 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)

$$g'' = g''_A = \tfrac{1}{2}\cosh^{-1}[n_0\kappa t_{\mathrm{opt}}] \tag{68}$$

defined at high occupation or long times (i.e., while $n_0\kappa t_{\mathrm{opt}} \geqslant 1$). This is dependent on a target time t_{opt} (which was taken to be $t_{\mathrm{opt}} = 3t_{\mathrm{coh}}$ in the calculations of [25]), and the initial Bose occupation $n_0 = n'_0$. The useful simulation times obtainable with this method are also shown in figures 3 and 4, and table 1. Their dependence on t_{opt} has been calculated here, and is shown in figure 5.

5.4.3. The work [7] of Drummond and Deuar. In section 5.3 of the above article, some preliminary results for the dynamics of a one-mode, undamped, gainless system (with $n_0 = 9$ particles on average) were shown. The drift gauge (59), and a constant imaginary diffusion gauge $g'' = 1.4$ were used.

6. Convergence issues

6.1. General

A subtle issue with many phase-space distributions is the possibility of so-called boundary term errors. These can arise when the tails of a distribution (say $G(\alpha, \beta, \Omega)$) do not fall off fast enough as the boundaries of phase space are approached (in the case here, as $|\alpha|$, $|\beta|$, or $|\Omega| \to \infty$). It is possible for this to lead to a bias in means of random variables even in the infinite sample limit, if parts of the distribution which have a non-negligible effect are never sampled.

Some numerical indicators have been developed [36] that allow one to search for symptoms of these errors using the numerical data. The most useful of these indicators is sudden appearance of spiking in observable estimates, where the onset of this spiking tends to come earlier as more trajectories are added. Such spikes usually occur when a single trajectory samples a long power-law tail. *Results obtained after first spiking are suspect*.

In our experience [16], another indicator can be obtained by performing two simulations with sample sizes $\mathcal{S}$ differing by an order of magnitude. If a *statistically significant difference (e.g., 2σ) in observable predictions* occurs, the simulation is suspect.

In all cases with which we are familiar, symptoms of potential convergence problems have been apparent already in the equations of motion. They have been of two kinds [16][4].

6.2. Divergence symptoms of the first kind: movable singularities

This is a symptom apparent in the drift parts of the equations of motion.

It has been found [36] that a systematic boundary term error is often associated with the presence of so-called *movable singularities*. These are trajectories (usually of measure zero) in the deterministic parts of the equations that diverge in a finite time. The presence

[4] Of course, other ways of achieving a divergence may occur.

of a bias in this situation can be understood by considering that the effect of an infinitesimal number of divergent trajectories may be nonzero. Some cases where this occurs with a positive-P simulation have been investigated by Gilchrist *et al* [36]. The bias has been shown to disappear when the movable singularities are removed using drift gauges [6]. See also [16], chapter 6.

Consider a set of generic stochastic equations

$$\mathrm{d}v = A_v(\mathbf{v})\,\mathrm{d}t + \sum_k B_{vk}(\mathbf{v})\,\mathrm{d}W_k(t), \tag{69}$$

in complex variables $\mathbf{v} = \{v\}$ free to explore the whole complex plane. If, in the limit $|v| \to \infty$, deterministic growth of $|v|$ is exponential or slower, then the trajectory cannot reach infinity in a finite time by deterministic processes, and moving singularities are ruled out. Exponential growth in v occurs when $\mathrm{Re}[A_v/v]$ is a positive constant, so we conclude that a condition sufficient to rule out moving singularities is that

$$\lim_{|v|\to\infty} \mathrm{Re}\left[\frac{A_v}{v}\right] \tag{70}$$

converges for all variables v.

6.3. Divergence symptoms of the second kind: noise-weight divergence

This is a symptom which arises from a combination of the noise behaviour and the form of the quantities which are averaged to obtain observable estimates.

A classic and well-known example of this convergence problem occurs for the present single-mode system (13) when using an un-normalized Bargmann coherent state kernel

$$\widehat{\Lambda}(\alpha,\beta) = \|\alpha\rangle\langle\beta^*\|. \tag{71}$$

(Compare to the gauge-P kernel (5)). The (non)convergence of this model has been considered by Carusotto and Castin [37], and also in [16], section 6.2.2.

One finds that the Ito equations of motion are

$$\mathrm{d}\alpha = \mathrm{i}\alpha\sqrt{\mathrm{i}\kappa}\,\mathrm{d}W_1 \tag{72a}$$

$$\mathrm{d}\beta = \beta\sqrt{\mathrm{i}\kappa}\,\mathrm{d}W_2. \tag{72b}$$

To calculate estimates for an observable $\widehat{O}$, we need to evaluate

$$\frac{\mathrm{Tr}[\widehat{O}\widehat{\rho}]}{\mathrm{Tr}[\widehat{\rho}]} = \frac{\langle \mathrm{Tr}[\widehat{\Lambda}\widehat{O}]\rangle_\mathrm{s}}{\langle \mathrm{Tr}[\widehat{\Lambda}]\rangle_\mathrm{s}} = \frac{\langle f(\alpha,\beta)\,\mathrm{e}^{\alpha\beta}\rangle_\mathrm{s}}{\langle \mathrm{e}^{\alpha\beta}\rangle_\mathrm{s}} \tag{73}$$

for some appropriate $f(\alpha,\beta)$ which depends on the details of $\widehat{O}$. The equations of motion can be formally solved to give

$$n(t) = \alpha(t)\beta(t) = n(0)\exp[-\sqrt{\kappa}\xi^-(t) + \mathrm{i}\sqrt{\kappa}\xi^+(t)], \tag{74}$$

with the variance t Gaussian random variables $\xi^\pm$ defined similarly to (31*c*)

$$\xi^\pm = \frac{1}{\sqrt{2}}\int_{s=0}^{t}[\mathrm{d}W_1(s) \pm \mathrm{d}W_2(s)]. \tag{75}$$

