Chapter 6

Boundary term errors and their removal

As has been mentioned previously, when one attempts to simulate non-separable physical models using stochastic equations derived via phase-space distributions, instabilities in the equations are common. These can give rise to systematic biases in the observable estimates (3.14) using a finite number of samples. Sometimes these biases abate at large sample number (Such a situation is considered in detail in Appendix A), while in other cases the biases remain even in the limit $\mathcal{S} \to \infty$.

The latter have been called *boundary term errors* because the root of the biasing lies in discarding nonzero boundary terms in integrations involving the distribution $P(C)$ (and some factors) over C. Rigorous results regarding how and when these biases arise in practical simulations are few, and many researchers while aware of phase-space distribution methods, are either 1) unaware of the potential for biases, or 2) know that biases exist and because of this avoid using phase-space distribution methods altogether. With the aim of reducing this confusion, in Sections 6.1 and 6.2, the mechanisms by which these biases come about are investigated and classified for general phase-space distributions. To clarify the conditions under which boundary term errors may be a problem, some symptoms that can be checked for by inspection of the stochastic equations and observable estimators are pointed out.

It will be seen in Chapters 7 and 9 that simulations of interacting Bose gases

can contain boundary term errors (at least when using coherent state kernels), and a method to overcome this problem is developed in Section 6.3. Subsequently, this un-biasing procedure is demonstrated on the two rather different examples from the literature where boundary term errors have been seen when using a positive P representation: Two-boson absorption (Section 6.4), and a single-mode laser model (Section 6.5). Removal of possible or confirmed biases in the interacting Bose gas is carried out in Sections 7.3 and 9.2, respectively.

6.1 Boundary term errors of the first kind

6.1.1 Origin and general form

During the derivation of the Fokker-Planck equation (from the original master equation), there is a point between (3.41) and (3.42) where integration by parts is performed, and boundary terms discarded. The correspondence between full quantum evolution and FPE (and hence the final stochastic equations) occurs *only* if these discarded boundary terms are zero. Otherwise, there is a discrepancy between quantum mechanics and what is calculated by the stochastic equations in the limit of $\mathcal{S} \to \infty$. This has been termed a boundary term error. Let us see what exact form these discarded boundary terms take.

The partial integration is made on a master equation that already contains all kernel gauge expressions (4.19), (4.12) and/or (4.2) as well as the basic nonzero terms arising from directly applying the operator correspondences (3.39) to the master equation. In any case, all these terms can be incorporated into the formalism of (3.41). With the real variable set $C = \{C_j\}$, when going from (3.41) to (3.42) the discarded boundary terms are found to be (after some algebra)

$$\widehat{\mathcal{B}}(C) = \sum_{lj} \int \left[\widehat{\mathcal{B}}_j^{\text{int}}(C) \right]_{\min[C_j]}^{\max[C_j]} \prod_{p \neq j} dC_p \tag{6.1a}$$

where

$$\widehat{\mathcal{B}}_j^{\text{int}}(C) = P(C) T_{lj}^{(1)}(C) \widehat{\Lambda}(C) + \frac{1}{2} P(C) \sum_k T_{ljk}^{(2)} \frac{\partial \widehat{\Lambda}(C)}{\partial C_k} - \frac{1}{2} \widehat{\Lambda}(C) \sum_k \frac{\partial}{\partial C_k} \left(P(C) T_{lkj}^{(2)}(C) \right). \tag{6.1b}$$

The usual definite evaluation notation $[f(v)]_a^b = \lim_{v\to b} f(v) - \lim_{v\to a} f(v)$ has been used. Note that the kernel is still an operator, and hence so is the boundary term expression $\widehat{\mathcal{B}}$. Every matrix element of the boundary term operator should be zero for the correspondence between quantum mechanics and the FPE to be exact. Such matrix elements could in principle be evaluated by inserting the elements of $\widehat{\Lambda}$ and $\partial\widehat{\Lambda}/\partial C_k$, which can usually be worked out exactly, into (6.1). This will be attempted for the gauge P representation in Section 6.1.5, and a toy example of the calculation of $\widehat{\mathcal{B}}$ is given in Section 6.2.2.

While (6.1) is nontrivial, there are generally a variety of ways in which it could turn out to be zero.

1. Especially for an open phase space, one expects that $P(C)$ will tend towards zero at the boundary, and if this rate is faster than the growth of the multiplying factors ($T^{(\bullet)}(C)$ times the matrix elements of $\widehat{\Lambda}$), then the terms in (6.1) will be zero.

2. Even if individually the terms in (6.1) do not tend to zero, but the distribution $P(C)$ times multiplying factors is an even function of C_j, then the values evaluated at $\max[C_j]$ and $\min[C_j]$ may cancel each other. A typical example of this is when a particular $C_j = \theta$ is the phase of a complex variable — any function of θ evaluated at the limits of its domain $[-\pi, \pi]$ is the same at both limits, and any boundary terms in θ sum to zero.

3. Also, there may be a cancellation between separate terms in the expression (6.1).

Conversely, the boundary terms can easily turn out to be nonzero in an open phase-space if $P(C)$ drops off too slowly as $|C_k|$ grows (typically — if the tails of $P(C)$ obey a power law), or, in a partly or fully closed phase-space, when $P(C) \neq 0$ at the boundary.

6.1.2 Symptoms indicating boundary term errors

Unfortunately, it is not generally possible to directly evaluate the boundary terms (6.1) and check correctness, because this involves knowing at least the far tail behavior (in an open phase space) of $P(C)$ as a function of time. In a fully or partly closed phase-space one needs to know the exact behavior near the boundary. These require detailed knowledge about aspects of the full solution to the master equation — generally not achievable in non-trivial systems, although power law tails of $P(C)$ can be considered a bad sign.

What is really needed are some symptoms of likely boundary term errors that could be looked for in the stochastic equations before simulation. Since boundary term errors arise when the distribution falls off too slowly as the (possibly open) boundaries of phase space are approached, at least four (there may be more) broad classes of features come to mind that indicate the probable presence of boundary term errors of the first kind.

1. The first such symptom, fairly well known, is the presence of so-called MOVING SINGULARITIES. In a set of deterministic equations, these manifest themselves as solutions (usually a set of measure zero) that diverge in a finite time. Previous studies [64, 66] have shown that when moving singularities are present in the deterministic parts of stochastic equations, boundary term errors usually result. In particular, it was seen that

   - In a number of studied systems, moving singularities in equations derived using the positive P representation occur whenever systematic errors are seen in observables calculated with estimators polynomial in the system variables.
   - In these systems, when the moving singularities are removed with the aid of stochastic gauges (as explained in Section 6.3), the systematic errors are removed as well[66].

   The presence of pathological behaviour when moving singularities are present can be grasped intuitively: Due to the stochasticity of the equations, and

the continuous nature of the noise, one expects that after small times, the distribution of trajectories (at least the far tails of it) will extend over all of phase space. In particular into those phase-space regions where the solutions that diverge in a finite time occur. While the divergent trajectories are usually a set of measure zero, an infinitesimal proportion of divergent trajectories may have a finite effect on the observable means, on the general basis that $0 \times \infty$ does not necessarily equal zero. However, because the number of such trajectories is infinitesimal in comparison with the rest, this effect can never be estimated by a finite sample of them, hence the systematic errors.

Moving singularities may arise either in a dynamic fashion (e.g. in an equation $dv = v^2\,dt$, $v(t) = v(0)/[1 - tv(0)]$, and a divergence occurs at $t = 1/v(0)$), or due to divergent drift terms for finite phase-space variables (e.g. $dv = dt/(1 - v)$ at $v = 1$, or $dv = dt/(1 - t)$ at $t = 1$). In the latter situation the divergent trajectories may even be a set of nonzero measure.

2. Secondly, one should also be wary of NOISE DIVERGENCES — i.e. instabilities in the noise matrices B. For some small set of trajectories the Wiener increments $dW(t)$ will conspire to act approximately as a constant $(\Delta W) \approx \mathcal{O}\left(\sqrt{\Delta t}\right)$ over a time interval $\Delta t \gg dt$. This results in approximately an effective drift term proportional to B, and if such "drift terms" give rise to what effectively are moving singularities, then systematic errors will also result by an analogous mechanism as for instabilities of the drift terms.

3. Thirdly, Another indicator of possible boundary term errors is a DISCONTINUOUS DRIFT AND/OR NOISE MATRIX B. This may cause an effect that can not be simulated stochastically, as the discontinuity is only ever sampled by a set of trajectories of measure zero. In particular, there may be a region near such a boundary that contains only an infinitesimal portion of total trajectories, but those trajectories have large impact on the observable estimates.

4. Fourthly, if the INITIAL DISTRIBUTION IS TOO BROAD, boundary term errors may be present from the outset. An example would be a Lorentzian ini-

tial distribution of variable $v(0)$ if the Fokker-Planck equation contains terms quadratic in v. Another example is the single-mode laser model of Section 6.5 with non-delta-function initial conditions.

If any of the above symptoms is present, boundary term errors may be suspected. Conversely, if the symptoms are absent, it *appears likely* that no boundary term errors occur, but unfortunately rigorous results for general distributions are hard to come by. No boundary term errors were seen for any simulation that did not suffer from one or more of these symptoms.

There have also been developed several numerical indicators (described in detail in [64]) that one can use to check for the presence of boundary term errors from a finite set of trajectories.

6.1.3 Moving singularities in a complex phase-space

For a set of stochastic equations (e.g. such as (3.45) or (4.90)) it is relatively easy to rule out boundary term symptoms 3 and 4 from the previous section (by inspection). Moving singularities and noise divergences (symptoms 1 and 2) are a little more slippery, but practical conditions for their absence can be stated. Simulations in this thesis are with complex variable sets on an open phase space, so here a practical condition for absence of moving singularities will be derived only for this particular (though commonly occurring) kind of phase-space.

In an open complex phase space, barring symptoms 3 and 4, moving singularities cannot occur if the both the drift and noise matrices of the stochastic equations contain no radial components that lead to super-exponential growth. This is because then all variables remain finite at finite times, and so at finite times never reach the "boundary" at modulus infinity. Since $z = ze^{2i\pi}$ for any complex z, tangential evolution does not lead to divergences, because boundary terms evaluated at the boundaries of the polar variable ($\angle z = 0$ or 2π) are equal, and thus cancel in (6.1a).

More precisely, for a set of stochastic equations

$$dz_j = A_j(C)\,dt + \sum_k B_{jk}(C)dW_k(t), \tag{6.2}$$

in complex variables z_j ($C = \{z_j\}$, including, possibly z_0), this gives the *Condition:*

- Moving singularities or noise divergences will not occur provided that the limits

$$\lim_{|z_j|\to\infty} \frac{A_j}{|z_j|} \qquad \text{and} \qquad \lim_{|z_j|\to\infty} \frac{B_{jk}}{|z_j|} \tag{6.3}$$

  converge for all $z_j \in C$ and all k, and that symptoms 3 and 4 of Section 6.1.2 are ruled out (by inspection).

6.1.4 Generic presence in many-mode simulations

Off-diagonal kernels appear necessary to allow all quantum states to be described by a positive real distribution that can be sampled stochastically (see Section 3.2.3). In such a case, the independent "ket" and "bra" parameters of the kernel will (if complex) almost always enter the kernel in an analytic way. Then, to ensure a positive propagator, requires in turn that the complex derivatives be chosen as in Section 3.4.3, preserving the analytic nature of the formulation into the stochastic equations.

The presence of moving singularities in (deterministic) dynamical equations for an analytically continued phase-space is a well-studied problem for both ordinary and partial differential nonlinear evolution equations. According to the Painleve conjecture (see, e.g. [79, 80, 81]), this is equivalent to non-integrability - and is therefore generic to many (in fact, all but a set of measure zero) complex dynamical equations with large numbers of degrees of freedom. Just the sort one obtains when simulating many-body systems.

Kernel gauges are, however, not restricted by the above mechanisms to be analytic. This is because they arise as arbitrary (hence, possibly non-analytic) multipliers to operators on the kernel, rather than as a result of applying analytic operator identities (3.39) to it. Because of this freedom of gauge choice, it has been conjectured by P. D. Drummond[56] that breaking the analyticity of the complex equations using stochastic gauges is a necessary (though not sufficient just by itself) condition to remove boundary term errors. This agrees with previous work[66, 1], and also with the gauges used to remove boundary term errors in this thesis.

