

Quantum dynamics of long-range interacting systems using the positive- P and gauge- P representations

S. Wüster,^{1,2,3,*} J. F. Corney,⁴ J. M. Rost,¹ and P. Deuar⁵

¹Max Planck Institute for the Physics of Complex Systems, Nöthnitzer Strasse 38, 01187 Dresden, Germany

²Department of Physics, Bilkent University, Ankara 06800, Turkey

³Department of Physics, Indian Institute of Science Education and Research, Bhopal, Madhya Pradesh 462 023, India

⁴School of Mathematics and Physics, University of Queensland, Brisbane QLD 4072, Australia

⁵Institute of Physics, Polish Academy of Sciences, Aleja Lotników 32/46, Pl-02-668 Warsaw, Poland

(Received 12 March 2017; published 13 July 2017)

We provide the necessary framework for carrying out stochastic positive- P and gauge- P simulations of bosonic systems with long-range interactions. In these approaches, the quantum evolution is sampled by trajectories in phase space, allowing calculation of correlations without truncation of the Hilbert space or other approximations to the quantum state. The main drawback is that the simulation time is limited by noise arising from interactions. We show that the long-range character of these interactions does not further increase the limitations of these methods, in contrast to the situation for alternatives such as the density matrix renormalization group. Furthermore, stochastic gauge techniques can also successfully extend simulation times in the long-range-interaction case, by making use of parameters that affect the noise properties of trajectories, without affecting physical observables. We derive essential results that significantly aid the use of these methods: estimates of the available simulation time, optimized stochastic gauges, a general form of the characteristic stochastic variance, and adaptations for very large systems. Testing the performance of particular *drift* and *diffusion* gauges for nonlocal interactions, we find that, for small to medium systems, drift gauges are beneficial, whereas for sufficiently large systems, it is optimal to use only a diffusion gauge. The methods are illustrated with direct numerical simulations of interaction quenches in extended Bose-Hubbard lattice systems and the excitation of Rydberg states in a Bose-Einstein condensate, also without the need for the typical frozen gas approximation. We demonstrate that gauges can indeed lengthen the useful simulation time.

DOI: [10.1103/PhysRevE.96.013309](https://doi.org/10.1103/PhysRevE.96.013309)

I. INTRODUCTION

First-principles treatments of many-body quantum problems are notoriously difficult due to the exponential increase of Hilbert space dimension with the number of system components. Tackling this complexity is an outstanding goal of theoretical physics. Long-range interactions usually amplify the difficulties, because they can break symmetries or frustrate many of the tricks used to reduce the Hilbert space to manageable sizes. For example, the entanglement area laws can be broken [1].

If limited precision is acceptable, stochastic phase-space methods are a promising contender compared to other standard methods such as Monte Carlo or entangled pair states [2–4], especially for higher-dimensional systems. The computational complexity of phase-space methods tends to scale only linearly or quadratically in system size and to be largely independent of dimensionality. Originating from quantum optics [5], these methods have also been successful for cold degenerate gases [6–17] including fermions [18,19] and spin systems [20–22]. They are particularly suited to relatively dilute boson systems deep in the quantum regime.

Apart from a few small forays [4,23], previous work with phase-space methods has been limited to systems with contact interactions. However, long-range interactions have recently become important in the dilute quantum regime due to advances in the production of ultracold Rydberg atoms

[24,25], dipolar atomic gases [26,27], and molecular gases [28], the physics of which cannot be captured with a contact interaction alone. Motivated by this experimental progress, we develop here tools for the application of phase-space methods to systems with long-range interactions. These tools could then go substantially beyond the initial exploration of the positive- P method to a Rydberg system, reported in [29], which required a simplified interaction model.

The price paid for use of phase-space methods is that the dynamics is tractable usually only over short time scales. This limitation arises from the nonlinear amplification of stochastic fluctuations in individual trajectories, leading to a phase-space distribution that is too broad [30–33]. Hence, a central requirement in practice is to estimate the time limits and if possible to extend them. Fortunately, one can exploit the overcompleteness of the basis used to keep the distribution as compact as possible and to stabilize trajectories. Simulation schemes exploiting this feature have been termed gauge- P methods [10,11,23,31,34–38]. Two different types of gauges are available. *Drift* gauges modify the drift term of the underlying Fokker-Planck equation for the phase-space distribution, allowing, e.g., a modification of nonlinearities in the evolution equation at the expense of introducing a fluctuating trajectory weight. *Diffusion* gauges alter the noise terms in the stochastic evolution equations and can be used to vary the relative fraction of noise affecting the phase or amplitude of complex variables.

These stochastic gauges can become particularly powerful when it is possible to determine the best choice of gauge for a given simulation. For problems that rely essentially on single

*sebastian@iiserb.ac.in

bosonic modes, these best choices have been found [23,37]. For the many-mode case to date, optimization has only been considered insofar as a collection of single modes can offer some guidance, in situations where the interactions are weak or short range.

In this work, we adapt the prescriptions and approaches that are successful for short-range interactions to long-range interactions, and we show that they still give useful extensions of simulation time. To this end, we calculate the evolution of a characteristic variance that can serve to estimate the effect of a specific gauge, analogous to the approach of [37] for the single-mode case. Though quite an involved expression, it provides the essential basis for the application of gauge- P techniques to problems with long-range interactions. Estimates of the available simulation time and optimal adaptive gauge choices then follow in a relatively straightforward way.

The article is organized as follows: In Sec. II we describe the class of many-body quantum mechanical models that are being considered here. The general idea behind the positive- P and gauge- P methods is then briefly reviewed in Sec. III, and the evolution equations for nonlocal interactions derived. In Sec. IV, which contains the core of our analysis, we derive and analyze the characteristic noise variances and optimize gauges. Based on this, Sec. V presents estimates of the available simulation time.

All these results are then applied to two test cases in Sec. VI: an interaction quench in an extended (long-range) Bose-Hubbard model and the excitation dynamics of mobile Rydberg atoms within a Bose-Einstein condensate (BEC). These examples are oriented towards long-range interacting ultracold atomic physics, such as BECs of dipolar atoms [26], polar molecules [28], or Rydberg dressed condensates [39–47]. Following this, Sec. VII presents some further results that facilitate simulation of very large systems: a more efficient representation of the noise, and an optimization of pure diffusion gauges. In Sec. VIII, we report also on some initial explorations of more general nonlocal gauges as a starting point for future work. The Appendices contain various technical details, such as the Stratonovich corrections (Appendix A), the complete derivation of the characteristic variances (Appendices B and C), and supporting material on nonlocal noises (Appendix D) and gauges (Appendix E).

II. LONG-RANGE INTERACTING SYSTEMS

The class of Hamiltonian operators with which we work in this article is represented by

$$\hat{H} = \sum_{nm} \left[\hat{a}_n^\dagger \omega_{nm} \hat{a}_m + \frac{1}{2} \hat{a}_n^\dagger \hat{a}_m^\dagger W_{nm} \hat{a}_n \hat{a}_m \right]. \quad (1)$$

The indices label elements of the chosen single-particle basis $|m\rangle$ and the operator \hat{a}_m creates a particle in that basis state. Single-particle energies and linear mode couplings are included in ω_{nm} , while W_{nm} parametrizes generic two-body interactions. This potential W is assumed to be real here. For Hermiticity, $\omega_{nm} = \omega_{mn}^*$.

We now introduce two physical model systems, the Hamiltonian of which has the form (1). For these examples, we show simulation results in Sec. VI.

A. Extended Bose-Hubbard model

One Hamiltonian of the form Eq. (1) is that of an extended (long-range interacting) Bose-Hubbard model. In a cold atom context, $|m\rangle$ is then the localized, lowest-band, Wannier function at a single site of an optical lattice. The coupling $\omega_{nm} = J(\delta_{n+1,m} + \delta_{n-1,m})$ describes particle tunneling between neighboring lattice sites with amplitude J . Many cases of long-range interactions can be captured by the potential

$$W(|\mathbf{x} - \mathbf{y}|) = -\frac{C_6}{(|\mathbf{x} - \mathbf{y}|^a + \epsilon^a)^b}. \quad (2)$$

Possible experimental realizations include optical lattices filled with polar atoms [26] or Rydberg dressed atoms [48–51]. The W_{nm} in (1) can then be obtained from Eq. (2) by use of overlap integrals involving Wannier basis functions. However, here we simply use $W_{nm} = W(r)$, where $r = |\mathbf{x} - \mathbf{y}|$ and $\mathbf{x}_n, \mathbf{x}_m$ are discrete lattice points. We will later present simulation results for $a = 2$ and $b = 3$, corresponding to a softened van der Waals potential. Alternatively we could have used $a = 6$ and $b = 1$ to represent Rydberg-dressed potentials that include the cutoff due to blockade effects [40]. The contact part of the interaction plays a prominent mathematical role in the following, despite the presence of long-range forces. For a homogeneous interaction, we abbreviate it with

$$W_{nm} = W_0. \quad (3)$$

With the many-body model above, we will consider controlled interaction quenches to illustrate our results in Sec. VIA. These are readily realized with, e.g., Rydberg atom interactions due to the possibility of time-dependent external control.

B. Rydberg excitations in a Bose gas

As a second example, we will consider a Bose gas where atoms are continuously coherently excited from a ground state $|g\rangle$, in which interactions are neglected, to a Rydberg state $|e\rangle$ with strong long-range interactions between the atoms [52–57]. Initially all atoms are in the ground state. With more atoms getting excited, the so-called dipole blockade develops [52–54], where the simultaneous excitation of nearby atoms is strongly suppressed.

For this system we use a continuum formulation of the Hamiltonian (1). To this end, we define field operators via $\hat{a}_n = \hat{\Psi}(\mathbf{x}_n) \sqrt{dV}$, $\hat{a}_m = \hat{\Psi}(\mathbf{y}_m) \sqrt{dV}$, $W_{nm} = W(\mathbf{x}_n - \mathbf{y}_m)$, and $\omega_{nm} = -(\delta_{n+1,m} + \delta_{n-1,m} - 2\delta_{n,m}) / (2m dV^2)$. Here, $\hat{\Psi}(\mathbf{x})$ is a field operator creating an atom at a location \mathbf{x} , dV is the (physically small) volume corresponding to one discrete location \mathbf{x}_n on the numerical lattice. We set $\hbar = m = 1$ for simplicity, unless stated otherwise.

We then extend our model to include two fields describing atoms in components $|g\rangle$ and $|e\rangle$. Only the latter shall experience the long-range interactions. Rydberg excitation and dynamics often happen much faster than atomic motion, which justifies the so-called “frozen gas” approximation where atomic motion is neglected. With the coherent coupling amplitude κ between the components we then

have

$$\hat{H} = \frac{\kappa}{2} \int d\mathbf{x} (\hat{\Psi}_e^\dagger(\mathbf{x})\hat{\Psi}_g(\mathbf{x}) + \hat{\Psi}_g^\dagger(\mathbf{x})\hat{\Psi}_e(\mathbf{x})) + \frac{1}{2} \int d\mathbf{x} \int d\mathbf{y} \hat{\Psi}_e^\dagger(\mathbf{x})\hat{\Psi}_e^\dagger(\mathbf{y})W(\mathbf{x}-\mathbf{y})\hat{\Psi}_e(\mathbf{x})\hat{\Psi}_e(\mathbf{y}). \quad (4)$$

The crucial difference from the extended Bose-Hubbard model is that the number of atoms in the component experiencing the interactions is no longer conserved. We will assume the same interaction potential W as in the previous model (2).

For short times, Eq. (4) represents an adequate model (see, e.g., [58]). However, corrections due to atomic motion can easily become relevant [59–65], in which case the kinetic-energy term

$$\hat{H}_{\text{kin}} = - \int d\mathbf{x} \left(\hat{\Psi}_e^\dagger(\mathbf{x}) \frac{\nabla^2}{2m} \hat{\Psi}_e(\mathbf{x}) + \hat{\Psi}_g^\dagger(\mathbf{x}) \frac{\nabla^2}{2m} \hat{\Psi}_g(\mathbf{x}) \right) \quad (5)$$

must be reinstated to Eq. (4), precluding the use of a spin model. This reduces the range of methods available, but does not constitute a problem for phase-space methods.

With or without motion, we will later be interested in correlation function dynamics at the onset of a dipole blockade: Assume all atoms are initially condensed in the ground state. The coupling κ then acts for a given time, transferring population into the excited state, which has interactions between atoms. The interaction energy of atoms in the excited state can become large enough to suppress transition to that state after an initial one. As a result, there is only one atom excited per “blockade volume.” The Rydberg atom autocorrelation function $g^{(2)}$ has a pronounced dip for radii less than the blockade radius r_b and the frequency of incomplete Rabi oscillations increases to $\kappa_{mb} = \sqrt{N_b}\kappa$, where N_b is the number of ground state atoms within the blockade volume and κ is the single atom Rabi frequency [66].

III. GAUGE- P AND POSITIVE- P DESCRIPTIONS

If we consider M single particle modes $|m\rangle$, populated with up to N particles, the number of many-body basis states of our problem scales like M^N , an astronomical number for all but the smallest systems. The first-principles simulation of the Hamiltonian (1) is thus a severe challenge. We now briefly review how the dynamics can be tackled with the gauge- P distribution, by stochastically sampling the many-body state, rather than representing it exactly. For an in-depth explanation we refer to the extensive literature on the positive- P [2–5,7,10,12,13,30,32,67–71] and gauge- P methods [8,9,11,21,23,31,34,36–38,72–74].

A. Phase-space representation

First, we expand the system’s density matrix $\hat{\rho}$ in an overcomplete operator basis of coherent states:

$$\hat{\rho}(t) = \int d^{2M}\alpha \int d^{2M}\beta \int d^2\Omega P_G(\alpha, \beta, \Omega, t) \left[\Omega \frac{|\alpha\rangle\langle\beta^*|}{\langle\beta^*|\alpha\rangle} \right]. \quad (6)$$

The operator kernel in this expression is $\hat{\Lambda} = \Omega|\alpha\rangle\langle\beta^*|/\langle\beta^*|\alpha\rangle$, where we have made use of many-mode coherent states

[75] $|\alpha\rangle$ defined by $\hat{a}_n|\alpha\rangle = \alpha_n|\alpha\rangle$, with $\alpha_n \in \mathbb{C}$. The function P_G is the “gauge- P ” distribution, which can be chosen positive and real for any density matrix $\hat{\rho}$, similarly to the positive- P distribution $P_+(\alpha, \beta, t)$, which is the special case that occurs when the global weight Ω is chosen to be constant.

For unitary dynamics, on which we will focus here, the density matrix obeys the von Neumann equation

$$\frac{d}{dt}\hat{\rho} = -\frac{i}{\hbar}[\hat{H}, \hat{\rho}]. \quad (7)$$

Dissipative environments that would add, for example, Lindblad terms to Eq. (7) and produce a master equation can be straightforwardly included.

Due to the doubling of the phase-space dimension for quantum systems seen above, it will be convenient to introduce the following notation: We collect all $2M+1$ stochastic fields into a vector $\underline{v} = (\alpha, \beta, \Omega)$ whose components v_μ are labeled by greek indices μ, ν, \dots , as opposed to the M modes which are labeled by roman indices n, m, j, \dots . Continuing this distinction for vectors in general, boldface, underlined vectors will have $2M$ components [76]; without the underline they have M components. Similarly a double underline is used to distinguish $2M \times 2M$ matrices from the $M \times M$ matrices that they are usually composed of. Components of extended matrices have greek subscripts that run over $2M$ values.

For a sufficiently well bounded [77] distribution $P_G(\alpha, \beta, \Omega)$, the master equation can be converted into a Fokker-Planck equation (FPE) of the generic form,

$$\frac{\partial P_G}{\partial t} = - \sum_{\mu} \frac{\partial}{\partial v_{\mu}} A_{\mu} P_G + \frac{1}{2} \sum_{\mu\nu} \frac{\partial}{\partial v_{\mu}} \frac{\partial}{\partial v_{\nu}} D_{\mu\nu} P_G, \quad (8)$$

by virtue of the operator correspondences [23,78]:

$$\hat{a}_n \hat{\rho} \leftrightarrow \alpha_n P_G, \quad \hat{a}_n^\dagger \hat{\rho} \leftrightarrow \left(\beta_n - \frac{\partial}{\partial \alpha_n} \right) P_G, \quad (9a)$$

$$\hat{\rho} \hat{a}_n \leftrightarrow \left(\alpha_n - \frac{\partial}{\partial \beta_n} \right) P_G, \quad \hat{\rho} \hat{a}_n^\dagger \leftrightarrow \beta_n P_G, \quad (9b)$$

$$0 \leftrightarrow P_G + \frac{\partial}{\partial \Omega} \Omega P_G. \quad (9c)$$

Crucially, Eq. (8) is equivalent to the set of coupled stochastic differential equations (SDEs),

$$\frac{dv_{\mu}}{dt} = A_{\mu} + \sum_{\nu} B_{\mu\nu} \xi_{\nu}(t), \quad (10)$$

if the *diffusion matrix* \underline{D} is positive semidefinite [3,79,80]. These SDEs have to be interpreted in the Itô form of stochastic calculus [80]. The $\xi_{\nu}(t)$ are independent real noise fields, defined by a zero mean $\overline{\xi_{\nu}(t)}$ and correlations $\overline{\xi_{\mu}(t)\xi_{\nu}(t')} = \delta_{\mu,\nu}\delta(t-t')$, where \overline{f} denotes the stochastic average of f . They are usually implemented with independent Gaussian noises at each time step Δt with a variance of $1/(\Delta t dV)$. The “noise matrix” \underline{B} must satisfy $\underline{D} = \underline{B}\underline{B}^T$. The matrix

square root $\underline{D}^{1/2}$ is a viable noise matrix in most cases of interest, provided that it is self-transposed ($\underline{D}^{1/2} = [\underline{D}^{1/2}]^T$).

All quantum expectation values (observables) can be found in principle by averaging [81] over the set of independent stochastic trajectories. Expectation values of normal ordered operator products take the form [31]

$$\langle \hat{a}_n^\dagger \hat{a}_m^\dagger \hat{a}_l^\dagger \dots \hat{a}_p \hat{a}_q \hat{a}_r \dots \rangle = \frac{\Omega \beta_n \beta_m \beta_l \dots \alpha_p \alpha_q \alpha_r \dots + \Omega^* \alpha_n^* \alpha_m^* \alpha_l^* \dots \beta_p^* \beta_q^* \beta_r^* \dots}{\Omega + \Omega^*}. \quad (11)$$

The discussion above also applies to the positive- P distribution, in which case we can set $\Omega \rightarrow 1$ and simplify the calculation considerably [32].