For any observable estimate, we must estimate $\langle \mathrm{e}^{\alpha\beta}\rangle_\mathrm{s}$ using our samples of $n(\xi^+,\xi^-)$. The distribution of this can be explicitly evaluated

$$\begin{aligned}\langle \mathrm{e}^{\alpha\beta}\rangle_\mathrm{s} &= \int_{-\infty}^{\infty} P(\xi^+)P(\xi^-)\,\mathrm{e}^{\alpha\beta}\,\mathrm{d}\xi_1\,\mathrm{d}\xi_2 \\ &\propto \int_{-\infty}^{\infty} \exp\left\{n(0)\,\mathrm{e}^{-\sqrt{k}\xi^-}\,\mathrm{e}^{\mathrm{i}\sqrt{k}\xi^+} - \frac{(\xi^-)^2}{2t} - \frac{(\xi^+)^2}{2t}\right\}\mathrm{d}\xi^-\mathrm{d}\xi^+.\end{aligned} \tag{76}$$

This distribution is divergent as $\xi^- \to -\infty$ because the $\mathrm{e}^{-\sqrt{k}\xi^-}$ factor in the exponent always beats the (relatively) weak Gaussian tail, which can only manage a quadratic $-(\xi^-)^2/2t$ drop-off in the exponent.

6.4. The stochastic model system

A more general feeling for the effect on convergence of the noise-weight relationship can be gained by considering the following simple equation for a complex variable v:

$$\mathrm{d}v(t) = cv(t)^n\,\mathrm{d}W(t), \tag{77}$$

with real constants n and c. Its formal solution is

$$v(t) = \begin{cases} [v(0)^{1-n} + c(1-n)\xi(t)]^{\frac{1}{1-n}} & \text{if} \quad n \neq 1, \\ v(0)\exp[c\xi(t)] & \text{if} \quad n = 1, \end{cases} \tag{78}$$

where again $\xi(t) = \int_{s=0}^{t} \mathrm{d}W(s)$ is a Gaussian random variable of variance t, mean zero. One finds that

$$\langle v^m \rangle_{\mathrm{s}} \begin{cases} \text{converges } \forall\, m \text{ and } \forall\, t & \text{for} \quad n \leqslant 1 \\ \text{does not converge} & \text{for} \quad n > 1 \end{cases} \tag{79}$$

and

$$\langle \exp(v^m) \rangle_{\mathrm{s}} \begin{cases} \text{converges } \forall\, m < 2(1-n) \text{ and } \forall\, t & \text{for} \quad n < 1 \\ \text{converges for } m = 2(1-n) \text{ and for } t < \frac{1}{2c(1-n)^2} & \text{for} \quad n < 1 \\ \text{does not converge} & \text{for} \quad n \geqslant 1. \end{cases} \tag{80}$$

One can see from above that, barring special favourable circumstances[5], the following relationships will hold (the notation is as in (69)):

(1) If observable averages involve only expressions polynomial in the variables v, then schemes for which noise terms grow faster than *linearly* as $|v| \to \infty$ will be divergent. That is those where

$$\lim_{|v|\to\infty} \mathrm{Re}[B_v(\mathbf{v})/v] \tag{81}$$

is unbounded for any v. For example $B_v = cv^n$, where $n > 1$.

(2) If observable averages involve expressions exponential in variables v, then schemes for which noise terms grow faster than $\propto \sqrt{|v|}$ as $|v| \to \infty$ will be divergent. That is those where

$$\lim_{|v|\to\infty} |B_v(\mathbf{v})/\sqrt{v}| \neq 0. \tag{82}$$

The special case when the limit (82) is finite is also divergent, but only after some initial time period.

6.5. Single mode: diffusion gauge

How do the gauged methods compare to these divergence symptoms for the single-mode system? The diffusion-gauged method of section 3 has a finite formal solution (31) at all times. For observables of finite order in $\widehat{a}^\dagger$ and $\widehat{a}$, only polynomial expressions of α and β need to be averaged. Such expressions scale as $\propto \exp[\text{factors} \times \xi_j(t)]$, and thus their stochastic averages are equivalent to integrals like $\int P(\xi)\exp[\text{factors} \times \xi]\,\mathrm{d}\xi$ which are convergent due to the Gaussian form of $P(\xi)$.

[5] For example, a topological barrier which prevents any trajectories reaching $|v| \to \infty$, given the right initial conditions.

6.6. Single mode: drift gauge

Equations (46) with the drift gauge (47) can be formally solved

$$\log n(t) = \log n(0) - \gamma t - \mathrm{e}^{-g''}\sqrt{\kappa}[\xi^-(t) - \mathrm{i}\xi^+(t)] \tag{83}$$

$$\log \Omega(t) = -\mathrm{e}^{-g''}\sqrt{\kappa}\int_{s=0}^{t} \mathrm{Im}[n(t)][\mathrm{d}W_1(s) - \mathrm{i}\,\mathrm{d}W_2(s)]. \tag{84}$$

(taking $\Omega(0) = 1$). From (10), observables require the averaging of quantities like $\Omega f(\alpha, \beta)$, which include the factor $|\Omega(t)|$. This takes the form

$$|\Omega(t)| = \exp\left\{-\sqrt{\kappa}\,\mathrm{e}^{-g''}\int_{s=0}^{t} |n(0)|\,\mathrm{e}^{-\gamma s}\,\mathrm{e}^{-\sqrt{\kappa}\,\mathrm{e}^{-g''}\xi^-(s)}\sin[\angle n(0) + \sqrt{\kappa}\xi^+(s)]\,\mathrm{d}W^+(s)\right\}, \tag{85}$$

where $\mathrm{d}W^+ = (\mathrm{d}W_1 + \mathrm{d}W_2)/\sqrt{2}$.

