Monte Carlo sampling of negative-temperature plasma states

John A. Krommes* and Sharadini Rath†
Plasma Physics Laboratory, Princeton University, P.O. Box 451, Princeton, New Jersey 08543, USA
(Received 12 July 2002; published 9 June 2003)

A Monte Carlo procedure is used to generate *N*-particle configurations compatible with two-temperature canonical equilibria in two dimensions, with particular attention to nonlinear plasma gyrokinetics. An unusual feature of the problem is the importance of a nontrivial probability density function $\mathcal{P}_0(\varphi)$, the probability of realizing a set φ of Fourier amplitudes associated with an ensemble of uniformly distributed, independent particles. This quantity arises because the equilibrium distribution is specified in terms of φ, whereas the sampling procedure naturally produces particle states Γ; φ and Γ are related via a gyrokinetic Poisson equation, highly nonlinear in its dependence on Γ. Expansion and asymptotic methods are used to calculate $\mathcal{P}_0(\varphi)$ analytically; excellent agreement is found between the large-*N* asymptotic result and a direct numerical calculation. The algorithm is tested by successfully generating a variety of states of both positive and negative temperature, including ones in which either the longest- or shortest-wavelength modes are excited to relatively large amplitudes.

DOI: 10.1103/PhysRevE.67.066402 PACS number(s): 52.65.Pp, 02.50.Ng, 05.10.Ln

I. INTRODUCTION

In the present paper, we will show how to generate *N*-particle configurations compatible with two-temperature canonical spectral equilibria in two dimensions. The method could be used for testing nonlinear gyrokinetic simulations of plasmas; more generally, the problem raises interesting issues involving equilibrium statistical mechanics and asymptotic analysis.

Particle simulations (often called molecular-dynamics experiments in the literature on neutral fluids) are widely used for studying the equilibrium and nonequilibrium behavior of nonlinear systems [1–4]. Such methods are quite natural when the discreteness of the many-body system is important, but particle simulations have also been employed as viable alternatives to more conventional spectral or finite-difference approaches to the integration of hyperbolic partial differential equations in the continuum approximation. This is particularly true in plasma physics [3], where collective nonlinear, neutral-fluid-like phenomena—essentially independent of the details of the microscopic velocity distribution—are often driven by wave-particle interactions strongly dependent on such details. The conservative nature of the Vlasov or closely related gyrokinetic equation [5] is difficult to handle with conventional techniques because phase-space elements can be stretched and otherwise distorted to scales that are arbitrarily small (in the absence of collisions [6]). Particle simulation deals with this situation efficiently by following the particle or gyrocenter trajectories exactly in an electromagnetic field that is coarse grained to a finite resolution in space.

Recent interest in plasma physics has focused on nonequilibrium gyrokinetic phenomena [7,8]. The nonlinear gyrokinetic equation [5], appropriate for low-frequency, long-wavelength fluctuations, is derivable from the magnetized Vlasov equation by Hamiltonian transformations[1] [9] that analytically remove the rapid gyromotion in favor of appropriately defined gyrocenters whose characteristic equations of motion are the generalizations of the familiar drift equations to the case of finite ion gyroradius. In the quasineutrality condition that replaces Poisson's equation, a key role is played by the polarization charge density of the ions (see Appendix A), proportional to the *z* component of the vorticity of the cross-field $\mathbf{E}\times\mathbf{B}$ velocity. The resulting nonlinear equations display phenomena closely related to the well known, unusual behavior of two-dimensional (2D) neutral fluids [10]. For example, the Hasegawa-Mima equation [8,11], an important limit [9] of the gyrokinetic system that is appropriate for adiabatic electrons and vanishing ion temperature, possesses both an energy- and an enstrophy-related invariant and therefore displays all of the interesting phenomena associated with two-temperature[2] canonical (Gibbsian) equilibria [10] of the (finitely truncated) set of Fourier amplitudes of the electrostatic potential φ (or ion gyrocenter density, to within a simple wave-number-dependent scale factor). In the presence of dissipation or antidissipation (e.g., when nonadiabatic electron response is reinstated), those invariants are broken, but the tendency of the nonlinear terms to relax fluctuations to thermal equilibrium is still manifested through the dual cascades [12]—of energy *E*, toward the long wavelengths, and of (potential) enstrophy Ω, toward the short wavelengths—that can be excited [13] for forcing at intermediate wave numbers.

A standard test in the development of gyrokinetic particle simulations should therefore be to check that the Hasegawa-

*Electronic address: krommes@princeton.edu
†Present address: 20 Akashganga, Ganeshkhind, Pune 411 007, India.

[1]A review of some of the formal aspects of the derivation of the nonlinear gyrokinetic Poisson system can be found in Appendix C of Ref. [8].

[2]The ''temperatures'' referred to here should not be confused with the thermal particle temperature associated with a heat bath. See Sec. II for a review of two-parameter spectral equilibria.

Mima limit relaxes to the two-temperature canonical spectrum based on the initial values of energy and enstrophy. Somewhat surprisingly, this has not been done, although the analogous exercise is standard for Vlasov codes [3], where one verifies the thermal-equilibrium wave-number spectrum $\langle \delta E^2 \rangle(\mathbf{k})/8\pi = \frac{1}{2}T/(1+k^2\lambda_D^2)$ (here, $\delta\mathbf{E}$ is the electric field fluctuation, T is the thermal temperature, and λ_D is the Debye length), and has also been performed for simulations of the full gyrokinetic equation with finite $k_\parallel$ [14]. (The thermal-equilibrium fluctuation properties of the gyrokinetic system are discussed in Refs. [15–17].) One explanation for this omission involves the difficulty of achieving random initial particle distributions with specified E and Ω. It is, of course, possible to begin with an arbitrary initialization (e.g., particle positions that are completely independent and uniformly distributed, or alternatively a ''quiet start'' [18]), calculate the associated E and Ω and the predicted equilibrium spectrum, then check for relaxation toward that equilibrium, and that certainly provides a nontrivial test of the code. However, reliance on just one or two standard initializations and the arbitrariness of any particular initial state imply that one has no systematic way of exploring the extremes of the E-Ω space, including, in particular, the interesting regimes of negative temperature. For example, it is noteworthy that the simplest random 2D initialization, in which the particle positions are sampled from a homogeneous, correlation-free distribution, has mean energy and enstrophy compatible with a canonical distribution with *positive* (and equal) temperatures. (We will derive this result in Sec. II C.)

In addition to the gyrokinetic initialization problem, the generation of negative-temperature particle states is interesting in its own right as a nontrivial problem in statistics. Therefore, in this paper, we propose [46] and theoretically analyze a method of generating 2D particle-state realizations of canonical equilibria for arbitrary values of the invariants (equivalently, for arbitrary temperatures). The (Monte Carlo) procedure employs the Markov-chain algorithm of Metropolis *et al.* ([19], henceforth called MRRTT) originally (and still [20]) used for investigating the thermodynamic properties of dense liquids. Although the present application is essentially straightforward, it does not appear to have been previously used in this particular context. The subtlety in the calculation is that in the present case the random variables that are canonically distributed are the Fourier amplitudes, whereas one desires realizations of particle positions, which are related to the Fourier amplitudes via a nontrivial nonlinear functional dependence. An additional complication is that typically there are many more particles than retained Fourier amplitudes, so the relation between the particle states and the potentials is many to one. In the standard application to statistical mechanics, on the other hand, the natural variables of the Gibbs distribution are just the particle phase-space coordinates themselves; no functional relation is involved. Thus, in conventional MRRTT the successive states of the Markov chain are used for the calculation of analytically intractable ensemble averages over a highly non-Gaussian distribution. In the present case, the equilibrium spectrum is purely Gaussian in the Fourier amplitudes, so analytical calculations of arbitrary (static) moments of the amplitudes are straightforward; however, the determination of compatible particle realizations is nontrivial.

For notational convenience, we will denote by φ both the field $\varphi(\mathbf{x})$ and the set of $\mathcal{M}$ retained Fourier amplitudes $\varphi \equiv \{\varphi_{\mathbf{k}(\nu)} | \nu = 1,...,\mathcal{M}\}$, where ν is an arbitrary labeling of the wave numbers.

The necessity for dealing simultaneously with two sets of random variables, namely, the particle states Γ and the Fourier amplitudes φ, means that a nontrivial transition probability enters the formalism. In standard Monte Carlo algorithms that deal just with particle probabilities, the basic function $T_0(\Gamma'|\Gamma)$ describes the probability of proposing state Γ', given initial state Γ. Usually this function is just a constant. However, in the present application we begin with a particle state Γ [which corresponds to the set of Fourier amplitudes $\varphi(\Gamma)$] and, effectively, propose an *amplitude* state $\varphi'(\Gamma')$. The relevant transition probability $\mathcal{T}_0(\varphi'|\varphi)$ can be related to the probability density function (PDF) $\mathcal{P}_0(\varphi)$ of achieving a particular set of Fourier amplitudes by distributing particles uniformly on a lattice; this function depends nontrivially on φ. We devote considerable effort in calculating $\mathcal{P}_0$ asymptotically and in analyzing its role in the appropriately modified Monte Carlo algorithm.

The organization of this paper is as follows. We devote Sec. II to a review of the form and properties of the standard two-temperature canonical equilibria. In Sec. III, we review the algorithm of MRRTT and describe how to generalize it for application to the initialization problem. In Sec. IV, we generate a variety of two-temperature particle states in order to demonstrate the viability of the method. We summarize and discuss our results in Sec. V. Various details are relegated to the Appendixes. In Appendix A, we present for completeness a brief derivation of the Hasegawa-Mima equation using gyrokinetic methods. In Appendix B, we derive a formula for $\mathcal{P}_0(\varphi)$ and present a few exact results. Finally, we devote Appendix C to a presentation of various approximate calculations of $\mathcal{P}_0$ valid for a large number of particles.

II. TWO-TEMPERATURE EQUILIBRIA

In essence, we are concerned with statistical sampling from a particular, somewhat unusual canonical probability distribution. In this section, we provide the necessary background material. Our conventions for Fourier transforms are introduced in Sec. II A (see also Appendix A of Ref. [8]). Then in Sec. II B, we introduce the two-temperature equilibria relevant to the Hasegawa-Mima problem and review their properties.

A. Fourier transform conventions

We work in a 2D box of sides L_x and L_y (area $V = L_x L_y$), on which periodic boundary conditions are imposed. The electrostatic potential is resolved onto a rectangular lattice of $M_{\rm tot} = M_x \times M_y$ points; e.g., $x_j = j\Delta x$, $\Delta x \doteq L_x/M_x$, $j = 0, 1,..., M_x - 1$ (we use the symbol $\doteq$ for definitions). In practice, we consider a square box with $L_x = L_y = L$, $M_x = M_y$; however, we sometimes retain the x or y labels for pedagogical purposes. The area of the fundamental

cell is $\Delta V=\Delta x\Delta y$. The associated Fourier components obey, e.g., $k_{n_x}\equiv k_x=n_x\delta k_x$, where $\delta k_x\equiv k_{\min}\doteq 2\pi/L_x$. Wavenumber magnitudes are denoted as $k\doteq|\mathbf{k}|=(k_x^2+k_y^2)^{1/2}$. For arbitrary function $A(\mathbf{x})$, the fundamental Fourier conventions are

$$A_{\mathbf{k}}=V^{-1}\int d\mathbf{x}e^{-i\mathbf{k}\cdot\mathbf{x}}A(\mathbf{x}), \tag{1a}$$

$$A(\mathbf{x})=\sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{x}}A_{\mathbf{k}}. \tag{1b}$$

For points $\mathbf{x_j}$ on the lattice, we interpret

$$\int d\mathbf{x}\cdots\equiv\sum_{\mathbf{j}}\Delta V\cdots \tag{2}$$

and pair (1) becomes the discrete Fourier transform, with $\mathbf{k_n}\cdot\mathbf{x_j}=2\pi(n_xj_x/M_x+n_yj_y/M_y)$. [We write sets of Cartesian integers as boldfaced vectors, e.g., $\mathbf{n}\doteq(n_x,n_y)$.] We sometimes write $A_{\mathbf{k_n}}\equiv A_{\mathbf{n}}$; one has

$$A_{-\mathbf{n}}=A_{\mathbf{M}-\mathbf{n}}=A_{\mathbf{n}}^*, \tag{3}$$

the last equality holding for functions that are real in $\mathbf{x}$ space. Although all operations on the lattice are discrete, we often find the integral form of Eq. (2) to be a convenient shorthand.

B. Two-temperature canonical equilibria

We consider a conservative nonlinear system of coupled Fourier modes $\psi_{\mathbf{k}}$. In the derivation of such equations from continuum equations locally nonlinear in $\mathbf{x}$ space, the mode coupling arises from the Fourier convolution theorem and thus involves an infinite number of Fourier amplitudes. We consider instead a system truncated to a finite number $\mathcal{M}$ of amplitudes. This corresponds to the actual situation in the simulations and is also required theoretically in order that a conventional statistical dynamics can be introduced [10]. Generally, the truncation is spherical, $k_{\min}\leqslant k\leqslant k_{\max}$, so $\mathcal{M}<M_{\rm tot}$. It is assumed that under such truncations two constants of the motion, the energy $\hat{E}$ and the enstrophy $\hat{\Omega}$, are preserved. The hat denotes a function of the underlying random Fourier amplitudes: $\hat{E}\equiv\hat{E}(\psi)$ and $\hat{\Omega}\equiv\hat{\Omega}(\psi)$, where the braces denote the collection of all retained modes: $\psi\equiv\{\psi_{\mathbf{k}(\nu)}|\nu=1,...,\mathcal{M}\}$. Functions without hats will denote the ensemble average, e.g., $E\doteq\langle\hat{E}\rangle$. It is then well known [8,10,21] that the real and imaginary parts of the $\psi_{\mathbf{k}}$ can be used as independent variables in standard statistical-mechanics arguments that predict relaxation of arbitrary perturbations to realizations drawn from a microcanonical ensemble. In practice, the Gibbs distribution is used more frequently:

$$P(\psi)=Z^{-1}\exp[-\alpha\hat{E}(\psi)-\beta\hat{\Omega}(\psi)], \tag{4}$$

where Z is the appropriate normalization integral. The parameters α and β serve as inverse temperatures for energy and enstrophy, respectively, and are functions of the ensemble averaged E and Ω. For some ratios of Ω/E, either α or β can be negative, as we will review in detail shortly. In states of negative α, the longest-wavelength modes are excited to relatively large levels; in states of negative β, the shortest-wavelength modes are so excited.

In the application to Hasegawa-Mima dynamics, we may choose ψ to be the appropriately dimensionless electrostatic potential φ. The Hasegawa-Mima equation [11] is briefly rederived for completeness in Appendix A; it is

$$(1-\nabla_\perp^2)\partial_t\varphi(\mathbf{x},t)+V_*\partial_y\varphi+\mathbf{V}_E\cdot\boldsymbol{\nabla}(-\nabla_\perp^2\varphi)=0, \tag{5}$$

where V_* is the diamagnetic drift velocity (considered to be a constant in this approximation), $\mathbf{V}_E\doteq\mathbf{E}\times\hat{\mathbf{z}}$, $\mathbf{E}\doteq-\boldsymbol{\nabla}\varphi$, and perpendicular means with respect to the constant magnetic field $\mathbf{B}=B\hat{\mathbf{z}}$. We consider the 2D (in $\mathbf{x}_\perp$) version of this equation, although 3D generalizations are possible. Upon Fourier transformation in space, Eq. (5) becomes

$$(1+k^2)\partial_t\varphi_{\mathbf{k}}+i\omega_*(k_y)\varphi_{\mathbf{k}}=\frac{1}{2}\sum_{\mathbf{k}+\mathbf{p}+\mathbf{q}=\mathbf{0}}\hat{\mathbf{z}}\cdot(\mathbf{p}\times\mathbf{q})\times\left(\frac{\chi_{\mathbf{q}}^*-\chi_{\mathbf{p}}^*}{1+\chi_{\mathbf{k}}}\right)\varphi_{\mathbf{p}}^*\varphi_{\mathbf{q}}^*, \tag{6}$$

where $\mathbf{k}\equiv\mathbf{k}_\perp$, $\omega_*(k_y)\doteq k_yV_*$, and $\chi_{\mathbf{k}}\doteq k_\perp^2$. This equation is conservative. The more physically relevant model, in which both ω_* and χ are replaced by complex quantities including dissipative effects, is called the Terry-Horton equation [22,23]. See Sec. V for further remarks on that equation.