Upon reaching Eq. (10), we have reduced the generally intractable quantum-many body problem in an N^M dimensional Hilbert space to the seemingly much easier solution of a system of $2M$ coupled complex SDEs. Whether the problem in this new form is tractable depends on the number of stochastic realizations required to obtain converged results for Eq. (11). Therein lies the fundamental limitation of stochastic phase-space methods: After a certain simulated time t_{sim} , the noise in the simulation can become nonlinearly amplified or the phase-space distribution too broad [30,31], which prevents convergence of the averages. As a consequence, these stochastic methods are only useful for understanding properties of the system that manifest themselves before t_{sim} . Fortunately t_{sim} is clearly heralded in simulation results by excessive sampling error, possibly in the form of spikes.

Fortunately, the above derivation of the equations of motion contains several mathematical degrees of freedom, that do not affect the simulated physics. These degrees of freedom are termed “gauges.” Under favorable conditions, they can be exploited to limit the growth of noise and obtain simulations that give results for longer times t_{sim} . The following options for gauges are available: First, we can use the replacement rule (9c) to add functions of our choice to the *drift vector* A_μ for the evolution of the α and β . These functions are the *drift gauges*. As a tradeoff, their use induces a stochastic evolution of the global trajectory weight Ω . Second, we can transform the noise matrix \underline{B} via $\underline{B}' = \underline{B} \underline{Q}$ with an orthogonal matrix \underline{Q} such that $\underline{Q} \underline{Q}^T = I$. If \underline{B} had fulfilled the diffusion condition $\underline{D} = \underline{B} \underline{B}^T$, then clearly so does \underline{B}' . An orthogonal matrix \underline{Q} can be written as $\underline{Q} = \exp \underline{G}$, where \underline{G} is antisymmetric ($\underline{G}^T = -\underline{G}$). It has been shown that the real part of \underline{G} merely induces inconsequential resummations of the noise, whereas the imaginary part redistributes the noise between amplitude and phase components of the stochastic fields and can have strong influence on the available simulation times [23,31,34,36]. The matrices \underline{G} or \underline{Q} are called *diffusion gauges*.

Different gauges yield different stochastic equations of motion, whose solutions have different convergence properties. Nonetheless, the evolution still corresponds to one and the same physical density matrix, and averages such as (11) with physical meaning remain unique. Thus gauges affect the

mathematical but not the physical properties of our problem, just as their namesake in electrodynamics.

B. Direct form of the equations of motion

Now we apply this formalism to the lattice system described by Hamiltonian (1). In the positive- P representation (without gauges), the diffusion matrix is block diagonal with separate blocks for the α and β variables [3,23], i.e.,

$$\underline{D} = \begin{pmatrix} D^{(\alpha)} & 0 \\ 0 & D^{(\beta)} \end{pmatrix},$$

$$D_{nm}^{(\alpha)} = -i W_{nm} \alpha_n \alpha_m, \quad (12)$$

$$D_{nm}^{(\beta)} = i W_{nm} \beta_n \beta_m.$$

Realistic two-body interactions are symmetric, $W_{nm} = W_{mn}$, which leads to symmetric roots $\sqrt{W} = W^{1/2} = [\sqrt{W}]^T$. Then, the simplest decomposition of the noise matrix \underline{D} gives the following equations:

$$\frac{d\alpha_n}{dt} = -i \sum_m [\omega_{nm} \alpha_m + \alpha_n W_{nm} \alpha_m \beta_m] - i \sqrt{i} \alpha_n \sum_j [\sqrt{W}]_{nj} \xi_j^{(1)}(t), \quad (13a)$$

$$\frac{d\beta_n}{dt} = i \sum_m [\omega_{nm} \beta_m + \beta_n W_{nm} \alpha_m \beta_m] + \sqrt{i} \beta_n \sum_j [\sqrt{W}]_{nj} \xi_j^{(2)}(t). \quad (13b)$$

The noises $\xi_j^{(1)}$ and $\xi_j^{(2)}$ are independent, and $[\sqrt{W}]_{nj}$ are components of the matrix square root \sqrt{W} of the interaction potential.

For the gauge- P representation, with stochastic gauge freedoms included, it will be convenient to use the matrix notation described at the beginning of Sec. III A to concisely deal with the 2×2 block nature of the problem. Vectors are assumed to be column vectors unless explicitly transposed, and we use the notation $\text{diag}[\mathbf{v}]$ to assemble a square diagonal matrix with vector \mathbf{v} on the diagonal. Following [18,82], we introduce the following:

$$\underline{\boldsymbol{\gamma}} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \quad \underline{\omega} = \begin{pmatrix} -\omega & 0 \\ 0 & \omega \end{pmatrix}, \quad \underline{W} = \begin{pmatrix} -W & 0 \\ 0 & W \end{pmatrix},$$

$$\underline{\mathbf{n}} = \begin{pmatrix} \mathbf{n} \\ \mathbf{n} \end{pmatrix} = \begin{pmatrix} (\dots, \alpha_j \beta_j, \dots)^T \\ (\dots, \alpha_j \beta_j, \dots)^T \end{pmatrix}, \quad \underline{\mathbf{A}} = i \underline{\omega} \underline{\boldsymbol{\gamma}} + i \underline{\Gamma} \underline{W} \underline{\mathbf{n}},$$

$$\underline{\boldsymbol{\xi}} = \begin{pmatrix} \boldsymbol{\xi}^{(1)} \\ \boldsymbol{\xi}^{(2)} \end{pmatrix}, \quad \underline{\Gamma} = \text{diag}(\underline{\boldsymbol{\gamma}}), \quad \underline{B} = \sqrt{i} \underline{\Gamma} \underline{S},$$

$$\underline{S} = \begin{pmatrix} -i\sqrt{W} & 0 \\ 0 & \sqrt{W} \end{pmatrix}. \quad (14)$$

A particularly useful quantity that will recur is the (complex) mode occupation

$$n_m = n'_m + i n''_m = \alpha_m \beta_m. \quad (15)$$

Using these definitions, we can write the gauge- P representation equations of motion

$$\frac{\partial}{\partial t} \underline{\boldsymbol{\gamma}} = \underline{\mathbf{A}} + \underline{\mathbf{B}}'(\underline{\boldsymbol{\xi}} - \underline{\mathbf{g}}), \quad \frac{\partial}{\partial t} \Omega = \Omega \underline{\mathbf{g}} \cdot \underline{\boldsymbol{\xi}}, \quad (16a)$$

$$\underline{\mathbf{B}}' = \underline{\mathbf{B}} \underline{\mathbf{O}} = \sqrt{i} \underline{\Gamma} \underline{\mathbf{S}} \underline{\mathbf{O}}, \quad \underline{\mathbf{g}} = \begin{pmatrix} \mathbf{g}^{(1)} \\ \mathbf{g}^{(2)} \end{pmatrix}, \quad (16b)$$

as per the formalism outlined in [10,11,23,31,34,36–38].

The *drift gauge* $\underline{\mathbf{g}}$ is a yet unspecified function of the variables $\underline{\boldsymbol{\gamma}}$. We also have already inserted an arbitrary orthogonal matrix $\underline{\mathbf{O}} = \exp[\underline{\mathbf{G}}]$ with $\underline{\mathbf{O}} \underline{\mathbf{O}}^T = \underline{\mathbf{1}}$, the *diffusion gauge*, into the diffusion term.

IV. OPTIMIZED GAUGES

For practical applications of the available stochastic gauges, we introduce more specific forms and then optimize them, similarly to how it was done for local interactions in [32,37].

A. Intuitive drift gauge form

Here, we introduce a specific form of $\underline{\mathbf{g}}$ and describe its effect on the equations of motion. The objective of $\underline{\mathbf{g}}$ is to control the tendency of the interaction term to create so-called moving instabilities and boundary term errors in the distribution P_G [23,31,36]. Let

$$g_\mu = \sqrt{i} \sum_{\nu\sigma} O_{\nu\mu} S_{\sigma\nu} f_\sigma, \quad (17)$$

where \mathbf{f} is a yet unspecified vector that becomes the gauge. In matrix notation this reads $\underline{\mathbf{g}} = \sqrt{i} \underline{\mathbf{O}}^T \underline{\mathbf{S}}^T \underline{\mathbf{f}}$. Using the properties of $\underline{\mathbf{O}}$ and $\underline{\mathbf{S}}$, such as $\underline{\mathbf{S}} \underline{\mathbf{S}}^T = \underline{\mathbf{S}} \underline{\mathbf{S}} = \underline{\mathbf{W}}$, one can see that the new equations of motion are

$$\begin{aligned} \frac{\partial}{\partial t} \gamma_\mu &= i \sum_\nu \omega_{\mu\nu} \gamma_\nu + i \sum_\nu \gamma_\mu W_{\mu\nu} (n_\nu - f_\nu) \\ &+ \sqrt{i} \sum_{\nu\sigma} \gamma_\mu S_{\mu\nu} O_{\nu\sigma} \xi_\sigma, \end{aligned} \quad (18a)$$

$$\frac{\partial}{\partial t} \Omega = \sqrt{i} \Omega \sum_{\nu\sigma\lambda} f_\lambda O_{\sigma\nu} S_{\lambda\sigma} \xi_\nu = \sqrt{i} \Omega \underline{\mathbf{f}}^T \underline{\mathbf{S}} \underline{\mathbf{O}} \underline{\boldsymbol{\xi}}. \quad (18b)$$

Using the following notation for the real and imaginary part of complex numbers: $z = z' + iz''$, we can now pick for instance $f_\lambda = in'_\lambda$ to remove the imaginary part of the stochastic density $\underline{\mathbf{n}}$ from the nonlinear interaction term of Eq. (16), or $f_\lambda = n_\lambda - |n_\lambda|$ to replace $\underline{\mathbf{n}}$ by its modulus. Both choices reduce the tendency of the nonlinear term in Eq. (16) to result in exponentially diverging trajectories [31].

In the following sections, we will make the choice

$$\underline{\mathbf{f}} = i \underline{\mathbf{n}}'' \quad (19)$$

for the drift gauge. This is a form that was found to be particularly useful for the dynamics of the contact-interacting gas in previous work, giving the best extensions of simulation time [37].

Most of the efficient numerical time-propagation schemes for stochastic differential equations require the use of the Stratonovich form. The necessary corrections to convert

Eq. (18) from the Itô to the Stratonovich form are described in Appendix A.

B. Ensemble spread and characteristic variance

To use the diffusion gauge beneficially, we have to know how a given choice for the matrix $\underline{\mathbf{O}}$ affects the noise induced spread of trajectories in the ensemble as time progresses. The trajectory spread determines the simulation time available. A characteristic variance $\mathcal{V}_1 = \frac{1}{2}(\text{var}[\ln|\Omega\alpha|] + \text{var}[\ln|\Omega\beta|])$ was used for the single mode case in the past to obtain quite accurate estimates of the simulation time t_{sim} [37]. The quantities $\Omega\gamma$ appearing in \mathcal{V}_1 are the estimators that appear in Eq. (11) for expectation values of the one-point correlation function. They are fundamental to most quantities that one might want to calculate. The logarithm is used in anticipation of an exponential increase in noise variances throughout the simulation, as is observed in practice. The first task here, and a keystone of subsequent analysis, is to generalize this quantity for our purposes.

A many mode generalization of \mathcal{V}_1 is

$$\mathcal{V} = \frac{1}{2M} \sum_\mu^{2M} \text{var}[\ln|\Omega\gamma_\mu|]. \quad (20)$$

The stochastic uncertainty of fields in all modes contributes to \mathcal{V} , hence this simple scalar quantity can signal when the simulation becomes too noisy to extract useful information via Eq. (11). Even if just the variance of γ_μ for a specific mode μ becomes too large, this will soon contaminate all the remaining modes through the linear coupling terms $\propto \omega$ in the equations of motion. Note that \mathcal{V} does not have any physical interpretation, and indeed must not, since it should be gauge dependent.

For the “standard” drift gauge (19), the approximate time evolution of Eq. (20) is calculated from Eq. (18), in Appendix B. The calculation neglects the kinetic $\omega_{\mu\nu}$ part of the drift, as was done for the two-mode case in [37], expecting the predominant contribution to the increase of \mathcal{V} to come from the noise terms and nonlinearities. The remaining terms conserve the expectation value of the local density $\langle n_m \rangle$, so that $\mathcal{V}(t)$ depends only on their initial values $n_m(0)$. The result is

$$\begin{aligned} \mathcal{V}(t) &= \mathcal{V}(0) + \frac{1}{2M} \left\{ \frac{t}{2} \text{Tr}[\underline{\mathbf{S}} \underline{\mathbf{O}} \underline{\mathbf{O}}^\dagger \underline{\mathbf{S}}^\dagger] \right. \\ &+ t \text{Tr}[\text{Im}\{\{\underline{\mathbf{S}} \underline{\mathbf{O}} \underline{\mathbf{O}}^\dagger \underline{\mathbf{S}}^\dagger\} \underline{\mathbf{N}}''\}] \\ &+ M \text{Tr}[\text{Re}\{\{\underline{\mathbf{S}} \underline{\mathbf{O}} \underline{\mathbf{O}}^\dagger \underline{\mathbf{S}}^\dagger\} \underline{\mathbf{Q}}\}] \left. \right\}, \end{aligned} \quad (21a)$$

$$\begin{aligned} Q_{\mu\nu} &= \frac{1}{2} \text{Re}[\overline{n_\mu(0)n_\nu^*(0)}(\exp[t P_{\mu\nu}] - 1 - t P_{\mu\nu})/P_{\mu\nu}] \\ &+ t \overline{[n''_\mu(0)n''_\nu(0)]} \end{aligned} \quad (21b)$$

$$P_{\mu\nu} = [\underline{\mathbf{F}} \underline{\mathbf{S}} \underline{\mathbf{O}} \underline{\mathbf{O}}^\dagger \underline{\mathbf{S}}^\dagger \underline{\mathbf{F}}]_{\mu\nu}. \quad (21c)$$

This is central to most of the rest of the paper. A useful $2M \times 2M$ matrix is

$$\underline{\mathbf{F}} = \begin{pmatrix} \mathbf{1} & \mathbf{1} \\ \mathbf{1} & \mathbf{1} \end{pmatrix}, \quad (22)$$

and $[\cdot\cdot\cdot]_{\mu\nu}$ denotes the component $\mu\nu$ of the matrix product within square brackets. The auxiliary diagonal matrix $\underline{\underline{N}}$ is

$$\underline{\underline{N}} = \underline{\underline{N}}' + i\underline{\underline{N}}'' = \text{diag}[\overline{\underline{\mathbf{n}}(0)}] \quad (23)$$

and has the initial density on the diagonal.

The first term of Eq. (21a) is the contribution from noise placed directly into the amplitudes $\underline{\mathbf{y}}$. The third term, involving $\underline{\underline{Q}}$, arises from noise accumulated in the weight Ω . The remaining term, involving $\underline{\underline{N}}''$, comes from an accumulation of noise in Ω due to any initial fluctuations in $\underline{\mathbf{n}}''(0)$ and is less important. For single-mode repulsive contact interactions we have $W_{nm} = (g/dV)\delta_{nm} > 0$, and thus Eq. (21a) reduces to the known results [37].

The variance Eq. (21) forms the essential basis for most other results of this article along with Eq. (33), and may additionally contain useful information for future work on stochastic gauges. The calculation is involved but indispensable to (i) determine simulation time limitations that help to assess whether the methods here are applicable to a given physical problem, (ii) assess whether gauges are expected to outperform the direct application of the (ungauged) positive- P method, (iii) operate diffusion gauges in analogy to the single-mode case, and (iv) possibly enable future work to fully harness the large number of gauge degrees of freedom available in the long-range interacting many-mode case.

C. Global diffusion gauge

To make use of Eq. (21) to choose a good gauge, we have to specify the form of the diffusion gauge matrix $\underline{\underline{Q}}$. In the single-mode case [34,37] available simulation times for the gauge- P method can be substantially improved using $\underline{\underline{Q}} = -\cosh(a)\sigma_3 - i\sinh(a)\sigma_1$, where a is an adjustable real *gauge parameter* and σ_i are the standard Pauli matrices. The underlying mechanism is a shift of the numerical noise from the amplitude of $n = \alpha\beta$ to its phase, which less directly leads to deleterious noise amplification. The simplest generalization to many modes is to use one global value of the parameter a , giving the *global diffusion gauge*

$$\begin{aligned} \underline{\underline{Q}} &= \begin{pmatrix} \cosh a \mathbb{1} & -i \sinh a \mathbb{1} \\ i \sinh a \mathbb{1} & \cosh a \mathbb{1} \end{pmatrix} = \exp(g), \\ g &= \begin{pmatrix} 0 & -ia \mathbb{1} \\ ia \mathbb{1} & 0 \end{pmatrix}. \end{aligned} \quad (24)$$

For a nonuniform system (in a trap, etc.), the obvious adaptation is to have a vector \mathbf{a} of gauge parameters with spatially varying values $a_n = a(\mathbf{x}_n)$. Such a gauge could most easily be implemented piecewise if the length scale on which the density varies is larger than the characteristic length scale of the potential W . Then each region would correspond to a block like Eq. (24), optimized using densities inside the block. A local gauge with arbitrarily varying $a(\mathbf{x})$ seriously complicates analysis, because there is an interplay between the length scales of the density variation and of the interaction. We will comment on several such more involved choices for the gauge in Sec. VIII.