For large negative $\xi^-(s)$, the Gaussian drop-off of $P(\xi^-(s))$ is insufficient by itself to directly prevent the divergence of this factor. However, the situation is more subtle than in the simple model equations, due to the presence of an oscillatory stochastic term in the integral. Hence the convergence of the observable averages is still an open question for the drift-gauged method, despite the absence of movable singularities.

Lack of convergence has been shown for the 'simple coherent' stochastic wavefunction method [37], which has been discussed in section 5.4.1.

6.7. Many-modes

Regarding movable singularities in the many-mode equations (11), let us first compare to the movable singularity condition (70). One sees that

(1) The drift terms dependent on $\omega_{\mathbf{nn}}$ and $\gamma_{\mathbf{n}}$ as well as the noise terms dependent on κ lead to only exponential growth in the moduli of $\alpha_{\mathbf{n}}$ and $\beta_{\mathbf{n}}$, and so cannot cause movable singularities.

(2) The mode-mixing terms dependent on $\omega_{\mathbf{nm}}$, where $\mathbf{n} \neq \mathbf{m}$, lead to drift of the form

$$\begin{aligned} d|\alpha_{\mathbf{n}}| &= |\alpha_{\mathbf{m}}||\omega_{\mathbf{nm}}|\sin(\angle\omega_{\mathbf{nm}} + \angle\alpha_{\mathbf{m}} - \angle\alpha_{\mathbf{n}})\,\mathrm{d}t + \cdots \\ &\leqslant |\alpha_{\mathbf{m}}||\omega_{\mathbf{nm}}|\mathrm{d}t + \cdots. \end{aligned} \tag{86}$$

and of similar form for $|\beta_{\mathbf{n}}|$. These terms lead to behaviour like a linear matrix differential equation for the radial evolution of the coherent state amplitudes. Such equations do not diverge in finite time, their solutions being finite linear combinations of exponentials, hence these terms by themselves cannot lead to movable singularities either.

(3) Only the nonlinear drift terms can cause movable singularities. With no drift gauge ($\mathcal{G}_k = 0$) as in the positive-P or diffusion-gauge-only equations, these lead (when dominant) to evolution of the form

$$\begin{aligned} \mathrm{d}|\alpha_{\mathbf{n}}| &= \kappa|\alpha_{\mathbf{n}}|^2|\beta_{\mathbf{n}}|\sin(\angle\alpha_{\mathbf{n}} + \angle\beta_{\mathbf{n}}) + \cdots \\ \mathrm{d}|\beta_{\mathbf{n}}| &= -\kappa|\beta_{\mathbf{n}}|^2|\alpha_{\mathbf{n}}|\sin(\angle\alpha_{\mathbf{n}} + \angle\beta_{\mathbf{n}}) + \cdots. \end{aligned} \tag{87}$$

These *can* cause either $|\alpha_{\mathbf{n}}|$ or $|\beta_{\mathbf{n}}|$ to diverge at finite time, for some trajectories. For example, if $|\beta_{\mathbf{n}}|\sin(\angle\alpha_{\mathbf{n}} + \angle\beta_{\mathbf{n}}) = K$ conspire to be approximately constant and nonzero over the relevant timescale, then

$$|\alpha_{\mathbf{n}}(t)| \approx \frac{|\alpha_{\mathbf{n}}(t_0)|}{1 - \kappa K|\alpha_{\mathbf{n}}(t_0)|(t - t_0)}, \tag{88}$$

which diverges at time $t_{\rm div} = t_0 + 1/\kappa K|\alpha_{\mathbf{n}}(t_0)|$. Of course the precise condition 'K is constant over the time t_0 to $t_{\rm div}$' will only occur for a set of trajectories of measure zero, but this is typical of movable singularities.

(4) The equations that use the drift gauges (59) do not contain the terms (87), and no movable singularities occur.

As for noise-weight divergences, the situation appears to be similar to the single-mode cases. The formal solution for the weight is

$$\Omega(t) = \exp\left\{-\mathrm{i}\sqrt{\mathrm{i}\kappa}\int_{s=0}^{t} \mathrm{e}^{-g''(s)} \sum_{\mathbf{n}} \mathrm{Im}[n_{\mathbf{n}}(s)] \times \text{noise increments}(s)\right\}, \tag{89}$$

and contains an exponent containing exponentials of Gaussian random variables $\xi_k^-(s)$ as part of the $n_{\mathbf{n}}(s)$. This is indicative of possible divergences.

6.8. The 'what happened to the divergence?' puzzle

It appears that the observable averages using the drift gauge (47) have the potential to be non-convergent even in the single-mode case. Why then are the numerical simulations of section 5 well behaved for such a long time, showing no sign of bias? Similarly, no systematic error was seen in simulations using the 'simple coherent' stochastic wavefunction scheme [8], despite the subsequent proof of its divergence in [37].

The detailed investigation of this is beyond the scope of this paper. The simplest explanation is simply that the distribution tails are sufficiently convergent to eliminate boundary terms, while still having a large (perhaps infinite) variance in some observables. It is possible that the appearance of large statistical uncertainty masks any systematic errors that may occur in the $\mathcal{S} \to \infty$ limit. This is plausible because long distribution tails certainly give rise to large phase-space excursions and thus huge statistical uncertainty, whether or not systematic biases in the limit $\mathcal{S} \to \infty$ are present. Another reason for this lack of bias may be some type of special symmetry properties in the equations or the observables calculated.