The quadratic invariants of the Hasegawa-Mima equation are

$$\begin{pmatrix}\hat{E}(\varphi)\\ \hat{\Omega}(\varphi)\end{pmatrix}=\sum_{\mathbf{k}}\begin{pmatrix}1\\ k^2\end{pmatrix}\hat{E}_{\mathbf{k}}, \tag{7a}$$

where

$$\hat{E}_{\mathbf{k}}\doteq\tfrac{1}{2}(1+k^2)|\hat{\varphi}_{\mathbf{k}}|^2. \tag{7b}$$

The term 1 in the factor $(1+k^2)$ describes the adiabatic response of the electrons, which stream rapidly along the magnetic field lines and tend to short out charge fluctuations. If that term is ignored (equivalently, if one considers the short-wavelength limit [47]), the resulting equation is formally identical to the two-dimensional Navier-Stokes equation, with $\nabla_\perp^2\varphi$ playing the role of the z component of the vorticity.

Quantities (7) are invariant when $\mathbf{k}$ is summed over all modes out to ∞. They remain invariant if one removes from the convolution sum in Eq. (6) all triad interactions having a leg with magnitude larger than some $k_{\max}$; that corresponds to the spherical truncation mentioned earlier. Later, we will consider other truncations and/or weightings. All of those can be embraced by introducing a non-negative weight function $w_{\mathbf{k}}$ and generalizing Eq. (7a) to

$$\begin{pmatrix} \hat{E}(\varphi) \\ \hat{\Omega}(\varphi) \end{pmatrix} = \sum_{\mathbf{k}} w_{\mathbf{k}} \begin{pmatrix} 1 \\ k^2 \end{pmatrix} \hat{E}_{\mathbf{k}}. \tag{8}$$

With this definition, the effective number of modes is $\mathcal{M} = \Sigma_{\mathbf{k}} w_{\mathbf{k}}$. For modes interior to the boundaries of the truncated $\mathbf{k}$ space, we consider only $w_{\mathbf{k}}=1$; however, the weightings of the edge or corner points may differ from 1, as we will discuss later.

The thermal-equilibrium wave-number spectrum can now be shown to be

$$E_{\mathbf{k}} \doteq \langle \hat{E}_{\mathbf{k}} \rangle = \frac{1}{2} \left(\frac{1}{\alpha + \beta k^2} \right). \tag{9}$$

It is worth giving the derivation of this result in detail in order to discuss and justify an annoying factor of 2. With r and i denoting real and imaginary parts and $\varphi^{(\nu)} \equiv \varphi_{\mathbf{k}^{(\nu)}}$, one has by definition

$$\langle \hat{E}_{\mathbf{k}} \rangle = \int \prod_{\nu=1}^{\mathcal{M}'} d\varphi_r^{(\nu)} d\varphi_i^{(\nu)} (1+k^2) |\varphi_{\mathbf{k}}|^2 \times Z^{-1} \exp\left(-\sum_{\mathbf{k}}{}' \, w'_{\mathbf{k}} (1+\overline{k^2})(\alpha + \beta \overline{k^2}) |\varphi_{\mathbf{k}}|^2 \right); \tag{10}$$

of course, $|\varphi_{\mathbf{k}}|^2 = \varphi_{k,r}^2 + \varphi_{k,i}^2$. We have observed that $\varphi_{\mathbf{k}}$ and $\varphi_{-\mathbf{k}}$ are not independent, since by reality of $\varphi(\mathbf{x})$ one has $\varphi_{-\mathbf{k}} = \varphi_{\mathbf{k}}^*$. Therefore, the factor of $\frac{1}{2}$ in definition (7b) of the invariants has been eliminated in the exponent of Eq. (10) by replacing the unrestricted wave-number sum in Eq. (7a) by the appropriate $\Sigma'_{\mathbf{k}}$ over the half space of $\mathcal{M}' \doteq \frac{1}{2}\mathcal{M}$ independent modes. (We allow for the possibility that the appropriate weight function $w'_{\mathbf{k}}$ for $\Sigma'_{\mathbf{k}}$ may differ from $w_{\mathbf{k}}$ on the boundary of the $\mathbf{k}$ space.) Result (9) then follows readily, at least for the interior modes, by performing a simple Gaussian integral. (Note that the partition function Z factors.)

The discrete nature of the Fourier transform introduces subtlety into the evaluation of the invariants. It is numerically convenient to work with a square truncation, i.e., to sum over all retained modes, as this eliminates a time-consuming test to determine whether a mode should be included. However, because of the symmetry properties (3) the special modes $(n_x, n_y) \in \{(0,0), (\frac{1}{2}M_x, 0), (0, \frac{1}{2}M_y), (\frac{1}{2}M_x, \frac{1}{2}M_y)\}$ are real. Since we are interested in fluctuations, we exclude the (0, 0) mode; however, the other modes must be counted appropriately. Now since the imaginary parts of those modes vanish identically, those parts are not available as independent coordinates for the canonical distribution. That, thus, has the schematic form (with x and y referring here to the real and imaginary parts of any $E_{\mathbf{k}}$)

$$\mathcal{P}^{(\mathcal{M})} \sim \exp\left[-\left(\frac{x_1^2 + y_1^2}{2\sigma_1^2} \right) - \left(\frac{x_2^2}{2\sigma_2^2} \right) \right], \tag{11}$$

where the subscript 1 refers to an interior mode, 2 refers to a boundary mode, $x_1^2 + y_1^2 = \hat{E}_{\mathbf{k}_1}$, $\sigma_1^2 \doteq [4(\alpha + \beta k_1^2)]^{-1}$, and σ_2 must be determined such that the spectrum $E_{\mathbf{k}}$ is a smooth function of $\mathbf{k}$. Since $E_{\mathbf{k}_1} = \langle (x_1^2 + y_1^2) \rangle = 2\sigma_1^2$ and $E_{\mathbf{k}_2} = \langle x_2^2 \rangle = \sigma_2^2$, one deduces that the real modes must be weighted by a factor of $\frac{1}{2}$ in the summation Σ' that defines the invariants of the finite, discrete system. The corresponding weight function $w'_{\mathbf{k}}$ is thus asymmetric when considered over the entire square; however, the reflection symmetries on the lattice guarantee that the four corner points are identical and that edge points are identical to appropriate points on the facing edge. Theoretically, then, one can replace $\Sigma'_{\mathbf{k}} w'_{\mathbf{k}}$ by $\frac{1}{2}\Sigma_{\mathbf{k}} w_{\mathbf{k}}$, where $w_{\mathbf{k}}$ corresponds to a weighting of the interior modes by 1, the corner points by $\frac{1}{4}$ and the other edge points by $\frac{1}{2}$. We will call this the *weighted truncation*. Note that numerically it is more efficient to employ the asymmetric weights $w'_{\mathbf{k}}$; however, the $w_{\mathbf{k}}$'s are easier to work with theoretically.[3]

Spectrum (9) is identical to that for the 2D neutral fluid, which has been discussed in depth by Kraichnan [25]. For our later work in Sec. III, it is necessary to record here[4] the portion of Kraichnan's analysis concerning the allowable values of α and β and the relation between $\{E, \Omega\}$ and $\{\alpha, \beta\}$ in a notation that emphasizes the discrete nature of the spectrum and allows for arbitrary weight functions $w_{\mathbf{k}}$. It is useful to couch the relations in terms of the energy and enstrophy per mode,

$$\overline{E} \doteq E/\mathcal{M}, \quad \overline{\Omega} \doteq \Omega/\mathcal{M}. \tag{12}$$

It is also useful to define the dimensionless parameters

$$\overline{\alpha} \doteq \alpha \overline{E}, \quad \overline{\beta} \doteq \beta \overline{E}, \tag{13}$$

and the ratio

$$\hat{\alpha} \doteq \alpha/\beta = \overline{\alpha}/\overline{\beta}. \tag{14}$$

(This new use of the hat should cause no confusion in context.) One can then write definitions (7) in the form

$$\begin{pmatrix} \overline{E} \\ \overline{\Omega} \end{pmatrix} = \frac{1}{2\beta} \left\langle \begin{pmatrix} 1 \\ k^2 \end{pmatrix} \left(\frac{1}{\hat{\alpha} + k^2} \right) \right\rangle_{\mathbf{k}}, \tag{15}$$

where the notation $\langle \cdots \rangle_{\mathbf{k}}$ denotes the average over the discrete, truncated wave-number spectrum:

$$\langle A \rangle_{\mathbf{k}} \doteq \frac{1}{\mathcal{M}} \sum_{\mathbf{k}} w_{\mathbf{k}} A_{\mathbf{k}}. \tag{16}$$

It is useful to define the ratio of enstrophy to energy, a dimensionless quantity that should be thought of as the square of a (dimensionless) wave number κ,

$$\kappa^2 \doteq \Omega/E = \overline{\Omega}/\overline{E}. \tag{17}$$

[3]Figures illustrating this discussion can be found in Ref. [24].

[4]A more concise version of this discussion can be found in Sec. 3.7.2 of Ref. [8].

Then, upon adding and subtracting $\hat{\alpha}$ to the numerator of expression (15) for $\overline{\Omega}$, one finds the convenient expression

$$\kappa^2 = (2\overline{\beta})^{-1} - \hat{\alpha}, \tag{18}$$

where

$$\overline{\beta} = \frac{1}{2}\left\langle \frac{1}{\hat{\alpha}+k^2} \right\rangle_{\mathbf{k}}. \tag{19}$$

Equation (18) can be used to prove an important constraint that will be very useful in the subsequent Monte Carlo calculations. Define

$$\overline{U} \doteq 2(\alpha\overline{E} + \beta\overline{\Omega}). \tag{20}$$

One has

$$\overline{U} = 2\beta\overline{E}(\alpha/\beta + \overline{\Omega}/\overline{E}) \tag{21a}$$

$$= 2\overline{\beta}(\hat{\alpha} + \kappa^2). \tag{21b}$$

Upon using Eq. (18), one finally finds

$$\overline{U} = 1. \tag{22}$$

This is a simple generalization of the result that for a 1D PDF of the form $P(x) \propto \exp(-\alpha x^2)$, one has $2\alpha\langle x^2\rangle = 1$. Its importance is that for a general non-Gaussian distribution one will have $\overline{U} \neq 1$, so the approach of $\overline{U}$ to 1 can be used as a convenient (and, in practice, very sensitive) diagnostic of the convergence of the Monte Carlo Markov chain to the desired asymptotic equilibrium distribution; see Sec. III B.

The parameter space can now be analyzed by demanding that $\overline{E}$, $\overline{\Omega}$, and $E_{\mathbf{k}}$ be non-negative. Considered as a function of $\hat{\alpha}$, $E_{\mathbf{k}}$ is singular at $\hat{\alpha} = -k_{\min}^2$ and $\hat{\alpha} = -k_{\max}^2$, and one can determine that the region $-k_{\max}^2 < \hat{\alpha} < -k_{\min}^2$ is forbidden since one or more of the $E_{\mathbf{k}}$ would be negative. To analyze the behavior in the vicinity of $\hat{\alpha} = -k_{\min}^2$ we write $\hat{\alpha} = -k_{\min}^2 + \epsilon/\mathcal{M}$. Then, $\overline{\beta} = O(\epsilon^{-1}) \to +\infty$ as $\epsilon \to 0_+$. For fixed $\overline{E}$, which we always assume in considering the various limiting cases, one sees that also $\beta \to +\infty$. The behavior of $\overline{\alpha} \doteq \alpha\overline{E}$ follows from $\overline{\alpha} = \hat{\alpha}\overline{\beta} = (-k_{\min}^2 + \epsilon/\mathcal{M})\overline{\beta} \approx -k_{\min}^2\overline{\beta} \to -\infty$. Also, from Eq. (18) $\kappa^2 \to -\hat{\alpha} \to k_{\min}^2$. Symmetrical behavior ensues in the vicinity of $\hat{\alpha} = -k_{\max}^2 - \epsilon/\mathcal{M}$, with the roles of α and β as well as $k_{\min}$ and $k_{\max}$ reversed. The other interesting points are $\hat{\alpha} = 0$ and $\hat{\alpha} = \pm\infty$. Define the special wave numbers k_a and k_b according to

$$k_a^2 \doteq \langle k^{-2}\rangle_{\mathbf{k}}^{-1}, \tag{23a}$$

$$k_b^2 = \langle k^2\rangle_{\mathbf{k}}. \tag{23b}$$

[That $k_b \geqslant k_a$ is a consequence of a Schwartz inequality applied to the identity $\langle (k^2)(k^{-2})\rangle = 1$.] Then at $\hat{\alpha} = 0$ one finds $\overline{E} = (2\beta k_a^2)^{-1}$, $\overline{\Omega} = (2\beta)^{-1}$, and from Eq. (17) $\kappa^2 = k_a^2$. As $\hat{\alpha} \to +\infty$ one obtains $\overline{E} = (2\alpha)^{-1}$, $\overline{\Omega} = (2\alpha)^{-1}k_b^2$, and $\kappa^2 = k_b^2$. Since for fixed $\overline{E}$ α remains finite, we see that

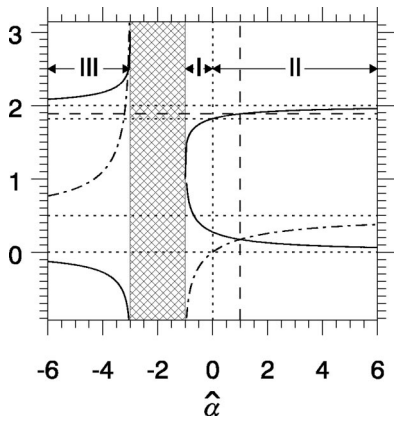

FIG. 1. Important quantities for two-temperature equilibria, plotted vs $\hat{\alpha} \doteq \alpha/\beta$. The crosshatched region is the forbidden zone $-k_{\max}^2 < \hat{\alpha} < -k_{\min}^2$. For the purpose of illustration, $k_{\min} = 1$ and $k_{\max} = \sqrt{3}$ are used in this and the following figure. Solid line, $\kappa^2(\hat{\alpha})$; dash-dotted line, $\overline{\alpha}(\hat{\alpha})$; triple dash-dotted line, $\overline{\beta}(\hat{\alpha})$. The dashed lines indicate the special noise case $\hat{\alpha} = 1$, $\kappa^2 = \kappa_1^2$. The horizontal dotted lines indicate, from bottom to top, the special cases $\kappa^2 = (0, \frac{1}{2}, k_a^2, k_b^2)$. The vertical dotted line indicates the boundary $\hat{\alpha} = 0$ between regimes I and II.