Returning to the global choice Eq. (24), the quantities appearing in Eq. (21) simplify and are

$$\begin{aligned} \underline{\underline{Q}} \underline{\underline{Q}}^\dagger &= \begin{pmatrix} \mathbb{1} \cosh 2a & -i \mathbb{1} \sinh 2a \\ i \mathbb{1} \sinh 2a & \mathbb{1} \cosh 2a \end{pmatrix}, \\ \underline{\underline{P}} &= 2e^{-2a} \begin{pmatrix} U & U \\ U & U \end{pmatrix}. \end{aligned} \quad (25)$$

1. Noise minimization and gauge choice

Assuming the form Eq. (24), an optimal parameter a can be found by minimizing $\mathcal{V}(t_{\text{opt}})$, where t_{opt} is the chosen target time, typically the longest time for which simulation results are required [37]. The analytical minimization of Eq. (21) is only tractable with several simplifications. We proceed in a similar fashion as was done in [37] for a single-mode case. We expand the variance Eq. (21) to second order in t :

$$\begin{aligned} \mathcal{V}(t) &= \mathcal{V}(0) + \frac{t}{2} U_0 \cosh(2a) + t e^{-2a} I_1 + \frac{t^2}{2} e^{-4a} I_2, \\ I_1 &\equiv \sum_{kk'} U_{kk'} \overline{n_k''(0) n_{k'}''(0)} \approx \int d\mathbf{x} d\mathbf{y} \overline{\rho_0''(\mathbf{x}) \rho_0''(\mathbf{y})} U(\mathbf{x} - \mathbf{y}), \\ I_2 &\equiv \sum_{kk'} U_{kk'}^2 \text{Re}[\overline{n_k(0) n_{k'}^*(0)}] \\ &\approx \int d\mathbf{x} d\mathbf{y} \overline{\rho_0(\mathbf{x}) \rho_0^*(\mathbf{y})} U(\mathbf{x} - \mathbf{y})^2, \end{aligned} \quad (26)$$

where, for the continuum version of the integrals, we define the complex density $\rho_0(\mathbf{x}) = \rho_0'(\mathbf{x}) + i\rho_0''(\mathbf{x}) = n(\mathbf{x}_n)/dV$. Here, we have also introduced a semipositive definite ‘‘rectified interaction’’ matrix

$$U = \sqrt{W} \sqrt{W}^\dagger \quad (27)$$

that plays an important role and will often reappear later. In general, its diagonal elements $U_0 = U_{nn}$ are not equal to those of W which are W_0 , but they can be related, and are always positive real [83]. The potential Eq. (2) used for the examples in Sec. VI *usually* has all eigenvalues negative though not always, particularly if periodic boundary conditions occur.

It is instructive to consider the limit of contact interactions, defining their strength g via $U_0 \rightarrow |g|/dV$. The integrals I_j in (26) then become

$$\begin{aligned} I_1 &\rightarrow M U_0 \overline{[n''(0)]^2} \sim |g| (\rho_0'')^2 V, \\ \text{and } I_2 &\rightarrow M U_0^2 \overline{|n(0)|^2} \sim \left[\frac{g^2 |\rho_0|^2}{dV} \right] V. \end{aligned} \quad (28)$$

We see that the I_j are extensive quantities, growing with system size, whether measured by mode count M or volume V .

We now wish to minimize $\mathcal{V}(t_{\text{opt}})$ at a given time t_{opt} . Following [37], we obtain analytically tractable results in two limits: *weight dominated*,

$$a_{\text{opt}} \approx \frac{1}{6} \ln \left(\frac{4t_{\text{opt}} I_2}{U_0} \right), \quad (29)$$

when the I_1 term is small, i.e., n_0 and t_{opt} are large and *direct noise dominated*,

$$a_{\text{opt}} \approx \frac{1}{4} \ln \left(1 + \frac{4I_1}{U_0} \right), \quad (30)$$

when the I_2 term is small, i.e., n_0 or t_{opt} is small. Then we follow the same interpolation procedure that was successful in [37] to give an overall best estimate

$$a_{\text{approx}} = \frac{1}{6} \ln \left(\frac{4I_2 t_{\text{opt}}}{U_0} + \left[1 + \frac{4I_1}{U_0} \right]^{3/2} \right). \quad (31)$$

2. Adaptive diffusion gauge

Single mode work has shown that the performance can sometimes be improved further, by adapting the gauge parameter a in a time-dependent manner during the simulation, leading to an $a = a(t)$. In that case, one determines $a(t)$ at each time step according to Eq. (31) using $t_{\text{opt}} = t_{\text{fin}} - t$, where t_{fin} is the desired final time of the simulation and the values $n_0(\mathbf{x})$ appearing in Eq. (31) are changed to the time-dependent $n(\mathbf{x}, t)$. We then obtain

$$a_{\text{adaptive}} = \frac{1}{6} \ln \left(\frac{4I_2[n(t)](t_{\text{fin}} - t)}{U_0} + \left[1 + \frac{4I_1[n(t)]}{U_0} \right]^{3/2} \right). \quad (32)$$

The Stratonovich correction Eq. (A1) becomes more involved when a depends on the phase-space variables γ_μ ; details of the derivation are given in Appendix A 2.

V. AVAILABLE SIMULATION TIMES

Before attempting to simulate physics with the methods presented here, the crucial question is whether one expects the physical effects of interest to take place prior to materialize before the time t_{sim} when the noise level overwhelms the simulation. As in earlier work [37], we define this moment via $\mathcal{V}(t_{\text{sim}}) = 10$. Based on the optimized gauges from the previous section, we will determine t_{sim} for the gauge- P method in this section.

Gauges do not always offer an advantage over the technically much simpler plain positive- P method. In some other cases the use of diffusion gauges without drift gauge is preferable. To allow an *a priori* assessment, we will also derive t_{sim} for those methods. This requires a rederivation of the characteristic variance for those cases, similar to Sec. IV B, which is presented in the following.

A. Positive- P and diffusion gauged characteristic variance

Let us consider the case without drift gauges ($\underline{\mathbf{g}} = \underline{\mathbf{f}} = 0$), which includes the plain positive- P representation as a special case ($\underline{\mathbf{Q}} = \underline{\mathbf{I}}$), and will also be useful in Sec. VI B 1. The approximate time evolution of \mathcal{V} is calculated similarly to Eq. (21) from Eq. (16), taking $\Omega = 1$ and $\underline{\mathbf{f}} = 0$. Details can be found in Appendix C. We obtain

$$\begin{aligned} \mathcal{V}^{(P)}(t) = \mathcal{V}(0) + \frac{1}{2M} \left\{ \frac{t}{2} \text{Tr}[\underline{\underline{S}} \underline{\underline{O}} \underline{\underline{O}}^\dagger \underline{\underline{S}}^\dagger] + \text{Tr}[\underline{\underline{W}} \underline{\underline{Q}} \underline{\underline{W}}] + t^2 \text{Tr}[\underline{\underline{W}} \underline{\underline{C}}^{(0)} \underline{\underline{W}}] \right. \\ \left. - \frac{t^2}{2} \text{Tr}[\underline{\underline{W}} \underline{\underline{N}}' \underline{\underline{F}} \underline{\underline{W}}] - \text{Im}(\underline{\underline{S}} \underline{\underline{O}} \underline{\underline{O}}^\dagger \underline{\underline{S}}^\dagger \underline{\underline{N}}'^*) \underline{\underline{F}} \underline{\underline{W}}] - 2t \text{Tr}[\underline{\underline{W}} \underline{\underline{C}}^{(0)}] \right\} \end{aligned} \quad (33a)$$

$$\tilde{\mathbf{Q}}_{\mu\nu} = \text{Re} \left[\frac{\overline{n_\mu(0)n_\nu^*(0)}}{P_{\mu\nu}^2} \left(\exp[t P_{\mu\nu}] - 1 - t P_{\mu\nu} - \frac{t^2 P_{\mu\nu}^2}{2} \right) / P_{\mu\nu}^2 \right]. \quad (33b)$$

In the above very general case, two initial state covariances appear,

$$C_{\mu\nu}^{(0)} = \text{covar}[n_\mu''(0), n_\nu''(0)], \quad \text{and} \quad \tilde{C}_{\mu\nu}^{(0)} = \text{covar}[n_\mu''(0), \ln|\gamma_\nu(0)|], \quad (34)$$

which are, however, zero in most cases.

The first term of Eq. (33a) is again the direct noise contribution placed into the amplitudes $\underline{\mathbf{y}}$. The next term involving $\underline{\underline{Q}}$ is the main noise amplification, due to the nonlinearity working on the direct noise. The term involving $\underline{\underline{N}}'$ is a cross correlation between these, while the remaining terms are due to amplification of any initial fluctuations in $\mathbf{n}''(0)$.

For the case of plain positive- P with $\underline{\underline{Q}} = \underline{\underline{I}}$, Eq. (33) reduces to the following:

$$\begin{aligned} \mathcal{V}^{(P)}(t) = \mathcal{V}(0) + \frac{1}{2M} \left\{ t \text{Tr}[U] - t^2 \text{Tr}[W N' W] + 2t^2 \text{Tr}[W C^{(0)} W] - 2t \text{Tr}[W \tilde{C}^{(0)}] \right. \\ \left. + \sum_{mkk'} W_{mk} \text{Re} \left[\frac{\overline{n_k(0)n_{k'}^*(0)}}{2} [e^{2U_{kk'}t} - 1 - 2U_{kk'}t - 2(U_{kk'}t)^2] \right] W_{k'm} / U_{kk'}^2 \right\}. \end{aligned} \quad (35)$$

We have used $M \times M$ matrices here.

The initial covariances in (35) are

$$C_{kk'}^{(0)} = \text{covar}[n_k''(0), n_{k'}''(0)], \quad \text{and} \quad \tilde{C}_{kk'}^{(0)} = \text{covar} \left[n_k''(0), \left(\ln \left| \frac{\beta_{k'}(0)}{\alpha_{k'}(0)} \right| \right) \right]. \quad (36)$$

The known single-mode results [32,37] for repulsive interaction are recovered from Eq. (35) with the replacement $M = 1$, $W = (g/dV)\delta_{nm} > 0$, and $\text{var}[n''(0)] = 0$ [84].

B. Useful simulation time (positive P)

As has been justified in detail in [32], a characteristic variance of $\mathcal{V} \approx 10$ sets the ultimate limit of usefulness in practice. It gives the variance of the *logarithm* of a typical observable estimator, and exponentials of Gaussian random variables acquire intractably long distribution tails once this variance reaches a value of $\mathcal{O}(10)$. The observable estimators are largely of such a form.

Analytical results are only easily obtainable with several simplifications. We proceed in a similar fashion as was done in [32] for a single-mode case. First, we assume deterministic initial conditions so that the initial covariances $C^{(0)}$ and $\tilde{C}^{(0)}$ can be neglected. Then, we expand the variance Eq. (35) to third order in t . Unlike for (26), here this is the first order that includes the important noise amplification contribution:

$$\begin{aligned} \mathcal{V}^{(P)}(t) - \mathcal{V}(0) &= +\frac{t}{2}U_0 - \frac{t^2}{2}I_1^{(P)} + \frac{t^3}{3}I_2^{(P)}, \\ I_1^{(P)} &\equiv \frac{1}{M} \sum_{kk'} W_{kk'}^2 \overline{n'_k} \approx \frac{1}{V} \int d\mathbf{x} d\mathbf{y} \overline{\rho'_0(\mathbf{x})} W(\mathbf{x}-\mathbf{y}), \\ I_2^{(P)} &\equiv \frac{1}{M} \sum_{kk'k''} U_{kk'} W_{kk''} W_{k'k''} \text{Re}[\overline{n_k(0)n_{k'}^*(0)}] \\ &\approx \frac{1}{V} \int d\mathbf{x} d\mathbf{y} d\mathbf{z} \overline{\rho_0(\mathbf{x})\rho_0^*(\mathbf{y})} \\ &\quad \times U(\mathbf{x}-\mathbf{y})W(\mathbf{x}-\mathbf{z})W(\mathbf{y}-\mathbf{z}). \end{aligned} \quad (37)$$

The rectified interaction U and its diagonal value $U_0 = U_{jj} \geq 0$ appear again.

To get an idea of the order of magnitude of these quantities, note that in the limit of contact interactions (as before, strength g , $U_0 \rightarrow |g|/dV$, and density ρ_0) we have

$$\begin{aligned} I_1^{(P)} &\rightarrow W_0^2 \overline{(n'(0))} \sim \frac{g^2 \rho'_0}{dV}, \\ \text{and } I_2^{(P)} &\rightarrow U_0 W_0^2 \overline{|n(0)|^2} \sim \frac{|g|^3 |\rho_0|^2}{dV}. \end{aligned} \quad (38)$$

The integrals I_j are thus intensive quantities, independent of system volume, unlike those encountered in the corresponding gauge- P derivation (26). However, they are dependent on the numerical lattice spacing dV , growing as lattice spacing drops. This is a manifestation of the known feature that simulation time improves as the lattice becomes more coarse.

Consider now the useful case of coherent state (nonentangled) initial conditions, with $n''(0) = 0$ and no starting variance [$\mathcal{V}(0) = 0$]. We seek to determine when $\mathcal{V}^{(P)} = 10$ is reached.

The most important regime is when the last term in Eq. (37), that comes from noise amplification, dominates the variance. By using the estimates Eq. (38) to make analogy with the contact-interacting case we can get some bearings. This regime is seen to occur when

$$t|g|\rho_0 \gtrsim \mathcal{O}(1), \quad \text{i.e., } t \gtrsim \frac{1}{|\mu|}, \quad (39)$$

for chemical potentials $\mu \approx g\rho$ that are typical for the superfluid regime of ultracold gases. So, requiring $\mathcal{V}^{(P)} = 10$

in this regime, one readily obtains

$$\begin{aligned} t_{\text{sim}} &\approx \frac{3}{[I_2^{(P)}]^{1/3}} \approx 3V^{1/3} \left[\int d\mathbf{x} d\mathbf{y} d\mathbf{z} \overline{\rho_0(\mathbf{x})\rho_0^*(\mathbf{y})} \right. \\ &\quad \left. \times U(\mathbf{x}-\mathbf{y})W(\mathbf{x}-\mathbf{z})W(\mathbf{y}-\mathbf{z}) \right]^{-1/3}. \end{aligned} \quad (40)$$

This is a remarkably simple expression. It is roughly independent of the system size, and reduces to the known single-mode result $t_{\sim} \approx 2.5dV^{1/3}/[g\rho_0^{2/3}]$ in that limit [32].

The opposite regime of short times or small chemical potential has evolution still dominated by the direct noise. Then the first term in Eq. (37) dominates and we have

$$t_{\text{sim}} \approx \frac{20}{U_0}. \quad (41)$$

The situation with dominant $I_1^{(P)}$ term does not cover a significant range of parameters.

C. Useful simulation time (gauge P)

With an optimal gauge chosen as discussed in Sec. IV, we can proceed to further estimate the useful simulation time t_{sim} for the gauge- P representation. We follow the same approach as in Sec. VB and [37]. We take again coherent state initial conditions [$n''(0) = 0$], and $\mathcal{V}(0) = 0$. Then, I_1 from Eq. (26) vanishes, and we substitute the optimal gauges in the two regimes (29) and (30) into Eq. (26) and impose the condition $\mathcal{V}(t_{\text{sim}}) = 10$.

In the weight-dominated regime, which applies for large systems or not-too-small densities, one obtains the important result that

$$t_{\text{sim}} \approx \frac{8}{\sqrt{U_0}\sqrt{I_2}} \quad (42)$$

(assuming also $\cosh 2a \approx e^{2a}/2$). This reduces to the results for the single-mode contact interacting case. In particular, for M contact-interacting modes,

$$t_{\text{sim}} \sim \frac{8\sqrt{dV}}{g\sqrt{\rho_0}} \times \frac{1}{M^{1/4}}. \quad (43)$$

In the direct-noise dominated regime that is relevant for short times or small occupations,

$$t_{\text{sim}} \approx \frac{20}{U_0}, \quad (44)$$

which roughly applies when it is shorter than Eq. (42), i.e., for $I_2 \lesssim U_0^2/40$. Notably, it is the same as Eq. (41) for positive P . This criterion is usually only met for very small occupations n_0 .

VI. EXAMPLE SIMULATIONS

We now proceed to demonstrate the utility of these methods with some exemplary applications. The physical models to which we will apply the method have been presented in Sec. II. For both examples, we consider for simplicity a one-dimensional (1D) system with periodic boundary conditions.

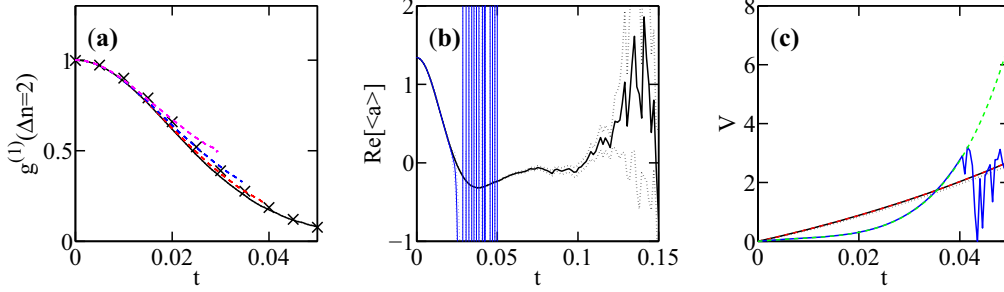


FIG. 1. (a) Dephasing of intersite coherence after a sudden interaction quench in the extended Bose-Hubbard model, using the gauge- P method: $J = 0$ (solid black), $J = 2$ (dashed red), $J = 5$ (dashed black), $J = 10$ (dashed magenta). All lines terminate at those times where simulations become intractable. (\times) Comparison with the analytical result for $J = 0$. (b) The on-site mean field $\langle a \rangle$ for $J = 0$ from gauge- P is shown as solid black line, dotted lines indicating the sampling error. For comparison the blue line shows a corresponding positive- P simulation. Both use 10^5 trajectories. (c) Characteristic variance \mathcal{V} directly from the $J = 0$ simulation, gauge P for a fixed $a = 0.76$ (solid black), gauge P with adaptive diffusion gauge as in Eq. (31) (dotted black), positive P (blue). The corresponding analytical expression Eq. (21) (dashed red) and Eq. (35) (dashed green) give the exact variance since they are based on the complete Hamiltonian for this example. Note the different time axis from (b).

A. Extended Bose-Hubbard model

As promised in Sec. II A, we first consider long-range interactions that are suddenly switched on (interaction quench) in an extended Bose-Hubbard model. This is realistic for engineered interactions in cold atom systems, e.g., with Rydberg dressing. For an initial coherent state, interactions will dephase the different Fock-state components, and coherence between neighboring sites will be lost. At much longer time scales, the coherence may revive due to the limited number of relevant frequencies involved. Such collapse-and-revival sequences are well understood in the case of local interactions [85,86].