6.9. Summary

Divergences of the moving singularity type may be present in many-mode (but not single-mode) simulations using the diffusion-gauge-only method of section 3, while divergences of the noise-weight type may be present in simulations with the drift-gauged method described in section 4. However, no bias of any kind was seen in the single-mode simulations (or the two-mode simulations, as shall be seen below).

Hence, the simulations appear to give correct results, but numerical indicators such as spiking and ensemble-size dependence should be rigorously monitored in all calculations.

7. Two coupled modes

We now look at the behaviour of a two-mode system, to investigate how the statistical behaviour seen for the single mode is affected once coupling between modes is present. This is a simple enough system that investigation of several examples gives meaningful insight into the general situation.

7.1. The model

The system consists of two orthogonal modes labelled 1 and 2 with inter-particle interactions in each mode, and Rabi coupling between them. No damping will be considered for simplicity.

The coupling frequency will be restricted to be real. Time units are chosen so that the nonlinear interaction frequency is $\kappa = 2$, and a transformation is made to an interaction picture in which the linear mode self-energies $\hbar\omega_{jj}\widehat{n}_j$ are moved into the Heisenberg evolution of operators. The interaction picture Hamiltonian is then

$$\widehat{H} = \hbar\omega_{12}\left[\widehat{a}_1^\dagger\widehat{a}_2 + \widehat{a}_2^\dagger\widehat{a}_1\right] + \hbar\sum_{j=1}^{2}\widehat{a}_j^{\dagger 2}\widehat{a}_j^2. \tag{90}$$

The Rabi frequency is ω_{12} (in scaled time units). Stochastic equations are as in (2.2). Physically this model can represent, for example, two internal boson states coupled by an EM field, or two trapped condensates spatially separated by a barrier. This approximation has been widely used to investigate the quantum behaviour of BECs in two-state systems [38].

Let us consider two kinds of initial conditions which broadly represent the two kinds of situations generically occurring in all many-mode simulations

(i) Case 1: coupling between modes of widely differing occupation.
(ii) Case 2: coupling between modes of similar occupation.

In a many-mode calculation, adjacent modes typically behave like case 2, since if a field model is well resolved by the lattice, then physical properties (e.g., density, and hence mode occupation) should not change much over the distance between neighbouring lattice points. Long distance coupling will tend to behave like case 1.

7.2. Case 1: coupling to a vacuum mode

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 (conserved in time) was varied.

At low particle number, the Rabi oscillations dominate the Hamiltonian, and particles oscillate between the modes, without much phase collapse. At high particle number n_0, on the other hand, phase collapse dominates mode 1, suppressing also the coherent transfer of particles to mode 2.

The particular values chosen to simulate were

$$n_0 = \{1, 17, 200, 1500, 10^4\}. \tag{91}$$

Simulation times were assessed using the calculated uncertainties in the two observables: (1) the fraction of particles in the (initially empty) mode 2:

$$p_2 = \frac{\langle\widehat{n}_2\rangle}{\overline{N}}, \tag{92}$$

where $\widehat{n}_j = \widehat{a}_j^\dagger\widehat{a}_j$, and (2) the local normalized second-order correlation functions

$$g_j^{(2)}(t,t) = \frac{\langle:\widehat{n}_j(t):\rangle}{\langle\widehat{n}_j(t)\rangle} = \frac{\langle\widehat{n}_j^2\rangle}{\langle\widehat{n}_j\rangle^2} - \frac{1}{\langle\widehat{n}_j\rangle}. \tag{93}$$

The second-order correlations quantify the amount of (instantaneous) bunching/antibunching in the boson field. These 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 $|\psi\rangle = \sin\theta|0\rangle + \cos\theta|n\rangle$, $g^{(2)} = \left(1 - \frac{1}{n}\right)/\cos^2\theta$.

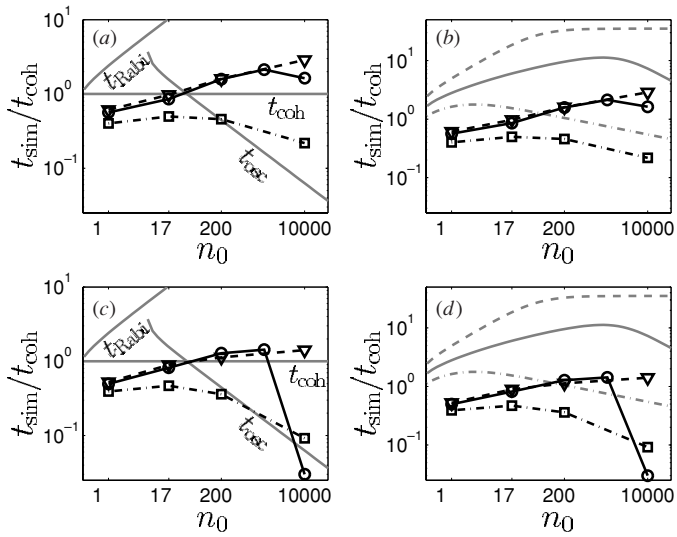

Figure 7. Useful simulation times $t_{\rm sim}$ for a mode coupled to vacuum as in section 7.2. Calculated simulation times are shown as data points, with the symbols denoting gauge used. '□': positive P; '○': drift gauge (59) and diffusion gauge (94); '▽': drift and diffusion gauges (59) and (58) with best target time parameters $\omega_{12}t_{\rm opt} = \{2.5, 0.5, 0.5, 0.075\}$ for $n_0 = \{1, 17, 200, 10^4\}$, respectively. Dependence of $t_{\rm sim}$ on $t_{\rm opt}$ can be found in [16], figure 8.6. 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 the Rabi oscillation period $t_{\rm Rabi} = 2\pi/\omega_{12}$, while subplots (*b*) and (*d*) compare to expected simulation times for a single mode using the empirical fits of table 2 and (95). The expected $t_{\rm sim}$ are plotted as light lines. Dotted: positive P; solid: drift and diffusion gauges with $t_{\rm opt} = 0$; dashed: with optimum $t_{\rm opt}$ choice.