Diffusion gauges, while possibly non-analytic, do not appear useful for removal of boundary term errors of the first kind because they appear only at the FPE-Langevin stage, after any nonzero boundary term operators have been already discarded.

6.1.5 Boundary terms in the gauge P representation

Let us consider the gauge P representation. The kernel (5.4), expanded in the complete Fock number state basis for each subsystem is $\widehat{\Lambda} = \Omega \otimes_k \widehat{\Lambda}_k$, with

$$\widehat{\Lambda}_k = \sum_{n\widetilde{n}} \frac{\alpha_k^n \beta_k^{\widetilde{n}}}{\sqrt{n!\widetilde{n}!}} \left|n\right\rangle_k \left\langle\widetilde{n}\right|_k e^{-\alpha_k\beta_k}, \tag{6.4}$$

as in (5.53). A master equation leads via (5.7), without invoking gauges, to terms $T_{lj}^{(1)}$ and $T_{ljk}^{(2)}$ polynomial (of low order) in α_k and β_k. Use of a weighting gauge will lead to some terms $T^{(\bullet)} \propto \Omega$. Drift gauges could generally lead to any form of the corresponding coefficients $T^{(\bullet)}$, but let us *assume* that the drift gauges $\mathcal{G}_k$, etc. are polynomial in α_k, β_k, α_k^*, β_k^*, and autonomous (i.e. not dependent on Ω). Under this assumption, the terms $T^{(\bullet)}$ due to drift gauges will be polynomial in all the variables, and at most $\propto \Omega^2$ (Diffusion coefficient in Ω due to standard drift gauge introduction is $\propto \Omega^2\mathcal{G}_k^2$).

Combining this together, the (matrix elements of the) multipliers of $P(C)$ appearing in the boundary term expression (6.1) will at most scale as :

$$\leq \propto \Omega^3 \prod_k \alpha_k^{a_k} \beta_k^{b_k} e^{-\alpha_k\beta_k}, \tag{6.5}$$

where the coefficients take on all values $a_k \geq 0$ and $b_k \geq 0$ for various matrix elements.

In an open phase space of complex variables z_j (as here), the boundary lies at $|z_j| \to \infty$. The multipliers of P(C) in (6.1) scale in these far tails as

$$\leq \propto \quad |\Omega^3| \prod_k |\alpha_k|^{a_k} |\beta_k|^{b_k} e^{c_k|\alpha_k||\beta_k|}, \tag{6.6}$$

where $c_k = -e^{i\angle\alpha_k} e^{i\angle\beta_k}$, can have positive real part. This implies that if one wants to be sure that no boundary term errors are present, (assuming any drift gauges $\mathcal{G}_k$

are polynomial in α_k, β_k, or their conjugates, and independent of Ω), then: $P(C)$ should decay in the far tails no slower than

$$P(C) < \propto 1/|\Omega^{4+\widetilde{b}_k}|, \tag{6.7a}$$

$$P(C) < \propto \exp\left(-\widetilde{c}_k|\alpha_k|^{\widetilde{a}_k}\right), \tag{6.7b}$$

$$P(C) < \propto \exp\left(-\widetilde{c}_k|\beta_k|^{\widetilde{a}_k}\right), \tag{6.7c}$$

where $\widetilde{b}_k > 0$, $\widetilde{c}_k > 0$ and $\widetilde{a}_k > 1$. (And there should be no time-dependent discontinuities in the drift or noise equations).

This does not preclude that by some canceling of terms, a less favorable scaling of the tails may also become acceptable, but certainly if P scales like (6.7) or better in the far tails, boundary term errors (of the first kind) will be absent in a gauge P representation simulation. Note in particular that from (6.7):

1. Power law tails of $P(C)$ in the $\boldsymbol{\alpha}, \boldsymbol{\beta}$ part of phase-space will invariably lead to boundary term errors. For some matrix elements of $\widehat{\mathcal{B}}$ for which the coefficient of $|n\rangle_k \langle\widetilde{n}|_k$ in (6.4) is of high enough order, the boundary terms will become nonzero. This is consistent with the conclusions drawn by Gilchrist *et al*[64] that power law tails in a positive P distribution are an indicator of boundary term errors.

2. Exponential decay of $P(C)$ with $|\alpha_k|$ or $|\beta_k|$ is also not good enough in general. This is because the coefficient of (say) $|\alpha_k|$ in the exponents of (6.7) is not constant, but $\propto |\beta_k|$ — which may be any arbitrarily large value. Only a faster-than-exponential decay will be able to overcome this when $|\beta_k|$ is large. However, Gilchrist *et al*[64] found no boundary term errors in a single-mode anharmonic oscillator (6.9) simulation for which the far tail behaviour of P_+ appeared to be exponential, and no indications of boundary term errors in this system are seen here in Chapter 7 either. This may be because the single-mode system is a special case — certainly in Chapter 8 it is seen that when acting on several coupled modes, a similar Hamiltonian *can* lead to boundary term errors.

6.2 Boundary term errors of the second kind

Biases that do not abate for $\mathcal{S} \to \infty$ may also arise via a different mechanism when the estimator functions to be averaged to obtain the expectation value of a given observable grow too rapidly in phase space.

6.2.1 Mechanism and overview

Working at the level of the distribution $P(C)$, expectation values of observables are given by the real part of (3.13). When one takes a number $\mathcal{S}$ of samples of this distribution, the estimator (3.14) is used, but if the real parts of the integrals in (3.13) do not converge then this finite sample estimate does not approach the quantum mechanical value even in the limit $\mathcal{S} \to \infty$. Systematic errors arising from this source will be called here boundary term errors of the second kind, and have very different properties to those of the first kind. Explicitly, from (3.13) and by imposing Hermiticity, the integrals that must converge are

$$\int P(C) \quad \mathrm{Re}\left\{\mathrm{Tr}\left[\widehat{O}\left(\widehat{\Lambda}(C) + \widehat{\Lambda}^{\dagger}(C^{*})\right)\right]\right\} \quad dC \tag{6.8a}$$

$$\int P(C) \qquad \mathrm{Re}\left\{\mathrm{Tr}\left[\widehat{\Lambda}(C)\right]\right\} \qquad dC, \tag{6.8b}$$

for any particular observable $\widehat{O}$.

One can see that even when (6.8b) is convergent, errors may occur for some observables for which only (6.8a) diverges. Given a particular distribution, divergences may come about due to the matrix elements of either the kernel $\widehat{\Lambda}$, the operator $\widehat{O}$, or both growing too rapidly as the boundaries of phase space are approached. Observables that involve an infinite sum of moments polynomial in the system variables (such as the parity in the gauge P representation, whose estimate is given by the sum (5.15)) may be particularly prone to these boundary term errors, because the dependence of $\mathrm{Tr}\left[\widehat{O}\widehat{\Lambda}\right]$ on the phase-space variables will grow particularly rapidly.

Boundary term errors of the second kind depend on the equations of motion only by their effect on the distribution of samples $P(C)$, and so do not show up as moving singularities or noise divergences in the equations.

A sobering point is that, in theory, one could always come up with some observables that grow fast enough as the boundary of phase space is approached to overcome whatever far tail behaviour occurs in $\widehat{\Lambda}(C)P(C)$. Such an observable could be defined as an infinite sum of polynomial moments whose terms grow appropriately rapidly. Thus, boundary term errors of the second kind are likely to occur for *some* observable estimates for any phase-space distribution simulation method, for any model. In practice, fortunately, it suffices to make the estimators of only all observables *of interest* convergent.

6.2.2 Example: un-normalized positive P representation

A good example of boundary term errors of the second kind occurs for the "anharmonic" oscillator single-mode Hamiltonian, when using an un-normalized coherent state kernel. This Hamiltonian[1] is

$$\widehat{H} = \hbar \widehat{a}^{\dagger 2} \widehat{a}^2, \tag{6.9}$$

the single-mode analogue of a repulsive two-body potential in the interaction picture[2]. Let us consider the kernel

$$\widehat{\Lambda} = ||\alpha\rangle\langle\beta^*||, \tag{6.10}$$

off-diagonal in the un-normalized Bargmann coherent states (5.1). It is convenient to change to logarithmic variables $n_L = \log(\alpha\beta)$, and $m_L = \log(\alpha/\beta)$, then the basic operator correspondences are, from (5.6),

$$\widehat{a}\widehat{\Lambda} = \quad \alpha\widehat{\Lambda} = \quad e^{(n_L+m_L)/2}\widehat{\Lambda} \tag{6.11a}$$

$$\widehat{a}^\dagger\widehat{\Lambda} = \quad \tfrac{\partial}{\partial\alpha}\widehat{\Lambda} = \quad e^{-(n_L+m_L)/2}\left(\frac{\partial}{\partial n_L} + \frac{\partial}{\partial m_L}\right)\widehat{\Lambda} \tag{6.11b}$$

$$\widehat{\Lambda}\widehat{a}^\dagger = \quad \beta\widehat{\Lambda} = \quad e^{(n_L-m_L)/2}\widehat{\Lambda} \tag{6.11c}$$

$$\widehat{\Lambda}\widehat{a} = \quad \tfrac{\partial}{\partial\beta}\widehat{\Lambda} = \quad e^{(m_L-n_L)/2}\left(\frac{\partial}{\partial n_L} - \frac{\partial}{\partial m_L}\right)\widehat{\Lambda}. \tag{6.11d}$$

It is also convenient to split the variables into real and imaginary parts: $n_L = n'_L + in''_L$, and $m_L = m'_L + im''_L$. For a dynamical evolution model with no coupling to the

[1] In some appropriate time units.

[2] That is, with terms linear in $\widehat{a}$ and $\widehat{a}^\dagger$ transformed away. Compare with (2.17).

environment, utilizing the analytic nature of the kernel as explained in Section 3.4.3, one obtains the FPE (provided no boundary term errors of the first kind are present)

$$\begin{aligned}\frac{\partial P}{\partial t} &= \left\{ 2\frac{\partial}{\partial m_L''} + \frac{\partial^2}{\partial {n_L'}^2} + \frac{\partial^2}{\partial {n_L''}^2} + \frac{\partial^2}{\partial {m_L'}^2} + \frac{\partial^2}{\partial {m_L''}^2} \right. \\ &\quad \left. - \frac{\partial^2}{\partial n_L' \partial m_L''} - \frac{\partial^2}{\partial m_L'' \partial n_L'} - \frac{\partial^2}{\partial n_L'' \partial m_L'} - \frac{\partial^2}{\partial m_L' \partial n_L''} \right\} P. \end{aligned} \tag{6.12}$$

This gives the Ito stochastic equations (with standard square root form of $B_0 = \sqrt{D}$)

$$dn_L = i\sqrt{2i}\,[dW(t) - i d\widetilde{W}(t)] = 2d\eta(t) \tag{6.13a}$$

$$dm_L = i\sqrt{2i}\,[dW(t) + i d\widetilde{W}(t)] - 2i\,dt = 2i[d\eta^*(t) - dt]. \tag{6.13b}$$

$dW(t)$ and $d\widetilde{W}(t)$ are the usual independent Wiener increments, and the complex noise $d\eta(t)$ defined here obeys

$$\langle d\eta(t) \rangle_{\text{stoch}} = 0 \tag{6.14a}$$

$$\langle d\eta^*(t) d\eta(t') \rangle_{\text{stoch}} = \delta(t - t')\, dt^2 \tag{6.14b}$$

$$\langle d\eta(t) d\eta(t') \rangle_{\text{stoch}} = 0. \tag{6.14c}$$

The FPE (or stochastic equations) can be solved, and in the case of coherent state initial conditions (with initial amplitude α_0)

$$P(n_L, m_L, 0) = \delta^2(m_L - 2i\angle\alpha_0)\ \delta^2(n_L - 2\log|\alpha_0|), \tag{6.15}$$

the distribution at time t is

$$\begin{aligned} P(n_L, m_L, t) &= \delta^2\Big(m_L - \Big[i(n_L^* - 2\log|\alpha_0|) - 2it + 2i\angle\alpha_0\Big]\Big) \\ &\quad \times \frac{1}{2\pi t} \exp\left(-\frac{|n_L - 2\log|\alpha_0|\,|^2}{4t}\right). \end{aligned} \tag{6.16}$$