Since the expression for the characteristic variance (21) was derived for $\omega_{nm} = 0$, we shall consider $J = 0$ first. We later demonstrate that for small nonzero J the tools derived from (21) still offer useful guidance. For $J = 0$, a state $|\mathbf{n}\rangle = |n_1, n_2, n_3, \dots\rangle$ with exactly n_i bosons on site i is an eigenstate of the Hamiltonian and the time evolution can be found analytically. Let us consider the case where the initial quantum state is a uniform coherent state on each site such as for a Bose-Einstein condensed gas $|\Psi(t=0)\rangle = |\phi, \phi, \dots, \phi\rangle$, where $\phi \in \mathbb{C}$ is the BEC order parameter and $|\phi\rangle = \exp(-|\phi|^2/2) \sum_n \phi^n |n\rangle / \sqrt{n!}$. Thus $|\Psi(t=0)\rangle = \sum_n C_n |\mathbf{n}\rangle$, with $C_n = \exp(-|\phi \cdot \phi|/2) \phi^{\sum n_i} / \sqrt{\prod n_i!}$. The state approximates the true state of a BEC in a lattice deeply in the superfluid regime. For our demonstration we compare gauge- P simulations for $J = 0$ with the analytical results and then consider various values of J in gauge P only.

In Fig. 1 we show the evolution of the intersite coherence, $g^{(1)}(\Delta n) = \sum_n \langle a_n^\dagger a_{n+\Delta n} \rangle / \sqrt{n_n n_{n+\Delta n}}$, the on-site mean field $\langle \hat{a}_n \rangle$, and the characteristic variance \mathcal{V} for an interaction quench on $M = 6$ sites spread on a periodic 1D interval of length $L = 2$. The potential Eq. (2) is described by $C_6 = -32$, $\epsilon = 1$, $a = 2$, $b = 3$ and the initial state is $|\phi|^2 = \rho dV = 1.2$. J is varied over $J = 0, 2, 5, 10$.

To know in advance up to which times we can simulate, we determine I_2 from Eq. (26) using the potential Eq. (2) [87]. Due to the small number of modes in this example, replacing the integral in I_2 by the corresponding discrete sum is more accurate, yielding $t < t_{\text{sim}} = 0.14$ using Eq. (42). A comparison with Fig. 1(b) shows that this indeed describes the

available time for the observable $\langle a_n \rangle$ quite well, for which the stochastic average is structurally similar to the characteristic variance (20) underlying the estimate for t_{sim} . In contrast, the sampling of the intersite coherence $g^{(1)}$ becomes intractable some time before t_{sim} , around the maximal time shown in panel (a). This is common for higher order quantities. Being interested in $g^{(1)}$, we thus have chosen $t_{\text{opt}} = 0.05$ in the formula Eq. (31), which then gives $a = 0.7$.

Finally, panel (c) numerically verifies Eq. (21). It also shows that for the present model with conserved mean occupation of all the sites, the adaptive diffusion gauge $a(t, n)$ yields no significant improvement of the variance compared to a fixed gauge a . The probable reason is the constant mean mode occupation for this example. We expect that there exist cases with time-varying occupation where the adaptive gauge outperforms the fixed one.

Using the ungauged positive- P method we can only get results for a much shorter time as shown in Fig. 1(b). The estimate Eq. (40) $t \lesssim 0.07$ is of the right magnitude, but a little high.

B. Coulomb blockade

We now consider the excitation of atoms within a homogeneous 1D BEC to long-range interacting Rydberg states, as was outlined in Sec. II B. Establishing an interaction blockade within the excited state fraction will correspond to nontrivial correlation function dynamics. To see this within the accessible time interval up to t_{sim} , let us consider an echo sequence as used in [57,88–90]: Atoms are excited from the ground to the Rydberg state with a Rabi frequency κ during a time interval $\tau/2$, then de-excited again using $\kappa \rightarrow -\kappa$ for a time $\tau/2$. This yields an additional time scale τ which can be chosen such that if it is short enough, it can fit within t_{sim} . During this time, we are then interested in the time evolution of the Rydberg-Rydberg correlation function

$$g^{(2)}(r) = \frac{\langle \Psi_e^\dagger(x) \Psi_e^\dagger(x+r) \Psi_e(x+r) \Psi_e(x) \rangle}{\rho_e(x) \rho_e(x+r)}, \quad (45)$$

with $\rho_e(x) = \langle \Psi_e^\dagger(x) \Psi_e(x) \rangle$. Due to our homogeneous initial conditions, we can additionally average the expression (45)

over x . The evolution equations are Eq. (18) with the drift gauge Eq. (19) and global diffusion gauge Eq. (24), adapted for two fields and the coupling κ in Eq. (4). Namely,

$$\begin{aligned} \frac{\partial}{\partial t} \gamma_{e,\mu} = & i \sum_{\nu} \omega_{\mu\nu} \gamma_{e,\nu} \mp i \frac{\kappa}{2} \gamma_{g,\mu} + i \sum_{\nu} \gamma_{e,\mu} W_{\mu\nu} n'_{e,\nu} \\ & + \sqrt{i} \sum_{\nu\sigma} \gamma_{e,\mu} S_{\mu\nu} O_{\nu\sigma} \xi_{\sigma}, \end{aligned} \quad (46a)$$

$$\frac{\partial}{\partial t} \gamma_{g,\mu} = i \sum_{\nu} \omega_{\mu\nu} \gamma_{g,\nu} \mp i \frac{\kappa}{2} \gamma_{e,\mu}, \quad (46b)$$

$$\frac{\partial}{\partial t} \Omega = i \sqrt{i} \Omega \sum_{\nu\sigma\lambda} n''_{e,\lambda} O_{\sigma\nu} S_{\lambda\sigma} \xi_{\nu} \quad (46c)$$

with $\hbar = m = 1$ and kinetic motion, if present, given by $\omega = \frac{1}{2m} \nabla^2$. The sign \mp refers to $-$ for α and $+$ for β variables.

It would also be interesting to follow the entire development of the blockade dip for larger times, and many-body Rabi oscillations in the saturated regime. However, we found the time scale for development of a complete blockade out of reach for all physical and gauge parameters tested here.

1. Echo sequence

Without interactions, after completing the echo sequence at time τ , all the atoms would have returned to the ground state. With interactions, in contrast, the induced dephasing makes this reversal incomplete. We model this scenario with parameters $\tau = 0.18, \kappa = 3, C_6 = -5.96 \times 10^7, \epsilon = 12.5$ and initially $N = 500$ atoms in state $|g\rangle$ between $x = -L$ and $x = L$ for $L = 50$. These parameters are chosen for demonstration,

to yield an excitation blockade that only allows a few atoms in state $|e\rangle$ within the domain $[-L, L]$. Figure 2 shows some of the results. Except where indicated, we invoke the frozen gas approximation, in which the ω terms in Eq. (46) are neglected.

We find that, counterintuitively, the remnant excited state population after a portion of the de-excitation pulse develops *bunching* correlations as seen in Fig. 2(e). We confirmed this behavior using the ω expansion [91,92], which allows a qualitative calculation of the excited state population N_e and correlations $g^{(2)}$ as an expansion in ωt . The leading order results, shown in Fig. 2 for comparison, reveal qualitative agreement with the exact gauge P with some quantitative differences. Preliminary calculations of the kind shown in Fig. 2 were the first sign of bunching correlations after an echo sequence discovered by us. This prompted our more detailed investigation using exact diagonalization for small Rydberg ensembles, published in [93] and experimentally confirmed in [94].

A good *a priori* estimate of gauge parameters is not so straightforward here like in the last section, because the guiding expressions that we have derived depend on the density of the interacting species $\rho_e = N_e/2L$, which is strongly time dependent as seen in Fig. 2(a). First, using $N_e \approx 8$ from Fig. 2(a) as a worst-case estimate, we obtain for the positive- P method $t_{\text{sim}} = 0.21$ using Eq. (40), while for gauge P we find $t_{\text{sim}} = 0.32$ using Eq. (42). These are roughly consistent with an empirical simulation time of $t_{\text{sim}} = 0.2$, but the numerics do not bear out the predicted modest advantage of the gauge P here. We also note that the optimum fixed gauge found empirically, $a = 0.55$, is somewhat smaller than that predicted by Eq. (31), which is $a_{\text{approx}} = 0.66$ for $t_{\text{opt}} = 0.135$. These

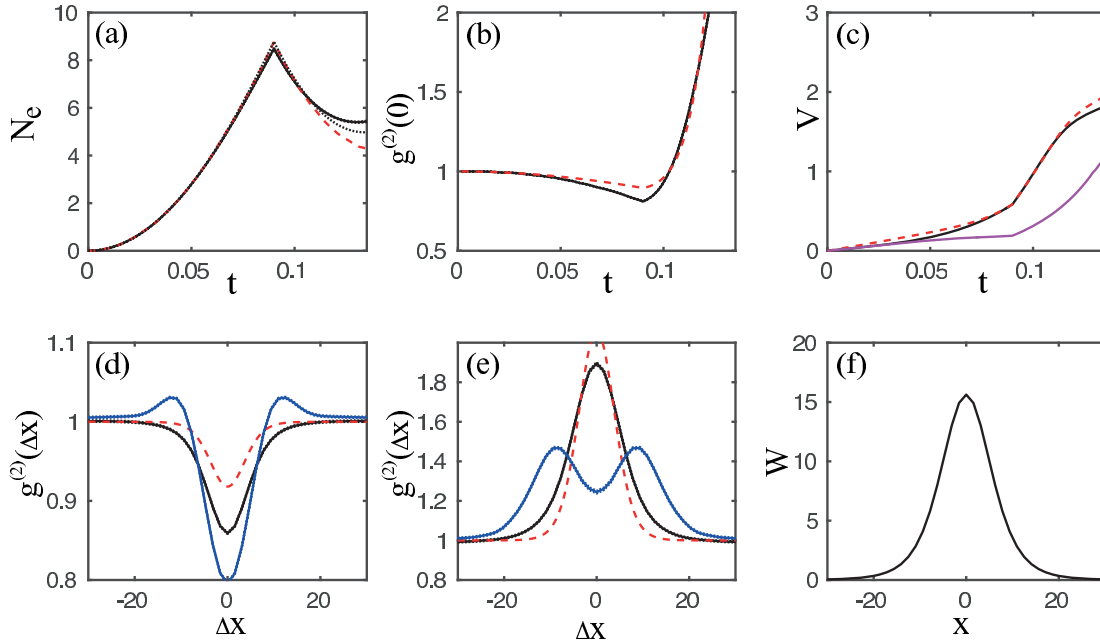


FIG. 2. Rydberg excitation echo sequence as described in the text. (a) Excited state population N_e : gauge P with adaptive gauge (solid black), mean-field (dotted black), deviating only at the latest times, ω -expansion (dashed red). (b) Rydberg-Rydberg correlation function at the origin $g^{(2)}(0,t)$, lines as in (a). (c) Characteristic variance \mathcal{V} , adaptive gauge Eq. (31) (solid black), fixed gauge $a = 0.55$ (dashed red), positive- P method (solid magenta). (d) Spatial correlations $g^{(2)}(\Delta x, t)$ at $t = 0.08$ and (e) $t = 0.12$, lines as in (a), additionally, gauge P with atomic motion (solid blue), i.e., $\omega = \frac{1}{2m} \nabla^2$. (f) Interaction potential Eq. (2).

differences do not come as a complete surprise, since our calculations neglect the strong time dependence of n_e .

While the frozen Rydberg gas dynamics presented so far can be alternatively calculated using exact diagonalization, the quantum mechanical inclusion of motion typically renders that approach intractable, and would require the use of classical or quantum-classical methods [59–65,95–97]. Motion however poses no additional challenge to the gauge- P method, which we now illustrate by dropping the frozen gas approximation and considering atomic motion as outlined in Sec. II B, with the full Eq. (46). Simulated echo sequences including motion are added to Fig. 2. We used an exemplary mass of $m = 4 \times 10^{-3}$, chosen to produce a visible effect but not too short t_{sim} . As one can see, the change is significant.

There are many scenarios in which the frozen Rydberg gas approximation is not applicable such as [59–65,95–97]. Other possible realizations of our long-range interacting model with relevant atomic motion could be a two component polar BEC, where the long-range interactions between one of these components are dramatically enhanced with a Feshbach resonance [26,98] or a Rydberg dressed BEC [48–51].

VII. METHODS FOR VERY LARGE SYSTEMS

The sets of stochastic differential equations, Eqs. (13) and (16a), are usually quite straightforward to implement for small and medium size systems up to $M \sim \mathcal{O}(10-1000)$. However, for a larger system two problems arise: the direct numerical implementation of the equations as written may become too inefficient, and the drift gauge may cease to be beneficial. We discuss both of these issues, and possible solutions, below.

A. A more efficient treatment of the interaction for large systems

For larger systems, the direct application of Eqs. (13) and (16a) by summing over $m = 1, \dots, M$ becomes a problem for two reasons: (i) integrating a single time step takes $\mathcal{O}(M^2)$ operations (a separate summation of $\sum_m W_{nm} \alpha_m \beta_m$ and $\sum_j [\sqrt{W}]_{nj} \xi_j$ for each lattice site n), and (ii) if M gets really large, e.g., of the order of $M \gtrsim \mathcal{O}(10^5)$ in a three-dimensional system, calculating the $M \times M$ matrix \sqrt{W} at the beginning of a calculation becomes time and memory consuming.

However, the computational work could be substantially reduced by taking advantage of the fact that physical potentials $W(\mathbf{z})$ depend only on the vector difference $\mathbf{z} = \mathbf{x} - \mathbf{y}$. The overwhelming part of information in the two-index matrix W_{nm} is redundant: $W_{nm} = W_{(n-m) \bmod M, 0} = W_{q0}$ with $q = (n - m) \bmod M$, an index that embodies only the relative displacement on the numerical lattice.

Consider the discrete Fourier transform (DFT) of this translated interaction potential $W_{q0} = W(\mathbf{z}_q)$ on a numerical lattice with periodic boundary conditions:

$$\tilde{W}_{\tilde{m}} = dV \sum_q W_{q0} e^{-i\mathbf{k}_{\tilde{m}} \cdot \mathbf{z}_q} = dV e^{i\mathbf{k}_{\tilde{m}} \cdot \mathbf{x}_m} \sum_n W_{nm} e^{-i\mathbf{k}_{\tilde{m}} \cdot \mathbf{x}_n}. \quad (47)$$

The last expression brings us back to the two-coordinate form W_{nm} for all values of m . Here, $\mathbf{k}_{\tilde{m}}$ are the standard wave vectors on the numerical lattice. Since the potential W_{nm} is even, that is $W_{n,(n+d) \bmod M} = W_{n,(n-d) \bmod M}$ for any d due to 1D periodicity (analogously in higher dimensions), the transform \tilde{W} will be real. The interaction drift term for $d\alpha_n/dt$ in Eq. (13) can then be written as an inverse DFT,

$$-i\alpha_n \frac{1}{V dV} \sum_{\tilde{m}} \tilde{W}_{\tilde{m}} \tilde{n}_{\tilde{m}} e^{i\mathbf{k}_{\tilde{m}} \cdot \mathbf{x}_n}, \quad (48)$$

where $V = MdV$ is the total volume, tildes will indicate k space, and

$$\tilde{n}_{\tilde{m}}(t) = dV \sum_p \alpha_p(t) \beta_p(t) e^{-i\mathbf{k}_{\tilde{m}} \cdot \mathbf{x}_p} \quad (49)$$

is the DFT of the density $n_p(t) = \alpha_p(t) \beta_p(t)$ for a given trajectory at a given time. This way, we can see that the deterministic evolution can be carried out by storing the interaction vector $\tilde{W}_{\tilde{m}}$ of just size M from the beginning of a simulation, and by carrying out two DFTs per trajectory and time step to obtain \tilde{n} . This has only $M \log M$ computational cost with “fast Fourier transform” methods.

Remarkably, the noise terms can be dealt with in a related, though more involved, way that follows an approach reported in [23]. Denoting the noise term for $d\alpha_n/dt$ as $X_n = \sqrt{-i} \alpha_n \sum_j [\sqrt{W}]_{nj} \xi_j^{(1)}(t)$, the diffusion condition Eq. (12) requires simply that

$$\overline{X_n X_m} = -i W_{nm} \alpha_n \alpha_m \quad (50)$$

(equal times will be assumed in this section). Then, if we rewrite the noise X_n in terms of new scaled and Fourier-transformed noise quantities \tilde{Y} as

$$X_n = \sqrt{-i} \alpha_n \frac{1}{V} \sum_{\tilde{p}} \tilde{Y}_{\tilde{p}} e^{i\mathbf{k}_{\tilde{p}} \cdot \mathbf{x}_n}, \quad (51)$$

it can be shown by simple substitution that the condition

$$\overline{\tilde{Y}_{\tilde{p}} \tilde{Y}_{\tilde{q}}} = V \tilde{W}_{\tilde{p}} \delta_{\tilde{p}, -\tilde{q}} \quad (52)$$

on the new noises is sufficient to satisfy the required diffusion condition Eq. (50). If we are able to construct a noise field $\chi_{\tilde{p}}$ with the somewhat atypical correlation properties

$$\overline{\chi_{\tilde{p}} \chi_{\tilde{q}}} = \delta_{\tilde{p}, -\tilde{q}}, \quad (53)$$

then the choice $\tilde{Y}_{\tilde{p}} = \chi_{\tilde{p}} \sqrt{V \tilde{W}_{\tilde{p}}}$ will satisfy all that is required. Such a noise can indeed be constructed, as is described in Appendix D.

With it, the noise term Eq. (51) can be written as an appropriate inverse DFT:

$$X_n = \sqrt{\frac{-i}{V}} \alpha_n \sum_{\tilde{p}} \chi_{\tilde{p}}(t) \sqrt{\tilde{W}_{\tilde{p}}} e^{i\mathbf{k}_{\tilde{p}} \cdot \mathbf{x}_n}. \quad (54)$$

This is also implementable with computational cost $M \log M$ per time step. The linear term $\sum_m \omega_{nm} \alpha_m$ does not usually constitute an efficiency problem, so we leave it as is. Applying the above, using independent noise fields $\chi^{(1)}$ and $\chi^{(2)}$, the final

more efficient equations in the positive- P representation are

$$\begin{aligned} \frac{d\alpha_n}{dt} = & -i \sum_m \omega_{nm} \alpha_m - i\alpha_n \frac{1}{dV} \\ & \times \sum_{\tilde{m}} \left[\frac{\tilde{W}_{\tilde{m}} \tilde{n}_{\tilde{m}}}{V} - \frac{\sqrt{idV}}{M} \chi_{\tilde{m}}^{(1)}(t) \sqrt{\tilde{W}_{\tilde{m}}} \right] e^{i\mathbf{k}_{\tilde{m}} \cdot \mathbf{x}_n}, \end{aligned} \quad (55a)$$

$$\begin{aligned} \frac{d\beta_n}{dt} = & i \sum_m \omega_{nm} \beta_m - i\beta_n \frac{1}{dV} \\ & \times \sum_{\tilde{m}} \left[\frac{\tilde{W}_{\tilde{m}} \tilde{n}_{\tilde{m}}}{V} - i \frac{\sqrt{idV}}{M} \chi_{\tilde{m}}^{(2)}(t) \sqrt{\tilde{W}_{\tilde{m}}} \right] e^{i\mathbf{k}_{\tilde{m}} \cdot \mathbf{x}_n}. \end{aligned} \quad (55b)$$

B. Drift gauges and system size

Looking at the simulation time estimate for the gauge- P distribution in the contact-interacting case Eq. (43), one sees that after intensive quantities like g and ρ have been factored out, a disadvantageous reduction with the system size M remains. This is a feature that has been also noted previously. It arises because the single weight variable collects noise from the entire system [37], while the limit $\mathcal{V} \lesssim 10$ remains unforgiving as system size grows. In contrast, positive- P or only diffusion gauged calculations do not show this behavior. For this reason, though they are consistently less advantageous for single modes, they may be better when the system size becomes large.