$\beta \to 0_+$ as $\hat{\alpha} \to +\infty$. The point $\hat{\alpha} = -\infty$ is obtained continuously from $\hat{\alpha} = +\infty$ as β passes continuously through 0 from above. One thus identifies three regimes:

$$\begin{aligned}
&\text{regime I} \quad k_{\min}^2 \leqslant \kappa^2 \leqslant k_a^2,\\
&\qquad -k_{\min}^2 \leqslant \hat{\alpha} \leqslant 0,\\
&\qquad -\infty \leqslant \overline{\alpha} \leqslant 0,\\
&\qquad \infty \geqslant \overline{\beta} \geqslant (2k_a^2)^{-1};\\
&\text{regime II} \quad k_a^2 \leqslant \kappa^2 \leqslant k_b^2,\\
&\qquad 0 \leqslant \hat{\alpha} < \infty,\\
&\qquad 0 \leqslant \overline{\alpha} \leqslant 1/2,\\
&\qquad (2k_a^2)^{-1} \geqslant \overline{\beta} \geqslant 0;\\
&\text{regime III} \quad k_b^2 \leqslant \kappa^2 \leqslant k_{\max}^2,\\
&\qquad -\infty < \hat{\alpha} \leqslant -k_{\max}^2,\\
&\qquad 1/2 \leqslant \overline{\alpha} \leqslant \infty,\\
&\qquad 0 \geqslant \overline{\beta} \geqslant -\infty.
\end{aligned}$$

The qualitative features of this behavior are summarized in Figs. 1 and 2. Those figures should not be used for precise quantitative work, since they actually plot the approximation obtained by assuming that the spectrum is dense and spherically truncated [25]:

$$\mathcal{M} \approx \pi(k_{\max}^2 - k_{\min}^2), \tag{24a}$$

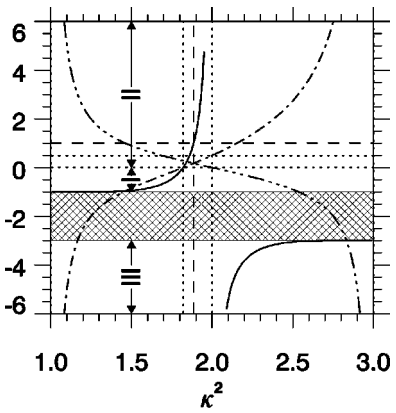

FIG. 2. Important quantities for two-temperature equilibria, plotted vs $\kappa^2 \doteq \Omega/E$. Crosshatched region: forbidden zone $-k_{\max}^2 < \hat{\alpha} < -k_{\min}^2$. Solid line, $\hat{\alpha}(\kappa^2)$; dash-dotted line, $\overline{\alpha}(\kappa^2)$; triple dash-dotted line, $\overline{\beta}(\kappa^2)$. The dashed lines indicated the special noise case $\hat{\alpha}=1$, $\kappa^2=\kappa_1^2$. The vertical dotted lines indicate, from left to right, k_a^2 and k_b^2. The horizontal dotted lines indicate, from bottom to top, the special cases $\overline{\beta}(k_b^2)=0$ and $\overline{\alpha}(k_a^2)=\frac{1}{2}$.

$$\overline{\beta} \approx \frac{1}{2}\ln\left(\frac{\hat{\alpha}+k_{\max}^2}{\overline{\alpha}+k_{\min}^2}\right) \Big/ (k_{\max}^2-k_{\min}^2), \tag{24b}$$

$$k_a^2 \approx \frac{k_{\max}^2-k_{\min}^2}{\ln(k_{\max}^2/k_{\min}^2)}, \tag{24c}$$

$$k_b^2 \approx \tfrac{1}{2}(k_{\min}^2+k_{\max}^2). \tag{24d}$$

Regime II is the most intuitively familiar regime of positive temperatures, bounded on the left by the enstrophy-equipartition state $\hat{\alpha}=0$, $\Omega_{\mathbf{k}}=\overline{\Omega}=(2\beta)^{-1}=k_a^2\overline{E}$ and on the right by the energy-equipartition state $\beta=0$, $E_{\mathbf{k}}=\overline{E}=(2\alpha)^{-1}=\overline{\Omega}/k_b^2$. Regime I corresponds to negative-α states; symmetrically, regime III corresponds to negative-β states. States with highly negative α have the longest-wavelength modes excited to high levels. For states with highly negative β, the excitation is concentrated at the shortest-wavelength modes.

The precise numerical values of the characteristic wave numbers k_a and k_b are not accessible from the continuum approximations (24); they must be determined numerically as a function of the number of discrete modes. We generally consider $M_x=M_y=2^m$. In Tables I–III, we tabulate the values of k_a, k_b, and the additional important wave number κ_1 [defined later; see Eq. (37b)] normalized to $k_{\min}$ for representative m's and the truncations defined by

$$\text{spherical} \quad w_{\mathbf{k}}=\begin{cases} 1, & 1\leqslant k\leqslant M_x/2 \\ 0, & \text{otherwise,} \end{cases}$$

$$\text{square} \quad w_{\mathbf{k}}=\begin{cases} 1 & \begin{cases} -M_x/2\leqslant k_x\leqslant M_x/2 \\ -M_y/2\leqslant k_y\leqslant M_y/2 \\ \mathbf{k}\neq\mathbf{0} \end{cases} \\ 0, & \text{otherwise,} \end{cases}$$

$$\text{weighted} \quad w_{\mathbf{k}}=\begin{cases} d_{\mathbf{k}} & \begin{cases} -M_x/2\leqslant k_x\leqslant M_x/2 \\ -M_y/2\leqslant k_y\leqslant M_y/2 \\ \mathbf{k}\neq\mathbf{0} \end{cases} \\ 0, & \text{otherwise,} \end{cases}$$

where

$$d_{\mathbf{k}} \doteq \begin{cases} 1 & (\text{interior point}) \\ 1/2 & (\text{edge but not corner}) \\ 1/4 & (\text{corner}). \end{cases} \tag{25}$$

For the spherical truncation one has $k_{\max}/k_{\min}=\frac{1}{2}M_x=2^{m-1}$; for the other truncations $k_{\max}$ is $\sqrt{2}$ times larger.

For given $\overline{E}$ and $\overline{\Omega}$, the associated α and β are determined as follows. We replace the set $\{\overline{E},\overline{\Omega}\}$ by $\{\overline{E},\kappa^2\}$. Relation (18),

$$\kappa^2=\kappa^2(\hat{\alpha})=[2\overline{\beta}(\hat{\alpha})]^{-1}-\hat{\alpha}, \tag{26}$$

can be inverted (numerically, in practice) to give $\hat{\alpha}(\kappa^2)$. The function $\overline{\beta}(\kappa^2)$ [Eq. (19)] is then known from Eq. (18) as

$$\overline{\beta}(\kappa^2)=\tfrac{1}{2}[\hat{\alpha}(\kappa^2)+\kappa^2]^{-1}. \tag{27}$$

The function $\overline{\alpha}(\kappa^2)\doteq\alpha\overline{E}$ follows from

$$\overline{\alpha}(\kappa^2)=\hat{\alpha}(\kappa^2)\overline{\beta}(\kappa^2). \tag{28}$$

Finally, the absolute inverse temperatures follow from Eqs. (13).

TABLE I. Important parameters for a discrete spectrum in the spherical truncation. The parenthesized numbers are the relative differences between the discrete results and the continuum approximation described by Eqs. (24). All k's are normalized to $k_{\min}$.

m	No. of modes		k_a		k_1		$\overline{\alpha}$		k_b		$k_{\max}$
2	12	(−0.2146)	1.309	(0.123 55)	1.517	(0.038 896)	0.477	(−0.003 653 6)	1.528	(0.035 098)	2
3	48	(−0.018 252)	2.069	(0.124 04)	2.764	(0.033 887)	0.431	(−0.009 457 7)	2.828	(0.030 776)	4
4	196	(0.009 797 6)	3.546	(0.097 497)	5.219	(0.017 103)	0.318	(−0.012 419)	5.617	(0.014 927)	8
5	796	(0.006 414 7)	6.312	(0.074 295)	9.496	(0.007 694 5)	0.172	(−0.009 888 3)	11.270	(0.005 868 3)	16
6	3208	(0.001 823 3)	11.477	(0.058 483)	16.768	(0.003 280 7)	0.072	(−0.005 591 1)	22.603	(0.001 554 6)	32
7	12 852	(0.000 997 66)	21.157	(0.048 752)	29.489	(0.002 064 6)	0.026	(−0.003 904)	45.230	(0.000 662 21)	64
8	51 432	(0.000 713 81)	39.437	(0.041 868)	52.501	(0.001 531)	0.008	(−0.003 003 5)	90.476	(0.000 897 74)	128

TABLE II. Important parameters for a discrete spectrum in the square truncation.

m	No. of modes		k_a		k_1		$\overline{\alpha}$		k_b	$k_{\max}$
2	24	(-0.5)	1.624	$(-0.094\,167)$	2.028	$(-0.221\,62)$	0.478	(0.017 304)	2.041	(-0.2254)	2.828
3	80	(-0.25)	2.505	$(-0.071\,519)$	3.588	$(-0.196\,49)$	0.438	$(-0.045\,849)$	3.674	$(-0.206\,51)$	5.657
4	288	(-0.125)	4.149	$(-0.061\,972)$	6.472	$(-0.156\,84)$	0.343	(0.1001)	6.940	$(-0.178\,57)$	11.314
5	1088	(-0.0625)	7.216	$(-0.060\,305)$	11.487	(-0.1246)	0.204	(0.160 34)	13.472	$(-0.158\,58)$	22.627
6	4224	$(-0.031\,25)$	12.960	$(-0.062\,585)$	20.013	$(-0.107\,32)$	0.093	(0.198 25)	26.536	$(-0.146\,88)$	45.255
7	16 640	$(-0.015\,625)$	23.768	$(-0.066\,453)$	34.858	$(-0.101\,05)$	0.035	(0.217 27)	52.664	$(-0.140\,58)$	90.510
8	66 048	$(-0.007\,812\,5)$	44.214	$(-0.070\,68)$	61.545	(-0.1)	0.012	(0.227 81)	104.920	$(-0.137\,32)$	181.019

Instead of specifying $\overline{E}$ and $\overline{\Omega}$, it is often desirable in the applications to specifically select a particular temperature regime, e.g., a state with highly negative α. Then, one first *specifies* κ instead of determining it by computing the ratio of known values of $\overline{E}$ and $\overline{\Omega}$; one then proceeds as before to determine $\overline{\alpha}$ and $\overline{\beta}$. The ratio $\overline{\Omega}/\overline{E}=\Omega/E$ is immediately given by κ^2; if one requires absolute values of E, Ω, α, or β, either $\overline{E}$ or $\overline{\Omega}$ must be additionally specified.

C. Equal-temperature equilibrium

One situation in which $\overline{E}$ and $\overline{\Omega}$ are known arises when one considers the noise in the common initialization in which particles (of finite number N) are distributed uniformly and independently in the box. The particle density as a function of continuous position $\mathbf{x}$ is $n(\mathbf{x})=\Sigma_\ell \delta(\mathbf{x}-\mathbf{x}^{(\ell)})$ (the superscript ℓ denotes an arbitrary labeling of the particles), where on the lattice one must interpret $\delta(\mathbf{x_j})$ as $\mathbf{\Delta}(\mathbf{x_j})$, $\mathbf{\Delta}(\mathbf{x})$ being the periodic δ function $\mathbf{\Delta}(\mathbf{x_j})\doteq V^{-1}\Sigma_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{x_j}}$ obeying $\mathbf{\Delta}(\mathbf{0})=(\Delta V)^{-1}$. The mean particle density is $\overline{n}=V^{-1}\int d\mathbf{x} n(\mathbf{x})=N/V$ and the power spectrum of the density fluctuation $\delta n_{\mathbf{k}}\doteq n_{\mathbf{k}}-\overline{n}$ is readily determined to be

$$\left\langle \left| \frac{\delta n_{\mathbf{k}}}{\overline{n}} \right|^2 \right\rangle = \begin{cases} N^{-1} & (\mathbf{k}\neq\mathbf{0}) \\ 0 & (\mathbf{k}=\mathbf{0}). \end{cases} \tag{29}$$

In subsequent formulas, it will be understood that we consider only $\mathbf{k}\neq\mathbf{0}$.

In Hasegawa-Mima dynamics the above density can be identified with the density of ion gyrocenters. From the gyrokinetic Poisson equation for adiabatic electrons (Appendix A),

$$(1-\nabla^2)\varphi = n(\mathbf{x})/\overline{n}, \tag{30}$$

one finds $\langle|\delta\varphi_{\mathbf{k}}|^2\rangle=(1+k^2)^{-2}N^{-1}$, so the quantity $E_{\mathbf{k}}\doteq\frac{1}{2}(1+k^2)\langle|\delta\varphi_{\mathbf{k}}|^2\rangle$ [see Eq. (7b)] is

$$E_{\mathbf{k}}=\frac{1}{2}\left(\frac{1}{1+k^2}\right)\frac{1}{N}. \tag{31}$$

Upon comparing result (31) with Eq. (9), one sees that the spectrum of uniform random particle noise has energy and enstrophy compatible with a canonical (and Gaussian) equilibrium potential distribution with equal and positive inverse temperatures

$$\alpha_1=\beta_1=N. \tag{32}$$

(Here, the 1 subscript reminds one that for this state $\hat{\alpha}=\alpha_1/\beta_1=1$.) However, it is very important to realize that the full PDF of such noise is *noncanonical* and *non-Gaussian*. That is, the probability density $\mathcal{P}_0(\varphi)$ of realizing a set of Fourier amplitudes φ is not proportional to

$$\exp\left(-\sum_{\mathbf{k}} N(1+k^2)\hat{E}_{\mathbf{k}}\right), \tag{33}$$

even though the mean $E_{\mathbf{k}}$ (proportional to the second moment of the potential) is correctly predicted by the Gaussian result. The true $\mathcal{P}_0$ is much more complicated because of higher-order correlations arising from the complicated nonlinear form of the Poisson equation considered as a function of the particle positions, and this observation has important consequences for our subsequent application of the Monte Carlo method. The correct $\mathcal{P}_0(\varphi)$ is computed in detail in Appendixes B and C. Here, we merely point out that the

TABLE III. Important parameters for a discrete spectrum in the weighted truncation.

m	No. of modes		k_a		k_1		$\overline{\alpha}$		k_b	$k_{\max}$
2	15	(-0.2)	1.421	(0.034 988)	1.777	$(-0.111\,48)$	0.483	(0.007 093 6)	1.789	$(-0.116\,12)$	2.828
3	63	$(-0.047\,619)$	2.298	(0.012 321)	3.270	$(-0.118\,22)$	0.448	(0.0239)	3.343	$(-0.127\,85)$	5.657
4	255	$(-0.011\,765)$	3.952	$(-0.015\,118)$	6.151	(-0.1128)	0.353	(0.066 577)	6.570	$(-0.132\,32)$	11.314
5	1023	$(-0.002\,932\,6)$	7.031	$(-0.035\,521)$	11.201	$(-0.102\,24)$	0.210	(0.126 59)	13.083	$(-0.133\,55)$	22.627
6	4095	$(-0.000\,732\,6)$	12.785	$(-0.049\,804)$	19.765	$(-0.096\,126)$	0.095	(0.174 23)	26.137	$(-0.133\,87)$	45.255
7	16 383	$(-0.000\,183\,12)$	23.603	$(-0.059\,945)$	34.638	$(-0.095\,342)$	0.035	(0.203 03)	52.261	$(-0.133\,95)$	90.510
8	65 535	(-4.5777×10^{-5})	44.057	$(-0.067\,39)$	61.344	$(-0.097\,068)$	0.012	(0.220 02)	104.514	$(-0.133\,97)$	181.019

difference between $\mathcal{P}_0$ and a Gaussian distribution is intimately related to the difference between the random-phase approximation and the assumption of Gaussian statistics. That is, the characteristic function (moment generating function, or Fourier transform of the PDF[5]) for the random variable $\psi \doteq \sin\theta$, where θ is distributed uniformly in the interval $[0, 2\pi)$, is $P_k = J_0(k)$, where J_0 is the ordinary Bessel function of the first kind. Such a variable has variance $\langle \delta\psi^2 \rangle = \frac{1}{2}$. The characteristic function for a Gaussian with variance $\frac{1}{2}$ is $\exp(-\frac{1}{4}k^2)$. Both this function and $J_0(k)$ behave as $1 - \frac{1}{4}k^2 + O(k^4)$ for small k, but differ in the terms of $O(k^4)$ and higher. In other words the logarithm of P_k is the cumulant generating function [26]:

$$\ln P_k = \sum_{l=1}^{\infty} \frac{(-ik)^l}{l!} C_l . \tag{34}$$

One readily obtains, for example, $C_2 = \frac{1}{2}$ and $C_4 = -\frac{3}{8}$; in general, the random-phase approximation has nonvanishing cumulants of all even orders. If ψ were Gaussian, on the other hand, all cumulants higher than the second would vanish. Unfortunately, ''random phase'' is often used synonymously with ''Gaussian'' [8]. Although in many applications (especially to homogeneous turbulence [27]) this does not matter, in general it is misleading. In the present application, the difference is crucial.