The drift and diffusion gauge scheme using (59) and (58) was considered. Comparison was also made to the positive-P ($g''_j = \mathcal{G}_k = 0$), and to the special case of $t_{\rm opt} = 0$ in the diffusion gauge which is then

$$g''_{\mathbf{n}} = \tfrac{1}{6}\log[1 + 4n''_{\mathbf{n}}(t)^2]. \tag{94}$$

This may become nonzero after spread in the n''_j from the initial $n''_j = 0$ occurs. In each run, $\mathcal{S} = 2\times 10^5$ trajectories were used, and useful simulation precision taken to occur at such a time $t_{\rm sim}$ when 10% or smaller relative uncertainty in an observable could be obtained using 10^6 trajectories.

In figure 7, simulation times are compared to physical timescales and expected values based on single-mode expressions from table 2. An example simulation is shown in figure 8. Note that for gauged simulations using (59) and (94), single-mode simulations led to the $t_{\rm sim}$ empirical fitting parameters

$$\{c_0, \ldots, c_4\} = \left\{1, 800 \pm 260, 3.6^{+\infty}_{-2.3}, 1.2 \pm 0.2, 0.42 \pm 0.03\right\}. \tag{95}$$

One point to note is that $t_{\rm sim}$ was based on the moment when the relative error in a quantity g_j or p_j 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(*f*). Good accuracy can often be obtained between peaks for much longer times than shown in figure 7(*c*) or (*d*), up to about the same simulation time as worked out based on p_2. This is especially evident for $n_0 = 10^4$.

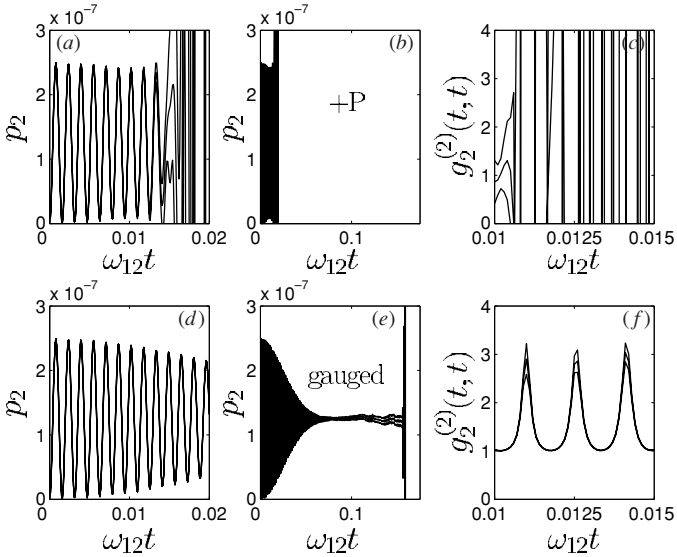

Figure 8. Coupling to vacuum mode as in section 7.2, $n_0 = 10^4$. Subfigures (*a*)–(*c*) show results with positive-P simulations, whereas (*d*)–(*f*) show results with combined drift gauge (59) and diffusion gauge (58) using target time $\omega_{12}t_{\rm opt} = 0.1$. The triple lines indicate mean and error bars.

7.3. Case 2: coherent mixing of two identical modes

The system starts initially with identical coherent states of mean particle number n_0 in both modes. The two-mode state is separable. This time simulations were carried out with constant particle number $n_0 = 100$, but the coupling frequency ω_{12} was varied.

At low frequency $\omega_{12} \ll n_0 = 100$, phase collapse local in each mode dominates, and phase oscillations in each mode occur with period $t_{\rm osc} = \pi/n_0$, while at high frequency $\omega_{12} \gg n_0 = 100$, the inter-mode coupling dominates and phase oscillations for each mode occur with period $t_{\rm Rabi} = 2\pi/\omega_{12}$. One expects that for weak coupling, the two modes should behave largely as two independent single modes of section 5.

The particular values chosen to simulate were

$$\omega_{12} = \{5000, 500, 50, 5, 0.5, 0.05, 0.005, 0.0005\}. \tag{96}$$

The simulation time $t_{\rm sim}$ was assessed here based on 10% uncertainty in $|G^{(1)}(0, t)|$ (for either mode), as in the single-mode case. There were $\mathcal{S} = 10^4$ trajectories per simulation.

In figure 9, simulation times are compared to physical timescales and expected values based on single-mode expressions using table 2. An example simulation is shown in figure 10.

Note that the $t_{\rm opt} = 0$ gauge (94) performed better than any $t_{\rm opt} > 0$ gauge for all ω_{12} values tried here. The dependence of $t_{\rm sim}$ on $t_{\rm opt}$ is shown in figure 8.8 of [16].

7.4. Analysis

The above examples have not by any means been a comprehensive assessment of gauge performance for general cases of the model (90), since only a few parameter regimes have been explored. Still, several aspects of the situation when modes are coupled have been seen.