1. Diffusion gauge only

To address these issues, we first evaluate how a global diffusion gauge Eq. (24) performs in long-range interacting systems without any drift gauges. We proceed the same way as in Secs. IV C 1 and V C and [37]. The variance Eq. (33) is expanded to third order in t , with gauge Eq. (24), and deterministic initial conditions $C^{(0)} = \tilde{C}^{(0)} = 0$. The result is

$$\mathcal{V}^{(P)}(t) = \mathcal{V}(0) + \frac{tU_0}{2} \cosh 2a - \frac{t^2}{2} I_1^{(P)} + \frac{t^3}{3} e^{-2a} I_2^{(P)} \quad (56)$$

with the same integrals as in Eq. (37). The $I_1^{(P)}$ term is irrelevant as it does not depend on a . The optimum gauge then is readily shown to be simply

$$a_{\text{opt}} \approx \frac{1}{4} \ln \left(\frac{4t_{\text{opt}}^2 I_2^{(P)}}{3U_0} + 1 \right), \quad (57)$$

in agreement with known single-mode results. From the estimates Eq. (38), we see that this is largely independent of M , as hoped.

The simulation time, requiring $\mathcal{V}(t_{\text{sim}}) = 10$, can be found in the usual two regimes: The noise amplification dominated regime, when n and t_{sim} are large enough, has

$$t_{\text{sim}} \approx \frac{O(4)}{[U_0 I_2^{(P)}]^{1/4}}. \quad (58)$$

This is a different and more advantageous scaling as compared to the plain positive- P Eq. (40). The diffusion gauged value

of t_{sim} is longer when

$$I_2^{(P)} \gtrsim 0.03 U_0^3. \quad (59)$$

Notably, $I_2^{(P)}$ is an integral involving third powers of the potential, so Eq. (59) is primarily a condition on the density, requiring it to be sufficiently high. Substituting the estimates Eq. (38), one has $n_0 \gtrsim 0.2$ for the lattice occupations. This result does not depend on system size. The direct-noise dominated regime has the same estimate Eq. (41) as in the other cases.

2. Diffusion gauges vs drift gauges

With the different t_{sim} of drift and diffusion-only gauged calculations, a pertinent question is when should the drift gauge be added to the diffusion gauge? Comparing Eqs. (42) and (58), diffusion only calculations last longer when

$$I_2^{(P)} \lesssim \frac{U_0}{16} I_2. \quad (60)$$

For contact interactions in a uniform system when substituting Eqs. (28) and (38) this reduces to the surprisingly simple expression

$$M \gtrsim 16. \quad (61)$$

For long-range interactions, condition Eq. (60) should be used instead of Eq. (61). Also there it is clear, however, that for large enough systems, say $M \gg 100$, the diffusion gauge only approach will be better.

The simulations of Sec. VIB were an example of a sufficiently large system, with $M = 64$, $I_2^{(P)} \approx 2500$, $U_0 = 15.624$, and $I_2 \approx 1500$, where drift gauges cease to offer a huge advantage.

VIII. GENERAL MANY-MODE DIFFUSION GAUGES

We have seen so far, that the global gauge ansatz (24) can already provide significant advantages as well as technical challenges in the implementation. However, it only exploits a small fraction of the degrees of freedom for the many-body case that are inherent in the full $2M \times 2M$ matrix \underline{Q} . In this section we present our initial explorations of *nonlocal many-mode* gauges. These are intended as starting point for future research, beyond the scope of this article.

For many-mode problems with contact potentials $W = W_0 \mathbb{1}$ as considered in previous work, a *local diffusion gauge*

$$\underline{Q} = \begin{pmatrix} \text{diag}(\cosh a_n) & -i \text{diag}(\sinh a_n) \\ i \text{diag}(\sinh a_n) & \text{diag}(\cosh a_n) \end{pmatrix} \quad (62)$$

is a feasible choice that already goes beyond (24). Since \underline{Q} and \underline{S} in Eq. (18b) commute in this case, each gauge parameter a_m can be individually associated with the mode m . For sufficiently small intermode coupling, one can then use an optimized gauge $a_m(t)$ based on single-mode results that depend on the stochastic occupation n_m of this particular mode m . Often this gives good results [23].

However, for a nonlocal potential W , the matrices \underline{Q} and \underline{S} no longer commute and hence the a_n would have to be viewed as a gauge parameter of the noise source ξ_n . In that case, it is not clear to what degree a naive application of single-mode results

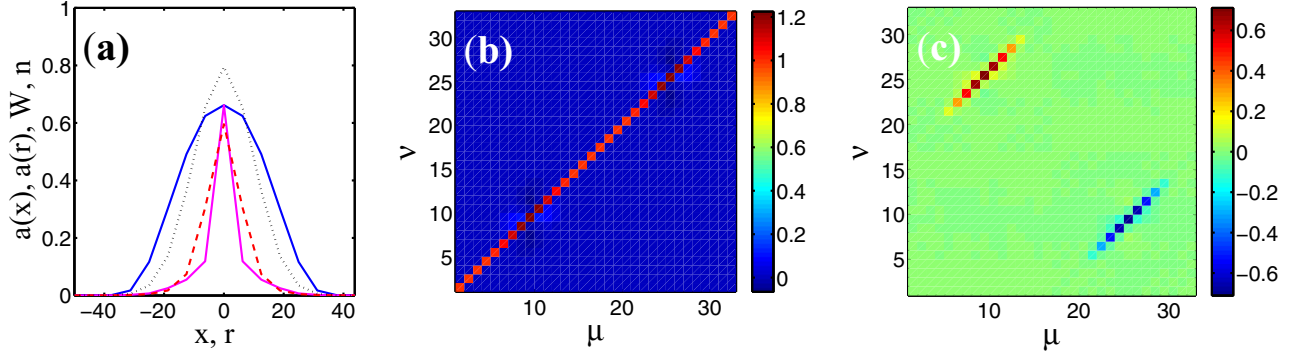


FIG. 3. A numerical optimization of the nonlocal diffusion gauge Eq. (63). (a) Physical conditions are given by the potential $W(r)$ (red-dashed) and density $n(x)$ (black dotted) described in the text, shown here in arbitrary units. We then show the resulting spatial and nonlocal variations of the diffusion gauge using $a(x)$ (blue solid) and $[a(r)]$ (magenta solid), respectively, as described in the text. (b) $\text{Re}[O_{\mu\nu}]$. (c) $\text{Im}[O_{\mu\nu}]$. The initial condition (i) described in the text was used.

is still possible. In the remainder of this section we describe two approaches aimed at harnessing more of the gauge freedom of $\underline{\underline{Q}}$: an analytical one and a numerical one.

A. Analytical nonlocal diffusion gauges

Finding the matrix $\underline{\underline{Q}}$ which globally minimizes the variance for a given interaction potential and simulation time is a difficult task, analytically and numerically. Here, we investigate the ansatz

$$\underline{\underline{Q}} = \begin{pmatrix} \cosh A & -i \sinh A \\ i \sinh A & \cosh A \end{pmatrix}, \quad (63)$$

where the hyperbolic functions are defined in terms of matrix exponentials, and A is a symmetric $M \times M$ matrix. The matrix $\underline{\underline{Q}}$ is orthogonal for all choices of A and reduces to (24) for $\overline{A} = a\mathbb{1}$, with scalar a .

Now let us consider the subset of nonlocal interactions for which $W = U$, i.e., $\sqrt{W} = \sqrt{W}^\dagger = \text{Re}[\sqrt{W}]$. We take A real and the starting densities deterministic: $\mathbf{n}(0) = \overline{\mathbf{n}}(0)$. Then, inserting $\underline{\underline{Q}}$ into the variance expression Eq. (21) and simplifying, we obtain the following for short times:

$$\begin{aligned} \mathcal{V} &= \frac{t}{2M} \text{Tr}[W \cosh(2A)] + t \mathbf{n}''^T(0) \sqrt{W} e^{-2A} \sqrt{W} \mathbf{n}''(0) \\ &+ \frac{t^2}{2} \text{Tr}[\sqrt{W} e^{-2A} \sqrt{W} N' \sqrt{W} e^{-2A} \sqrt{W} N'] \\ &+ (N' \rightarrow N''). \end{aligned} \quad (64)$$

We used the matrices

$$N' = \text{diag}[\mathbf{n}'(0)], \quad N'' = \text{diag}[\mathbf{n}''(0)]. \quad (65)$$

A minimum of the scalar quantity \mathcal{V} in the space of matrices A is given by the solution of $\partial\mathcal{V}/\partial A_{ij} = 0$.

As described in more detail in Appendix E, if \mathbf{n} is homogeneous and real, we can obtain the matrix equation

$$W = W \exp(-4A) + 4Mt[\sqrt{W} N' W N' \sqrt{W}] \exp(-6A). \quad (66)$$

Equation (66) has the same structure as the single mode version in [37]. Hence, following [37], we can solve this equation separately for cases where the $\exp(-4A)$ term or $\exp(-6A)$

term on the right hand side is dominant, and then suggest a simple interpolation equation:

$$A = \frac{1}{6} \ln\{4Mt_{\text{opt}} \sqrt{W}^{-1} N' W N' \sqrt{W} + \mathbb{1}\}, \quad (67)$$

where \ln denotes a matrix logarithm. We present Eq. (67) for completeness, as for the cases explored, it did not prove more useful than the much simpler adaptive global gauge of Sec. IV C.

We defer an initial test of the performance of (67) to the next section.

B. Numerical nonlocal diffusion gauges

Another route by which nonlocal diffusion gauges might offer an advantage over local adaptive ones is the numerical minimization of Eq. (21) with respect to an arbitrary orthogonal matrix O for given initial mode occupations $n_m(0)$, t and W . We carried out such a minimization for the same potential as in Sec. VIB 1 with spatial atom-density $n(x) = \bar{n} \exp[-x^2/2\sigma^2]$, $\sigma = 10$, $\bar{n} = 0.5$, and $t_{\text{opt}} = 0.15$. A small number of modes $M = 16$ covering the range $x \in [-L, L]$ with $L = 50$ allowed reasonably fast optimization. In practice the degree of freedom for the conjugate gradient optimization routine [99] was a real, antisymmetric matrix g such that $O = \exp(g)$ as in Eq. (24).

We employed a variety of different initial conditions: (i) local global gauge (24) with a guessed, (ii) nonlocal gauge (63) with $A_{kl} = \text{const}$, (iii) $\underline{\underline{Q}} = \exp(g)$ with random g_{kl} , (iv) $g = 0$, (v) the analytical solution Eq. (67). For comparison with local diffusion gauges Eq. (62), we extracted from the optimized $\underline{\underline{Q}}$ a local diffusion gauge parameter $a(x_\mu) = \sinh^{-1}[\text{Im}[\underline{\underline{Q}}_{M+\mu, \mu}]]$ for $1 \leq \mu \leq M$, as well as the nonlocal shape $a(r_\mu) = \sinh^{-1}[\text{Im}[\underline{\underline{Q}}_{M, \mu}]]$, both shown in Fig. 3 for the initial condition (i). For comparison, a pure local diffusion form Eq. (62) would contain nonzero elements only exactly on the dominant diagonal features described by $g(x)$ as seen in Figs. 3(b) and 3(c), while the magenta line for $g(r)$ in Fig. 3(a) would only have the one nonzero element at $r = 0$.

All initial conditions led to quite similar results as shown, with a significant spatial dependence of the gauge parameter $a(x)$ due to the inhomogeneous density $n(x)$, and small nonlocal tails visible in $a(r)$. The overall decrease in variance

TABLE I. Summary of main analytic results.

Result	Positive P	Gauge P
Evolution equations for nonlocal interactions	basic (13) large system (55)	general gauge P (18) standard drift gauge (19)
Characteristic variances	(35)	all gauges (21) diffusion gauges only (33)
Optimized global gauges		all gauges (31) diffusion gauge only (57)
Simulation time estimate	(40)	all gauges (42) diffusion gauge only (58) gauge comparison (60)

of Eq. (21) compared to a local diffusion gauge was by a factor of 2.3 in the case shown, with weak nonlocality. We also used the numerical optimization scheme to confirm that (67) is indeed an optimal choice if $n(x)$ is uniform and times are short. Even for the inhomogeneous test case described above it already outperforms the global gauge by a factor of 2.1, close to the performance of the completely optimized result. The expression (67) also roughly captures the off-diagonal shape of gauge parameters found by the numerical optimization in those cases.

From this reduction of the characteristic variance \mathcal{V} , one can make one important general conclusion—that there is indeed some potential for gains from a nonlocal optimization of the gauge. Since we just skimmed the surface of gauges more general than Eq. (24), this indicates an extensive future route for possible improvement in our simulation methods, beyond the scope of the present article.

IX. SUMMARY

We have provided the necessary apparatus for carrying out stochastic positive- P and gauge- P simulations for systems of long-range interacting bosons. Our focus has been on directions relevant for ultracold gases such as Rydberg and dipolar atoms, as this is presently a burgeoning field of experimental and theoretical investigation. A summary of our main analytical results is given in Table I for easy reference.

Our most important results are the dynamical equations with a long-range potential, the estimates for the available simulation times t_{sim} , and simple formulas for optimal diffusion gauge parameters a_{opt} . The simulation times reveal whether a particular problem is amenable to calculation using this method, while the gauge formulas give a straightforward means of improving simulation times beyond those directly available from the positive- P method. The underlying general expressions for the noise evolution (the characteristic variances \mathcal{V}) constitute a good starting point for future investigations, including the development of truly nonlocal diffusion gauges. The results of Sec. VIII B confirm that nonlocal gauges can offer advantages in principle.

Comparing with the previously known results for contact interactions, we can make several statements: stochastic gauges continue to improve simulation times. For this we use a global gauge that is optimized with the help of nonlocal integrals

such as I_2 and $I_2^{(P)}$. The resulting expressions for the optimal gauge, simulation time, and characteristic variance that we have derived bear structural similarity to the contact interacting case, and reduce to them in the appropriate limit. For large systems with long-range interaction, special care must be taken to avoid a quadratic scaling with system size and memory problems, but remarkably, this can be achieved using Fourier transformed interactions and noise fields. This technique makes simulations with, say, $M = 10^5$ – 10^6 possible, which would otherwise be well out of reach.

Furthermore, we have confirmed here a suspicion that arose already for contact-interacting gases, that very large systems with $\gtrsim \mathcal{O}(100)$ modes have different stochastic gauge properties than small to medium systems. Drift gauges become inefficient for large systems because of the accumulation of noise from all modes into the one global weight Ω . In that case, diffusion gauges should be used, and continue to provide improvement over the plain positive- P treatment for many cases.

Finally, we have illustrated our techniques by modeling an interaction quench in the extended Bose-Hubbard model and the excitation of Rydberg states in a Bose-Einstein condensate, using an echo sequence. Notably, the inclusion of atomic motion in such a calculation introduces no additional significant complications here, in contrast to many other methods used to study Rydberg excitations. The main arena where we expect the methods developed here to offer advantages over alternatives (mainly density-matrix renormalization methods) are time-dependent quantum many-body problems in two or three spatial dimensions.

ACKNOWLEDGMENTS

It is a pleasure to thank André Ribeiro de Carvalho, Karen Kheruntsyan, and Murray Olsen for comments. We also gratefully acknowledge useful discussions with Cenap Ates, Thomas Pohl, and Jovica Stanojevic. P.D. acknowledges support from National Science Centre (Poland) Grant No. 2012/07/E/ST2/01389.

APPENDIX A: STRATONOVICH CORRECTIONS

The evolution equations (18a) and (18b) are to be interpreted in Itô stochastic calculus. Most numerical implementations that go beyond simple and inefficient first-order

methods require the use of the Stratonovich form. This includes the adaptive 8th and 9th order Runge-Kutta method [100] employed for all simulations in this article using the XMDS package used here [101] (see [102,103] for the most recent version). For equations with variable-dependent noise such as the ones derived here, the Stratonovich correction can be complicated, especially when the gauges themselves are a function of the variables. Hence, it is useful to provide the end results here.

If one decomposes the variables in $\bar{v} = (\boldsymbol{\gamma}, \Omega)$ into all the independent variables (either $\text{Re}[v]_\mu$ and $\text{Im}[v]_\mu$ or v and v^*), labeling them as u_j with $j = 1, \dots, 2(2M+1)$, then one can apply the usual formula [3,80] for the Stratonovich correction,

$$S_{u_j} = -\frac{1}{2} \sum_{\mu} \sum_{k=1}^{2(2M+1)} B_{k\mu} \frac{\partial}{\partial u_k} B_{j\mu}, \quad (\text{A1})$$

for a real random process with Itô equation

$$\frac{du_j}{dt} = A_j(\bar{v}) + \sum_{\mu} B_{j\mu}(\bar{v}) \xi_{\mu}. \quad (\text{A2})$$

The Stratonovich form of the equation is then $\frac{du_j}{dt} = A_j(\bar{v}) + \sum_{\mu} B_{j\mu}(\bar{v}) \xi_{\mu} + S_{u_j}(\bar{v})$.