The precise energy per mode,

$$\overline{E} = \frac{1}{2N} \left\langle \frac{1}{1+k^2} \right\rangle_{\mathbf{k}} , \tag{35}$$

must be computed numerically as a function of the wavenumber cutoffs, but it is important to note that the normalized quantities,

$$\overline{\alpha}_1 = \overline{\beta}_1 = \left\langle \frac{1}{1+k^2} \right\rangle_{\mathbf{k}} , \tag{36}$$

are of the order of unity (independent of N). The associated κ corresponds to that point κ_1 in Fig. 1 where the curves $\overline{\alpha}(\kappa^2)$ and $\overline{\beta}(\kappa^2)$ intersect, namely,

$$\hat{\alpha} = 1, \tag{37a}$$

$$\kappa_1^2 = \frac{1}{2} \left\langle \frac{1}{1+k^2} \right\rangle_{\mathbf{k}}^{-1} - 1. \tag{37b}$$

The values of κ_1 and the associated $\overline{\alpha}_1 = \overline{\beta}_1 = [2(1 + \kappa_1^2)]^{-1}$ are tabulated in Tables I–III as functions of m.

[5]An introduction (with references) to statistical methods can be found in Ref. [8].

III. DETERMINING AND SAMPLING FROM THE PARTICLE PROBABILITY DISTRIBUTION

By definition, generating a particle state means sampling from the N-particle PDF $P^{(N)}(\Gamma)$. Two questions arise: (1) What *is* $P^{(N)}(\Gamma)$? (2) *Given* $P^{(N)}(\Gamma)$, how does one construct an appropriate sampling algorithm? As a trivial example, suppose that the particles are to be distributed independently and uniformly in a box of volume $V = L^d$, where d is the number of spatial dimensions. Independence means $P^{(N)}(\Gamma) = \Pi_\ell P^{(1)}(\mathbf{x}^{(\ell)})$, and homogeneity implies $P^{(1)}(\mathbf{x}) = V^{-1}$, so $P^{(N)}(\Gamma) = V^{-N}$. A straightforward sampling algorithm consists of initializing each Cartesian component of the $\mathbf{x}^{(\ell)}$ with random numbers drawn from a distribution uniform in $[0,L)$.

Unfortunately, as we have remarked, the present problem differs from the conventional one of equilibrium statistical mechanics in that we are given $\mathcal{P}^{(\mathcal{M})}(\varphi)$ (the PDF of the Fourier components) rather than $P^{(N)}(\Gamma)$ (the PDF of the particles). The $\varphi_{\mathbf{k}}$'s and $\mathbf{x}^{(\ell)}$'s are related via the Fourier transform of the gyrokinetic Poisson equation (30),

$$(1 + \mathbf{k}^{(\nu)2}) \varphi_{\mathbf{k}^{(\nu)}} = \frac{1}{N} \sum_{\ell=1}^{N} e^{-i\mathbf{k}^{(\nu)} \cdot \mathbf{x}^{(\ell)}} \quad (\nu = 1,...,\mathcal{M}), \tag{38}$$

where the factor of N^{-1} is equal to $(\overline{n}V)^{-1}$, the V arising from the Fourier transform convention (1a). This complicated nonlinear relation is a system of $2\mathcal{M}$ real equations involving dN Cartesian positions, where in the present calculation $d = 2$. For $2\mathcal{M} = dN$ or $\mathcal{M} = N$ one expects that it should be possible, in general, to invert this relation and thus determine $P^{(N)}(\Gamma)$ in terms of the Jacobian of transformation (38). However, it is very unusual that $\mathcal{M} = N$. Usually the desire for low sampling noise dictates $N \gg \mathcal{M}$, so the system is underdetermined; many particle (micro)states are compatible with a given set of Fourier amplitudes (macrostate). Furthermore, even if the inversion were possible, the resulting $P^{(N)}(\Gamma)$ would be extremely complicated, so a suitable sampling algorithm would probably not be apparent.

As we have suggested, it is possible to avoid these difficulties by employing a Monte Carlo technique. However, before turning to that we will discuss an alternative possible procedure that, although flawed, provides additional insight and motivation.

A. An impractical but instructive procedure

The idea is based on two observations: (i) usually $N \gg \mathcal{M}$; (ii) $\mathcal{P}^{(\mathcal{M})}(\varphi)$ is Gaussian. Consider dividing the particle population into G independent groups of $\mathcal{M}$ particles each, choosing N such that $N = G\mathcal{M}$. Then, consider the G systems

$$(1 + \mathbf{k}^{(\nu)2}) \psi^{(g)}_{\boldsymbol{\kappa}^{(\nu)}} = \frac{1}{N} \sum_{\ell \in g} e^{-i\mathbf{k}^{(\nu)} \cdot \mathbf{x}^{(\ell)}}$$

$$(\nu = 1,...,\mathcal{M}; \quad g = 1,...,G), \tag{39}$$

where $\Sigma_{\ell \in g}$ means sum over the $\mathcal{M}$ particles in group g. If one constructs $\varphi_{\mathbf{k}}=\Sigma_{g=1}^{G}\psi_{\mathbf{k}}^{(g)}$, then $\varphi_{\mathbf{k}}$ obeys

$$(1+k^2)\varphi_{\mathbf{k}}=\sum_{g=1}^{G}\left(\frac{1}{N}\sum_{\ell\in g}e^{-i\mathbf{k}\cdot\mathbf{x}^{(\ell)}}\right) \tag{40a}$$

$$=\frac{1}{N}\sum_{\ell}e^{-(i\mathbf{k}\cdot\mathbf{x}^{(\ell)})}. \tag{40b}$$

Therefore, if the $\mathbf{x}^{(\ell)}$ exist the Poisson equation will be satisfied. Furthermore, the variance of $\varphi_{\mathbf{k}}$ is related to the variance $\sigma_k^2\doteq\langle|\psi_{\mathbf{k}}|^2\rangle$ by

$$\langle|\varphi_{\mathbf{k}}|^2\rangle=\sum_{g,g'}\langle\psi_{\mathbf{k}}^{(g)}\psi_{\mathbf{k}}^{(g')*}\rangle \tag{41a}$$

$$=\sum_{g}\langle|\psi_{\mathbf{k}}^{(g)}|^2\rangle \tag{41b}$$

$$=G\sigma_{\mathbf{k}}^2, \tag{41c}$$

where we exploited the assumption that the systems are independent. Therefore, the variance of $\psi_{\mathbf{k}}$ is related in a simple way to the desired variance of $\varphi_{\mathbf{k}}$.

The proposal is now to sample, for each g, a collection of $\mathcal{M}$ $\psi_{\mathbf{k}}$'s from an appropriate PDF (independent of g) whose variance is $\sigma_{\mathbf{k}}^2$. (Appropriate care must be taken to satisfy the reality conditions.) This is easy with the aid of standard numerical software packages. Then solve the nonlinear system to obtain $\mathcal{M}$ $\mathbf{x}^{(\ell)}$'s. Those $\mathbf{x}^{(\ell)}$'s are guaranteed to be compatible with the given variance of the $\varphi_{\mathbf{k}}$.

If the PDF for the $\psi_{\mathbf{k}}^{(g)}$ were Gaussian, then since the sum of Gaussian random variables is again Gaussian the $\varphi_{\mathbf{k}}$ would be Gaussian and we would have constructed a valid particle realization. Unfortunately, Gaussian $\psi_{\mathbf{k}}^{(g)}$ are not permitted. The modulus of Eq. (39) obeys

$$(1+k^2)|\psi_{\mathbf{k}}^{(g)}|=N^{-1}\left|\sum_{\ell\in g}\exp(-i\mathbf{k}\cdot\mathbf{x}^{(\ell)})\right|\leqslant G/N. \tag{42}$$

If $\psi_{\mathbf{k}}^{(g)}$ were sampled from a Gaussian distribution, there would be a finite probability of obtaining a ψ such that $|\psi|>G/N$. For such ψ, the solution of Eq. (38) will not exist. One must therefore sample from a bounded PDF. If that obeys appropriate constraints, one can appeal to the central limit theorem to argue that $\varphi_{\mathbf{k}}$ is asymptotically Gaussian for $G\gg 1$. In practice, this may not be the regime of interest, since we often consider relatively small numbers of particles per Fourier mode. Furthermore, practical difficulties are certain to ensue. For example, the usual methods for solving nonlinear systems such as Eq. (39) involve some sort of functional iteration. However, there are no guarantees that such iteration will converge unconditionally. Nonconvergence may occur because of a poor initial guess, the existence of multiple solutions, and/or degeneracies associated with the regular nature of the wave-number lattice. Although one might think of solutions for each of these difficulties, it is clear that the proposed method is at best cumbersome and difficult to fully automate. Moreover, it suffers from the conceptual disadvantage that the realizations it generates are statistically only approximately valid for fixed, finite N.

Fortunately, the deficiencies of the procedure suggest a more fruitful line of approach. What is needed is a way of selecting N-particle states that are as random as possible consistent with the desired Gaussian statistics on the Fourier amplitudes, while guaranteeing unconditional convergence to an acceptable sequence of states. These criteria are satisfied by the Monte Carlo algorithm of MRRTT.

B. The MRRTT algorithm

Monte Carlo procedures [20,28,29] are often introduced from the point of view of the evaluation of multidimensional integrals. Although we will not need to evaluate such integrals explicitly, the application of computing integrals does provide fundamental motivation. Thus, an integral $I\doteq\int_{\mathcal{D}}d\mathbf{x}\,g(\mathbf{x})$ over a domain $\mathcal{D}$ whose volume is $V\doteq\int_{\mathcal{D}}d\mathbf{x}$ can be interpreted as an ensemble average over a PDF $f(\mathbf{x})$ that is uniform over the domain: $f(\mathbf{x})=V^{-1}$, $I=V\int_{\mathcal{D}}d\mathbf{x}\,g(\mathbf{x})f(\mathbf{x})=V\langle g\rangle$, where $\langle g\rangle\approx n^{-1}\Sigma_{i=1}^{n}g(\mathbf{x}_i)$, the $\mathbf{x}_i$ being sampled from $f(\mathbf{x})$. In statistical mechanics, the prototypical application is to the computation of the ensemble average of some quantity $Q(\Gamma)$ in the canonical ensemble,

$$\langle Q\rangle=Z^{-1}\int d\Gamma\,Q(\Gamma)\exp[-H(\Gamma)/T], \tag{43}$$

where $Z\doteq\int d\Gamma\exp[-H(\Gamma)/T]$. The difficulties with straightforward Monte Carlo evaluation of this integral by sampling from a uniform distribution, i.e., by identifying $g(\mathbf{x})\rightarrow Q(\Gamma)\exp[-H(\Gamma)/T]$, are twofold. First, the integrand g varies rapidly with Γ and will be exponentially small for almost all random points, so a possibly prohibitively large number of points would have to be sampled to ensure accuracy. Second, the method requires the explicit numerical value of the partition function Z, which can be very large and difficult to evaluate.

A better procedure is to devise a way of sampling directly from the canonical distribution $P(\Gamma)=Z^{-1}\exp[-H(\Gamma)/T]$; then $\langle Q\rangle\approx n^{-1}\Sigma_{i=1}^{n}Q(\mathbf{x}_i)$. (This is a special case of so-called importance sampling, as defined and discussed, for example, in Ref. [28].) The algorithm of MRRTT accomplishes this by defining a Markov chain that is guaranteed to converge asymptotically to $P(\Gamma)$.

In reviewing the algorithm, we follow the lucid exposition of Kalos and Whitlock [28]. The elegant technique of MRRTT is based on the fundamental Chapman-Kolmogorov equation for Markov processes, which can be written for a PDF f that depends on a discrete timelike variable n and a continuous spacelike coordinate or set of (abstract or generalized) coordinates X as

$$f_{n+1}(X)=\int dY\,T(X|Y)f_n(Y). \tag{44}$$

Here, T is an arbitrary conditional probability density. If we introduce the transition probability S according to

$$T(X|X') = S(X|X') + \left(1 - \int dY S(Y|X)\right)\delta(X - X'), \tag{45}$$

one is led to the conventional master equation [30]

$$f_{n+1}(X) = \int dY S(X|Y) f_n(Y) + \left(1 - \int dY S(Y|X)\right) f_n(X). \tag{46}$$

Here, the term $S(X|Y)$ describes the probability of leaving the state Y; the parenthesized term is the conditional probability of remaining in the state. Note that S is not a true conditional probability density since $\int dY S(Y|X) \neq 1$.

It is easy to see that if an asymptotic distribution exists, namely, $f_{n+1} = f_n = f$, then

$$\int dY S(X|Y) f(Y) = \int dY S(Y|X) f(X). \tag{47}$$

This is satisfied by the detailed-balance condition

$$S(X|Y) f(Y) = S(Y|X) f(X). \tag{48}$$

The MRRTT algorithm and its variants correspond to particular convenient choices of $S(X|Y)$.

Specifically, the algorithm proceeds by *proposing* a transition from state Y to a new state Y' generated from an arbitrary conditional probability $T_0(Y'|Y)$. The proposed state is then tested against an acceptance criterion q and conditionally accepted ($X = Y'$) or rejected ($X = Y$) in such a way that detailed balance is satisfied. One has [28]

$$S(X|Y) = A(X|Y) T_0(X|Y), \tag{49}$$

where A is the acceptance probability. We will follow MRRTT in choosing

$$A(X|Y) = \min(1, q(X|Y)), \tag{50}$$

where

$$q(X|Y) \doteq \frac{T_0(Y|X) f(X)}{T_0(X|Y) f(Y)}. \tag{51}$$

Usually an algorithm is chosen such that

$$T_0(X|Y) = T_0(Y|X) \tag{52}$$

(although we will discuss a generalization in Sec. III C 1). In that case,

$$q(X|Y) \to f(X)/f(Y). \tag{53}$$

One can then summarize the algorithm as follows. When the given probability density at the new proposed state is larger than that at the old state [$q(Y'|Y) > 1$], the new state is accepted unconditionally. Otherwise, the state is accepted with probability $q = f(Y')/f(Y)$. If one writes

$$f(Y) \propto e^{-W(Y)}, \tag{54}$$

which is always possible for real W since $f(Y)$ is a PDF (hence non-negative), then

$$q = e^{-\Delta W}, \tag{55}$$

where $\Delta W \doteq W(Y') - W(Y)$; thus proposed states with lower "energy," $\Delta W < 0$, are accepted unconditionally. It is easily shown that choice (50) satisfies detailed balance, and asymptotic theorems on Markov chains guarantee that f_n converges to f for reasonable T_0.