In this toy model, one can actually check if there really were no boundary terms of the first kind, by substituting (6.16), and (6.10) into the expression for $\widehat{\mathcal{B}}$. The boundary term operator $\widehat{\mathcal{B}}$ is given by the integrals (6.1) of operators $\widehat{\mathcal{B}}_j^{\text{int}}$ composed of P, $\widehat{\Lambda}$ and coefficients $T^{(\bullet)}$, with individual terms of $\widehat{\mathcal{B}}$ evaluated at the boundaries of phase space. In our case these boundaries are $C_j \to -\infty$ and $C_j \to \infty$ for all $j = 1, \ldots, 4$, and one requires all the matrix elements of $\widehat{\mathcal{B}}$ to be zero. By comparison of (6.12) to (3.43), and labeling variables as $n_L' = C_1$, $n_L'' = C_2$, $m_L' = C_3$, and

$m_L'' = C_4$, and also setting the term counter $l = 1$, the nonzero $T^{(\bullet)}$ coefficients in the $\widehat{B}_j^{\text{int}}$ are $T_{14}^{(1)} = -2$, $T_{1jj}^{(2)} = 2$ for all j, and $T_{114}^{(2)} = T_{141}^{(2)} = T_{123}^{(2)} = T_{132}^{(2)} = -2$. Also, the matrix elements of $\widehat{\Lambda}$ (in a Fock basis $|n\rangle$) can be written (from the definition (5.1) of $||\alpha\rangle$) as

$$\langle \widetilde{n}| \widehat{\Lambda} |n\rangle = \frac{1}{\sqrt{n!\widetilde{n}!}} \exp\Big[\, n_L(n+\widetilde{n})/2 + m_L(n-\widetilde{n})/2 \, \Big]. \tag{6.17}$$

So:

1. The $T^{(\bullet)}$ coefficients are constants.

2. The matrix elements of $\widehat{\Lambda}$ and $\partial\widehat{\Lambda}/\partial z$ are exponential in both variables $z = n_L$ or m_L, with some constant positive or negative coefficient.

3. The distribution P is Gaussian in real and imaginary parts of both variables (n_L around mean $2\log|\alpha_0|$), m_L around mean $2i(\angle\alpha_0 - t)$).

One can see that all matrix elements of all the integrands ($\langle \widetilde{n}| \mathcal{B}_j^{\text{int}} |n\rangle$) will be dominated by the Gaussian decay of $P(C)$ in the far tails for all n and $\widetilde{n}$. For $j = 1, 2$, after integration over m_L and a (real or imaginary) component of n_L, the expression to be evaluated at large n_L will be Gaussian, hence zero in the far tail limit. Also, for the terms ($j = 3, 4$) in which one integrates over n_L and a component of m_L, the expression to be evaluated will be a Dirac delta function — also zero in the far tail limit. Hence,

$$\widehat{\mathcal{B}} = 0. \tag{6.18}$$

Thus it has been confirmed that there are no boundary term errors of the first kind.

So far it seems like this might be a very good kernel to try even for many-body locally interacting Bose gases, which have the same kind of interparticle potential term. The advantages would be that there are no boundary term errors of the first kind, no moving singularities (compare to Section 6.1.4) and even no nonlinear terms in the stochastic equations (compare to (5.17)).

Having the luxury of knowing $P(C)$ for this toy problem, let us see what happens when one tries to calculate an observable. Directly from (6.10) and (5.3):

$$\text{Tr}\left[\widehat{\Lambda}\right] = \exp(\alpha\beta) = \exp\left[e^{n_L}\right]. \tag{6.19}$$

The essential denominator term (6.8b) in any observable calculation, is then (using (6.16) and defining $n_L = n'_L + in''_L$),

$$\begin{aligned} & \int \mathrm{Re}\left\{\exp\left[e^{n_L}\right]\right\} \frac{1}{4\pi t} e^{-|n_L - 2\log|\alpha_0||^2/4t} d^2 n_L \\ \propto & \int \exp\left[e^{n'_L}\cos n''_L\right] \cos\left(e^{n'_L}\sin n''_L\right) e^{-(n'_L - 2\log|\alpha_0|)^2/4t} e^{-(n''_L)^2/4t}\, d^2 n_L \end{aligned} \tag{6.20}$$

For n_L with large positive real part, the factor from $\mathrm{Tr}\left[\widehat{\Lambda}\right]$ dominates (it's an exponential of an exponential!), and since the integrand is symmetric in n''_L and nonzero, it must become unbounded, and the integral diverge.

A plot of the integrand $P\mathrm{Tr}\left[\widehat{\Lambda}\right]$ for several time values is shown in Figure 6.1. The unbounded behaviour at high n'_L values indicates that (6.8b) will diverge and boundary term errors (of the second kind) will be present. A comparison to the probability distribution of n'_L is also made there to indicate which contributions to the integral (6.8b) are sampled by a stochastic simulation.

This means that: *All observable estimates using* (3.14) *are suspect with this representation!*

6.2.3 In the gauge P representation

For the gauge P representation, $\mathrm{Tr}\left[\widehat{\Lambda}\right] = e^{z_0}$, and so the denominator integral (6.8b) becomes simply

$$\int P(C)\, e^{z'_0} \cos z''_0\, dC \tag{6.21}$$

in terms of $z_0 = z'_0 + iz''_0$. This will be convergent provided the far tails of the marginal probability distribution of z'_0 decay as $e^{-z'_0}$, or faster in the limit $|z'_0| \to \infty$.

From the estimator expression (5.13) for an arbitrary observable that is a polynomial function of $\widehat{a}$ and $\widehat{a}^\dagger$, one can see that the numerator integral (6.8a) will be of the form

$$\int P(C) \sum_j f_j(C)\, dC, \tag{6.22}$$

where the functions for each term f_j are

$$f_j(C) = \frac{1}{2}\, e^{z'_0} \left(\prod_k |\alpha_k|^{a_k} |\beta_k|^{b_k}\right) \cos\left(z''_0 \pm \theta + \sum_k [a_k \angle\alpha_k + b_k \angle\beta_k]\right), \tag{6.23}$$

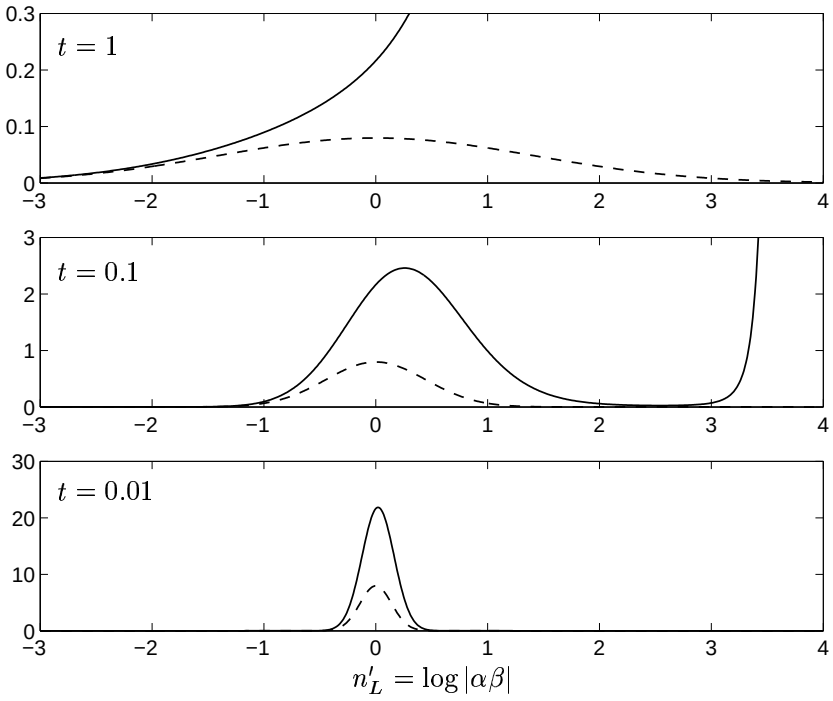

Figure 6.1: **Integrands appearing in the denominator** of observable estimates for anharmonic oscillator dynamics governed by the Hamiltonian (6.9). Values plotted are (SOLID LINE): $\int \mathrm{Tr}\left[\widehat{\Lambda}(C)\right] P(C)\, d^2 m_L$, evaluated (using (6.19) and (6.16)) at $n''_L = 0$. Coherent state initial conditions (6.15) were used with mean occupation number unity ($|\alpha_0| = 1$). To show what parts of the integral (6.8b) would be sampled by a stochastic simulation, the probability distribution of n_L evaluated at $n''_L = 0$ (i.e. $\int P(C) d^2 m_L$ evaluated at $n''_L = 0$) is also plotted as the DASHED LINE. Note the unbounded behaviour of the integrand at large positive n'_L values (which are not sampled).

---

with some non-negative integer coefficients a_k and b_k, which take on different values for different f_j. The behaviour of each such term as one approaches the boundaries of phase space, where the modulus of at least one variable tends to $+\infty$, is exponential in z'_0 (with unity coefficient), and polynomial in $|\alpha_k|$ and $|\beta_k|$.

This means that

- Boundary term errors can be expected for observables polynomial in $\widehat{a}_k$ and/or $\widehat{a}_k^\dagger$ if the tails of $P(C)$ decay as any power law in the $|\alpha_k|$, or $|\beta_k|$, or as $e^{-z'_0}$ or slower in z'_0.

Conversely, if the tails of $P(C)$ decay faster than power law in all $|\alpha_k|$ and $|\beta_k|$ and

faster than $e^{-z_0'}$ in the modulus of the weight, then boundary term errors *(at least of the second kind)* will not occur for any observable polynomial in the annihilation and creation operators on subsystems. That is, any observable given by (5.10), or a finite linear combination of such expressions.

Power law tails of $P(C)$ in $|\alpha_k|$ or $|\beta_k|$ do not immediately exclude all observables, only those whose estimators grow too rapidly. By inspection of (5.10), (5.13), one can see that an observable that is an a_kth polynomial of $\widehat{a}_k$ will lead to a term $f_j \propto e^{z_0'}|\alpha_k|^{a_k}$. If $P(C) \propto e^{-cz_0'}|\alpha_k|^{-\widetilde{a}_k}$ in the far tails, then the indefinite integral $\int P(C)f_j(C)dC$ will be $\propto e^{(1-c)z_0'}|\alpha_k|^{a_k-\widetilde{a}_k+1}$ (if $a_k-\widetilde{a}_k+1 \neq 0$), or $\propto e^{(1-c)z_0'}\log|\alpha_k|$ (otherwise) in this far tail region. When this is evaluated in the far tail limit, the integral will convergent if the decay of $P(C)$ is fast enough. Similarly for the relationship between $\widehat{a}_k^\dagger$ and β_k. Hence, one can state the conditions for expectation value estimates of a particular operator $\widehat{A}$ to not suffer from boundary term errors of the second kind: When ϵ_j are infinitesimal positive real constants, one requires

$$\widehat{A} \propto \widehat{a}_k^{a_k} \implies \lim_{|\alpha_k|\to\infty} P(C) \quad < \propto \quad \frac{1}{|\alpha_k|^{-(1+a_k+\epsilon_1)}} \tag{6.24a}$$

$$\widehat{A} \propto \widehat{a}_k^{\dagger\, b_k} \implies \lim_{|\beta_k|\to\infty} P(C) \quad < \propto \quad \frac{1}{|\beta_k|^{-(1+b_k+\epsilon_2)}} \tag{6.24b}$$

as well as

$$\lim_{z_0'\to\infty} P(C) \quad < \propto \quad e^{-(1+\epsilon_3)z_0'}. \tag{6.24c}$$

Non-polynomial observables defined as infinite sums (e.g. the parity (5.14)) can be more troublesome, because the functions $\mathrm{Tr}\left[\widehat{\Lambda}\widehat{O}\right]$ etc. that are to be sampled by (6.8a) can grow faster than polynomially as $|\alpha|_k, |\beta_k| \to \infty$. Conditions of the kind (6.24) can also be derived, but will not scale as a power law. For example since the parity estimator (5.15) has a numerator expression that scales $\propto e^{-2\mathrm{Re}\{\alpha_k\beta_k\}}$, one requires $P(C)$ to decay faster than $e^{2\mathrm{Re}\{\alpha_k\beta_k\}}$ in the far tails for an unbiased estimate of parity expectation values (remember that $\alpha_k\beta_k$ may have negative real part in highly non-classical parts of phase-space).