1. Uniform diffusion gauge

With a diffusion gauge \underline{O} independent of the variables, but allowing arbitrary drift gauges \underline{f} , we obtain

$$S_{\gamma_{\mu}} = -\frac{i}{2} \gamma_{\mu} W_{\mu\mu}, \quad (\text{A3})$$

$$S_{\Omega} = -\frac{i}{2} \Omega \left[\underline{f}^T \underline{W} \underline{f} + \sum_{\nu} \gamma_{\nu} \left[\underline{W} \frac{\partial \underline{f}}{\partial \gamma_{\nu}} \right]_{\nu} - i \sum_{\nu} \gamma_{\nu}^* \left[\underline{S} \underline{O} \underline{O}^{\dagger} \underline{S}^{\dagger} \frac{\partial \underline{f}^*}{\partial \gamma_{\nu}} \right]_{\nu} \right]. \quad (\text{A4})$$

The notation $[\underline{a}]_{\nu}$ is used for the ν component of the vector inside the brackets.

If we specify the drift gauge as $\underline{f} = i \underline{n}''$ from Eq. (19) we have used throughout this paper, one obtains

$$S_{\Omega} = \frac{\Omega}{4} \sum_{\nu} n_{\nu}^* [\underline{S} \underline{O} \underline{O}^{\dagger} \underline{S}^{\dagger} \underline{F}]_{\nu\nu}^* \quad (\text{A5})$$

with \underline{F} defined in Eq. (22). A case relevant for this paper is the uniform diffusion gauge Eq. (62) using a real \mathbf{a} , where one has

$$S_{\Omega} = \frac{\Omega}{2} \left(\sum_n n_n \sum_m [\sqrt{W}]_{nm} e^{-2a_m} [\sqrt{W}^{\dagger}]_{mn} \right)^*, \quad (\text{A6})$$

which reduces to

$$S_{\Omega} = \frac{\Omega}{2} \sum_n n_n^* e^{-2a} U_0 = \frac{\Omega U_0 e^{-2a}}{2} \int d\mathbf{x} \rho(\mathbf{x})^* \quad (\text{A7})$$

for the global gauge Eq. (24). The real rectified potential Eq. (27) appears again. The case of contact interactions is recovered when $W_{nm} = g \delta_{nm} / dV$. An ungauged positive- P simulation, or one with diffusion gauges that do not depend on the variables, has only the Eq. (A3) corrections.

2. Adaptive gauge Stratonovich correction

In the case of an adaptive diffusion gauge the Stratonovich correction becomes significantly more complicated, since the derivatives in Eq. (A1) now also act on the gauge parameter a . Taking the global form Eq. (24) but with the variable and time dependence of Eq. (32), after much algebra, one obtains

$$S_{\alpha(\mathbf{x})} = \frac{i}{2} W_0 \alpha(\mathbf{x}) + i \frac{\alpha(\mathbf{x}) e^{-6a}}{U_0} \times \left\{ \frac{(t_{\text{fin}} - t)}{3} [I_3(\mathbf{x}) - i e^{-2a} I_4(\mathbf{x})^*] + \frac{1}{2} \sqrt{1 + \frac{4I_1}{W_0}} (-i I_5(\mathbf{x}) + e^{-2a} I_6(\mathbf{x})^*) \right\}, \quad (\text{A8a})$$

$$S_{\beta(\mathbf{x})} = -\frac{i}{2} W_0 \beta(\mathbf{x}) - i \frac{\beta(\mathbf{x}) e^{-6a}}{U_0} \times \left\{ \frac{(t_{\text{fin}} - t)}{3} [I_3(\mathbf{x}) + i e^{-2a} I_4(\mathbf{x})^*] - \frac{1}{2} \sqrt{1 + \frac{4I_1}{W_0}} [(I_5(\mathbf{x}) + e^{-2a} I_6(\mathbf{x})^*)] \right\}, \quad (\text{A8b})$$

$$S_{\Omega} = \frac{\Omega U_0}{2} e^{-2a} \int d\mathbf{x} \rho^*(\mathbf{x}) - \frac{\Omega e^{-8a}}{U_0} \int d\mathbf{x} \rho''(\mathbf{x}) \times \left\{ -\frac{2i(t_{\text{fin}} - t)}{3} I_4(\mathbf{x})^* + \sqrt{1 + \frac{4I_1}{W_0}} I_6(\mathbf{x})^* \right\} \quad (\text{A8c})$$

$$I_3(\mathbf{x}) = \int d\mathbf{z} d\mathbf{y} \rho(\mathbf{y}) \rho(\mathbf{z})^* U(\mathbf{y} - \mathbf{z})^2 W(\mathbf{x} - \mathbf{y}), \quad (\text{A9a})$$

$$I_4(\mathbf{x}) = \int d\mathbf{z} d\mathbf{y} \rho(\mathbf{y}) \rho(\mathbf{z})^* U(\mathbf{y} - \mathbf{z})^2 U(\mathbf{x} - \mathbf{y}), \quad (\text{A9b})$$

$$I_5(\mathbf{x}) = \int d\mathbf{z} d\mathbf{y} \rho(\mathbf{y}) \rho''(\mathbf{z}) U(\mathbf{y} - \mathbf{z}) W(\mathbf{x} - \mathbf{y}), \quad (\text{A9c})$$

$$I_6(\mathbf{x}) = \int d\mathbf{z} d\mathbf{y} \rho(\mathbf{y}) \rho''(\mathbf{z}) U(\mathbf{y} - \mathbf{z}) U(\mathbf{x} - \mathbf{y}). \quad (\text{A9d})$$

APPENDIX B: CHARACTERISTIC VARIANCE: WITH DRIFT GAUGE

1. Logarithmic variables

To estimate Eq. (20), a first step is to turn the stochastic equations (differentials), Eqs. (18a) and (18b), into equations for $\ln|\Omega\gamma_{\mu}|$. We use the usual multivariate Itô formula [80,104] using the independent variables \bar{v} like in Eq. (A2), to give

$$\frac{d}{dt} \ln(\Omega\gamma_{\mu}) = \frac{i}{\gamma_{\mu}} [\underline{\omega} \underline{\boldsymbol{\gamma}}]_{\mu} + i [\underline{W} (\underline{n} - \underline{f})]_{\mu} - \frac{i}{2} [W_0 + \underline{f}^T \underline{W} \underline{f}] + \sqrt{i} [\underline{S} \underline{O} \underline{\xi}]_{\mu} + \sqrt{i} \underline{f}^T \underline{S} \underline{O} \underline{\xi}. \quad (\text{B1})$$

From this point on, we will neglect kinetic terms, setting $\underline{\omega} = 0$, as explained in Sec. [V A](#). Formally, Eq. [\(B1\)](#) can be integrated over time, with the result

$$\begin{aligned} \ln[\Omega(t)\gamma_\mu(t)] &= \ln[\Omega(0)\gamma_\mu(0)] + i \left[\underline{W} \int_0^t (\underline{n}(s) - \underline{f}(s)) ds \right]_\mu - i \frac{W_0}{2} t - \frac{i}{2} \int_0^t \underline{f}^T(s) \underline{W} \underline{f}(s) ds \\ &\quad + \sqrt{i} \left[\underline{S} \underline{O} \int_0^t \underline{\xi}(s) ds \right]_\mu + \sqrt{i} \int_0^t \underline{f}^T(s) \underline{S} \underline{O} \underline{\xi}(s). \end{aligned} \quad (\text{B2})$$

Next, we move from $\ln(\Omega\gamma_\mu)$ to $\ln(|\Omega\gamma_\mu|)$. It is helpful to specify the form of \underline{f} at this point, as the usual $\underline{f} = i\underline{n}''$. This has the convenient feature that the deterministic parts of $\ln(\Omega\gamma_\mu)$ are imaginary and do not contribute to $\ln|\Omega\gamma_\mu|$ since $\ln|z| = \text{Re}[\ln z]$. This is likely an important factor for why this is a useful gauge. With this choice of \underline{f} , we arrive at

$$\ln[|\Omega(t)\gamma_\mu(t)|] = \ln[|\Omega(0)\gamma_\mu(0)|] + \frac{\sqrt{i}}{2} \sum_v [\underline{S} \underline{O} - i \underline{S}^* \underline{O}^*]_{\mu v} \int_0^t \xi_v(s) ds + \frac{i\sqrt{i}}{2} \sum_{v\lambda} [\underline{S} \underline{O} + i \underline{S}^* \underline{O}^*]_{\lambda v} \int_0^t n''_\lambda(s) \xi_v(s). \quad (\text{B3})$$

As previously, the indices below square brackets denote the elements of the enclosed matrix or vector.

2. Variance assembly

The mean of Eq. [\(B3\)](#) is simply $\overline{\ln[|\Omega(t)\gamma_\mu(t)|]} = \overline{\ln[|\Omega(0)\gamma_\mu(0)|]}$, since $\int_0^t G(s) \xi_\sigma(s) ds = 0$ for any nonanticipating function G [\[80\]](#). The remaining averaging at $t = 0$ is over initial state noise, which is independent of any later dynamical noises ξ . The mean of the square is

$$\overline{\ln[|\Omega(t)\gamma_\mu(t)|]^2} = \overline{\ln[|\Omega(0)\gamma_\mu(0)|]^2} + \frac{i}{4} \sum_{vv'} [\underline{S} \underline{O} - i \underline{S}^* \underline{O}^*]_{\mu v} [\underline{S} \underline{O} - i \underline{S}^* \underline{O}^*]_{\mu v'} \int_0^t ds \int_0^t ds' \overline{\xi_v(s) \xi_{v'}(s')} \quad (\text{B4a})$$

$$- \frac{i}{4} \sum_{vv'\lambda\lambda'} [\underline{S} \underline{O} + i \underline{S}^* \underline{O}^*]_{\lambda v} [\underline{S} \underline{O} + i \underline{S}^* \underline{O}^*]_{\lambda' v'} \int_0^t ds \int_0^t ds' \overline{n''_\lambda(s) n''_{\lambda'}(s') \xi_v(s) \xi_{v'}(s')} \quad (\text{B4b})$$

$$- \frac{1}{2} \sum_{vv'\lambda\lambda'} [\underline{S} \underline{O} - i \underline{S}^* \underline{O}^*]_{\mu v} [\underline{S} \underline{O} + i \underline{S}^* \underline{O}^*]_{\lambda' v'} \int_0^t ds \int_0^t ds' \overline{n''_{\lambda'}(s') \xi_v(s) \xi_{v'}(s')}. \quad (\text{B4c})$$

We consider the three terms in sequence:

Term (B4a). From independence of noises, $\overline{\xi_v(s) \xi_{v'}(s')} = \delta_{vv'} \delta(s - s')$ so that

$$\int_0^t ds \int_0^t ds' \overline{\xi_v(s) \xi_{v'}(s')} = t \delta_{vv'}. \quad (\text{B5})$$

Hence

$$\text{Eq. (B4a)} = \frac{it}{4} [\underline{S} \underline{S}^T - (\underline{S} \underline{S}^T)^* - i \underline{S} \underline{O} \underline{O}^\dagger \underline{S}^\dagger - i (\underline{S} \underline{O} \underline{O}^\dagger \underline{S}^\dagger)^*]_{\mu\mu} = \frac{t}{2} [\underline{S} \underline{O} \underline{O}^\dagger \underline{S}^\dagger]_{\mu\mu}. \quad (\text{B6})$$

For the second line we note that $\underline{S} \underline{S}^T = \underline{W}$ and \underline{W} is real, as well as that $(\underline{S} \underline{O} \underline{O}^\dagger \underline{S}^\dagger)^* = (\underline{S} \underline{O} \underline{O}^\dagger \underline{S}^\dagger)^T$. The transpose is then irrelevant for the $[\]_{\mu\mu}$ element. Note that while $\underline{O} \underline{O}^T = \underline{1}$, this is not generally the case for $\underline{O} \underline{O}^\dagger$.

Term (B4b). The stochastic average is $\overline{n''_\lambda(s) n''_{\lambda'}(s') \xi_v(s) \xi_{v'}(s')}$. Now consider the following: if $s > s'$ then then $\xi(s)$ is uncorrelated with all the other terms, because of causality, and for the case of $n''(s)$, the nonanticipating nature of an Itô process. Hence, in this case we can write

$$\overline{n''_\lambda(s) n''_{\lambda'}(s') \xi_v(s) \xi_{v'}(s')} = \overline{n''_\lambda(s) n''_{\lambda'}(s') \xi_{v'}(s') \xi_v(s)} = 0 \quad \text{when } s > s' \quad (\text{B7})$$

since $\overline{\xi_v(s)} = 0$. The same argument can be made for $s' > s$, in which case a factor $\overline{\xi_{v'}(s')} = 0$ occurs. The remaining situation is $s = s'$, in which case the noises ξ are also uncorrelated with the n'' because of the nonanticipating nature. Hence, we can write

$$\begin{aligned} \int_0^t ds \int_0^{s'} ds' \overline{n''_\lambda(s) n''_{\lambda'}(s') \xi_v(s) \xi_{v'}(s')} &= \int_0^t ds \int_0^{s'} ds' \overline{n''_\lambda(s) n''_{\lambda'}(s')} \begin{cases} \overline{\xi_v(s) \xi_{v'}(s')} & \text{when } s = s' \\ 0 & \text{when } s \neq s' \end{cases} \\ &= \int_0^t ds \int_0^{s'} ds' \overline{n''_\lambda(s) n''_{\lambda'}(s')} \delta(s - s') \delta_{vv'} = \delta_{vv'} \int_0^t ds \overline{n''_\lambda(s) n''_{\lambda'}(s)} = \delta_{vv'} \mathcal{N}_{\lambda\lambda'}''. \end{aligned} \quad (\text{B8})$$

The shorthand expression

$$Q = \int_0^t ds \overline{n''_{\lambda}(s) n''_{\lambda'}(s)^T} \quad (\text{B9})$$

for the correlation matrix of integrated expectation values will be useful. Inserting Eq. (B8) into Eq. (B4b) gives

$$\begin{aligned} \text{Eq. (B4b)} &= \frac{1}{4} \sum_{\lambda\lambda'} [\underline{S} \underline{O} \underline{O}^\dagger \underline{S}^\dagger + (\underline{S} \underline{O} \underline{O}^\dagger \underline{S}^\dagger)^*]_{\lambda\lambda'} Q_{\lambda\lambda'} \\ &= \frac{1}{2} \text{Tr}[\text{Re}[\underline{S} \underline{O} \underline{O}^\dagger \underline{S}^\dagger] Q]. \end{aligned} \quad (\text{B10})$$

Term (B4c). Proceeding in the same way, one finds

$$\overline{n''_{\lambda'}(s') \xi_v(s) \xi_{v'}(s')} = \delta_{vv'} \int_0^t ds \overline{n''_{\lambda}(s)}. \quad (\text{B11})$$

and

$$\text{Eq. (B4c)} = \sum_{\lambda} (\text{Im}[\underline{S} \underline{O} \underline{O}^\dagger \underline{S}^\dagger]_{\mu\lambda} - W_{\mu\lambda}) \int_0^t ds \overline{n''_{\lambda}(s)}. \quad (\text{B12})$$

Gathering now all the terms and inserting them into Eqs. (B4) and (20) we obtain

$$\begin{aligned} \mathcal{V} &= \mathcal{V}(0) + \frac{1}{2M} \left\{ \frac{t}{2} \text{Tr}[\underline{S} \underline{O} \underline{O}^\dagger \underline{S}^\dagger] \right. \\ &\quad + M \text{Tr}[\text{Re}\{\underline{S} \underline{O} \underline{O}^\dagger \underline{S}^\dagger\} Q] \\ &\quad \left. + \sum_{\mu\lambda} (\text{Im}[\underline{S} \underline{O} \underline{O}^\dagger \underline{S}^\dagger]_{\mu\lambda} - W_{\mu\lambda}) \int_0^t ds \overline{n''_{\lambda}(s)} \right\}. \end{aligned} \quad (\text{B13})$$

3. Stochastic density evolution

To proceed further we must evaluate the expectation values involving \underline{n}'' . Applying the Itô formula for $n_m = \alpha_m \beta_m$, and again discarding the ω terms, leads to

$$\begin{aligned} \frac{dn_m}{dt} &= \sqrt{i} n_m \sum_{\nu\sigma} [S_{m\nu} + S_{m+M,\nu}] O_{\nu\sigma} \xi_{\sigma} \\ &\quad + i n_m \sum_{\nu,\nu',\sigma} S_{m\nu} P_{\nu\sigma} S_{m+M,\nu'} O_{\nu'\sigma}. \end{aligned} \quad (\text{B14})$$

Using the matrix \underline{F} defined in Eq. (A5), this can be written

$$\begin{aligned} \frac{dn_m}{dt} &= \sqrt{i} n_m [\underline{F} \underline{S} \underline{O} \underline{\xi}]_m + i n_m [\underline{W} (\underline{F} - \underline{I})]_{mm} \\ &= \sqrt{i} n_m [\underline{F} \underline{S} \underline{O} \underline{\xi}]_m. \end{aligned} \quad (\text{B15})$$

The last follows because $\underline{W} (\underline{F} - \underline{I})$ has all zeros on the diagonal. Moving to logarithmic variables and formally integrating the differential we obtain

$$n_{\mu}(t) = n_{\mu}(0) \exp\{\sqrt{i} [\underline{F} \underline{S} \underline{O} \underline{\xi}(t)]_{\mu}\}, \quad (\text{B16})$$

where

$$\zeta_{\mu}(t) = \int_0^t \xi_{\mu}(s) ds. \quad (\text{B17})$$

Since it is a sum of Gaussian random variables, ζ_{μ} is itself Gaussian distributed. It has correlations

$$\overline{\zeta_{\mu}(t) \zeta_{\nu}(t')} = \delta_{\mu\nu} \min[t, t'], \quad (\text{B18})$$

as found in [23].