In practice, the proposed state is usually generated by first selecting one particle randomly, then examining the consequences of changing its position by a random amount $\Delta \mathbf{x} = \lambda \boldsymbol{\xi}$, where the Cartesian components of $\boldsymbol{\xi}$ are sampled from a distribution uniform in $(-\frac{1}{2}, \frac{1}{2})$ and λ ($\leqslant L$) is a parameter that is arbitrary, in principle. That is,

$$T_0(X|Y) = \begin{cases} \lambda^{-d} & (|x_i - y_i| \leqslant \lambda) \\ 0 & \text{otherwise.} \end{cases} \tag{56}$$

(The average acceptance probability and thus the rate of convergence depend on λ; see later discussion.) The effective energy W is then evaluated at the proposed state Y' and the increment ΔW is computed. If $\Delta W < 0$, then the proposed state is accepted as the next state in the Markov chain. Otherwise, another random number p is drawn from a distribution uniform in $[0, 1)$. If $p < q$, where q is defined by Eq. (55), the state is accepted ($X = Y'$); otherwise, the old state becomes the next state in the chain ($X = Y$). The role of p is to ensure that states with $q < 1$ are accepted with probability q under a long-time average.

C. Application of MRRTT to particle initialization

The application of the MRRTT algorithm to the particle initialization problem introduces both theoretical and computational nuances.

1. Theoretical considerations

We will use the simple and efficient procedure described in the last paragraph of Sec. III B to generate a sequence of particle states Γ_i (and associated Fourier amplitudes φ_i). However, because the target PDF $\mathcal{P}^{(\mathcal{M})}(\varphi)$ is couched in terms of the Fourier amplitudes, not the particle state directly, one must be cautious. In particular, although for generating particle states one may choose $X = \Gamma$, $f(X) = P^{(N)}(\Gamma)$, the assertion $P^{(N)}(\Gamma) = \mathcal{P}^{(\mathcal{M})}(\varphi(\Gamma))$ is *not* correct because it overlooks the nontrivial, nonlinear, many-to-one relation between the random variables Γ and φ. We will now explain how to take that relation into account. For notational brevity, we will henceforth write $P^{(N)}(\Gamma) \equiv P(\Gamma)$, $\mathcal{P}^{(\mathcal{M})}(\varphi) \equiv \mathcal{P}(\varphi)$. The underlying transition probability associated with Γ states will be written T_0, while the one associated with φ states will be written $\mathcal{T}_0$.

It is important to realize that the particle states that are generated by the Markov chain have no dynamical significance. For example, they do not contain the specific pair correlations that are associated with the Coulomb interaction and that arise from the dynamical relaxation on the fast time

scale during which Debye shielding is setup. There is no physical significance to the averages of arbitrary thermodynamic quantities over the Γ_i or to relaxation rates toward converged spectral equilibria. Because the only information built into the calculation is (the single time) $\mathcal{P}(\varphi)$, one is allowed to average only functions of φ itself. Let $\mathcal{A}(\varphi)$ be such an arbitrary function. Most directly, one has

$$\langle \mathcal{A} \rangle = \int d\varphi \mathcal{P}(\varphi) \mathcal{A}(\varphi). \tag{57}$$

Alternatively, for a compatible $P(\Gamma)$, one can write

$$\langle \mathcal{A} \rangle = \int d\Gamma P(\Gamma) \mathcal{A}(\varphi(\Gamma)). \tag{58}$$

The fundamental constraint relating $\mathcal{P}(\varphi)$ and $P(\Gamma)$ is that

$$\mathcal{P}(\varphi) = \langle \delta(\varphi - \varphi(\Gamma)) \rangle \tag{59a}$$

$$= \int d\Gamma P(\Gamma) \delta(\varphi - \varphi(\Gamma)). \tag{59b}$$

Since our only goal is to determine particle states *compatible with* $\mathcal{P}(\varphi)$ (but not necessarily the result of physically realizable dynamics), we have wide latitude in choosing $P(\Gamma)$. We will argue that a reasonable choice is

$$P(\Gamma) = \frac{1}{C} \frac{\mathcal{P}(\varphi)}{\mathcal{P}_0(\varphi)}, \tag{60}$$

where C is a normalizing factor (the volume of the Γ space) and $\mathcal{P}_0$ is the probability density of realizing the value φ from a *uniformly distributed, statistically independent* collection of particles. Note that the explicit normalization constant C is never needed in the MRRTT algorithm.

To arrive at Eq. (60), we argue that since only φ averages are of interest, one can choose $P(\Gamma)$ such that it depends on Γ only through φ: $P(\Gamma) = F(\varphi)$ for some function F. In the spirit of information theory [31], this is the unique choice compatible with the lack of any further information or constraints. Without loss of generality, we can write

$$P(\Gamma) = \mathcal{P}(\varphi) / \mathcal{Q}(\varphi), \tag{61}$$

where $\mathcal{Q}(\varphi)$ is to be determined. Upon inserting representation (61) into Eq. (59b), one obtains

$$\mathcal{P}(\varphi) = \int d\Gamma \frac{\mathcal{P}(\varphi(\Gamma))}{\mathcal{Q}(\varphi(\Gamma))} \delta(\varphi - \varphi(\Gamma)) \tag{62a}$$

$$= \frac{\mathcal{P}(\varphi)}{\mathcal{Q}(\varphi)} C \int d\Gamma \frac{1}{C} \delta(\varphi - \varphi(\Gamma)), \tag{62b}$$

or, upon canceling $\mathcal{P}(\varphi)$ from both sides and rearranging,

$$\mathcal{Q}(\varphi) = C \langle \delta(\varphi - \varphi(\Gamma)) \rangle_0, \tag{63}$$

where $\langle \cdots \rangle_0$ means the average over the PDF $P_0(\Gamma) = C^{-1}$, i.e., over a distribution of uniformly distributed, statistically independent particles. Thus, $\mathcal{Q}(\varphi) = C\mathcal{P}_0(\varphi)$ and one recovers Eq. (60).

Form (60) is a generalization of the well-known result that if $y(x)$ is a monotonic function of x, then $|P_X(x)dx| = |P_Y(y)dy|$, or $P_X(x) = P_Y(y)/|dx/dy|$. This can be written as $P_X(x) = P_Y(y)/P_0(y)$, where $P_0(y) = |dx(y)/dy|$. By setting $P_X(x) = \text{const}$, one sees that $P_0(y)$ is the PDF of y associated with a uniform x distribution. In the present application, one may identify $x \to \Gamma$, $y \to \varphi$; however, one cannot simply introduce a Jacobian $\partial\Gamma/\partial\varphi$ because the relation between φ and Γ is not one to one. Form (60) reflects a particular, minimally constrained way of handling the underdeterminism.

So far we have concentrated on generating particle states Γ_i that are compatible with the given φ distribution. An alternate approach that leads one to the same PDF (60) is to directly consider a Markov chain of φ states. Now we identify $X \to \varphi$, $f(X) \to \mathcal{P}(\varphi)$. One has

$$q(\varphi' | \varphi) = \frac{\mathcal{T}_0(\varphi | \varphi') \mathcal{P}(\varphi')}{\mathcal{T}_0(\varphi' | \varphi) \mathcal{P}(\varphi)}, \tag{64}$$

where $\mathcal{T}_0(\varphi' | \varphi)$ is the conditional probability of achieving φ', given φ, that is associated with the underlying algorithm for generating new Γ states. The function $\mathcal{T}_0(\varphi' | \varphi)$ is nontrivial. However, one may use the definition of conditional probability to write in complete generality

$$\mathcal{T}_0(\varphi' | \varphi) = \mathcal{T}_0(\varphi', \varphi) / \mathcal{P}_0(\varphi), \tag{65}$$

where $\mathcal{P}_0(\varphi)$ is the PDF for realizing φ at any step in the chain. The great appeal of the MRRTT method is that the joint probability $\mathcal{T}_0(\varphi', \varphi)$ need never be computed explicitly since the ratio of the $\mathcal{T}_0$'s required in Eq. (51) can be written as

$$\frac{\mathcal{T}_0(\varphi | \varphi')}{\mathcal{T}_0(\varphi' | \varphi)} = \frac{\mathcal{T}_0(\varphi, \varphi') / \mathcal{P}_0(\varphi')}{\mathcal{T}_0(\varphi', \varphi) / \mathcal{P}_0(\varphi)} = \frac{\mathcal{P}_0(\varphi)}{\mathcal{P}_0(\varphi')}. \tag{66}$$

Thus, Eq. (64) becomes

$$q(\varphi' | \varphi) = \frac{\mathcal{P}(\varphi') / \mathcal{P}_0(\varphi')}{\mathcal{P}(\varphi) / \mathcal{P}_0(\varphi)}, \tag{67}$$

where all the potentials are to be computed in terms of the random particle positions.

From the point of view of generating φ statistics, $\mathcal{P}_0$ is not unique; one must provide some information about how the underlying Γ states are generated. The arguments leading to Eq. (63) show that the minimally biased choice for $\mathcal{P}_0$ is the PDF associated with an independent, uniform distribution of particles. This important function is considered in Appendixes B and C. There asymptotic methods are used to show that for large N

$$\mathcal{P}_0(\varphi) \propto \exp[-N\Psi(\varphi)], \tag{68}$$

where the exponent Ψ satisfies[6]

$$\Psi(\varphi)\approx\frac{1}{2}\sum_{\mathbf{k}}\ (\overline{\epsilon}_{\mathbf{k}}+\tfrac{1}{4}\overline{\epsilon}_{\mathbf{k}}^{2}) \tag{69}$$

for $\overline{\epsilon}_{\mathbf{k}}\doteq(1+k^2)^2|\varphi_{\mathbf{k}}|^2\ll 1$ [see Eqs. (C12) and (C69)] and is determined from an implicit algorithm for larger fluctuation levels [see Eq. (C66)]. In the limit $N\to\infty$, for which the noise level approaches 0, $\overline{\epsilon}$ is very small, the quartic (in $|\varphi|$) correction in Eq. (69) is negligible, and one sees that the result reduces to the Gaussian approximation (33); however, for finite N the result is substantially more complicated. Note that even though $\overline{\epsilon}_{\mathbf{k}}\ll 1$, $N\overline{\epsilon}_{\mathbf{k}}=O(1)$, so $\mathcal{P}_0$ is a nontrivial function that is not well approximated by one. (See Fig. 9 for a numerical confirmation of this remark.)

Although the true form of Ψ is involved, its qualitative role in the Monte Carlo algorithm can be understood by considering the lowest-order result

$$\mathcal{P}_0(\varphi)\propto\exp\left(-\sum_{\mathbf{k}}\ N(1+k^2)\hat{E}_{\mathbf{k}}\right). \tag{70}$$

From the point of view of the general method of MRRTT, which attempts to converge to a distribution function $f(X)$, the effective PDF in the present problem is

$$f(X)\propto\mathcal{P}(\varphi)/\mathcal{P}_0(\varphi) \tag{71a}$$

$$\approx\exp\left(-\sum_{\mathbf{k}}\ [(\alpha+\beta k^2)-N(1+k^2)]\hat{E}_{\mathbf{k}}\right). \tag{71b}$$

For the special case of uniform, independent states, for which we have shown in Sec. II C that $\alpha=\beta=N$, the lowest-order contribution from $\mathcal{P}_0$ cancels the $(\alpha+\beta k^2)$ term, leaving one with $f(X)\approx$ const and $q=1$. In this approximation all states are accepted, which demonstrates a necessary consistency: to the extent that the particle states can be considered to be Gaussian (N sufficiently large), the algorithm need "do no work."

2. Computational algorithm

In addition to the appearance of the reference distribution $\mathcal{P}_0$, the unusual elements in the present application are that the potentials are spatially *nonlocal* functions of the microscopic particle state Γ and that the Fourier spectrum is resolved only to a finite $k_{\max}$, whereas the particles may occupy positions distributed continuously in V. Given a proposed state Γ', we proceed (in principle) as follows. The particles are collected onto the nearest lattice point *for the purpose of computing the potentials*. The resulting density distribution is (discrete-)Fourier transformed and the potential is determined from the solution of Poisson's equation (which is trivial in Fourier space). The invariants $\hat{E}(\Gamma')$ and $\widehat{\Omega}(\Gamma')$ are then computed, as is the change

$$\Delta W=[\widehat{W}(\Gamma')-\widehat{W}_0(\Gamma')]-[\widehat{W}(\Gamma)-\widehat{W}_0(\Gamma)], \tag{72}$$

where $\widehat{W}_0(\Gamma)\doteq N\Psi(\boldsymbol{\varphi}(\Gamma))$ and

$$\widehat{W}(\Gamma)\doteq\alpha\hat{E}(\Gamma)+\beta\widehat{\Omega}(\Gamma). \tag{73}$$

Finally, the state Γ' is accepted or rejected according to the MRRTT criterion with $q=\exp(-\Delta W)$.

The procedure as just described is not yet optimal for machine computations since it seems to require a time-consuming 2D Fourier transform at each step in the chain (which can be very long). That is unnecessary, however, since only one particle is moved per step and Fourier transformation is a linear operation. Because we use a nearest-grid-point algorithm to collect the particles onto the spatial lattice, it is clear that on the lattice the density of the proposed state Γ' will either be identical to that of Γ or will differ from it by a deficiency of one particle at the original point $\mathbf{x_j}$ and an excess of one particle at $\mathbf{x_j'}$. We can therefore calculate $\varphi_{\mathbf{k}}'$ by adding to $\varphi_{\mathbf{k}}$ the potential associated with a test particle at $\mathbf{x_j'}$ and subtracting that associated with a test particle at $\mathbf{x_j}$. The potential increment due to a test particle at $\mathbf{x_j}$ is [cf. Eq. (38)]

$$\varphi_{\mathbf{k}}^{(\mathbf{j})}=[(1+k^2)N]^{-1}\exp(-i\mathbf{k}\cdot\mathbf{x_j}); \tag{74}$$

thus, the calculation of $\varphi_{\mathbf{k}}'$ involves the computation of just two complex exponentials, or two cosines and two sines. In principle, evaluation of such quantities must be done at each step. However, run time can be decreased at the expense of memory by computing all of the possible potentials (74) once at the beginning of the run and storing them. Since there are $M_x(\frac{1}{2}M_y+1)$ independent complex Fourier modes and $M_{\rm tot}$ lattice points, one must store approximately $Y=2(\frac{1}{2}M_{\rm tot})(M_{\rm tot})=M_{\rm tot}^2$ real numbers. Usually, we consider $M_x=2^m$ for reasonably small m, so $Y(m)=2^{4m}$. Runs with $m=6$, or a 64×64 lattice, reside comfortably on desktop workstations. For the goal of testing gyrokinetic codes, there is no reason to work with larger grids.

In designing a satisfactory Monte Carlo run, it is important that the ratio of acceptances to rejections be neither too small nor too large. If the ratio is small, so that almost all states are rejected, then one gains little new information at each step, the steps are highly correlated, and the rate of convergence to the asymptotic distribution may be prohibitively slow. The same remarks pertain to the other extreme where almost all states are accepted. Common lore suggests an acceptance rate of about 50%. In principle, this rate can be calculated analytically as a function of the parameters, e.g., $\{\mathcal{M}, N, \lambda, \kappa, E\}$. Those results are somewhat tedious in detail, being expressable as infinite Fourier integrals over the characteristic function, which itself is known only as an infinite series of Bessel harmonics; our theoretical work in this area is incomplete. In practice, we proceed as follows. First, we choose a temperature regime by specifying an appropriate

[6]The factor of $\frac{1}{2}$ in Eq. (69) results from converting the sum of the $\mathcal{M}'\doteq\frac{1}{2}\mathcal{M}$ terms of $\ln\mathcal{P}_0(\varphi)$ [Eq. (B14)] to the sum over all $\mathcal{M}$ modes.

κ. Next, we set the overall height of the target spectrum, e.g., by forcing the intensity of the longest- or shortest-wavelength mode to be a specified percentage of the reference noise level. Finally, with the aid of several short trial runs we adjust λ, the size of the basic cell for the transition probability T_0, such that the rejection rate is $\approx 50\%$.

IV. EXAMPLES OF INITIALIZATION

The preceding considerations have been implemented in a computer code whose input consists of the parameters $\{m, N, \lambda, \kappa, E\}$ and whose output consists of a succession of particle states that sample the canonical ensemble (9).