6.2.4 Locally-interacting Bose gas calculations

Here the likelihood of boundary term errors of the second kind is considered for a many-mode interacting Bose gas model using the gauge P representation as described in Sections 5.3 and 5.6. By inspection of the dynamics and thermodynamics gauge P equations for the locally interacting Bose gas (5.17)–(5.21), and (5.50), one sees that the stochastic terms (assuming no gauges) take the forms

$$d\alpha_{\mathbf{n}} \quad \cdots + \quad c\,\alpha_{\mathbf{n}}\,dW_{\mathbf{n}} \tag{6.25a}$$

$$d\beta_{\mathbf{n}} \quad \cdots + \quad \widetilde{c}\,\beta_{\mathbf{n}}\,d\widetilde{W}_{\mathbf{n}}, \tag{6.25b}$$

with some complex constants c and $\widetilde{c}$, apart from possibly additional terms caused by finite temperature heat bath interactions (5.18)

$$d\alpha_{\mathbf{n}} \quad \cdots + \quad \sqrt{\gamma_{\mathbf{n}}\overline{n}_{\rm bath}(T)}d\eta_{\mathbf{n}} \tag{6.26a}$$

$$d\beta_{\mathbf{n}} \quad \cdots + \quad \sqrt{\gamma_{\mathbf{n}}\overline{n}_{\rm bath}(T)}d\eta^{*}_{\mathbf{n}}. \tag{6.26b}$$

In this Subsection, the far-tail regime of phase space where $|\alpha_{\mathbf{n}}|$ and/or $|\beta_{\mathbf{n}}|$ are large will be of interest, and there the terms (6.26) are negligible, and hence are ignored here.

Upon change of variables to $z_{\mathbf{n}} = \log\alpha_{\mathbf{n}}$ and $\widetilde{z}_{\mathbf{n}} = \log\beta^{*}_{\mathbf{n}}$, one finds by the rules of Ito calculus that these logarithmic variables will have stochastic evolution

$$dz_{\mathbf{n}} \quad \cdots + \quad c\,dW_{\mathbf{n}} \tag{6.27a}$$

$$d\widetilde{z}_{\mathbf{n}} \quad \cdots + \quad \widetilde{c}\,d\widetilde{W}_{\mathbf{n}}. \tag{6.27b}$$

This indicates that the *short time* evolution of $z_{\mathbf{n}}$ and $\widetilde{z}_{\mathbf{n}}$ is dominated by Gaussian noise. Thus at short times, (defining $z_{\mathbf{n}} = z'_{\mathbf{n}} + iz''_{\mathbf{n}}$, and $\widetilde{z}_{\mathbf{n}} = \widetilde{z}'_{\mathbf{n}} + i\widetilde{z}''_{\mathbf{n}}$)

$$P(C,t)\prod_{\mathbf{n}} d^2z_{\mathbf{n}}d^2\widetilde{z}_{\mathbf{n}} \propto P(C,0)\prod_{\mathbf{n}}\exp\left(-\frac{[z'_{\mathbf{n}}(t)-z'_{\mathbf{n}}(0)]^2}{2{\rm Re}\left\{c\right\}^2 t} - \frac{[\widetilde{z}'_{\mathbf{n}}(t)-\widetilde{z}'_{\mathbf{n}}(0)]^2}{2{\rm Re}\left\{\widetilde{c}^2\right\} t}\right) \tag{6.28}$$

Changing back to the coherent state amplitude variables,

$$P(C,t)\prod_{\mathbf{n}} d^2\alpha_{\mathbf{n}}d^2\beta_{\mathbf{n}} \propto P(C,0)\prod_{\mathbf{n}}\frac{1}{|\alpha_{\mathbf{n}}\beta_{\mathbf{n}}|}\exp\left\{-\frac{\left(\log\left|\frac{\alpha_{\mathbf{n}}(t)}{\alpha_{\mathbf{n}}(0)}\right|\right)^2}{2{\rm Re}\left\{c\right\}^2 t} - \frac{\left(\log\left|\frac{\beta_{\mathbf{n}}(t)}{\beta_{\mathbf{n}}(0)}\right|\right)^2}{2{\rm Re}\left\{\widetilde{c}\right\}^2 t}\right\} \tag{6.29}$$

The far tail decay as a Gaussian of a logarithm of $|\alpha_{\mathbf{n}}|$ or $|\beta_{\mathbf{n}}|$ is strong enough to overcome any polynomial of these variables if one moves far enough into the tails. (polynomials will be only exponential in the logarithm). So, there is no immediate reason to suspect boundary term errors of the second kind for polynomial observables here. (They cannot be ruled out by the above discussion, however, because at longer times the evolution of $P(C)$ is much more complicated and no proof for such a regime has been given).

On the other hand, (6.29) does not decay strongly enough to make an average over a parity estimator converge, since these latter can grow as fast as $\exp(2e^{z'_{\mathbf{n}}+\tilde{z}'_{\mathbf{n}}})$ (when $\alpha_{\mathbf{n}}\beta_{\mathbf{n}}$ is negative real), which grows much faster than (6.29) decays. One thus suspects that parity expectation values will not be estimated by (5.15) in an unbiased way for a gauge P representation of the locally interacting Bose gas model.

6.2.5 Unresolved issues

Lest one think that boundary term errors are fully understood, a variety of issues are still unclear (to the best of the author's knowledge). Some of these have to do with tradeoffs between the two kinds of error. For example:

- **The role of normalization:** In Section 6.2.2 it was seen that for a (single-mode) interacting Bose gas, the un-normalized positive P representation did not suffer boundary term errors of the first kind, but had severe problems with errors of the second kind when $\mathrm{Re}\{\alpha\beta\} \to +\infty$. On the other hand, the normalized positive P representation or the gauge P do not suffer from the second kind of error for most observables, but (as will be seen in Chapter 8) does have problems with moving singularities, and hence presumably the first kind of boundary term error, when $\mathrm{Re}\{\alpha\beta\} \to -\infty$.

  There appears to be a tradeoff between the two processes causing boundary term errors, which can involve the normalization of the kernel. Nevertheless, the normalization does not parameterize the tradeoff in a straightforward way, since the two kinds of boundary term error occur in different regions of phase-space.

- **The role of moving singularities:** Only the first kind of boundary term error appears during derivation of the FPE and stochastic equations while on the other hand, errors of the second kind can occur simply on a static initial distribution, which does not evolve in any way. For these reasons it appears that moving singularities in the deterministic part of the stochastic equations should be related to boundary term errors of the first kind.

  On the other hand, moving singularities do lead to pathological tail behaviour in $P(C)$ after a certain time t_{sing} (even without any noise terms, if the initial distribution is nonzero on any point of a trajectory that escapes to infinity at finite time.) Because the far tail behaviour of $P(C)$ changes in this qualitative way after t_{sing}, one should be able (at least in some cases) to find some observables that would be estimated in an unbiased way for $t < t_{\text{sing}}$, but suffer from boundary term errors of the second kind for $t > t_{\text{sing}}$. So it appears that moving singularities can also influence the second kind of boundary term error in some way, presumably by causing slowly decaying distribution tails.

A better understanding of the relationship between the two kinds of error might lead to a way to make an advantageous tradeoff that avoids both, at least in some cases.

6.3 Removal of the errors using kernel gauges

6.3.1 Conceptual justification

If one introduces a kernel gauge $\mathcal{F}$ (e.g. a drift gauge $\mathcal{G}_k$) using a gauge identity $\mathcal{J}\left[\widehat{\Lambda}\right] = 0$, defined by (4.1) as in Section 4.1, then zero is added to the master equation in the form $\int P(C)\mathcal{F}(C)\widehat{J}\left[\widehat{\Lambda}\right] dC$, as in (4.2). Introducing such a kernel gauge can have an effect on the presence of boundary term errors in two ways:

1. **Modification of the distribution:** A change in the FPE due to addition of extra gauge terms will cause a direct change to its solution $P(C, t)$. A good choice of gauge can lead to faster decaying tails, and hence directly to $\widehat{\mathcal{B}} = 0$ from evaluation of expression (6.1), and no boundary term errors of the first

kind. This is the easiest approach in practice, and will be considered in some detail in the next Subsection 6.3.2.

Less directly, this approach also has the potential to remove boundary term errors of the *second* kind. This is again because $P(C)$ changes, and if the far tail behaviour is made to decay faster, the observable expressions (6.8a) and (6.8b) may become convergent.

2. **Renormalization of boundary terms:** The expression (6.1b) gains extra gauge terms, which can have an effect of the final value of $\widehat{\mathcal{B}}$. During the derivation of the FPE in Section 3.4.1, these extra added terms to the master equation can be written in the $T^{(\bullet)}$ notation of (3.41) and also (6.1) for $\widehat{\mathcal{B}}$ as

$$T_0^{(0)} = \mathcal{F}J^{(0)} \tag{6.30a}$$

$$T_{0j}^{(1)} = \mathcal{F}J_j^{(1)} \tag{6.30b}$$

$$T_{0jk}^{(2)} = \mathcal{F}J_{jk}^{(2)}. \tag{6.30c}$$

One finds that the additional boundary terms introduced by this kernel gauge are (using (6.1)) are of the same general form as (6.1a), but with integration over some additional terms $\widehat{\mathcal{B}}^{\text{int}}_{\mathcal{J},j}$ on top of the usual $\widehat{\mathcal{B}}^{\text{int}}_{j}$. Using (6.1b), these new terms are

$$\widehat{\mathcal{B}}^{\text{int}}_{\mathcal{J},j}(C) = \tag{6.31}$$
$$\mathcal{F}(C)\left\{P(C)J_j^{(1)}(C) + \frac{1}{2}P(C)\sum_k T_{jk}^{(2)}\frac{\partial}{\partial C_k} - \frac{1}{2}\sum_k \frac{\partial}{\partial C_k}\left(P(C)T_{kj}^{(2)}(C)\right)\right\}\widehat{\Lambda}(C).$$

The point to be made is that while the identity $\mathcal{J}\left[\widehat{\Lambda}\right]$, which can be expressed in differential form as (4.1), is zero, the integrands $\widehat{\mathcal{B}}^{\text{int}}_{\mathcal{J},j}$ for the boundary term expression $\widehat{\mathcal{B}}$ have a different form, and are not necessarily zero at all. This then may open the possibility that if originally (with no gauges), the boundary terms were nonzero, an appropriate choice of gauge can cancel these, again leading to $\widehat{\mathcal{B}} = 0$, and no boundary term errors of the first kind.

6.3.2 Practical implementation using drift gauges

Let us take the direct approach of changing $P(C,t)$ to make the boundary terms $\widehat{\mathcal{B}} = 0$. A difficulty is that the solution of $P(C,t)$ is not known for non-trivial cases, so one cannot work with it directly. However, as pointed out in Section 6.1.2, moving singularities in the noiseless equations are very closely tied to boundary term errors of the first kind. Certainly in the sense that their presence indicates weakly decaying distribution tails, and hence boundary term errors. The drift part of stochastic equations can be modified practically at will using drift kernel gauges $\mathcal{G}_k$ (see Section 4.3), and in particular, the moving singularities can be removed in this way. As will be shown in examples of Sections 6.4 and 6.5, removal of these moving singularities does indeed appear to result in disappearance of the boundary term errors.

Given a kernel amenable to drift gauges (i.e. possessing a global weight Ω), and keeping in mind the symptoms of boundary term errors of the first kind outlined in Section 6.1.2 one wishes to choose drift gauges according to the following heuristic criteria:

1. Moving singularities that were present in the original (ungauged) stochastic equations are removed. This can be assessed by analysis of the deterministic parts of the stochastic equations. See examples in subsequent sections.