4. Stochastic density means and variances

The expectation value of Eq. (B16) can be written in terms of the probability distribution of Gaussian random variables ζ_{ν} that obey Eq. (B18):

$$\overline{n_{\mu}(t)} = \overline{n_{\mu}(0)} \int d^{2M} \underline{\zeta} \Pr[\underline{\zeta}, t] \exp\{\underline{\mathbf{K}}^{(\mu)} \cdot \underline{\zeta}\}. \quad (\text{B19})$$

Here, $\underline{\mathbf{K}}^{(\mu)}$ are vectors with elements $K_{\nu}^{(\mu)} = \sqrt{i} [\underline{F} \underline{S} \underline{O}]_{\mu\nu}$, and the probability distribution is

$$\Pr[\underline{\zeta}, t] = \prod_{\sigma=1}^{2M} \frac{1}{\sqrt{2\pi t}} \exp\left[-\frac{\zeta_{\sigma}^2}{2t}\right] = \frac{1}{(2\pi t)^M} \exp\left[-\frac{\underline{\zeta} \cdot \underline{\zeta}}{2t}\right]. \quad (\text{B20})$$

We can then write the expectation value as an integral over the quadratic form

$$\overline{n_{\mu}(t)} = \frac{\overline{n_{\mu}(0)}}{(2\pi t)^M} \int d^{2M} \underline{\zeta} \exp\{-\underline{\zeta}^T \underline{A}(t) \underline{\zeta} + \underline{\mathbf{K}}^{(\mu)} \cdot \underline{\zeta}\}, \quad (\text{B21})$$

where $\underline{A} = \underline{1}/2t$, whence $\det \underline{A} = (2t)^{-2M}$. Using the standard result that

$$\int d^n \mathbf{x} \exp\{-\mathbf{x}^T \mathbf{A} \mathbf{x} + \mathbf{K} \cdot \mathbf{x}\} = \sqrt{\frac{\pi^n}{\det \mathbf{A}}} \exp\left(\frac{1}{4} \mathbf{K}^T \mathbf{A}^{-1} \mathbf{K}\right), \quad (\text{B22})$$

the expectation value finally becomes

$$\begin{aligned} \overline{n_{\mu}(t)} &= \overline{n_{\mu}(0)} \exp\left\{\frac{t}{2} \underline{\mathbf{K}}^{(\mu)} \cdot \underline{\mathbf{K}}^{(\mu)}\right\} \\ &= \overline{n_{\mu}(0)} \exp\left\{\frac{it}{2} [\underline{F} \underline{W} \underline{F}^T]_{\mu\mu}\right\}. \end{aligned} \quad (\text{B23})$$

Note that for our form of matrices \underline{W} and \underline{F} in Eqs. (14) and (A5),

$$\underline{F} \underline{W} \underline{F}^T = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}. \quad (\text{B24})$$

Hence $\overline{n_{\mu}(t)} = \overline{n_{\mu}(0)}$ and

$$\int_0^t ds \overline{n''_{\lambda}(s)} = t \overline{n''_{\lambda}(0)}. \quad (\text{B25})$$

We still have to evaluate Q of Eq. (B9). Note that for $z, w \in \mathbb{C}$ we have $z'' w'' = \text{Re}(z w^* - z w)/2$. Thus

$$\begin{aligned} Q_{\lambda\lambda'} &= \overline{\int_0^t n''_{\lambda}(s) n''_{\lambda'}(s) ds} \\ &= \frac{1}{2} \text{Re} \left\{ \overline{\int_0^t [n_{\lambda}(s) n_{\lambda'}^*(s) - n_{\lambda}(s) n_{\lambda'}(s)] ds} \right\}. \end{aligned} \quad (\text{B26})$$

Using Eq. (B16) we can write

$$\begin{aligned} \overline{n_\lambda(t)n_{\lambda'}^*(t)} &= \overline{n_\lambda(0)n_{\lambda'}^*(0)} \exp\{\sqrt{i}[\underline{F}\underline{S}\underline{O}\underline{\zeta}(t)]_\lambda + \sqrt{-i}[\underline{F}\underline{S}^*\underline{O}^*\underline{\zeta}(t)]_{\lambda'}\} \\ &= \overline{n_\lambda(0)n_{\lambda'}^*(0)} \int d^{2M}\underline{\zeta} \Pr[\underline{\zeta}, t] \exp\{\underline{K}^{(\lambda, \lambda')} \cdot \underline{\zeta}\}. \end{aligned} \quad (\text{B27})$$

Now the symbol $\underline{K}^{(\lambda, \lambda')}$ is

$$\underline{K}_v^{(\lambda, \lambda')} = \sqrt{i}[\underline{F}\underline{S}\underline{O}]_{\lambda v} + \sqrt{-i}[\underline{F}\underline{S}^*\underline{O}^*]_{\lambda' v}. \quad (\text{B28})$$

To use Eq. (B22), we obtain

$$\begin{aligned} \underline{K}^{(\lambda, \lambda')} \cdot \underline{K}^{(\lambda, \lambda')} &= i[\underline{F}\underline{W}\underline{F}]_{\lambda\lambda} - i[\underline{F}\underline{W}\underline{F}]_{\lambda'\lambda'} \\ &\quad + 2[\underline{F}\underline{S}\underline{O}\underline{O}^\dagger\underline{S}^\dagger\underline{F}]_{\lambda\lambda'} \\ &= 2[\underline{F}\underline{S}\underline{O}\underline{O}^\dagger\underline{S}^\dagger\underline{F}]_{\lambda\lambda'}. \end{aligned} \quad (\text{B29})$$

We have used $\underline{F}\underline{W}\underline{F}^T = 0$ and the other obvious properties of \underline{F} , \underline{S} , \underline{O} . Thus

$$\overline{n_\lambda(s)n_{\lambda'}^*(s)} = \overline{n_\lambda(0)n_{\lambda'}^*(0)} \exp\{s[\underline{F}\underline{S}\underline{O}\underline{O}^\dagger\underline{S}^\dagger\underline{F}]_{\lambda\lambda'}\}. \quad (\text{B30})$$

Using the same methods we obtain

$$\begin{aligned} \overline{n_\lambda(s)n_{\lambda'}(s)} &= \overline{n_\lambda(0)n_{\lambda'}(0)} \exp\left\{\frac{is}{2}([\underline{F}\underline{W}\underline{F}]_{\lambda\lambda} \right. \\ &\quad \left. + [\underline{F}\underline{W}\underline{F}]_{\lambda'\lambda'} + 2[\underline{F}\underline{W}\underline{F}]_{\lambda\lambda'})\right\} \\ &= \overline{n_\lambda(0)n_{\lambda'}(0)}. \end{aligned} \quad (\text{B31})$$

Now we have to integrate over time. The result is Eq. (21b). We also make use of Eq. (B25) and the properties of \underline{W} and \underline{n} to find that

$$\begin{aligned} \sum_{\mu\lambda} W_{\mu\lambda} \int_0^t ds n_{\lambda'}''(s) &= t \sum_{\mu\lambda} W_{\mu\lambda} n_{\lambda'}''(0) = \sum_{\mu} [\underline{W}\underline{n}]_{\mu} \\ &= \sum_m [-W\underline{n} + W\underline{n}]_m = 0. \end{aligned} \quad (\text{B32})$$

Finally, substituting Eqs. (B32) and (21b) into Eq. (B13), we obtain the pivotal result, Eq. (21).

APPENDIX C: CHARACTERISTIC VARIANCE: WITHOUT DRIFT GAUGE

The case when there is no drift gauge ($\underline{f} = 0$) follows in the same general way as above in Appendix B, though with some additional elements. Summarizing the procedure, we begin with

$$\begin{aligned} \ln[|\gamma_\mu(t)|] &= \ln[|\gamma_\mu(0)|] + \frac{\sqrt{i}}{2} \sum_v [\underline{S}\underline{O} - i\underline{S}^*\underline{O}^*]_{\mu v} \\ &\quad \times \int_0^t \xi_v(s) ds - \sum_v W_{\mu v} \int_0^t n_v''(s). \end{aligned} \quad (\text{C1})$$

With the usual assumptions, $\underline{n}(t)$ continues to be described by Eq. (B16), and $\bar{n}_\mu = n_\mu(0)$. With the use of this and algebra

$$\begin{aligned} \mathcal{V}^{(P)} &= \mathcal{V}(0) + \frac{1}{2M} \left\{ \frac{t}{2} \text{Tr}[\underline{S}\underline{O}\underline{O}^\dagger\underline{S}^\dagger] + \text{Tr}[WQ^{(P)}W] \right. \\ &\quad \left. - 2t \sum_{\mu\lambda} W_{\mu\lambda} \text{covar}[\ln|\gamma_\mu(0)|, n_{\lambda'}''(0)] \right. \\ &\quad \left. - \sqrt{i} \sum_{\mu\lambda\lambda'} [\underline{S}\underline{O} - i\underline{S}^*\underline{O}^*]_{\mu\lambda} W_{\lambda'\mu} \int_0^t ds \overline{\zeta_\lambda(t)n_{\lambda'}''(s)} \right\}, \end{aligned} \quad (\text{C2})$$

in terms of a quantity

$$\mathcal{Q}_{\lambda\lambda'}^{(P)} = \int_0^t ds \int_0^t ds' \text{covar}[n_{\lambda}''(s), n_{\lambda'}''(s')] \quad (\text{C3})$$

that is similar to \mathcal{Q} .

To evaluate Eqs. (C2) and (C3), the two-time averages $\overline{n_{\lambda}''(s)n_{\lambda'}''(s')}$ and $\overline{\zeta_\lambda(t)n_{\lambda'}''(s)}$ are needed. From the construction Eq. (B17) we note that when $s' > s$, the random variable ζ can be split into completely independent early and late ($\widehat{\zeta}$) parts

$$\zeta_\mu(s') = \zeta_\mu(s) + \widehat{\zeta}_\mu(s' - s) \quad (\text{C4})$$

with $\widehat{\zeta}(v)$ having the same variances as $\zeta(v)$ given in Eq. (B18), except that it is independent because it consists of different building block noises $\xi(t)$. This way, after using Eqs. (B16) and (C4) one has

$$\begin{aligned} \int_0^t ds \int_0^t ds' \overline{n_\lambda(s)n_{\lambda'}^*(s')} &= 2 \int_0^t ds \int_s^t ds' \overline{n_\lambda(s)n_{\lambda'}^*(s')} \\ &= 2 \int_0^t ds \int_s^t ds' \overline{n_\lambda(s)n_{\lambda'}^*(s) \exp\{\sqrt{-i}[\underline{F}\underline{S}^*\underline{O}^*\widehat{\zeta}(s' - s)]_{\lambda'}\}} \\ &= 2 \int_0^t ds \overline{n_\lambda(s)n_{\lambda'}^*(s)} \int_s^t ds' \overline{\exp\{\sqrt{-i}[\underline{F}\underline{S}^*\underline{O}^*\widehat{\zeta}(s' - s)]_{\lambda'}\}}. \end{aligned} \quad (\text{C5})$$

The factor in the right-hand integral can be identified as a scaled expectation value from Eq. (B23), $\overline{n_{\lambda'}(s'-s)^*/n_{\lambda'}^*(0)}$, and substituting $s' = v + s$ one has

$$\begin{aligned} \int_0^t ds \int_0^t ds' \overline{n_{\lambda}(s)n_{\lambda'}^*(s')} &= 2 \int_0^t ds \overline{n_{\lambda}(s),n_{\lambda'}^*(s)} \left[\int_0^{t-s} dv \frac{n_{\lambda'}(v)}{n_{\lambda'}(0)} \right]^* \\ &= 2 \int_0^t ds \overline{n_{\lambda}(s),n_{\lambda'}^*(s)}(t-s) \\ &= 2n_{\lambda}(0)n_{\lambda'}^*(0) \int_0^t ds (t-s) \exp\{s[\underline{F}\underline{S}\underline{O}\underline{O}^\dagger\underline{S}^\dagger\underline{F}]_{\lambda\lambda'}\} \\ &= \frac{2n_{\lambda}(0)n_{\lambda'}^*(0)}{[\underline{F}\underline{S}\underline{O}\underline{O}^\dagger\underline{S}^\dagger\underline{F}]_{\lambda\lambda'}^2} (e^{t[\underline{F}\underline{S}\underline{O}\underline{O}^\dagger\underline{S}^\dagger\underline{F}]_{\lambda\lambda'}} - 1 - t[\underline{F}\underline{S}\underline{O}\underline{O}^\dagger\underline{S}^\dagger\underline{F}]_{\lambda\lambda'}). \end{aligned} \quad (C6)$$

Here, Eq. (B25) was used in the second line, and Eq. (B30) in the third. Similarly,

$$\int_0^t ds \int_0^t ds' \overline{n_{\lambda}(s)n_{\lambda'}(s')} = t^2 \overline{n_{\lambda}(0)n_{\lambda'}(0)} \quad (C7)$$

and $\int_0^t ds \int_0^t ds' \overline{n_{\lambda}''(s)n_{\lambda'}''(s')} = t^2 \overline{n_{\lambda}''(0)n_{\lambda'}''(0)}$ from Eq. (B25), so that using a form similar to Eq. (B26) one has

$$Q_{\lambda\lambda'}^{(P)} = \text{Re} \left[\frac{n_{\lambda}(0)n_{\lambda'}^*(0)}{[\underline{F}\underline{S}\underline{O}\underline{O}^\dagger\underline{S}^\dagger\underline{F}]_{\lambda\lambda'}^2} (e^{t[\underline{F}\underline{S}\underline{O}\underline{O}^\dagger\underline{S}^\dagger\underline{F}]_{\lambda\lambda'}} - 1 - t[\underline{F}\underline{S}\underline{O}\underline{O}^\dagger\underline{S}^\dagger\underline{F}]_{\lambda\lambda'}) - \frac{t^2}{2} \overline{n_{\lambda}(0)n_{\lambda'}(0)} \right] - t^2 \overline{n_{\lambda}''(0)n_{\lambda'}''(0)}. \quad (C8)$$

This can be rearranged to the form

$$Q_{\lambda\lambda'}^{(P)} = \tilde{Q}_{\lambda\lambda'} + t^2 \text{covar}[n_{\lambda}''(0),n_{\lambda'}''(0)] \quad (C9)$$

using \tilde{Q} from Eq. (33b). Consider now the quantity $\zeta_{\lambda}(t)n_{\lambda'}''(s)$ in the last line of Eq. (C2). Splitting ζ into early and late parts as $\zeta(t) = \zeta(s) + \tilde{\zeta}(t-s)$ like in Eq. (C4), one can see that $\tilde{\zeta}(t-s)$ must be uncorrelated with the earlier $n''(s)$. From this and Eq. (B16) one has

$$\overline{\zeta_{\lambda}(t)n_{\lambda'}''(s)} = \text{Im}\{\overline{n_{\lambda'}(0)\zeta_{\lambda}(s)e^{\sqrt{i}[\underline{F}\underline{S}\underline{O}\underline{O}^\dagger\underline{S}^\dagger\underline{F}]_{\lambda\lambda'}(s)}}\}. \quad (C10)$$

The right-most expectation value is a Gaussian integral similar to Eq. (B21), but the first moment of the random variable $\zeta(s)$. The matrix \underline{A} is now $\underline{1}/(2s)$ and we have $\underline{K}_{\sigma}^{(v')} = \sqrt{i}[\underline{F}\underline{S}\underline{O}\underline{O}^\dagger\underline{S}^\dagger\underline{F}]_{\lambda\lambda'}$. Here, the result from multivariate normal distributions that we need is

$$\begin{aligned} \int d^n \mathbf{x} x_{\mu} e^{-\mathbf{x}^T \mathbf{A} \mathbf{x} + \mathbf{K} \cdot \mathbf{x}} \\ = \frac{1}{2} [\mathbf{K}^T \mathbf{A}^{-1}]_{\mu} \sqrt{\frac{\pi^n}{\det \mathbf{A}}} \exp\left(\frac{1}{4} \mathbf{K}^T \mathbf{A}^{-1} \mathbf{K}\right). \end{aligned} \quad (C11)$$

With this one finds

$$\overline{\zeta_{\lambda}(t)n_{\lambda'}''(s)} = s \text{Im}\{\sqrt{i}[\underline{F}\underline{S}\underline{O}\underline{O}^\dagger\underline{S}^\dagger\underline{N}]_{\lambda\lambda'}\}, \quad (C12)$$

where $\underline{N} = \text{diag}[\underline{n}(0)]$ as defined before. The bottom line in Eq. (C2) is

$$-\frac{t^2}{2} \text{Tr}\{\underline{W}\underline{N}'\underline{F}\underline{W} - \text{Im}[\underline{S}\underline{O}\underline{O}^\dagger\underline{S}^\dagger\underline{N}^*]\underline{F}\underline{W}\}. \quad (C13)$$

Finally substituting this and Eq. (C9), the expression Eq. (33) is obtained after some manipulation to separate the mostly useless initial covariance $C^{(0)}$ defined in Eq. (34).

APPENDIX D: NOISE PERFORMANCE IMPROVEMENT FOR LARGE SYSTEMS

As described in Sec. VII A, a more efficient formulation of the gauge- P SDEs is possible using a noise field $\chi_{\tilde{p}}$ with correlation properties

$$\overline{\chi_{\tilde{p}}\chi_{\tilde{q}}} = \delta_{\tilde{p},-\tilde{q}}. \quad (D1)$$

Such a noise can indeed be constructed, by first dividing the k space of the problem into three regions: \mathcal{R} contains half of all the k -space vectors \mathbf{k} that possess a matching wave vector $-\mathbf{k}$ reflected across the origin. In three dimensions on a standard DFT lattice \mathcal{R} can consist for example of all wave vectors in the $k_x > 0$ half space, plus all vectors on the $k_x = 0$ plane with $k_y > 0$, plus all vectors on the k_z axis with $k_z > 0$. \mathcal{R}' contains all the negatives $-\mathbf{k}$ of the wave vectors $\mathbf{k} \in \mathcal{R}$. Finally, \mathcal{R}_0 contains the residual unpaired wave vectors. We always have $[\mathbf{k} = 0] \in \mathcal{R}_0$, but if the number of lattice points M_d in a particular dimension d is even, it will also contain the unpaired wave vectors possessing maximally negative values $k_d = -\pi L_d/M_d$ that occur in the DFT for box length L_d . With this division, the prescription for Eq. (D1) is

$$\chi_{\tilde{m}} = \begin{cases} \frac{1}{\sqrt{2}} [\xi_{(\tilde{m},1)} + i\xi_{(\tilde{m},2)}] & \text{if } \mathbf{k}_{\tilde{m}} \in \mathcal{R} \\ \frac{1}{\sqrt{2}} [\xi_{(-\tilde{m},1)} - i\xi_{(-\tilde{m},2)}] & \text{if } \mathbf{k}_{\tilde{m}} \in \mathcal{R}' \\ \xi_{(\tilde{m},0)} & \text{if } \mathbf{k}_{\tilde{m}} \in \mathcal{R}_0 \end{cases} \quad (D2)$$

The real noises ξ have the same properties as before, i.e., they are all independent with variance $\overline{\xi_{(\tilde{p},j)}(t)\xi_{(\tilde{q},j)}(t')} = \delta_{\tilde{p}\tilde{q}} \delta_{jj'} \delta(t-t')$.