- Broadly speaking, when local scattering within a mode dominates over the coupling between modes, the response of the system to local gauges is similar to what was seen

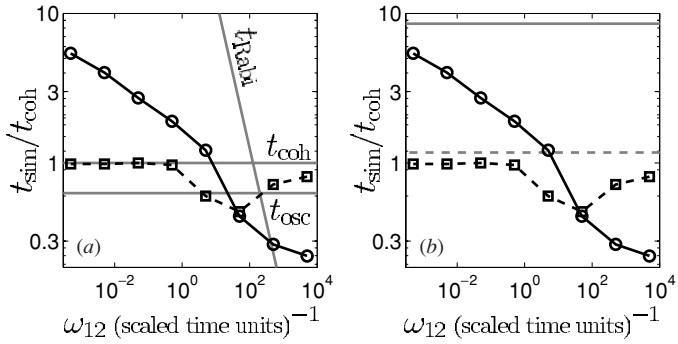

Figure 9. Useful simulation times $t_{\rm sim}$ for two identical modes undergoing coherent mixing as in section 7.3. Calculated simulation times are shown as data points, with the symbols denoting gauge used. '□': positive P; '○': drift gauge (59) and diffusion gauge (94). Subplot (*b*) compares to expected simulation times for a single mode using the empirical fits of table 2 and (95). These expected $t_{\rm sim}$ are plotted as light lines: dashed: positive P; solid: drift gauged.

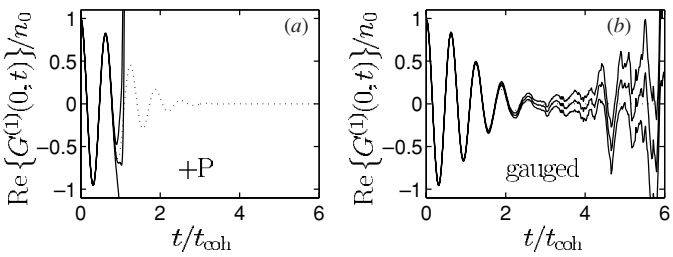

Figure 10. Mixing of two identical modes as in section 7.3, $\omega_{12} = 0.0005$. Subplot (*a*) shows results with a positive-P simulation, whereas (*b*) shows results with combined drift gauge (59) and diffusion gauge (94). The triple lines indicate mean and error bars.

for single modes. Large extensions of $t_{\rm sim}$ are found relative to the positive-P method to well beyond $t_{\rm coh}$ when $n \gg 1$. Little improvement is found when $n \lesssim \mathrm{O}(1)$. In this strong scattering regime, one does not need much inter-mode coupling ω_{12} to reduce the simulation time in absolute terms by a factor $\mathrm{O}(2-10)$, although $t_{\rm sim} > t_{\rm coh}$ is still obtained.

- When the inter-mode coupling dominates, the local diffusion gauges do not appear to be useful. They actually reduce simulation time as compared to the positive-P method, although several Rabi oscillation periods can always be simulated. In the Case 2 simulations, the transition between the strong and weak coupling behaviour appears to be at around

$$\omega_{12} \approx \mathrm{O}(\kappa n), \tag{97}$$

which is the point at which the expectation values of the coupling and two-body scattering energies in the Hamiltonian are approximately equal. This implies, as noted previously, that nonlocal gauges are likely to be more useful in these cases.

- The beneficial effect of choosing $t_{\rm opt} > 0$ seen for a single mode (see, e.g., figure 6) appears to be suppressed at intermediate mode occupations $n_0 \lesssim \mathrm{O}(200)$. However, the $t_{\rm opt} = 0$ diffusion gauge (94) has a marked beneficial effect even when two-body scattering is significant. This parameter range is for $\omega_{12} \ll n\kappa$, and $\mathrm{O}(1) \lesssim n_0 \lesssim \mathrm{O}(10^3)$. Simulation times obtained are smaller by about a factor of O(2) than those given for a single mode with

the same gauge. At higher occupations, the benefit gained with (94) abates but nonzero $t_{\rm opt}$ values appear to become useful again, and continue to provide strong improvements over positive-P simulations. (See, e.g. figure 7(*d*).) The gauge (94) is convenient also because there is no *a priori* parameter $t_{\rm opt}$.

At the level of the stochastic equations, the evolution of $\mathrm{d}\alpha_j$ can gather randomness from three sources.

(i) Directly from the local noise term $\propto \alpha_j \sqrt{\kappa}\,\mathrm{d}W_j$.
(ii) Indirectly from the local nonlinear term $\propto \kappa \alpha_j n_j$ that can amplify variation in the local noise term.
(iii) From the other mode through the coupling term $\propto \omega_{12}\alpha_{\neg j}$.

The drift gauges (59) neutralize source 2. The diffusion gauges (42), (58) or (94) suppress the direct noise source 1. No local gauge is good at suppressing the third source of randomness, however, because these fluctuations are 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. Combating such effects would require a nonlocal gauge.

Lastly, significantly more detail on two-mode simulations, including consideration of some other gauges, and the relationship between $t_{\rm opt}$ and $t_{\rm sim}$ can be found in chapter 8 of [16].

8. Many-mode example: uniform gas

We revisit the uniform gas system simulated in the companion paper [12]. This consists of a uniform one-dimensional gas of bosons with density ρ and inter-particle s-wave scattering length a_s. The lattice is chosen with a spacing $\Delta\mathrm{x} \gg a_s$ so that inter-particle interactions are effectively local at each lattice point, and the lattice Hamiltonian (1) applies. Periodic boundary conditions are assumed.

The initial state is taken to be a coherent wavefunction, which is a stationary state of the ideal gas with no inter-particle interactions (i.e., $a_s = 0$). The subsequent evolution is with constant $a_s > 0$, so that there is a disturbance at $t = 0$ when inter-particle interactions are rapidly turned on. Physically, this can correspond to the disturbance 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 [39, 40]. More discussion of the physics can be found in the companion paper [12].