2. No new moving singularities are introduced.

3. No noise divergences are introduced (in the $d\Omega = \cdots + \Omega \sum_k \mathcal{G}_k dW_k$ evolution).

4. No discontinuities in the drift or diffusion of any variables are introduced.

5. Given the above criteria 1–4 are satisfied (which can be assessed by inspection or analysis of the gauged stochastic equations), one wants to minimize the weight spread in the resulting simulation, to optimize efficiency.

   A broadly applicable rule of thumb is to **keep the gauges as small as possible while achieving the goal of removing boundary term errors.** This can be aimed for using several complementary approaches:

(a) Make the gauge functions zero in regions of phase space where no correction to the un-gauged drift is necessary.

(b) Minimize a relevant variance of the weights. This weight spread has been considered in detail in Section 4.3.4. Depending on the situation, the quantity $\sum_k |\mathcal{G}_k|^2$ (more generally) or $\sum_k \mathrm{Re}\{\mathcal{G}\}_k^2$ (for short time calculations of some observables) can be minimized, since the short-time growth of the variance of $|\Omega|$ or $\mathrm{Re}\{\Omega\}$ (respectively) has been shown in Section 4.3.4 to be proportional to these quantities.

(c) Ensure that drift gauges are zero at deterministic attractors in phase-space, so that no weight spread is accumulated when no significant evolution is taking place.

(d) Minimize the variances of estimators of observables of interest. That is, the variance of quantities $\mathrm{Tr}\left[\widehat{O}\widehat{\Lambda}\right]$ and $\mathrm{Tr}\left[\widehat{O}\widehat{\Lambda}^\dagger\right]$ appearing in the numerator of observable estimates (3.14). In this manner, one can optimize the simulation for the estimation of a particular observable $\widehat{O}$, rather than more generally as in point (a) above.

6. One should also avoid gauges with the following features:

(a) Gauges that are nonzero only in a rarely visited region of phase space. The problem is that good sampling of the gauged region is needed to make the correct change to observable estimates — otherwise, if the gauged region is not visited by any trajectories, one has practically the same simulation as without gauges! In a less extreme case where only a tiny (but nonzero) fraction of trajectories visit the gauged region, the correction from biased-ungauged to unbiased-gauged observable estimates is badly resolved, and the simulation results become very noisy. Furthermore, if the bias to be corrected is significant, then the change in the few gauged trajectories needs to be large to compensate for the bias, and hence the weight correction on those trajectories will become large as well. This leads to a large weight spread. It is much better to make small changes in many trajectories contribute to removal of the bias.

(b) Gauges rapidly changing in phase-space. These typically lead to stiff equations, and a gauged region of phase space that is badly sampled by reasonable-sized ensembles. This bad sampling arises when no single trajectory stays in the gauged region long, to such a degree that at a given time few (or, on average, possibly less than one) trajectories are sampling the gauged region, i.e. sampling the resulting observable estimate corrections. This results in similar problems as in the previous point (a).

(c) Non autonomous gauges (i.e. depending directly on the weight Ω). The problem with these is that they are usually much more difficult to analyze to check whether or not new moving singularities or noise divergences have been introduced.

(d) Gauges analytic in complex variables if the kernel is also analytic in these. As discussed in Section 6.1.4, it is desirable to break the analyticity of the stochastic equations to avoid moving singularities.

Note that an explicit variational minimization according to 5(b) or 5(d) could be in conflict with the recomemndations 6(a) or 6(b), in which case it is probably best to try the simulations and see which works better.

This technique appears to be broadly applicable, and only requires the recognition of what instabilities in the stochastic equations could lead to problems. It does not require detailed knowledge of what the boundary terms $\widehat{\mathcal{B}}$ are, nor of the behaviour of the far tails of $P(C)$, provided instabilities are removed.

6.3.3 Ito vs. Stratonovich moving singularities

There is a subtle ambiguity that can arise when using the method just presented: Should one remove moving singularities from the Ito or the Stratonovich form of the stochastic equations?

As indicated in Appendix B, when one numerically integrates the stochastic differential equations using a finite but small time step Δt, the exact form of the drift that is integrated can depend on details of the integration algorithm chosen. For

example, if one evaluates the derivatives using the variable values at the beginning of a time step (i.e. simply those calculated using the previous step), then one integrates the standard Ito form of the equations. However, with a semi-implicit method, where one first estimates the variables in the middle of the time step, before using *these* to calculate the derivative and actually advance the simulation, one must integrate the *Stratonovich* form of the stochastic differential equations. The difference between these is given by the Stratonovich correction to the drift (B.7). In practice, multiplicative[3] stochastic equations, such as those considered in this thesis, are more stable (and hence, efficient) using the semi-implicit method[82].

In some situations it may occur that moving singularities occur in the deterministic parts of the Stratonovich equations but not the Ito equations (or vice-versa). To arrive at practical guidelines on this issue, one can argue in the following way:

- Since both the explicit (Ito) and semi-implicit (Stratonovich) algorithm simulate the same underlying equations, then to be sure that no bias is being introduced, moving singularities should not occur in *either* algorithm. In fact, a whole family of hybrid Ito-Stratonovich algorithms are possible where derivatives are evaluated at arbitrary time values during the time step Δt, and the difference in the drift terms between two such algorithms is a multiple $\in [0, 2]$ of the Stratonovich correction (B.7) — See Appendix B. *None* of these algorithms should contain moving singularities, otherwise some kind of bias may be suspected.

- The Stratonovich correction for variable C_j (B.7) is highly dependent on the form of the noise matrix elements ($\propto \sum_{kl} B_{kl} \partial B_{jl} / \partial C_k$). If one finds that moving singularities are present for one algorithm but not for another, this indicates that the Stratonovich correction is growing very rapidly as one heads for the boundaries of phase space. This in turn indicates that the noise matrix elements must be growing in some appropriately rapid fashion as well — possibly rapidly enough to cause noise divergences.

[3] i.e. equations in which the noise terms are proportional to the variables, and hence of non-constant variance.

Taking the above points into consideration, it is conjectured here that:

1. In a simulation that does not suffer from boundary term errors, there must be no moving singularities in the deterministic part of either the Ito, Stratonovich, or any of the hybrid Ito-Stratonovich forms of the stochastic equations.

2. If a simulation has moving singularities for one of the family of algorithms (e.g. Ito), but not for some other member of this family, then noise divergences may be suspected as an additional cause of boundary term errors.

3. To eliminate biases caused by moving singularities using a drift gauge, one must remove moving singularities from *all* hybrid Ito-Stratonovich algorithms with the same gauge choice.

6.4 Removal example 1: Two-particle absorber

Two examples will now be worked through — those systems in which boundary term errors in positive P simulations have been reported in published work. Removal of boundary term errors using the guidelines of Section 6.3.2 will be demonstrated, and the procedure considered in some detail. The first example will be a single-mode two-boson absorber (in this section), while the second example of a single-mode laser will be treated in Section 6.5. These two examples demonstrate somewhat different aspects of boundary term errors and their removal.

This section is closely based on Section IV of the published article by Deuar and Drummond[66].

6.4.1 The single-mode model

Consider a single optical cavity mode driven by coherent radiation, and damped by a zero temperature bath that causes both one- and two-photon losses. For convenience, units of time are chosen so that the rate of two-photon loss is unity. The scaled one-photon loss rate is γ, and ε is the scaled (complex) driving field amplitude. Following

standard techniques, the system can be described with the master equation

$$\begin{aligned}\frac{\partial\widehat{\rho}}{\partial t} &= \left[\,\varepsilon\widehat{a}^{\dagger}-\varepsilon^{*}\widehat{a}\,,\,\widehat{\rho}\right]+\frac{\gamma}{2}(2\,\widehat{a}\widehat{\rho}\widehat{a}^{\dagger}-\widehat{a}^{\dagger}\widehat{a}\widehat{\rho}-\widehat{\rho}\widehat{a}^{\dagger}\widehat{a})\\ &\quad+\frac{1}{2}(2\,\widehat{a}^{2}\widehat{\rho}\widehat{a}^{\dagger 2}-\widehat{a}^{\dagger 2}\widehat{a}^{2}\widehat{\rho}-\widehat{\rho}\widehat{a}^{\dagger 2}\widehat{a}^{2})\;. \end{aligned} \tag{6.32}$$

This is also the master equation of two-particle absorption in an interaction picture where kinetic processes have has been relegated to the Heisenberg picture evolution of the operators.

Let us use the gauge P representation to convert this master equation into a set of stochastic equations. The single-mode kernel is written as

$$\widehat{\Lambda}(\;C=\{\alpha,\beta,\Omega\}\;)=\Omega||\alpha\rangle\langle\beta^{*}||e^{-\alpha\beta}. \tag{6.33}$$

Following the treatment of Sections 3.4, 4.3, and 5.3, and converting to Stratonovich form using (B.7), one arrives at the Stratonovich stochastic equations

$$d\alpha = [\,\varepsilon\; -\alpha(\alpha\beta+i\mathcal{G}+\{\gamma-1\}/2)]\,dt+i\alpha dW \tag{6.34a}$$

$$d\beta = [\,\varepsilon^{*}-\beta(\alpha\beta+i\widetilde{\mathcal{G}}+\{\gamma-1\}/2)]\,dt+i\beta d\widetilde{W} \tag{6.34b}$$

$$d\Omega = S_{\Omega}\,dt+\Omega\left[\mathcal{G}dW+\widetilde{\mathcal{G}}d\widetilde{W}\right]\quad. \tag{6.34c}$$

Here $S_{\Omega}dt$ is the appropriate Stratonovich correction term given by (B.7), which depends on the functional form of the particular gauges chosen.

With no gauge ($\mathcal{G}=\widetilde{\mathcal{G}}=0$), the positive P Stratonovich equations are recovered:

$$d\alpha = [\,\varepsilon\;\; -\alpha(\alpha\beta+\{\gamma-1\}/2)]\,dt+i\alpha dW \tag{6.35a}$$

$$d\beta = [\,\varepsilon^{*}-\beta(\alpha\beta+\{\gamma-1\}/2)]\,dt+i\beta d\widetilde{W}\;. \tag{6.35b}$$

This two-particle loss model can give either correct or incorrect results when treated with the usual positive P representation methods. Generally, problems only arise when any accompanying single-particle losses have small rates or when the number of bosons per mode is small (see Figure 6.3). It is a well-studied case, and a detailed treatment of the boundary term errors that appear can be found in [64]. It also has the merit that exact solutions can be readily found using other means. Analyzing this example gives a good check that the moving singularity

removal procedure of Section 6.3.2 does in fact eliminate boundary terms and leads to correct results. Several simplifications of this model will now be concentrated on, which correspond to existing literature.

6.4.2 Two-boson absorber

In the simplest form of this model, corresponding to $\gamma = \varepsilon = 0$, only two-boson absorption takes place. One expects that for some state $|\psi\rangle = \sum_{n=0}^{\infty} c_n |n\rangle$ in a Fock number state basis, all even boson number components will decay to vacuum, and all odd-numbered components will decay to $|1\rangle$, leaving a mixture of vacuum and one-boson states at long times.

The positive P representation has been found to give erroneous results [70, 83, 84, 85] accompanied by moving singularities [64]. There are power-law tails in the distribution that lead to boundary term errors — consistently with the conditions (6.7). The observable usually concentrated on in this system is the occupation number $\widehat{n} = \widehat{a}^\dagger \widehat{a}$, which is estimated by $\overline{n} = \langle \alpha\beta = \breve{n} \rangle_{\text{stoch}}$ in the positive P representation. The variable $\breve{n} = \alpha\beta$ has a convenient closed equation (Stratonovich),

$$d\breve{n} = -2\breve{n}(\breve{n} + i\mathcal{G}^+ - 1/2)\, dt + i\breve{n}\, dW^+ \tag{6.36}$$

with $dW^+ = (dW + d\widetilde{W})$, and $\mathcal{G}^+ = (\mathcal{G} + \widetilde{\mathcal{G}})/2$.

Let us examine the behavior of the above equation when $\mathcal{G}^+ = 0$, i.e., in the standard, un-gauged formulation. The deterministic part of the evolution has a repellor at $\breve{n} = 0$, and an attractor at $\breve{n} = 1/2$. The noise is real, finite, and of standard deviation $\sqrt{dt/2}$ at the attractor. The deterministic part of the evolution has a single trajectory of measure zero that can escape to infinity along the negative real axis,

$$\alpha(t) = -\beta(t) = \frac{1}{\sqrt{2(t_{\text{sing}} - t)}}, \tag{6.37}$$

where $t_{\text{sing}} = 1/2\alpha(0)^2 = -1/2\breve{n}(0)$. This moving singularity is a symptom of likely boundary term errors.