APPENDIX E: NONLOCAL GAUGE EXPRESSIONS

The purpose of this Appendix is to elaborate on how the results of Sec. VIII A were obtained. We desire a simpler version of Eq. (21) valid for short times. To this end, the exponential within each element λ, λ' of the matrix ($e^{i[\underline{E}\underline{S}\underline{O}\underline{O}^\dagger\underline{S}^\dagger\underline{F}]}_{\lambda\lambda'} - 1$) in Eq. (21) is expanded to second order in t . One obtains

$$\begin{aligned} \mathcal{V} = & \frac{t}{2M} \text{Tr}[\cosh(2A)W] \\ & + t \mathbf{n}''^T(0) \sqrt{W} e^{-2A} \sqrt{W} \mathbf{n}''(0) \\ & + \frac{t^2}{2} \text{Tr}[\sqrt{W} e^{-2A} \sqrt{W} N' \sqrt{W} \\ & \times e^{-2A} \sqrt{W} N' + (N' \rightarrow N'')], \end{aligned} \quad (\text{E1})$$

as written in Eq. (64), with the diagonal matrix N given by Eq. (65).

The more difficult task is to differentiate Eq. (E1) with respect to each element a_{ij} of the diffusion gauge matrix A . We make use of the expression for derivatives of matrix exponentials:

$$\frac{\partial \exp[A]}{\partial a_{ij}} = \int_0^1 d\tau \exp[(1-\tau)A] \frac{\partial A}{\partial a_{ij}} \exp[\tau A], \quad (\text{E2})$$

where $[\partial A / \partial a_{ij}]_{lm} = \delta_{li} \delta_{mj}$. With this expression we require

$$\begin{aligned} 0 = \frac{\partial \mathcal{V}}{\partial a_{ij}} = & \int_0^1 d\tau \left\{ \frac{t}{2M} \{ [\exp[2\tau A] W \exp[2(1-\tau)A]]_{ji} \right. \\ & - [\exp[-2\tau A] W \exp[-2(1-\tau)A]]_{ji} \} \\ & - 2t [\exp[-2\tau A] \sqrt{W} \mathbf{n}''(0) \mathbf{n}''^T(0) \\ & \times \sqrt{W} \exp[-2(1-\tau)A]]_{ji} \\ & - 2t^2 [\exp[2\tau A] \sqrt{W} N' \sqrt{W} e^{-2A} \sqrt{W} N' \sqrt{W} \\ & \times \exp[2(1-\tau)A] + \{N' \rightarrow N''\}]_{ji} \left. \right\}. \end{aligned} \quad (\text{E3})$$

Let us assume that A will be a function of \sqrt{W} only [105], and the initial mode occupation is constant and real, $n_0(x) = n_0$. Now matrices such as $\exp[2\tau A]$, N' and \sqrt{W} commute. The τ integration is then trivial, and we arrive at the matrix equation

$$W = W \exp(-4A) + 4Mt [\sqrt{W} N' W N' \sqrt{W}] \exp(-6A), \quad (\text{E4})$$

from which we can directly write Eq. (66) and then Eq. (67).

-
- [1] J. Eisert, M. Cramer, and M. B. Plenio, *Rev. Mod. Phys.* **82**, 277 (2010).
- [2] M. J. Steel, M. K. Olsen, L. I. Plimak, P. D. Drummond, S. M. Tan, M. J. Collett, D. F. Walls, and R. Graham, *Phys. Rev. A* **58**, 4824 (1998).
- [3] C. W. Gardiner and P. Zoller, *Quantum Noise* (Springer-Verlag, Berlin, Heidelberg, 2004).
- [4] P. D. Drummond, P. Deuar, T. G. Vaughan, and J. F. Corney, *J. Mod. Opt.* **54**, 2499 (2007).
- [5] P. D. Drummond and C. W. Gardiner, *J. Phys. A: Math. Gen.* **13**, 2353 (1980).
- [6] K. V. Kheruntsyan, M. K. Olsen, and P. D. Drummond, *Phys. Rev. Lett.* **95**, 150405 (2005).
- [7] C. M. Savage and K. V. Kheruntsyan, *Phys. Rev. Lett.* **99**, 220404 (2007).
- [8] I. Carusotto, Y. Castin, and J. Dalibard, *Phys. Rev. A* **63**, 023606 (2001).
- [9] I. Carusotto and Y. Castin, *J. Phys. B* **34**, 4589 (2001).
- [10] P. Deuar and P. D. Drummond, *Phys. Rev. Lett.* **98**, 120402 (2007).
- [11] P. D. Drummond, P. Deuar, and K. V. Kheruntsyan, *Phys. Rev. Lett.* **92**, 040405 (2004).
- [12] A. Polkovnikov, *Ann. Phys.* **325**, 1790 (2010).
- [13] P. Deuar, J. Chwedeńczuk, M. Trippenbach, and P. Ziń, *Phys. Rev. A* **83**, 063625 (2011).
- [14] P. Deuar, T. Wasak, P. Ziń, J. Chwedeńczuk, and M. Trippenbach, *Phys. Rev. A* **88**, 013617 (2013).
- [15] K. V. Kheruntsyan, J.-C. Jaskula, P. Deuar, M. Bonneau, G. B. Partridge, J. Ruauadel, R. Lopes, D. Boiron, and C. I. Westbrook, *Phys. Rev. Lett.* **108**, 260401 (2012).
- [16] R. J. Lewis-Swan and K. V. Kheruntsyan, *Nat. Commun.* **5**, 3752 (2014).
- [17] R. J. Lewis-Swan and K. V. Kheruntsyan, *Phys. Rev. A* **91**, 052114 (2015).
- [18] J. F. Corney and P. D. Drummond, *Phys. Rev. Lett.* **93**, 260401 (2004).
- [19] J. F. Corney and P. D. Drummond, *J. Phys. A: Math. Gen.* **39**, 269 (2006).
- [20] D. W. Barry and P. D. Drummond, *Phys. Rev. A* **78**, 052108 (2008).
- [21] R. Ng, E. S. Sørensen, and P. Deuar, *Phys. Rev. B* **88**, 144304 (2013).
- [22] J. Schachenmayer, A. Pikovski, and A. M. Rey, *Phys. Rev. X* **5**, 011022 (2015).
- [23] P. Deuar, Ph.D. thesis, The University of Queensland, 2005, [arXiv:cond-mat/0507023](https://arxiv.org/abs/cond-mat/0507023).
- [24] T. F. Gallagher, *Rydberg Atoms* (Cambridge University Press, Cambridge, England, 1994).
- [25] R. Löw, H. Weimer, J. Nipper, J. B. Balewski, B. Butscher, H. P. Büchler, and T. Pfau, *J. Phys. B* **45**, 113001 (2012).
- [26] A. Griesmaier, J. Werner, S. Hensler, J. Stuhler, and T. Pfau, *Phys. Rev. Lett.* **94**, 160401 (2005).
- [27] T. Lahaye, C. Menotti, L. Santos, M. Lewenstein, and T. Pfau, *Rep. Prog. Phys.* **72**, 126401 (2009).
- [28] L. D. Carr, D. DeMille, R. V. Krems, and J. Ye, *New J. Phys.* **11**, 055049 (2009).
- [29] H. Wu, Y. Li, Z.-B. Yang, and S.-B. Zheng, *Phys. Rev. A* **95**, 013842 (2017).
- [30] A. Gilchrist, C. W. Gardiner, and P. D. Drummond, *Phys. Rev. A* **55**, 3014 (1997).
- [31] P. Deuar and P. D. Drummond, *Phys. Rev. A* **66**, 033812 (2002).
- [32] P. Deuar and P. D. Drummond, *J. Phys. A: Math. Gen.* **39**, 1163 (2006).

- [33] I. Carusotto and Y. Castin, *Ann. Henri Poincaré* **4**, 783 (2003).
- [34] L. I. Plimak, M. K. Olsen, and M. J. Collett, *Phys. Rev. A* **64**, 025801 (2001).
- [35] M. R. Dowling, P. D. Drummond, M. J. Davis, and P. Deuar, *Phys. Rev. Lett.* **94**, 130401 (2005).
- [36] P. Deuar and P. D. Drummond, *J. Opt. B* **5**, S281 (2003).
- [37] P. Deuar and P. D. Drummond, *J. Phys. A: Math. Gen.* **39**, 2723 (2006).
- [38] P. Deuar and P. D. Drummond, *Comput. Phys. Commun.* **142**, 442 (2001).
- [39] L. Santos, G. V. Shlyapnikov, P. Zoller, and M. Lewenstein, *Phys. Rev. Lett.* **85**, 1791 (2000).
- [40] N. Henkel, R. Nath, and T. Pohl, *Phys. Rev. Lett.* **104**, 195302 (2010).
- [41] G. Pupillo, A. Micheli, M. Boninsegni, I. Lesanovsky, and P. Zoller, *Phys. Rev. Lett.* **104**, 223002 (2010).
- [42] J. Honer, H. Weimer, T. Pfau, and H. P. Büchler, *Phys. Rev. Lett.* **105**, 160404 (2010).
- [43] F. Maucher, N. Henkel, M. Saffman, W. Królikowski, S. Skupin, and T. Pohl, *Phys. Rev. Lett.* **106**, 170401 (2011).
- [44] S. Wüster, C. Ates, A. Eisfeld, and J. M. Rost, *New J. Phys.* **13**, 073044 (2011).
- [45] J. E. Johnson and S. L. Rolston, *Phys. Rev. A* **82**, 033412 (2010).
- [46] J. B. Balewski, A. T. Krupp, A. Gaj, S. Hofferberth, R. Löw, and T. Pfau, *New J. Phys.* **16**, 063012 (2014).
- [47] Y. Y. Jau, A. M. Hankin, T. Keating, I. H. Deutsch, and G. W. Biedermann, *Nat. Phys.* **12**, 71 (2016).
- [48] L. I. R. Gil, R. Mukherjee, E. M. Bridge, M. P. A. Jones, and T. Pohl, *Phys. Rev. Lett.* **112**, 103601 (2014).
- [49] T. Macri and T. Pohl, *Phys. Rev. A* **89**, 011402 (2014).
- [50] W. Li, C. Ates, and I. Lesanovsky, *Phys. Rev. Lett.* **110**, 213005 (2013).
- [51] J. Zeiher, R. van Bijnen, P. Schausz, S. Hild, J.-y. Choi, T. Pohl, I. Bloch, and C. Gross, *Nat. Phys.* **12**, 1095 (2016).
- [52] K. Singer, M. Reetz-Lamour, T. Amthor, L. G. Marcassa, and M. Weidemüller, *Phys. Rev. Lett.* **93**, 163001 (2004).
- [53] D. Tong, S. M. Farooqi, J. Stanojevic, S. Krishnan, Y. P. Zhang, R. Côté, E. E. Eyler, and P. L. Gould, *Phys. Rev. Lett.* **93**, 063001 (2004).
- [54] R. Heidemann, U. Raitzsch, V. Bendkowsky, B. Butscher, R. Löw, L. Santos, and T. Pfau, *Phys. Rev. Lett.* **99**, 163601 (2007).
- [55] T. A. Johnson, E. Urban, T. Henage, L. Isenhower, D. D. Yavuz, T. G. Walker, and M. Saffman, *Phys. Rev. Lett.* **100**, 113003 (2008).
- [56] P. J. Tanner, J. Han, E. S. Shuman, and T. F. Gallagher, *Phys. Rev. Lett.* **100**, 043002 (2008).
- [57] U. Raitzsch, V. Bendkowsky, R. Heidemann, B. Butscher, R. Löw, and T. Pfau, *Phys. Rev. Lett.* **100**, 013002 (2008).
- [58] S. Möbius, M. Genkin, S. Wüster, A. Eisfeld, and J.-M. Rost, *Phys. Rev. A* **88**, 012716 (2013).
- [59] S. Wüster, C. Ates, A. Eisfeld, and J. M. Rost, *Phys. Rev. Lett.* **105**, 053004 (2010).
- [60] S. Wüster, A. Eisfeld, and J. M. Rost, *Phys. Rev. Lett.* **106**, 153002 (2011).
- [61] K. Leonhardt, S. Wüster, and J.-M. Rost, *Phys. Rev. Lett.* **113**, 223001 (2014).
- [62] K. Leonhardt, S. Wüster, and J. M. Rost, *J. Phys. B* **50**, 054001 (2017).
- [63] N. Thaicharoen, A. Schwarzkopf, and G. Raithel, *Phys. Rev. A* **92**, 040701 (2015).
- [64] N. Thaicharoen, L. F. Gonçalves, and G. Raithel, *Phys. Rev. Lett.* **116**, 213002 (2016).
- [65] R. C. Teixeira, C. Hermann-Avigliano, T. L. Nguyen, T. Cantat-Moltrecht, J. M. Raimond, S. Haroche, S. Gleyzes, and M. Brune, *Phys. Rev. Lett.* **115**, 013001 (2015).
- [66] M. D. Lukin, M. Fleischhauer, R. Cote, L. M. Duan, D. Jaksch, J. I. Cirac, and P. Zoller, *Phys. Rev. Lett.* **87**, 037901 (2001).
- [67] P. D. Drummond and J. F. Corney, *Phys. Rev. A* **60**, R2661 (1999).
- [68] R. Ng and E. S. Sørensen, *J. Phys. A: Math. Theor.* **44**, 065305 (2011).
- [69] S. J. Carter, P. D. Drummond, M. D. Reid, and R. M. Shelby, *Phys. Rev. Lett.* **58**, 1841 (1987).
- [70] P. Deuar, *Phys. Rev. Lett.* **103**, 130402 (2009).
- [71] T. Świsłocki and P. Deuar, *J. Phys. B* **49**, 145303 (2016).
- [72] T. Aimi and M. Imada, *J. Phys. Soc. Jpn.* **76**, 113708 (2007).
- [73] P. Deuar, A. G. Sykes, D. M. Gangardt, M. J. Davis, P. D. Drummond, and K. V. Kheruntsyan, *Phys. Rev. A* **79**, 043619 (2009).
- [74] M. R. Dowling, M. J. Davis, P. D. Drummond, and J. F. Corney, *J. Comput. Phys.* **220**, 549 (2007).
- [75] The many-mode coherent states are a separable tensor product of coherent states $\otimes_{n=1}^M |\alpha_n\rangle$ with complex amplitude α_n for each of the single particle modes.
- [76] Or $2M + 1$ components if the weight Ω is included.
- [77] Almost all relevant states, including number states [106], initially have a well bounded distribution. However, as time evolves, the distribution may begin to violate this requirement.
- [78] J. J. Hope and M. K. Olsen, *Phys. Rev. Lett.* **86**, 3220 (2001).
- [79] A positive semidefinite diffusion matrix has all eigenvalues non-negative, and such a form for D can always be constructed here due to $\hat{\Lambda}$ being an analytic function of complex variables [5].
- [80] C. W. Gardiner, *Handbook of Stochastic Methods* (Springer-Verlag, Berlin, Heidelberg, 2003).
- [81] Pathological cases can occur if the distribution P_G does not decay sufficiently rapidly for large $|\alpha|, |\beta|$. This is discussed, e.g., in [23,30,31].
- [82] J. F. Corney and P. D. Drummond, *Phys. Rev. A* **68**, 063822 (2003).
- [83] Writing the potential W as an eigenvalue decomposition with orthogonal normalized eigenvectors $|j\rangle$ such that $\langle j|k\rangle = \delta_{jk}$ we have that $W = \sum_j \lambda_j |j\rangle\langle j|$ and hence $U = \sum_j |\lambda_j| |j\rangle\langle j|$. For cases when the potential W has all eigenvalues positive or all eigenvalues negative, one thus finds that $U_0 = |W_0|$. Otherwise, $U_0 > |W_0|$, and the relationship is more complicated.
- [84] The calculations in [32,37] had assumed a deterministic initial state such that $\overline{n(0)} = n(0)$, and thus did not contain the covariance contributions found here.
- [85] M. Greiner, O. Mandel, T. W. Hänsch, and I. Bloch, *Nature (London)* **419**, 51 (2002).
- [86] C. Kollath, A. M. Läuchli, and E. Altman, *Phys. Rev. Lett.* **98**, 180601 (2007).
- [87] I_1 is zero for our real coherent state initial conditions.
- [88] K. C. Younge and G. Raithel, *New J. Phys.* **11**, 043006 (2009).
- [89] U. Raitzsch, R. Heidemann, H. Weimer, B. Butscher, P. Kollmann, R. Löw, H. P. Büchler, and T. Pfau, *New J. Phys.* **11**, 055014 (2009).

- [90] J. V. Hernández and F. Robicheaux, *J. Phys. B* **41**, 195301 (2008).
- [91] J. Stanojevic and R. Côté, *Phys. Rev. A* **81**, 053406 (2010).
- [92] J. Stanojevic and R. Côté, *Phys. Rev. A* **80**, 033418 (2009).
- [93] S. Wüster, J. Stanojevic, C. Ates, T. Pohl, P. Deuar, J. F. Corney, and J. M. Rost, *Phys. Rev. A* **81**, 023406 (2010).
- [94] N. Thaicharoen, A. Schwarzkopf, and G. Raithel, *Phys. Rev. Lett.* **118**, 133401 (2017).
- [95] C. Ates, A. Eisfeld, and J. M. Rost, *New J. Phys.* **10**, 045030 (2008).
- [96] T. Amthor, M. Reetz-Lamour, S. Westermann, J. Denskat, and M. Weidemüller, *Phys. Rev. Lett.* **98**, 023004 (2007).
- [97] T. Amthor, M. Reetz-Lamour, C. Giese, and M. Weidemüller, *Phys. Rev. A* **76**, 054702 (2007).
- [98] T. Lahaye, T. Koch, B. Fröhlich, M. Fattori, J. Metz, A. Griesmaier, S. Giovanazzi, and T. Pfau, *Nature (London)* **448**, 672 (2007).
- [99] P. B. Blakie, Ph.D. thesis, University of Otago, 2001.
- [100] J. Wilkie and M. Çetinbaş, *Phys. Lett.* **A337**, 166 (2005).
- [101] G. Collett and P. Drummond, *Comput. Phys. Commun.* **142**, 219 (2001).
- [102] Our simulations were done with the high-level language xmds (online at www.xmds.org).
- [103] G. R. Dennis, J. J. Hope, and M. T. Johnsson, *Comput. Phys. Commun.* **184**, 201 (2013).
- [104] L. Arnold, *Stochastic Differential Equations: Theory and Applications* (Krieger, Malabar, FL, 1992).
- [105] Defined by a power series.
- [106] M. K. Olsen and A. S. Bradley, *Opt. Commun.* **282**, 3924 (2009).