A brief set of exploratory simulations were made with the diffusion gauge (42) (no drift gauge) to investigate whether simulation time can be extended. The case of $\rho = 100/\xi^{\rm heal}$ was picked, and target time $t_{\rm opt}$ was varied for the two cases $\Delta\mathrm{x} = \xi^{\rm heal}/2$ and $\Delta\mathrm{x} = 10\xi^{\rm heal}$. Note that in both cases the occupation per mode $\overline{n} = \rho\Delta\mathrm{x}$ is $\gg 1$. Simulations were with $M = 250$ and $M = 50$ lattice points (meaning 12 500 and 5×10^4 particles on an average), respectively, and $\mathcal{S} = 10^4$ trajectories in both cases.

Simulation times obtained are shown in figure 11, and correspond to the time of first spiking seen in estimates of the second-order spatial correlation function

$$\overline{g}^{(2)}(\mathbf{x_n}) = \frac{1}{M}\sum_{\mathbf{m}} \frac{\langle \widehat{a}^\dagger_{\mathbf{m}} \widehat{a}^\dagger_{\mathbf{m+n}} \widehat{a}_{\mathbf{m}} \widehat{a}_{\mathbf{m+n}} \rangle}{\langle \widehat{a}^\dagger_{\mathbf{m}} \widehat{a}_{\mathbf{m}} \rangle \langle \widehat{a}^\dagger_{\mathbf{m+n}} \widehat{a}_{\mathbf{m+n}} \rangle}. \tag{98}$$

The time scale used is

$$t_\xi = \frac{m(\xi^{\rm heal})^2}{\hbar} = \frac{\hbar}{2\rho g} \tag{99}$$

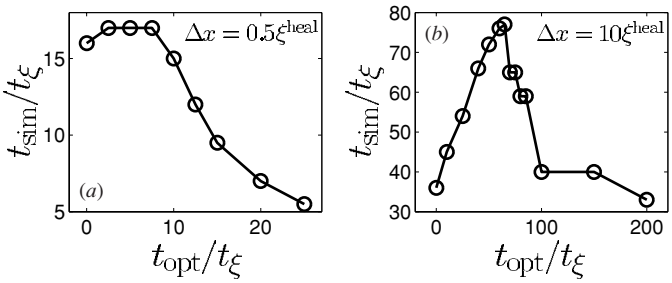

Figure 11. Simulation times as a function of the target time $t_{\rm opt}$ when using the diffusion gauge (42), and $\mathcal{G}_k = 0$. The standard positive-P method when $t_{\rm opt} = 0$. All simulations are of a $\rho = 100/\xi^{\rm heal}$ gas, but with differing lattice spacing Δx. $\mathcal{S} = 10^4$ trajectories, $M = 50$.

(the 'healing time'), which is approximately the time needed for the short-distance $O(\xi^{\rm heal})$ inter-atomic correlations to equilibrate after the disturbance. See, e.g. [12].

For comparison, gauged simulations of isolated lattice points with the same occupations $\overline{n} = \rho\Delta x$ gave much bigger improvements in $t_{\rm sim}$. Using (66) and diffusion gauge (42) data from table 2, we expect $t_{\rm sim} \approx 140 t_\xi$ and $t_{\rm sim} \approx 660 t_\xi$ for $\Delta x = 2\xi^{\rm heal}$ and $10\xi^{\rm heal}$, respectively. This is to be contrasted with the predictions of $t_{\rm sim} \approx 15 t_\xi$ and $t_{\rm sim} \approx 49 t_\xi$, respectively, for the standard positive-P method, which are also an accurate reflection of $t_{\rm sim}$ in the present many-mode case. This is consistent with what was seen in [12].

Clearly, the local gauges extend simulation time in these large uniform density systems, although the improvement is relatively smaller than for the single mode. Also, the numerical results are consistent with the analysis in section 3.4 which indicated in (45) that simulation time improvements would occur once $\Delta x \gtrsim O(\xi^{\rm heal})$. Lastly, it is noteworthy that the dependence of $t_{\rm sim}$ on $t_{\rm opt}$ shows qualitative similarities to the single-mode case of figure 5.

9. Conclusions

The positive-P representation method is capable of simulating many-body quantum dynamics of interacting Bose gases from first principles [12, 14, 15, 16]. A limiting factor is that precision is lost after a certain time $t_{\rm sim}$ due to sampling error, and this may be accompanied by systematic boundary term errors. This time can be short when some modes of a many-mode boson system are occupied by many particles. In that case, precision is lost before these highly occupied modes lose coherence due to phase diffusion [12].

Using the related gauge-P representation, local diffusion and drift gauges have been developed that greatly improve the useful simulation times at high occupation, most strongly for single-mode cases. The resulting simulation times have been investigated in some detail, and substantial improvement has been demonstrated in 1, 2, and many-mode ($M = 50$ and $M = 250$) simulations. For many of the single- and double-mode cases considered, full decoherence can be achieved while still retaining good precision.

Two gauge choices were proposed as being the most advantageous.

(i) Diffusion gauge (42) only.
(ii) Drift (59) and diffusion gauge (58).

The latter appears less broadly applicable because weight fluctuations accumulate from all modes, leading to a decrease in simulation time with increased system size, and because of

some doubt over its convergence properties. However, this latter method can lead to longer simulation times when there is one, or only several, dominant modes.

Considerations of convergence in section 6 have highlighted two routes that may lead to divergences: moving singularities and noise-weight considerations. Nevertheless, no sign of any bias is seen, an interesting situation which warrants further investigation.

Generally speaking, the local-gauge methods considered here give much better performance when the inter-particle scattering local to each mode is a stronger process than inter-mode coupling due to the kinetic evolution and/or external potentials. This is the case if sufficiently coarse lattices are used.

Finally, the investigation of local gauges carried out here may also be useful in developing more robust nonlocal gauges. These might lengthen simulation times for lattices denser than the healing length, beyond what is possible with either the standard positive-P method, or the local stochastic gauges.