Indeed, in the steady-state limit, all trajectories in a simulation will head toward $\breve{n} = 1/2$, with some remnant noise around this value being kept up by the stochastic

term. This noise does not affect the average, however, and the estimator of $\langle\widehat{n}\rangle$ becomes $\lim_{t\to\infty}\overline{n} = \lim_{t\to\infty}\langle\breve{n}\rangle_{\rm stoch} = 1/2$ in the many trajectories limit. Quantum mechanics, however, predicts that if we start from a state $\widehat{\rho}(0)$, the steady state will be

$$\lim_{t\to\infty}\langle\widehat{n}\rangle = \sum_{j=0}^{\infty}\langle 1+2j|\,\widehat{\rho}(0)\,|1+2j\rangle \ , \tag{6.38}$$

since all even occupations decay to vacuum, and all odd to $|1\rangle$. For a coherent state $|\alpha_0\rangle$ input, say, this will be

$$\lim_{t\to\infty}\langle\widehat{n}\rangle = \frac{1}{2}\left(1-e^{-2|\alpha_0|^2}\right) \ . \tag{6.39}$$

Thus one expects that the positive P simulation will give correct results only when $e^{|\alpha_0|^2}\gg 1$.

To correct the problem one has to change the phase-space topology in some way to prevent the occurrence of such moving singularities. It was found that a good gauge for a two-boson absorber nonlinearity in general is

$$\mathcal{G} = \widetilde{\mathcal{G}} = \mathcal{G}^{+} = i(\breve{n}-|\breve{n}|) \ . \tag{6.40}$$

This will be dubbed the "circular" gauge due to the creation of an attractor at $|\breve{n}| = 1/2$.

This gauge has the effect of replacing the $-2\breve{n}^2$ term in (6.36) with $-2\breve{n}|\breve{n}|$, which is always a restoring force, and so never leads to any escape to infinity in finite time. With the gauge (6.40), the Stratonovich equations become

$$d\breve{n} = \breve{n}(1-2|\breve{n}|)\,dt + i\breve{n}dW^{+} \ , \tag{6.41a}$$

$$d\alpha = -\alpha(|\breve{n}|-1/2)\,dt + i\alpha dW \ , \tag{6.41b}$$

$$d\Omega = \Omega\left\{\,[\breve{n}+(\breve{n}-|\breve{n}|)^2]\,dt + i(\breve{n}-|\breve{n}|)\,dW^{+}\right\} \ . \tag{6.41c}$$

How do these new equations measure up to the four symptoms of boundary term errors of the first kind outlined in Section 6.1.2? In logarithmic variables $\alpha_L = \log\alpha$, $n_L = \log\breve{n}$, and $z_0 = \log\Omega$, the Stratonovich system equations are

$$dn_L = (1-2|\breve{n}|)\,dt + i dW^{+} \tag{6.42a}$$

$$d\alpha_L = -(|\breve{n}|-1/2)\,dt + i dW \tag{6.42b}$$

$$dz_0 = [\breve{n}+(\breve{n}-|\breve{n}|)^2]\,dt + i(\breve{n}-|\breve{n}|)\,dW^{+}. \tag{6.42c}$$

One now has an attractor on the circle $|\breve{n}| = 1/2$, and a repellor at $\breve{n} = 0$, with free phase diffusion of n in the tangential direction. Once trajectories reach the attractor, only phase diffusion occurs. So:

1. *Moving singularities:* There are no moving singularities in n_L, since $|\breve{n}|$ always remains bounded. Also, since the deterministic evolution of the remaining independent variables α_L and z_0 depends only on $\breve{n}$ or $|\breve{n}|$ in such a way that $d\alpha_L$ and dz_0 remain finite for finite $\breve{n}$, then these variables never escape to infinity in finite time (being just a time integral of finite values of $\breve{n}$). Thus, no moving singularities occur.

2. *Noise divergences:* Similarly because the noise coefficient of n_L is constant, while the noise coefficients of α_L and z_0 depend only on $\breve{n}$, and in such a way that for finite $\breve{n}$ the noise magnitude remains finite — then one concludes that noise divergences are not present either.

3. *Discontinuous drift or noise:* Not present, by inspection.

4. *Initial distribution:* Initial coherent state has $P(0) = \delta^2(\alpha - \alpha_0)\delta^2(\beta - \alpha^*)\delta^2(\Omega - 1)$. There are no boundary terms at the outset with such an infinitely-peaked distribution.

How does this change in the Ito formulation? Using (B.7) on (6.42), one finds that the Ito and Stratonovich equations are identical for n_L and α_L, while the Ito weight equation is $dz_0 = (\breve{n} - |\breve{n}|)^2 dt + i(\breve{n} - |\breve{n}|)\, dW^+$. The above arguments apply without change to Ito, and to hybrid Ito-Stratonovich algorithms. So — no symptoms of boundary term errors remain in any formulation.

Let us see how the gauge P simulation compares with the positive P in practice — i.e. are boundary term errors *really* removed along with the moving singularities? Figure 6.2 compares results for an (exact) truncated number-state basis calculation, a positive P calculation, and a gauge calculation using the circular gauge (6.40) for an initial coherent state of $\alpha_0 = 1/\sqrt{2}$. Figure 6.3 compares steady-state values for exact, positive P, and gauge calculations for a wide range of initial coherent states. It is seen that the gauge calculation is correct to within the small errors due to

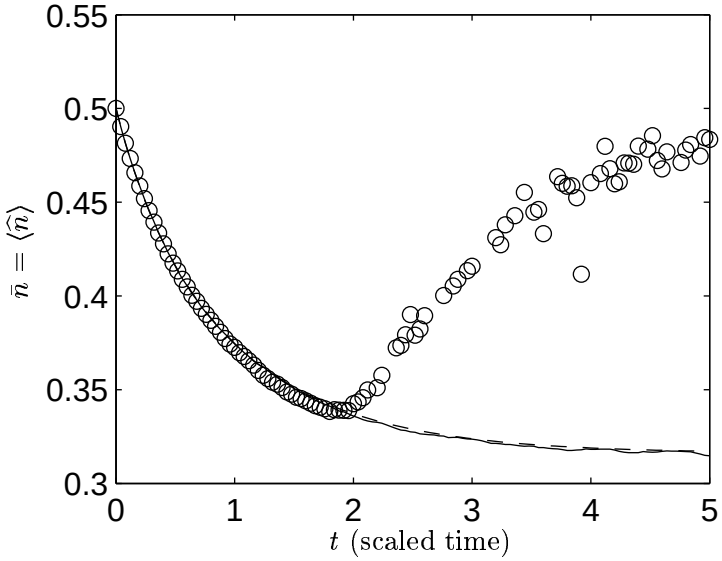

Figure 6.2: **Comparison of two-boson absorption simulations:** time-varying mode occupation. CIRCLES: positive P simulation; SOLID LINE: gauge simulation using circular gauge (6.40); DASHED LINE: exact calculation (truncated number-state basis). Simulation parameters: $\mathcal{S} = 4 \times 10^4$ trajectories; initial coherent state.

finite sample size, whereas the positive P calculation leads to severe errors at low occupations.

6.4.3 One- and two-boson absorber

If one now considers single-boson decay as well, but still with no driving, one expects that all states will decay to the vacuum on a time scale $1/\gamma$ on top of the two-particle losses, which produce decay of particle numbers on a timescale of 1 (although not necessarily right down to the vacuum.) When $\gamma \gg 1$, the single-particle decay dominates and nothing interesting is seen beyond an exponential decay of particle number. However, if $\gamma \lesssim 1$, one should first see a rapid decay to a mixture of vacuum and one-boson states due to the two-boson process, and then a slow decay of the one-boson state to the vacuum on a time scale of $t \approx 1/\gamma$.

In this case the positive P (Stratonovich) equations display different behavior depending on whether γ is above or below the threshold $\gamma = 1$. Below threshold, there is an attractor at $\breve{n} = (1-\gamma)/2$, and a repellor at $\breve{n} = 0$, while above threshold, the attractor is at $\breve{n} = 0$, and the repellor at $\breve{n} = -(\gamma-1)/2$. In either case, there is a

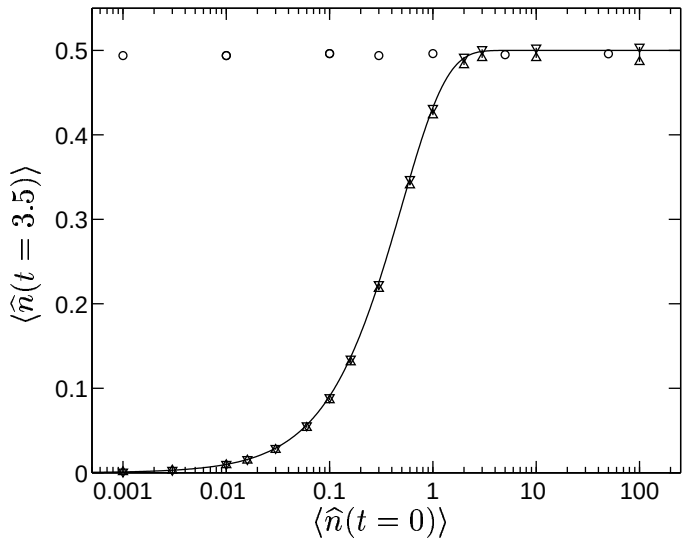

Figure 6.3: **Steady state expectation values of boson number** $\langle \widehat{n} \rangle$ obtained by gauge simulations (DOUBLE TRIANGLES) compared to exact analytic results (6.39) (SOLID LINE) and positive P simulations (CIRCLES) for a wide range of initial coherent states. Size of uncertainty in gauge results due to finite sample size is indicated by vertical extent of 'double-triangle' symbol. Steady state was observed to have been reached in all simulations by $t = 3.5$ or earlier (compare with Figure 6.2 and 6.4), hence this is the time for which the simulation data is plotted. $\mathcal{S} = 10^5$ trajectories

divergent trajectory along the negative real axis, which is again a moving singularity. It turns out that the steady state calculated with the positive P is erroneous while $\gamma < 1$, and that there are transient boundary term errors while $\gamma < 2$ [70]. The false steady state below threshold lies at the location of the attractor: $(1-\gamma)/2$.

Let us try to fix this problem using the same circular gauge (6.40) as before. The gauged equation for n_L is now

$$dn_L = (1 - \gamma - 2|\breve{n}|)\, dt + i dW^+ \,, \tag{6.43}$$

while the α and Ω evolution is unchanged. So, above threshold one is left with only an attractor at $\breve{n} = 0$, while below threshold one has a repellor at $\breve{n} = 0$ surrounded by an attracting circle at $|\breve{n}| = (1-\gamma)/2$. This phase space again has no moving singularities or noise divergences.

The results of simulations for the parameter $\gamma = 0.1$ are shown in Figure 6.4 ($\gamma \ll 1$ was chosen so that a system with two widely differing time scales is tested.) The gauge simulation tracks the exact results and avoids the false results of the positive P simulation. Note also that the gauge simulation remains efficient for a wide range of occupation numbers — from $\langle \widehat{n} \rangle \approx 100 \gg 1$, where the positive P is also accurate, to $\langle \widehat{n} \rangle \approx 0.1 \ll 1$ where it is totally incorrect.

6.4.4 Driven two-boson absorber

The other type of situation to consider is when there is a driving field as well as two-boson damping. In this Subsection the coherent driving ε is nonzero, but the single-particle loss rate is assumed negligible ($\gamma = 0$), since this process never causes any of the simulation problems anyway, but leaving it out simplifies analysis. Failure of the positive P representation method has been found in this limit as well [71], and is evident in Figure 6.5. The equation for $\breve{n}$ is no longer stand-alone in this case, and all three complex variables must be simulated as in (6.34c), the Ω equation being the same as in the undriven case (6.41).

A treatment of the singular trajectory problem with the same circular gauge (6.40) leads again to correct results, as seen in Figure 6.5.