The initial particle state is arbitrary, in principle. In addition to the uniform, independent initialization already mentioned, we also consider the well-known Fibonacci ''quiet start'' defined by (Ref. [18], and references therein)

$$x_r = \left(\frac{2r+1}{2N}\right) L, \tag{75a}$$

$$y_r = \alpha_{n-1} x_r \mod L, \tag{75b}$$

where $r = 0, 1, \ldots, N-1$; $n > 1$ is an arbitrary integer parameter; α_n is the nth Fibonacci number defined by

$$\alpha_0 = 0, \tag{76a}$$

$$\alpha_1 = 1, \tag{76b}$$

$$\alpha_n = \alpha_{n-1} + \alpha_{n-2}; \tag{76c}$$

and $N = \alpha_n$. To expedite easy comparison of the random and quiet starts, we generally choose N to be a Fibonacci number. Unless we specifically state otherwise, we use $n = 17$ ($N = 1597$).

The spectrum is resolved on a lattice corresponding to $m = 4$ ($M_x = M_y = 16$). This number of modes reasonably approximates a continuum in wave-number magnitude, but is sufficiently small that desktop workstation CPU time is not exorbitant. The results presented here correspond to a box size of $L = 43.3$ or $k_{\min} \approx 0.145$, $k_{\max} \approx 1.64$. These numbers are representative of other medium-size simulations of the Hasegawa-Mima and similar equations, but their precise values are not critical for the application of testing relaxation to appropriate equilibria.

In the following sections, we will present spectral information via 2D graphs of $E_{\mathbf{k}}$ vs $k \doteq |\mathbf{k}|$. On these graphs the noise level for the special uniform, independent particle state ''1''—i.e., $E_{\mathbf{k},1} \doteq [2N(1+k^2)]^{-1}$—is shown by a dashed line; the theoretically expected spectrum, $E_{\mathbf{k}} = [2(\alpha + \beta k^2)]^{-1}$, is shown by a solid curve. At any step in the Markov chain, the instantaneous levels $\hat{E}_{\mathbf{k}}$'s are indicated by a scatter plot of small plus signs. The averages of the $\hat{E}_{\mathbf{k}}$'s over the chain up to that point are indicated by a scatter plot of larger squares that in some cases is superimposed over the instantaneous data. It is convenient to measure time in units of cycles, where a cycle is defined to be N steps. (On the average, each particle is moved once in the course of 1 cycle.)

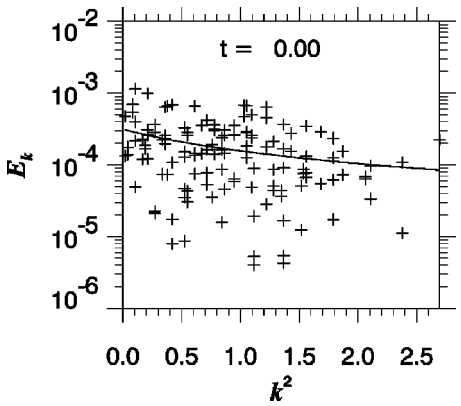

FIG. 3. Initial spectrum for a random start with $\alpha = \beta = N = 1597$.

It should be noted that if the number of particles N is too small relative to the number of retained Fourier amplitudes, it may not be possible to successfully generate arbitrary Fourier spectra (consider $N = 1$, for example). In such cases, the Monte Carlo algorithm fails to converge, or converges to unusual spectra with $\overline{U} \neq 1$. The time dependences of U and the running time average $\overline{U}$ are sensitive monitors of the convergence of the algorithm.

A. The reference noise spectrum

In the first experiment, we verify that the Monte Carlo code properly converges to the preferred noise spectrum with $\alpha = \beta = N$, specified by κ_1. (Upon referring to Table III, we see that $\kappa_1 = 6.151$ for the present parameters.) We show the initial random Fourier intensities for a random start in Fig. 3 and for a Fibonacci start in Fig. 4. In Figs. 5–7, we show snapshots from a Fibonacci start of the Markov chain at $t \approx 1$, 20, and 200, respectively. Convergence to the appropriate distribution is seen clearly with the expected $1/\sqrt{t}$ rate. The gross behavior of the instantaneous amplitudes in the final state is qualitatively similar to the initial scatter plot for the random start (Fig. 3), as it should be.

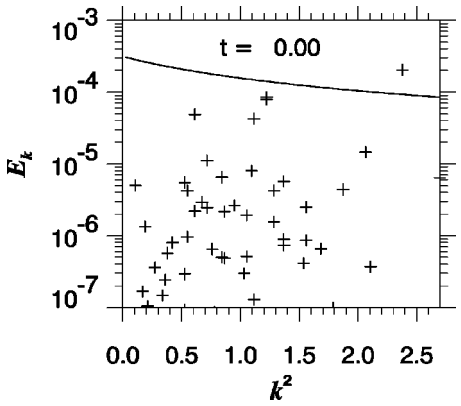

FIG. 4. Initial spectrum for a Fibonacci start with $\alpha = \beta = N = 1597$.

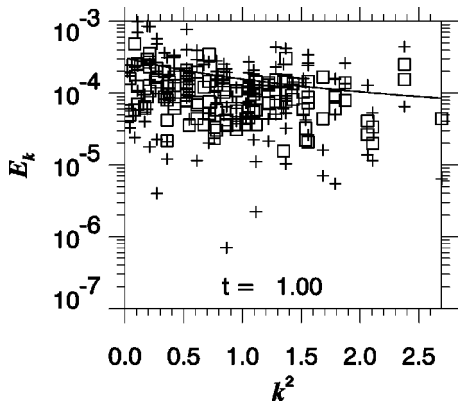

FIG. 5. Reference-case spectra for $t=1$.

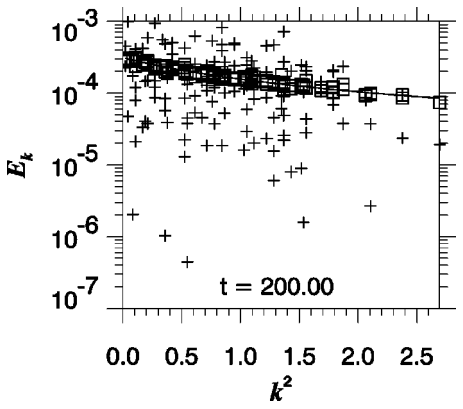

FIG. 7. Reference-case spectra for $t=200$.

Convergence to the proper value $\overline{\hat{U}}=1$ is illustrated in Fig. 8. Convergence is already recognizable from the fluctuating data in times of the order of several cycles. For this particular case, the running time average has saturated to its correct value after about 10 cycles. Convergence within about 10–25 cycles is typical for the runs presented here. Although not graphed here, $\hat{U}$ continues to fluctuate around one for the duration of the run; no tendency toward instability is seen.

Although we do not display the graphics here, it is easy to check that the particle positions have been thoroughly mixed after a small number of cycles. (Color coding can be used to emphasize that the particles have not just moved slightly from their initial positions.) A comparison of the final particle states with a typical set of random initial conditions shows no qualitative differences. The collected statistics verify that all states were accepted for this case, in agreement with the argument presented at the end of Sec. III C 1.

A scatter plot of the exponents U_0 for the previous run is shown in Fig. 9. Because the values are $O(1)$ but are not all equal, this figure emphasizes that $\mathcal{P}_0(\varphi)$ is a nontrivial function, as was remarked after Eq. (69).

In Fig. 10, we show that the algorithm has no trouble generating equal-temperature states with intensity one-tenth of the reference noise level.

B. Enstrophy equipartition

Another reference case of importance is the enstrophy-equipartition case $\hat{\alpha}=0$ $(\kappa=k_a)$. Although this is not a negative-temperature state (it corresponds to $\alpha=0$ or an infinite energy temperature), it is qualitatively similar to states of negative α (regime I) in that the long-wavelength modes are excited to relatively high levels. Working with a marginal state such as this affords a good consistency check of the software routines that relate α, β, and κ. Specifying a κ of 3.952 as indicated in Table III should lead to an α of 0, as indeed it does to within numerical roundoff error. For this case, we consider two absolute levels. Define the parameter f to be the ratio between the target intensity and the reference noise level for the longest-wavelength mode. We first consider $f=5$. The state after 400 cycles is shown in Fig. 11. As one expects, convergence is dominated by the time for the longest-wavelength modes to attain equilibrium. Note that this case demonstrates that there is no difficulty in exciting some modes to a superthermal level while suppressing others to a subthermal level.

Next we consider $f=10$, thereby doubling the target intensity from the previous case. As shown in Fig. 12, this run transiently attains a quasi-steady-state that well approximates the expected result (Fig. 13); however, after about 15 cycles an instability sets in. That this instability is of the longest-wavelength modes can be seen in Fig. 14.

Experience shows that such instabilities are associated with an inadequate number of particles. A precise theoretical

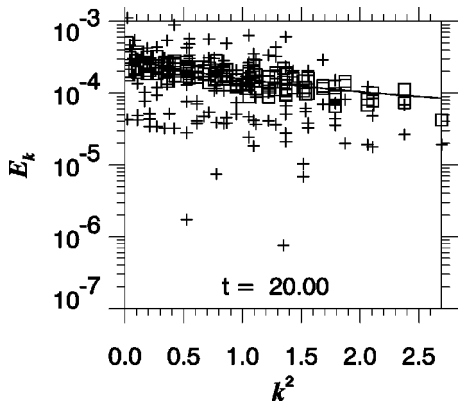

FIG. 6. Reference-case spectra for $t=20$.

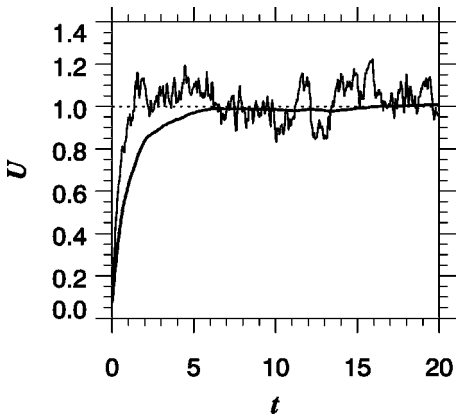

FIG. 8. Initial behavior of the convergence to the noise state. Thin solid line, data $\hat{U}\doteq 2(\alpha\hat{E}+\beta\hat{\Omega})$ sampled every 0.05 cycles; thick solid line, running time average.

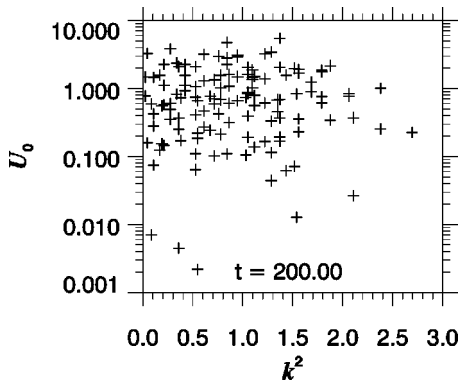

FIG. 9. Scatter plot of the exponents U_0 for the reference case plotted in Figs. 5–8.

criterion for the threshold of instability would be very valuable, but is left for future work. For lack of space, we will not show the successful result of stabilizing the present run by increasing N, but similar behavior is manifested by the runs presented in Secs. IV C and IV D. Note that cases with too small N need not always be unstable; sometimes the chain converges, but to states with $\overline{U} \neq 1$. We do not fully understand the significance of those states.

C. Negative α

Now we generate a true negative-temperature state by somewhat arbitrarily choosing $\kappa = 1.5$, a value deep in the negative-α regime I. Such equilibria have the longest-wavelength modes excited to relatively large levels, such as would (qualitatively) result from an inverse energy cascade. The spectrum after 4800 cycles is shown in Fig. 15. For these parameters, the longest-wavelength modes have not fully converged; however, there is no tendency toward instability.

D. Negative β

In the final experiment, we generate a state of negative β (regime III). We arbitrarily choose $\kappa = 10$. For the standard parameters $m = 4$, $n = 17$, the algorithm appears to be well converged at $t = 100$, but subsequently exhibits an instability

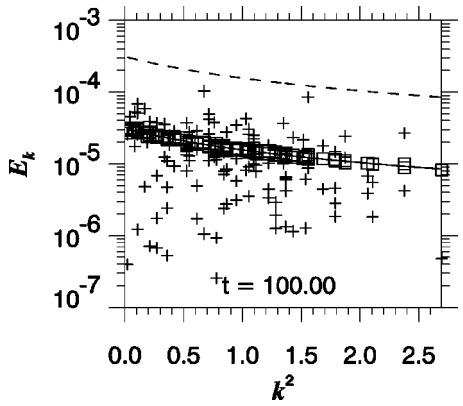

FIG. 10. Equal-temperature states with intensity one-tenth of the reference noise level (dashed line).

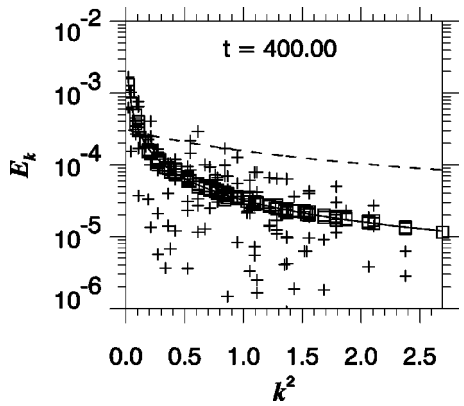

FIG. 11. Enstrophy-equipartition run ($\alpha = 0$) with $f = 5$ (snapshot at $t = 400$).

of the shortest-wavelength mode. Increasing n to 19 removes the instability; a well-converged spectrum is shown in Fig. 16.

V. DISCUSSION

The calculations presented here meld two of the principal avenues to the study of nonlinear phenomena in plasma systems that exhibit strongly fluidlike behavior: (gyro)fluid simulation with Fourier amplitudes and kinetic simulation with particles (or gyrocenters). Each has its strengths, but the two approaches should agree exactly in the inviscid, undriven limit in which collisional dissipation and the Landau resonance are ignored. Then the spectrum predicted by either approach should nonlinearly relax to the prediction of the appropriate canonical ensemble. The Hasegawa-Mima equation is arguably the simplest nonlinear equation with relevance to magnetized plasma physics. Nevertheless, its canonical behavior (for truncated Fourier spectra) is entirely nontrivial, including the existence of negative-temperature states. In this paper, we showed how to construct particle realizations compatible with those Fourier spectra by using a generalization of the well-known Monte Carlo algorithm of Metropolis *et al.* [19]. The numerical aspects of the calcula-

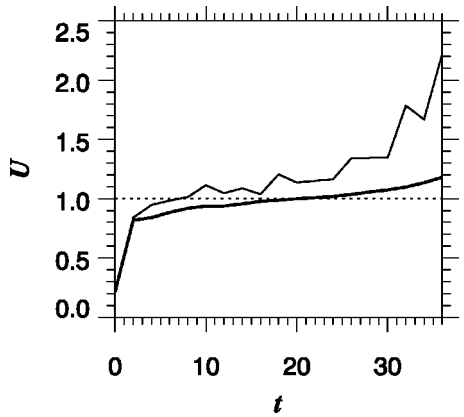

FIG. 12. Enstrophy-equipartition run with $f = 10$, demonstrating quasisaturation but a long-term instability. Thin line, data; thick line, running time average.

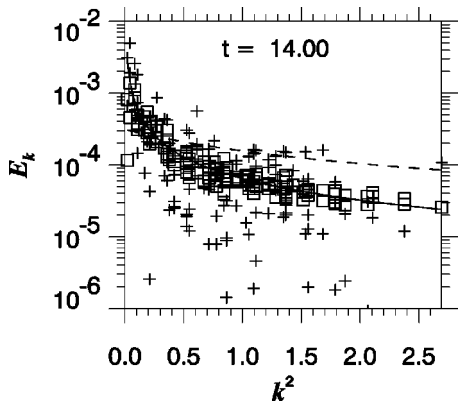

FIG. 13. Enstrophy-equipartition spectra in the quasisaturated regime of Fig. 12.