References

[1] Wigner E P 1932 *Phys. Rev.* **40** 749
[2] Husimi K 1940 *Proc. Phys. Math. Soc. Japan.* **22** 264
[3] Glauber R J 1963 *Phys. Rev.* **131** 2766
Sudarshan E C G 1963 *Phys. Rev. Lett.* **10** 277
[4] Chaturvedi S, Drummond P D and Walls D F 1977 *J. Phys. A: Math. Gen.* **10** L187–92
[5] Drummond P D and Gardiner C W 1980 *J. Phys. A: Math. Gen.* **13** 2353
[6] Deuar P and Drummond P D 2002 *Phys. Rev.* A **66** 033812
[7] Drummond P D and Deuar P 2003 *J. Opt. B-Quant. Semiclass. Opt.* **5** S281
[8] Carusotto I, Castin Y and Dalibard J 2001 *Phys. Rev.* A **63** 023606
[9] Feynman R P 1982 *Int. J. Theor. Phys.* **21** 467
[10] Wilson K G 1974 *Phys. Rev.* D **10** 2445
Ceperley D M 1995 *Rev. Mod. Phys.* **67** 279
[11] See, e.g. Ceperley D M 1996 *Lectures on quantum Monte Carlo* http://archive.acsa.uiuc.edu/Science/CMP/papers/cep96b/notes.ps
Ceperley D M 1995 *Rev. Mod. Phys.* **67** 279
[12] Deuar P and Drummond P D 2006 *J. Phys. A: Math. Gen.* **39** 1163 (See also [16])
[13] For a review of Gross–Pitaevskii equations as applied to BECs, see, e.g. Dalfovo F, Giorgini S, Pitaevskii L P and Stringari S 1999 *Rev. Mod. Phys.* **71** 463
[14] Drummond P D and Corney J F 1999 *Phys. Rev.* A **60** R2661
[15] Steel M J, Olsen M K, Plimak L I, Drummond P D, Tan S M, Collett M J, Walls D F and Graham R 1998 *Phys. Rev.* A **58** 4824
[16] In preparation
See also, Deuar P 2004 *PhD Thesis* The University of Queensland (*Preprint* cond-mat/0507023)
[17] Carter S J, Drummond P D, Reid M D and Shelby R M 1987 *Phys. Rev. Lett.* **58** 1841
[18] Drummond P D and Hardman A D 1993 *Europhys. Lett.* **21** 279
[19] Rosenbluh M and Shelby R M 1991 *Phys. Rev. Lett.* **66** 153
Drummond P D, Shelby R M, Friberg S R and Yamamoto Y 1993 *Nature* **365** 307
[20] Castin Y and Carusotto I 2001 *J. Phys. B: At. Mol. Opt. Phys.* **34** 4589
[21] Steel M J, Olsen M K, Plimak L I, Drummond P D, Tan S M, Collett M J, Walls D F and Graham R 1998 *Phys. Rev.* A **58** 4824
[22] Norrie A A, Ballagh R J and Gardiner C W 2005 *Phys. Rev. Lett.* **94** 040401
[23] Kinsler P and Drummond P D 1991 *Phys. Rev.* A **43** 6194
Kinsler P, Fernee M and Drummond P D 1993 *Phys Rev.* A **48** 3310–20
[24] Sinatra A, Lobo C and Castin Y 2002 *J. Phys. B: At. Mol. Opt. Phys.* **35** 3599
[25] Plimak L I, Olsen M K and Collett M J 2001 *Phys. Rev.* A **64** 025801
[26] Dowling M R, Drummond P D, Davis M J and Deuar P 2005 *Phys. Rev. Lett.* **94** 130401
[27] A treatment of these issues can be found in Leggett A J 2001 *Rev. Mod. Phys.* **73** 307
For a more detailed treatment see Weiner J, Bagnato V S, Zilio S and Julienne P S 1999 *Rev. Mod. Phys.* **71** 1
[28] Kokkelmans S J J M F, Milstein J N, Chiofalo M L, Walser R and Holland M J 2002 *Phys. Rev.* A **65** 053617

[29] Abrikosov A A, Gorkov L P and Dzyaloshinski I E 1963 *Methods of Quantum Field Theory in Statistical Physics* (New York: Dover)
[30] Louisell W H 1973 *Quantum Statistical Properties of Radiation* (New York: Wiley)
[31] See, e.g. Gardiner C W 1991 *Quantum Noise* (Berlin: Springer)
Gardiner C W 1983 *Handbook of Stochastic Methods* (Berlin: Springer)
[32] Lindblad G 1976 *Commun. Math. Phys.* **48** 119
[33] Greiner M, Mandel O, Hnsch T W and Bloch I 2002 *Nature* **419** 51
[34] Deuar P and Drummond P D 2001 *Comput. Phys. Commun.* **142** 442
[35] See, e.g. Dalfovo F, Giorgini S, Pitaevskii L P and Stringari S 1999 *Rev. Mod. Phys.* **71** 463, 481
[36] Gilchrist A, Gardiner C W and Drummond P D 1997 *Phys. Rev.* A **55** 3014
[37] Carusotto I and Castin Y 2003 *Laser Phys.* **13** 509
[38] See, e.g. references in Leggett A J 2001 *Rev. Mod. Phys.* **73** 307 (part VII)
[39] Donley E A *et al* 2002 *Nature* **417** 529
[40] Herbig J *et al* 2003 *Science* **301** 1510
Mukaiyama K *et al* 2003 *Science* **91** 210402 (ibid)
Dürr S, Volz T, Marte A and Rempe G 2004 *Phys. Rev. Lett.* **92** 020406
Dürr S, Volz T, Marte A and Rempe G 2004 *Phys. Rev. Lett.* **92** 040405