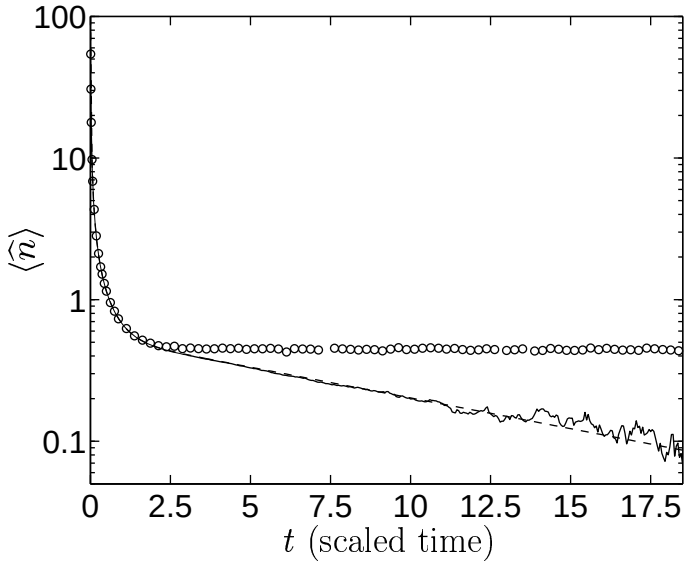

Figure 6.4: **Comparison of simulations for system with both one- and two-boson damping.** Relative single-boson damping strength $\gamma = 0.1$; CIRCLES: positive P simulation. SOLID LINE: simulation using the circular gauge (6.40); DASHED LINE: exact calculation (truncated number-state basis). Gauge simulation parameters: 100 000 trajectories; initial coherent state $|10\rangle$ with $\langle \widehat{n} \rangle = 100$ bosons.

6.4.5 Relevance to interacting Bose gas

The master equation (6.32) was first considered in the context of quantum optics. There, single-photon losses are usually large relative to two-photon processes, and mode occupations are high, so the boundary term errors have not been a practical problem in their original quantum optics regime. However, for other physical systems such as interacting Bose gases (which this thesis is most interested in), two-particle losses can be stronger, one expects a large number of modes to have low or practically zero occupation, and so the boundary term errors may be significant. For example, collisional many-body processes that remove particles in pairs from the modeled field $\widehat{\Psi}$, will lead to similar two-particle loss terms (2.24) in the master equation.

6.5 Removal example 2: Single-mode laser

Let us now consider the second quantum system for which systematic errors have been seen with the positive P representation. The mechanism by which boundary

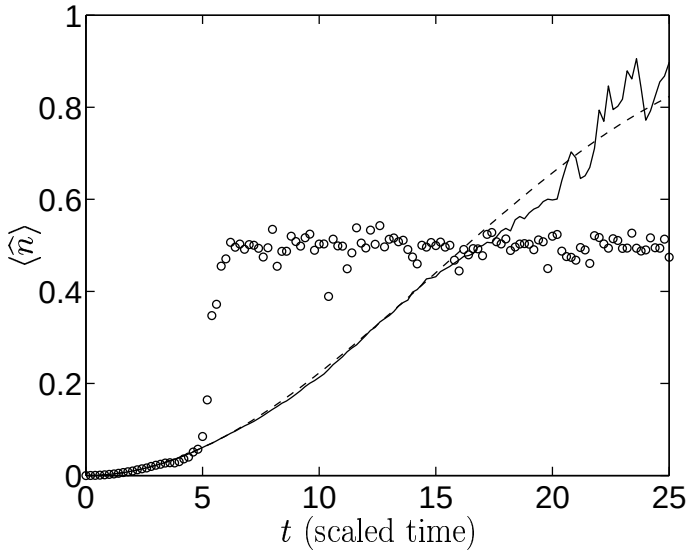

Figure 6.5: **Driven two-boson absorber** with $\varepsilon = 0.05$. CIRCLES: positive P simulation (1000 trajectories); SOLID LINE: circular gauge (6.40) simulation (100 000 trajectories); DASHED LINE: exact calculation (truncated number-state basis). Initial vacuum state.

term errors occur here is somewhat different than in the two-boson absorber. For two-boson damping, boundary term errors occur even when one chooses an optimal (i.e., compact) initial distribution to represent the starting state $|\alpha_0\rangle\langle\alpha_0|$, whereas for this laser model the systematic errors occur only if one starts with unusually broad initial distributions. Nevertheless, such broader distributions may arise during time-evolution from some other state, so consideration of what occurs for significant distribution broadness is also relevant to practical simulations.

This section is closely based on Section IV of the published article by Deuar and Drummond[66].

6.5.1 The laser model

For a simple model of a photonic or atomic laser, using the single-mode positive P representation,

$$\widehat{\Lambda}(\ C = \{\alpha, \beta\}\ ) = ||\alpha\rangle\langle\beta^*||e^{-\alpha\beta}, \tag{6.44}$$

one can derive[71, 64] the the Fokker-Planck equation

$$\frac{\partial P}{\partial t} = \left\{ \frac{\partial}{\partial \alpha}\left[\frac{\alpha\beta}{\mathcal{N}} - G\right]\alpha + \frac{\partial}{\partial \beta}\left[\frac{\alpha\beta}{\mathcal{N}} - G\right]\beta + 2Q\mathcal{N}\frac{\partial^2}{\partial\alpha\partial\beta}\right\} P, \tag{6.45}$$

where the average number of photons in the mode is $\overline{n} = \langle \alpha\beta \rangle_{\rm stoch}$, and $\mathcal{N}$ is a scaling parameter that equals the number of atoms in the gain medium. Both G, the gain parameter, and $Q \geq G/\mathcal{N}$, the noise parameter, are real and positive. One usually considers the scaled variables

$$\widetilde{\alpha} = \alpha/\sqrt{\mathcal{N}} \tag{6.46a}$$

$$\widetilde{\beta} = \beta/\sqrt{\mathcal{N}}, \tag{6.46b}$$

The stochastic Ito equations are then[71, 64]

$$d\widetilde{\alpha} = (G - \widetilde{\alpha}\widetilde{\beta})\,\widetilde{\alpha}\,dt + \sqrt{2Q}\,d\eta \tag{6.47a}$$

$$d\widetilde{\beta} = (G - \widetilde{\alpha}\widetilde{\beta})\,\widetilde{\beta}\,dt + \sqrt{2Q}\,d\eta^*, \tag{6.47b}$$

where $d\eta$ is a complex Wiener increment obeying $\langle d\eta^* d\eta \rangle_{\rm stoch} = dt$, $\langle d\eta \rangle_{\rm stoch} = \langle d\eta^2 \rangle_{\rm stoch} = 0$.

One is again usually interested in the (scaled) occupation of the lasing mode

$$\overline{n_{\rm sc}} = \left\langle n_{\rm sc} = \widetilde{\alpha}\widetilde{\beta} \right\rangle_{\rm stoch} = \langle \widehat{a}^\dagger \widehat{a} \rangle / \mathcal{N}, \tag{6.48}$$

i.e. the number of bosons in the lasing mode per each atom in the gain medium. Changing variables to $n_{\rm sc}$ and an auxiliary "phase-like" variable $z_2 = \widetilde{\alpha}/\widetilde{\beta}$, the FPE becomes

$$\frac{\partial P}{\partial t} = \left\{ 2\frac{\partial}{\partial n_{\rm sc}} \left[Q + n_{\rm sc}(n_{\rm sc} - G) \right] + 2Q \frac{\partial}{\partial z_2} \frac{n_{\rm sc}}{z_2} + 2Q \frac{\partial^2}{\partial n_{\rm sc}^2} n_{\rm sc} - 2Q \frac{\partial^2}{\partial z_2} \frac{z_2^2}{n_{\rm sc}} \right\} P. \tag{6.49}$$

The diffusion matrix is now decoupled

$$4Q \begin{bmatrix} n_{\rm sc} & 0 \\ 0 & -z_2^2/n_{\rm sc} \end{bmatrix}, \tag{6.50}$$

and so the two complex variables can evolve under independent noises. In fact the entire $n_{\rm sc}$ evolution decouples from z_2, and can be written as the closed Ito equation

$$dn_{\rm sc} = -2(n_{\rm sc} - a)(n_{\rm sc} - b)dt + 2\sqrt{Q n_{\rm sc}} dW \ , \tag{6.51}$$

where now the *real* Wiener increments (implemented as Gaussian noises) have variance $\langle dW^2 \rangle_{\rm stoch} = dt$, and the deterministic stationary points are

$$a = \tfrac{1}{2}\left(G + \sqrt{G^2+4Q}\right) \geq 0 \tag{6.52a}$$

$$b = \tfrac{1}{2}\left(G - \sqrt{G^2+4Q}\right) \leq 0. \tag{6.52b}$$

(In the Stratonovich calculus, $[\begin{smallmatrix} a \\ b \end{smallmatrix}] = \frac{1}{2}(G \pm \sqrt{G^2+2Q})$). One finds that the stationary point at a is an attractor, and at b there is a repellor. Defining $n_{\rm neg} = b - n_{\rm sc}$, as the distance (in the negative direction) from the repellor, one obtains

$$\frac{dn_{\rm neg}}{dt} = 2n_{\rm neg}(n_{\rm neg} + \sqrt{G^2+4Q}) + \text{noise} \ , \tag{6.53}$$

which shows that there is a singular trajectory escaping to infinity in finite time along the negative real axis $n_{\rm sc} < b$ — i.e. a moving singularity.

6.5.2 Initial conditions

Let us consider the case of vacuum initial conditions. A vacuum can be represented by the positive P representation

$$P_+(\widetilde{\alpha},\widetilde{\beta}) = \delta^2(\widetilde{\alpha})\,\delta^2(\widetilde{\beta}) \ , \tag{6.54}$$

but also by Gaussian distributions of any variance σ_0^2, around the above,

$$P^{(+)}(\widetilde{\alpha},\widetilde{\beta}) = \frac{1}{4\pi^2\sigma_0^4}\exp\left\{-\frac{|\widetilde{\alpha}|^2+|\widetilde{\beta}|^2}{2\sigma_0^2}\right\} \ . \tag{6.55}$$

It can be checked that this satisfies the vacuum positive P representation condition (3.34), giving another family of equivalent vacuum distributions different to (3.35), which was found in Section 3.3.6. Note: the distribution of $n_{\rm sc}$ is non-Gaussian, but has a standard deviation of $\sigma_{n_{\rm sc}} \approx \sqrt{2}\,\sigma_0^2$ in both the real and imaginary directions.

It has been found by Schack and Schenzle [71] that for this single-mode laser model, a positive P simulation of pumping from a vacuum will give correct answers if the usual δ-function initial condition (6.54) is used, but will have systematic errors if an initial condition (6.55) has a sufficiently large variance (see Figure 6.6).

This can be understood by inspection of the stochastic equation (6.51). If one has a sufficiently broad initial distribution, that some trajectories start with

$\mathrm{Re}\{n_{\mathrm{sc}}\} < b$, then the singular trajectory will be explored by the distribution. From this it appears that for all $\sigma_0 > 0$, some boundary term errors may be present. In practice, no errors were seen by Schack and Schenzle for small enough σ_0, presumably because their magnitude was below the signal-to-noise resolution obtained in those simulations — an interesting effect in itself. Additionally, even if initially no trajectories fall on the singular trajectory (e.g. some initial distribution that is exactly zero for $\mathrm{Re}\{n_{\mathrm{sc}}\} < b$), the region $\mathrm{Re}\{n_{\mathrm{sc}}\} < b$ may be subsequently explored due to the presence of the noise terms.

Apart from the obvious δ-function initial condition, one might want to try the canonical distribution (3.10) applicable generally to any density matrix. For optical laser models, this will not cause practical problems as then the variance is $\sigma_0^2 = 1/\mathcal{N}$, which for any realistic case will be very small (i.e., $\sigma_{n_{\mathrm{sc}}} \ll |b|$). Hence, boundary term errors would be negligible for a wide range of simulation times, based on the evidence that they were not seen [71] in this regime. Nevertheless, broader distributions may arise during time-evolution of some other state *to* a vacuum, so consideration of what occurs for significant σ_0 is also relevant to practical simulations.