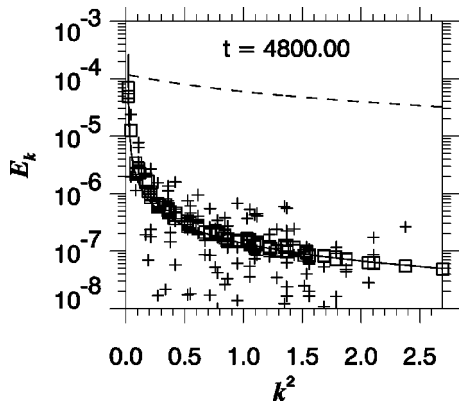

FIG. 15. A state of negative α. $\kappa=1.5$; $f=2.5$; $n=19$ ($N=4181$).

tion are straightforward. However, the calculation of the nontrivial PDF $P_0(\varphi)$ (the probability density of obtaining the set of Fourier amplitudes φ from a uniform distribution of statistically independent gyrocenters; see Appendix B) presents an interesting problem in asymptotics, as described in Appendix C.

A. Application to particle simulations

In the proposed application to particle simulations, issues arise concerning both the interpretation of averaging procedures as well as the practical implementation of the required fluid limit.

1. Averaging procedures

We first consider the interpretation of averages over the chain of states. In particular, we pose the following questions: (i) What is the role of a single microstate? (ii) Does an ensemble average yield any additional information?

To answer these questions, one must be precise about the distinction between time and ensemble averages. Let r label one of R realizations, each initialized by a random sample drawn from the canonical distribution with specified E and Ω. Of course, $\hat{E}_{\mathbf{k}}^{(r)}(t=0)\doteq\tilde{E}_{\mathbf{k}}^{(r)}\neq E_{\mathbf{k}}$. Each realization will evolve conserving its own $\hat{E}=\tilde{E}$ and $\hat{\Omega}$, and one expects that the action of the nonlinear terms will (on the average) readjust $\hat{E}_{\mathbf{k}}(t)$ to the canonical spectrum corresponding to $\hat{E}$ (*not* E). Within one realization, the only sensible average to perform is the time average; one expects

$$\overline{E_{\mathbf{k}}^{(r)}(t)}=\tilde{E}_{\mathbf{k}}^{(r)}, \tag{77a}$$

$$\overline{\hat{E}^{(r)}(t)}=\tilde{E}^{(r)}. \tag{77b}$$

Demonstration of correct time-averaged relaxation to a variety of positive- and negative-temperature canonical states is probably the most stringent test that can be performed on nonlinear simulation modules. Note that it is unnecessary to perform a new Monte Carlo run before each test of a simulation code or sequence of code updates. A few representative cases can be computed once and stored in disk files; they can then be used repeatedly in relaxation tests. Although those will be restricted in the numbers of particles N and Fourier amplitudes $\mathcal{M}$, that is probably not crucial because it is difficult to imagine software bugs that would not show up with a randomly chosen N and $\mathcal{M}$.

In order to demonstrate convergence to the specified E, a true ensemble average must be performed:

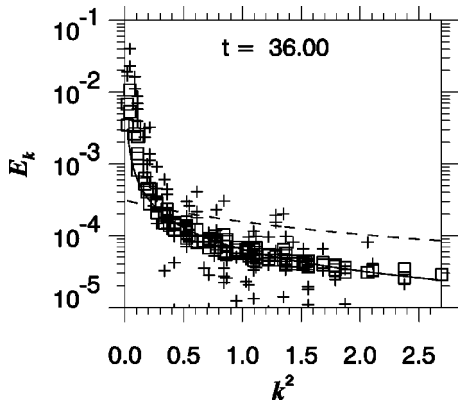

FIG. 14. Long-time spectra for the unstable case corresponding to Fig. 12. The long-wavelength amplitudes have drifted significantly above the target curve (solid line).

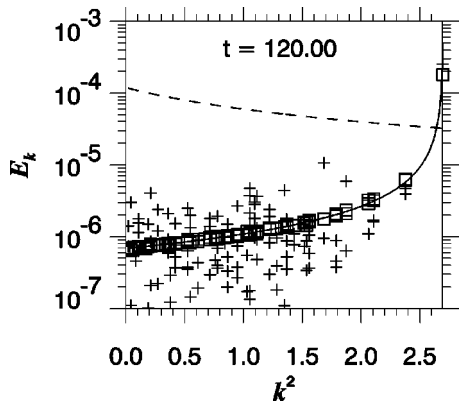

FIG. 16. A state of negative β. $\kappa=10$; $n=19$ ($N=4181$).

$$E=\langle\tilde{E}\rangle\approx\frac{1}{R}\sum_{r=1}^{R}\tilde{E}^{(r)}. \tag{78}$$

To reduce the variance of this calculation, one should not identify realizations with successive states along the Markov chain but rather with states separated by at least one correlation length. It is unlikely that such an experiment will be performed routinely because it involves many simulation runs and it is improbable that a bug causing erroneous convergence at this stage would not have been caught by the tests on individual realizations.

2. Practical considerations

The present paper presents the theory of generating particle microstates compatible with the nontrivial thermal equilibria of particular nonlinear *fluid* equations. Of course, a complete particle simulation does not approximately integrate a fluid equation, but rather a nonequilibrium kinetic equation. Therefore, achieving the inviscid, undriven fluid limit may not be trivial in practice.

In general, the physics contained in kinetic equations includes wave-particle interactions (Landau damping). To attain the fluid limit, one must turn off the Landau resonance by setting the parallel wave number $k_\parallel$ to zero. That is easy in unsheared slab geometry but not necessarily trivial in the presence of magnetic shear, which must be set to zero.

If one is to attain the Hasegawa-Mima limit, one must enforce adiabatic electrons and $T_i=0$. Strictly speaking, adiabatic electron response is not compatible with $k_\parallel=0$, so one cannot merely employ a two-species code with $k_\parallel=0$. Rather, the electron response must be built into the gyrokinetic Poisson equation (as in Appendix A), and only the ions should be integrated explicitly. Usually those ions are initialized by sampling from a Maxwellian velocity distribution. Setting $T_i=0$ in that distribution may cause problems; however, it should be adequate to simply use a small but nonzero ratio of T_i/T_e.

One should also note that thermal-equilibrium spectra apply only to homogeneous simulations (with, of course, no macroscopic linear drive or damping). Turning off linear drive is easy in two-scale formulations that incorporate the effects of background profile variations into constant or slowly varying parameters; it may be more difficult for global simulations with nonperiodic boundary conditions. In any event, such simulations are not homogeneous, so are not expected to attain true thermal equilibrium.

In modern simulation practice the full gyrokinetic equation is not solved directly. Rather, the so-called δf algorithm is employed, in which only the deviation δf from a Maxwellian is integrated explicitly (by the method of characteristics). A description of that intrinsically low-noise method, references to the original work, and some theoretical discussion of sampling noise can be found in Ref. [32]. Although the basic method would seem to be well suited to simulations of thermal-equilibrium fluid noise, the long-time behavior of δf simulations may be unstable in the collisionless limit (which is required for the Hasegawa-Mima spectra discussed in the present paper). The basic problem with that limit (unbounded increase of an entropylike function) was elucidated in Ref. [33], and a possible solution (involving the use of a numerical ''thermostat'') was advanced in Ref. [6]. However, further research on both theory and technique is required; therefore, we leave the demonstration of collisionless Hasegawa-Mima thermal-equilibrium spectra in a δf simulation to future work.

Note that a principal shortcoming of all Monte Carlo methods is the slow $1/\sqrt{t}$ rate of convergence. This limits to modest values the number of particles and Fourier amplitudes that can be treated. This need not be an issue for benchmarking purposes, however, particularly for the low-noise δf methods that work well with small numbers of particles.

B. Final remarks

Although our original motivation was the flexible initialization and robust testing of gyrokinetic simulation codes, the physics and algorithms we have discussed may be of more general interest. The δf algorithm mentioned in Sec. V A 2 is itself at core a Monte Carlo sampling technique [32] that can be used for integrating a variety of continuum partial differential equations, possibly unrelated to plasma physics. Gyrocenter motion has much in common with 2D turbulence as well as the dynamics of point vortices moving in two dimensions [34], which can be treated both with particle-simulation techniques [35,36], fluid approaches (Ref. [37], and references therein), and specially designed laboratory experiments (Ref. [38], and references therein). The way of handling the many-to-one relation between the particles and the Fourier amplitudes, as well as the statistical issues surrounding the determination of the basic PDF $P_0(\varphi)$, may be of interest in a variety of contexts, including the use of maximum-entropy methods in pattern recognition [39]. Finally, there should be no problem of principle in extending the calculations to problems with other numbers of invariants, such as the single energy invariant of the 3D Navier-Stokes equation, the single generalized invariant of the Terry-Horton equation (see the final paragraph of Appendix A), or the four invariants of the Hasegawa-Wakatani equation [8,40].

ACKNOWLEDGMENTS

We are grateful to W. W. Lee and J. Reynders for many informative discussions about the theory and numerical implementation of gyrokinetics. This work was supported by U.S. Department of Energy under Contract No. DE-AC02-76-CHO-3073.

APPENDIX A: GYROKINETIC DERIVATION OF THE HASEGAWA-MIMA EQUATION

In an attempt to make the manuscript reasonably self-contained, we present here a brief derivation of the Hasegawa-Mima equation using the gyrokinetic formalism [9]. Further details and discussion of the equation can be found in the original references [11,22], in Bowman's dissertation [23], and in Ref. [41].

In a set of dimensionless variables in which density is normalized to the mean density $\overline{n}$, lengths are normalized to the "sound radius" $\rho_s \doteq c_a/\omega_{ci}$ [where c_a is the sound speed $(T_e/m_i)^{1/2}$ and ω_{ci} is the ion gyrofrequency q_iB/m_ic], times are normalized to L_n/c_a (where L_n is the density scale length), and the electrostatic potential is normalized to $(T_e/e)(\rho_s/L_n)$, the continuity equation for the perturbed ion gyrocenter density n_i^G is

$$\frac{\partial n_i^G}{\partial t} + V_* \frac{\partial \varphi}{\partial y} + \mathbf{V}_E \cdot \boldsymbol{\nabla} n_i^G = 0. \tag{A1}$$

Here, the diamagnetic velocity $V_* \doteq (cT_e/eB)L_n^{-1}$ is unity in the present units, but is written symbolically for emphasis; the dimensionless $\mathbf{E}\times\mathbf{B}$ velocity is $\mathbf{V}_E \doteq \hat{\mathbf{z}}\times\boldsymbol{\nabla}\varphi$. The potential is determined by the *quasineutrality* condition, appropriate for low-frequency, long-wavelength fluctuations:

$$\nabla_\perp^2 \varphi = -(n_i^G - n_e^G). \tag{A2}$$

(One has $n_e^G \approx n_e$ because the electron gyroradius is very small.) The Laplacian term describes the ion polarization charge density $\rho^{\rm pol}$. [The conventional Laplacian in Poisson's original equation is $O(\lambda_D^2/\rho_s^2)$; this is small in the gyrokinetic ordering [15] and is neglected in the approximation of quasineutrality.] This is defined by the continuity equation

$$\partial_t \rho^{\rm pol} + \boldsymbol{\nabla}\cdot\mathbf{j}^{\rm pol} = 0, \tag{A3}$$

where the ion polarization current is $\mathbf{j}^{\rm pol} = n_i q_i \mathbf{V}^{\rm pol}$, with [42]

$$\mathbf{V}^{\rm pol} = \frac{1}{\omega_{ci}}\frac{\partial}{\partial t}\left(\frac{c\mathbf{E}_\perp}{B}\right). \tag{A4}$$

In the approximation of Hasegawa and Mima, the electron response is assumed to be adiabatic [48]:

$$n_e = \varphi. \tag{A5}$$

One then obtains the simplest form of the gyrokinetic Poisson equation

$$(1-\nabla_\perp^2)\varphi = n_i^G. \tag{A6}$$

If one substitutes this expression for n_i^G into the continuity equation (A1), one is led immediately to Eq. (5) of the text.

When the more realistic case of nonadiabatic electron response is considered, both the linear and nonlinear terms are modified. The resulting equation is called the Terry-Horton equation [22]. It conserves just one invariant $[\langle(\delta n_i^G)^2\rangle$, essentially the sum of the energy and enstrophy], so its equilibrium statistical mechanics differs from that of the Hasegawa-Mima equation and, in fact, is quite nontrivial. Nevertheless, we feel that our fundamental concerns of testing gyrokinetic simulations and exploring the generation of two-temperature equilibria are better served by concentrating on the simpler Hasegawa-Mima equation, so we do not consider the Terry-Horton equation further in this work.

APPENDIX B: CALCULATION OF THE FUNDAMENTAL PROBABILITY DENSITY FUNCTION $\mathcal{P}_0(\boldsymbol{\varphi})$

In this appendix, we will use the notation $\boldsymbol{\varphi}$ instead of φ to indicate the set of Fourier amplitudes (more specifically, the set of the real and imaginary parts of each $\varphi_{\mathbf{k}}$), reserving φ for the magnitude of $\boldsymbol{\varphi}$. The discussion in Sec. III shows that the quantity $\mathcal{P}_0(\boldsymbol{\varphi})$ plays a crucial role in the proposed Monte Carlo method. Again, $\mathcal{P}_0(\boldsymbol{\varphi})$ is the PDF for the Fourier amplitudes of the potentials arising from an ensemble of independent gyrocenters, each of which is distributed uniformly.

In the calculations to follow, we will prefix formula numbers from Abramowitz and Stegun [43] with AS, and prefix ones from Gradshteyn and Ryzhik [44] with GR.

1. General expression

To compute $\mathcal{P}_0$, one may recall the standard result, written first for a single real random variable $\tilde{\psi}$, that

$$P(\psi) = \langle \delta(\psi - \tilde{\psi})\rangle. \tag{B1}$$

Of course, when the $\langle\cdots\rangle$ average is expressed in terms of $P(\psi)$ itself, Eq. (B1) is a tautology. However, when the random properties of $\tilde{\psi}$ are expressed in terms of another underlying variable $X \equiv \tilde{x}$ whose density is $P_X(x)$, Eq. (B1) is nontrivial:

$$P(\psi) = \int dx P_X(x)\,\delta(\psi - \tilde{\psi}(x)). \tag{B2}$$

It is often convenient to work with the Fourier transform of this result, i.e., to compute the characteristic function

$$P_k = \int d\psi e^{-ik\psi} P(\psi) = \langle \exp[-ik\tilde{\psi}(\tilde{x})]\rangle. \tag{B3}$$

This average is analytically tractable if the relationship between $\tilde{\psi}$ and $\tilde{x}$ is sufficiently simple.