Incidentally, the anomalous results were discovered in [71] when $\sigma_0^2 = 1$ was chosen by the erroneous procedure of scaling the equations while not simultaneously scaling the canonical initial condition in α.

6.5.3 Drift gauged equations

Let us introduce the single-mode gauge P kernel (5.4) (compare with (6.44), which differs only by lack of the global weight factor e^{z_0}). Proceeding to introduce drift gauges as in Section 4.3 (for n_{sc} but not for z_2, as averages of this variable will not figure in what follows), one obtains the Ito stochastic equations

$$dn_{\mathrm{sc}} = 2n_{\mathrm{sc}}(G - n_{\mathrm{sc}})\,dt + 2Q\,dt + 2\sqrt{Qn_{\mathrm{sc}}}\,(dW - \mathcal{G}\,dt) \tag{6.56a}$$

$$dz_0 = -\frac{1}{2}\mathcal{G}^2\,dt + \mathcal{G}dW, \tag{6.56b}$$

as per the standard gauge formulation (4.90). It is convenient to define a transformed gauge function

$$\mathcal{G}_{(n)} = \mathcal{G}\sqrt{\frac{Q}{n_{\rm sc}}} \tag{6.57}$$

, which is also arbitrary and enters into the simulation as

$$dn_{\rm sc} = 2n_{\rm sc}(G - n_{\rm sc} - \mathcal{G}_{(n)})\,dt + 2Q\,dt + 2\sqrt{Qn_{\rm sc}}\,dW \tag{6.58a}$$

$$dz_0 = -\frac{n_{\rm sc}\mathcal{G}_{(n)}^2}{2Q}\,dt + \mathcal{G}_{(n)}\sqrt{\frac{n_{\rm sc}}{Q}}dW. \tag{6.58b}$$

6.5.4 Correcting for the moving singularity

Consider the (Ito) deterministic evolution of, $n'_{\rm sc}$, the real part of $n_{\rm sc} = n'_{\rm sc} + in''_{\rm sc}$,

$$\frac{dn'_{\rm sc}}{dt} = -2{n'_{\rm sc}}^2 + 2Gn'_{\rm sc} + 2Q + 2{n''_{\rm sc}}^2 - 2{\rm Re}\left\{n_{\rm sc}\mathcal{G}_{(n)}\right\}\ . \tag{6.59}$$

The moving singularity is due to the $-2{n'_{\rm sc}}^2$ leading term for negative values of $n'_{\rm sc}$. The drift gauge is now chosen according to the criteria below (as in Section 6.3.2, but in a different order):

1. As pointed out in Sections 4.3.4 and 6.3.2, it is desirable to keep the gauge terms to a minimum because whenever they act the weights of trajectories become more randomized. Thus, let us restrict ourselves to functions $\mathcal{G}_{(n)}$ that are only nonzero for $n'_{\rm sc} < 0$.

2. There should not be any discontinuities in the drift equations, as this might lead to bad sampling of some parts of the distribution and thus possibly to systematic biases. A further problem with discontinuous gauges is that the Stratonovich correction term for dz_0 cannot be calculated at the discontinuity, preventing use of the more stable semi-implicit numerical algorithm (See Appendix B).

   Hence, in particular, in light of the previous point 1., one should have $\lim_{n'_{\rm sc}\to 0}(\,\mathcal{G}_{(n)}\,) = 0$. For ease of analysis, let us start with a simple form for the gauge, $\mathcal{G}_{(n)} = c - \lambda n'_{\rm sc} + \lambda_y n''_{\rm sc}$. Continuity immediately implies $c = \lambda_y = 0$, hence

$$\mathcal{G}_{(n)} = \begin{cases} -\lambda n'_{\rm sc} & \text{if } \ n'_{\rm sc} < 0 \\ 0 & \text{if } \ n'_{\rm sc} \geq 0 \end{cases}\ , \tag{6.60}$$

3. The next necessary (and most important) condition, to remove moving singularities, is that the $-2{n'_{\rm sc}}^2$ term is canceled, hence:

$$\lambda \geq 1\,. \tag{6.61}$$

Using (B.7), the Stratonovich correction for $dn_{\rm sc}$ is $-Q\,dt$, and so the family of hybrid Ito-Stratonovich equations differ only in having this term a multiple of $Q\,dt$ in the range $[0, 2Q\,dt]$. This has no effect on the removal or not of the moving singularity.

4. Now, if $\lambda = 1$ there are no systematic errors, but the sampling error very quickly obscures any observable estimates because $n'_{\rm sc}$ still heads to $-\infty$ exponentially fast due to the $2Gn'_{\rm sc}$ term. This takes it into regions of ever increasing $|\mathcal{G}_{(n)}|$, and weights quickly become randomized. For slightly larger parameters λ, the $n'_{\rm sc}$ evolution takes trajectories to a point lying far into the negative $n'_{\rm sc}$ region where the two leading terms balance. Here the trajectories sit, and rapidly accumulate weight noise. It is clear that for an optimum simulation all stationary points of $n'_{\rm sc}$ in the nonzero gauge region must be removed. In this system the condition for this is

$$\lambda > 1 + \frac{G^2}{2Q}, \tag{6.62}$$

for the Stratonovich calculus, and $\lambda > 1 + G^2/4Q$ for Ito.

5. Are any new moving singularities or noise divergences introduced? The gauged equation for the imaginary part of $n_{\rm sc}$ is

$$dn''_{\rm sc} = 2n''_{\rm sc}[G + n'_{\rm sc}(\lambda - 2)]\,dt + 2\sqrt{Q|n_{\rm sc}|}\,\sin\left(\tfrac{1}{2}\angle n_{\rm sc}\right)\,dW, \tag{6.63}$$

which is attractive towards $n''_{\rm sc} = 0$ when $\lambda > 2$ in the gauged ($n'_{\rm sc} < 0$) region, and exponentially growing otherwise. In either case, no super-exponential escape occurs, and no moving singularities. The evolution of z_0 depends only on finite expressions of finite $n_{\rm sc}$, is thus an integration over the history of a trajectory, and so does not diverge in finite time. The Stratonovich correction for dz_0 is $\lambda\left(n'_{\rm sc} + n_{\rm sc} + |n_{\rm sc}|\right)/2$ in the gauged region, zero otherwise. Introduction

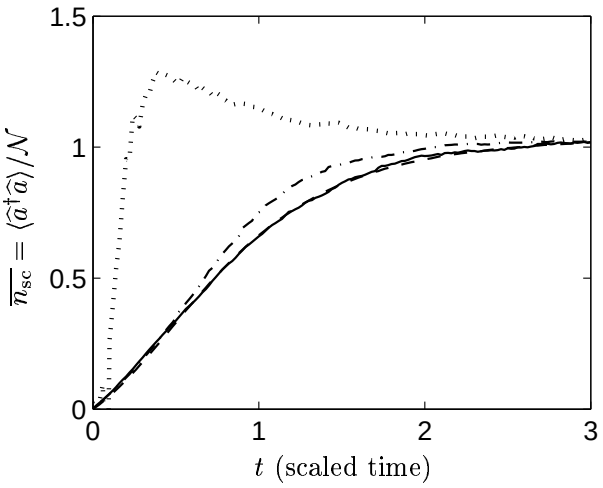

Figure 6.6: One-mode laser $G = 1$, $Q = 0.25$. DASHED LINE: (correct) positive P simulation with δ-function initial conditions (6.54) $\sigma_0^2 = 0$, and 100000 trajectories. DOTTED-DASHED LINE: erroneous positive P simulation with Gaussian initial conditions (6.55) $\sigma_0^2 = 0.1$ initially, and 100000 trajectories. DOTTED LINE: erroneous positive P simulation with $\sigma_0^2 = 1$, and 10000 trajectories. SOLID LINE: gauge calculation for $\sigma_0^2 = 0.1$ with $\lambda = 4$, which corrects the systematic error of the positive P. Only 4000 trajectories, so as not to obscure other data.

---

of such a term or its multiple does not lead to any moving singularities for z_0 evolution either. One concludes that no new moving singularities or noise divergences are introduced by the gauge (6.60).

The results for an example have been plotted in Figure 6.6. The parameters there were $G = 1$, $Q = 0.25$ (leading to $a \approx 1.1124$ and $b \approx -0.1124$ in Stratonovich formulation). The initial condition in the example was $\sigma_0^2 = 0.1$, which is already much larger than the canonical variance for physically likely parameters in a photonic laser. Typical values of $|n_{\rm sc}|$ initially will be of order $\sigma_{n_{\rm sc}} \approx 0.14 \gtrsim |b|$ here. A good choice of gauge was found to be $\lambda = 4$. The use of this gauge restores the correct results.

6.5.5 Non-optimal initial conditions

As one increases the spread of the initial distribution beyond $\sigma_{n_{\rm sc}} \approx |b|$, it becomes increasingly difficult to find a gauge that will give reasonable simulations. For example a wide variety of what seemed like promising gauges for $\sigma_0^2 = 0.3$ have been

tried (with the same values of parameters Q and G as in Figure 6.6), and none have come close to success. The problem is that while systematic errors are removed, large random noise appears and obscures whatever one is trying to calculate.

Trajectories that start off at a value of $n'_{\rm sc}$ lying significantly beyond b require a lot of modification to their subsequent evolution to (1) stop them from escaping to $-\infty$ and (2) move them out of the gauged region of phase space so that they do not accumulate excessive weight noise. If there are many of these, the trade-off between the gauge size and length of time spent in the gauged region does not give much benefit anymore. Some consolation is gained in knowing that results will at worst be noisy and unusable, rather than being systematically incorrect.

6.6 Summary of main boundary term results

The two mechanisms that lead to boundary term errors have been specified and investigated for general phase-space distribution methods: 1) Errors of the first kind, which are caused by discarding nonzero boundary terms in the partial integration step between the master and Fokker-Planck equations, and 2) Errors of the second kind, which can arise if the estimator of a particular observable for a given kernel grows too rapidly as the boundaries of phase-space C are approached, and may be present or absent depending on the moment chosen.

The exact expression for the first kind of boundary terms has been found: (6.1), and the integrals that must converge for moment calculations to avoid the second kind are given by (6.8). While these cannot in general be explicitly evaluated for non-toy models because the exact form of the distribution $P(C)$ at the boundaries of phase space is required, a simple example has been solved for instructive purposes in Section 6.2.2, and these boundary term expressions explicitly evaluated. For the gauge P representation, conditions on $P(C)$ (6.7) and (6.24) are found, which are sufficient to avoid boundary term errors of the first and second kind (respectively). In particular, for the gauge P representation:

1. The simulation will be free of boundary terms of the first kind when: A) Kernel drift gauges are polynomial in the system variables or their conjugates, and B)

the far tails of $P(C)$ decay faster than exponential in the coherent amplitudes and faster than power law in Ω.

2. Estimators of observables polynomial in the $\widehat{a}_k$ and $\widehat{a}_k^\dagger$ operators are found to be free of boundary term errors of the second kind provided that the far tails of $P(C)$ decay faster than a power law.

This is consistent with the results found by Gilchrist *et al*[64]. No indication that such biases will occur for interacting Bose gas simulations is found in Section 6.2.4.

In Section 6.1.2, four symptoms of boundary term errors of the first kind are pointed out. These are moving singularities, noise divergences, discontinuities, and excessively broad initial conditions. All four can be checked for by inspection and some analysis of the stochastic equations and initial condition, before any simulation. Explicit conditions (6.3) have been found for simulations on open complex phase-spaces, which are sufficient to ensure that neither moving singularities nor noise divergences occur. It is also noted that by the Painleve conjecture, boundary term errors of the first kind are likely to be generically present in ungauged many-mode simulations based on analytic kernels.

All the above results have been collected here with the aim of systematizing and reducing the confusion on the topic of boundary term errors. Subsequently, Section 6.3.2 introduces a broadly-applicable heuristic method by which the symptoms of boundary term errors of the first kind can be removed using drift kernel gauges (and hence, presumably, also the boundary term errors as such). This removal method is applied to the two cases in the literature where systematic bias has been reported when using the positive P representation: two-boson absorption, and a single-mode laser model. It is found that the systematic bias is indeed removed from simulation results by this method for these examples.