To apply this procedure to the present problem, we write formally

$$\mathcal{P}_0(\boldsymbol{\varphi}) = \langle \delta(\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}})\rangle, \tag{B4}$$

where the ensemble average is to be taken over the ensemble of independent, uniformly distributed gyrocenters. More explicitly,

$$\mathcal{P}_0(\boldsymbol{\varphi}) = \left\langle \prod_{m=1}^{\mathcal{M}'} \delta(\varphi_r^{(m)} - \tilde{\varphi}_r^{(m)})\,\delta(\varphi_i^{(m)} - \tilde{\varphi}_i^{(m)}) \right\rangle \tag{B5a}$$

$$= \int \frac{d\mathbf{p}^{(1)}}{(2\pi)^2}\frac{d\mathbf{p}^{(2)}}{(2\pi)^2}\cdots\frac{d\mathbf{p}^{(\mathcal{M}')}}{(2\pi)^2} \times \prod_{m=1}^{\mathcal{M}'} e^{i\mathbf{p}^{(m)}\cdot\boldsymbol{\varphi}^{(m)}} \langle e^{-i\mathbf{p}^{(m)}\cdot\tilde{\boldsymbol{\varphi}}^{(m)}}\rangle, \tag{B5b}$$

where

$$\mathbf{p}^{(m)} \doteq \begin{pmatrix} p_r^{(m)} \\ p_i^{(m)} \end{pmatrix}, \quad \boldsymbol{\varphi}^{(m)} \doteq \begin{pmatrix} \varphi_r^{(m)} \\ \varphi_i^{(m)} \end{pmatrix}, \tag{B6}$$

and m ranges over all $\mathcal{M}'$ independent Fourier modes: $\boldsymbol{\varphi}^{(m)} \equiv \boldsymbol{\varphi}_{\mathbf{k}^{(m)}}$. If one defines, for any $\mathbf{k}$, $a_{\mathbf{k}} \doteq [(1+k^2)N]^{-1}$, then the gyrokinetic Poisson equation that relates the potentials to the random gyrocenter positions is, from Eq. (38), $\tilde{\varphi}_{\mathbf{k}} = a_{\mathbf{k}} \Sigma_{j=1}^{N} e^{-i\mathbf{k}\cdot\tilde{\mathbf{x}}^{(j)}}$. One thus has

$$e^{-i\mathbf{p}\cdot\tilde{\boldsymbol{\varphi}}} = \exp\left(-ia_{\mathbf{k}} p_r \sum_{j=1}^{N} \cos(\mathbf{k}\cdot\tilde{\mathbf{x}}^{(j)}) \right) \times \exp\left(ia_{\mathbf{k}} p_i \sum_{l=1}^{N} \sin(\mathbf{k}\cdot\tilde{\mathbf{x}}^{(l)}) \right) \tag{B7}$$

or

$$e^{-i\mathbf{p}^{(m)}\cdot\tilde{\boldsymbol{\varphi}}^{(m)}} = \prod_{j=1}^{N} \mathcal{R}_m^{(j)}(w_r^{(m)}) \mathcal{I}_m^{(j)}(w_i^{(m)}), \tag{B8}$$

where $\mathbf{w} \doteq a\mathbf{p}$ and

$$\mathcal{R}_m^{(j)}(w_r^{(m)}) \doteq \sum_{n_r=-\infty}^{\infty} J_{n_r}(w_r^{(m)}) e^{-in_r\pi/2} e^{in_r \mathbf{k}^{(m)}\cdot\tilde{\mathbf{x}}^{(j)}}, \tag{B9a}$$

$$\mathcal{I}_m^{(j)}(w_i^{(m)}) \doteq \sum_{n_i=-\infty}^{\infty} J_{n_i}(w_i^{(m)}) e^{in_i \mathbf{k}^{(m)}\cdot\tilde{\mathbf{x}}^{(j)}}. \tag{B9b}$$

Thus, the Fourier transform of $\mathcal{P}_0(\boldsymbol{\varphi})$ is

$$\mathcal{P}_0(\mathbf{p}^{(1)},\ldots,\mathbf{p}^{(\mathcal{M}')}) = \left\langle \prod_{m=1}^{\mathcal{M}'} \prod_{j=1}^{N} \mathcal{R}_m^{(j)}(w_r^{(m)}) \mathcal{I}_m^{(j)}(w_i^{(m)}) \right\rangle. \tag{B10}$$

Note that, in general, the ensemble average does not commute with the product symbols.

To simplify Eq. (B10), we first recall that the gyrocenters are independent. Therefore,

$$\mathcal{P}_0(\mathbf{p}^{(1)},\ldots,\mathbf{p}^{(\mathcal{M}')}) = \prod_{j=1}^{N} \left\langle \prod_{m=1}^{\mathcal{M}'} \mathcal{R}_m^{(j)}(w_r^{(m)}) \mathcal{I}_m^{(j)}(w_i^{(m)}) \right\rangle. \tag{B11}$$

It is not hard to see that for uniformly distributed gyrocenters the Fourier amplitudes are independent. Then, $\mathcal{P}_0(\mathbf{p}^{(1)},\ldots,\mathbf{p}^{(\mathcal{M}')}) = \Pi_{m=1}^{\mathcal{M}'} P_0(\mathbf{p}^{(m)})$, where

$$P_0(\mathbf{p}) \doteq \prod_{j=1}^{N} \langle \mathcal{R}^{(j)}(w_r) \mathcal{I}^{(j)}(w_i) \rangle \tag{B12a}$$

$$= \langle \mathcal{R}^{(1)}(w_r) \mathcal{I}^{(1)}(w_i) \rangle^N, \tag{B12b}$$

since each gyrocenter has identical statistics. One has, for any gyrocenter,

$$\langle \mathcal{R}(w_r) \mathcal{I}(w_i) \rangle = \sum_{n_r,n_i} J_{n_r}(w_r) e^{-in_r\pi/2} J_{n_i}(w_i) \times \langle \exp[i(n_r+n_i)\mathbf{k}\cdot\tilde{\mathbf{x}}] \rangle \tag{B13a}$$

$$= \sum_{n} J_n(w_r) e^{in\pi/2} J_n(w_i) \tag{B13b}$$

$$= J_0(w), \tag{B13c}$$

where in the last step we employed Graf's addition theorem (AS 9.1.79). Finally, then,

$$\mathcal{P}_0(\boldsymbol{\varphi}) = \prod_{m=1}^{\mathcal{M}'} P_0(\boldsymbol{\varphi}^{(m)}), \tag{B14}$$

where (now writing $\boldsymbol{\varphi}$ instead of $\boldsymbol{\varphi}^{(m)}$ for convenience)

$$P_0(\boldsymbol{\varphi}) \doteq \int \frac{d\mathbf{p}}{(2\pi)^2} e^{i\mathbf{p}\cdot\boldsymbol{\varphi}} J_0^N(ap) \tag{B15a}$$

$$= \frac{1}{2\pi} \int_0^{\infty} p\,dp\, J_0(\varphi p) J_0^N(ap), \tag{B15b}$$

where $\varphi \doteq |\boldsymbol{\varphi}|$. Note that $P_0(\boldsymbol{\varphi})$ is the PDF for the *two-component vector* $\boldsymbol{\varphi}$ of real and imaginary parts; it is normalized such that $\int d\boldsymbol{\varphi} P_0(\boldsymbol{\varphi}) = \int d\varphi_r d\varphi_i P_0(\boldsymbol{\varphi}) = 1$. However, since by symmetry the result depends only on φ, we will frequently quote the magnitude PDF $P_0(\varphi)$, where $P_0(\varphi) = 2\pi\varphi P_0(\boldsymbol{\varphi})$. If one changes variables to

$$\overline{p} \doteq p/(1+k^2) \tag{B16}$$

and defines

$$\overline{\varphi} \doteq (1+k^2)\varphi \tag{B17}$$

(note that $\overline{\varphi} \equiv n_i^G$), then one obtains the final result

$$\overline{P}_0(\overline{\varphi};N) = (2\pi\overline{\varphi}) \overline{P}_0(\overline{\boldsymbol{\varphi}};N), \tag{B18a}$$

$$\overline{P}_0(\overline{\boldsymbol{\varphi}};N) \doteq \int_0^{\infty} \overline{p}\,d\overline{p} J_0(|\overline{\varphi}|\overline{p}) J_0^N(\overline{p}/N). \tag{B18b}$$

It is useful to note that result (B18b) can also be written in the interesting form

$$\overline{P}_0(\overline{\boldsymbol{\varphi}};N) = \left\langle \delta\left(\overline{\boldsymbol{\varphi}} + \sum_{j=1}^{N} \overline{\mathbf{k}}_j \right) \right\rangle_{\overline{\mathbf{k}}}, \tag{B19}$$

where the $\overline{\mathbf{k}}$'s are effective wave vectors (not to be confused with the $\mathbf{k}$ labels of the Fourier amplitudes) satisfying

$$|\overline{\mathbf{k}}_j| = N^{-1}, \tag{B20}$$

and $\langle \cdots \rangle_{\overline{\mathbf{k}}}$ denotes an average over all possible orientations of the $\overline{\mathbf{k}}$'s.

2. Exact results

The geometrical interpretation (B19) leads immediately to some interesting exact results. First, one has the important conclusion that integral (B18b) vanishes identically if $\overline{\varphi} > 1$, for in this case it is impossible to form a closed $(N+1)$-sided planar figure (possibly with crossed lines) with N of the sides constrained by Eq. (B20):

$$\overline{P}_0(\overline{\varphi}) \equiv 0 \quad (\overline{\varphi} > 1). \tag{B21}$$

Next, for $N=1$ and $N=2$ formula (B18b) can be computed exactly. For $N=1$, it is clear from the geometrical interpretation that the integral vanishes unless $|\overline{\boldsymbol{\varphi}}| = 1$; since the result is independent of orientation one concludes that

$$\overline{P}_0(\overline{\varphi};1) = \delta(\overline{\varphi} - 1). \tag{B22}$$

Consistently, this is just the joint PDF for a pair of random variables $\psi_1 \doteq \cos\theta$ and $\psi_2 \doteq \sin\theta$ for θ distributed uniformly in $[0,\ 2\pi)$: $P(\psi_1,\psi_2) = (2\pi)^{-1}\delta(\psi - 1)$, where $\psi \doteq (\psi_1^2 + \psi_2^2)^{1/2}$. This is just the situation described by the case $N=1$, where we identify θ with the position of the single gyrocenter and $\{\psi_1,\psi_2\}$ with the real and imaginary parts of the potential.

For $N=2$, the effective wave vectors are constrained to form a triangle of area $A(\overline{\boldsymbol{\varphi}},\overline{\mathbf{k}}_1,\overline{\mathbf{k}}_2)$, where

$$A(\mathbf{k}_0,\mathbf{k}_1,\mathbf{k}_2) = \tfrac{1}{2}|\mathbf{k}_0 \times \mathbf{k}_1|\,\delta_{\mathbf{k}_0+\mathbf{k}_1+\mathbf{k}_2,\mathbf{0}}. \tag{B23}$$

In the present case one has $k_1 = k_2 = \frac{1}{2}$, so from the simple geometry of an isosceles triangle one obtains

$$A = \tfrac{1}{4}\overline{\varphi}(1 - \overline{\varphi}^2)^{1/2}. \tag{B24}$$

Since it is well known[7] that

$$\langle \delta(\mathbf{k}_0 + \mathbf{k}_1 + \mathbf{k}_2)\rangle_{\mathbf{k}_1,\mathbf{k}_2} = [(2\pi)^2 A(\mathbf{k}_0,\mathbf{k}_1,\mathbf{k}_2)]^{-1}, \tag{B25}$$

one finally obtains

$$\overline{P}_0(\overline{\varphi};2) = \begin{cases} \dfrac{2}{\pi(1-\overline{\varphi}^2)^{1/2}} & (\overline{\varphi} \leqslant 1) \\ 0 & (\overline{\varphi} > 1). \end{cases} \tag{B26}$$

This result can also be obtained directly from formula (B18b); see GR 6.522.11.

For $N \geqslant 3$, it does not appear possible to obtain integral (B18c) in closed form. For modest N, numerical evaluation is feasible. It is convenient to drop some numerical factors from Eq. (B18b) and thus to write

$$P_0(\overline{\varphi}) = (2\pi\overline{\varphi})\left(\frac{N}{\pi}\right) I, \tag{B27}$$

[7]Result (B25) is employed frequently in the reduction of wavenumber convolutions arising in the statistical theory of 2D homogeneous turbulence; see Ref. [8], Appendix A.

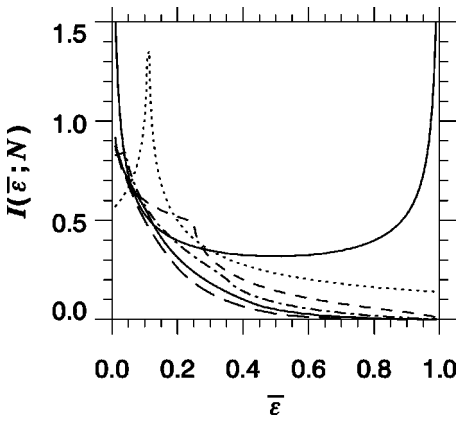

FIG. 17. Numerical evaluation of $\overline{I}(\overline{\epsilon};N)$. Solid line, $N=2$; dotted line, $N=3$; short dashed line, $N=4$; dash-dotted line, $N=5$; dash-triple-dotted line, $N=6$; long dashed line, $N=7$. Each curve contains 200 line segments.

where

$$I(\epsilon;N) \doteq \frac{1}{2}\int_0^\infty p\,dp\,J_0(\sqrt{\epsilon}p)J_0^N(p/\sqrt{N}). \tag{B28}$$

Here $\epsilon \doteq N\overline{\epsilon}$, where $\overline{\epsilon} \doteq \overline{\varphi}^2$. As we will verify later, $I(\epsilon;N)$ is normalized such that

$$I(0;\infty) = 1. \tag{B29}$$

Because of the rapid oscillations of $J_0(\sqrt{\epsilon}p)$ for $\epsilon \gg 1$, straightforward adaptive numerical methods based on grid refinement fail. Therefore, following the guidance in the documentation for [45], we first evaluate the partial integrals I_s defined by integrating the integrand of Eq. (B28) between $j_{0,s}$ and $j_{0,s+1}$, where $j_{0,s}$ is the sth zero of $J_0(\sqrt{\epsilon}p)$ (except that $j_{0,0}=0$). Then, we consider the sequence $\{S_n\}$ of partial sums $S_n \doteq \Sigma_{s=0}^n I_s$ and accelerate the convergence of that sequence by means of Shanks' transformation. The results for $N=2$ through $N=7$ are shown in Fig. 17.

The results for $N=6$ and $N=7$ suggest the approach to a limiting function (at least for $\overline{\epsilon} \ll 1$), which we will show is $e^{-\epsilon}$. To address the case of large N (the usual case in practice) a variety of asymptotic methods may be employed. We turn to those in the next Appendix.

APPENDIX C: ASYMPTOTICS OF P_0

We now develop various asymptotic analyses of the fundamental PDF $\mathcal{P}_0$. In addition to the central importance of $\mathcal{P}_0$ to the Monte Carlo procedure, the asymptotic analysis is interesting in its own right, and comparison of a variety of approaches provides important cross checks on the calculations.

1. The limit $N \to \infty$

It is simplest to begin by considering the limit $N \to \infty$. Note that

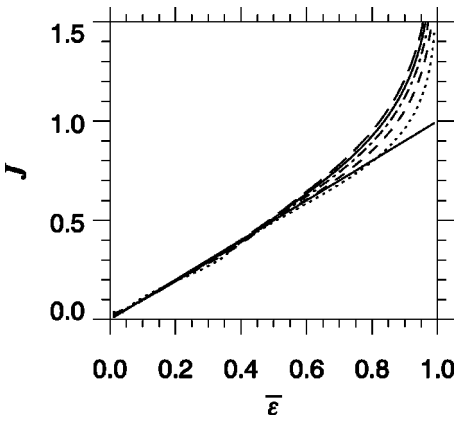

FIG. 18. Numerical evaluation of $J \doteq -\ln I/N$. Solid line, reference case $J=\bar{\epsilon}$ [Eq. (C3)]; dotted line, $N=5$; short dashed line, $N=6$; dash-dotted line, $N=7$; dash-triple-dotted line, $N=8$; long dashed line, $N=9$.

$$\lim_{N\to\infty} J_0^N(p/\sqrt{N}) = \exp(-\tfrac{1}{4}p^2). \tag{C1}$$

(This can be understood as a consequence of the central limit theorem or can be proven directly.) Then,

$$I(\epsilon;\infty) = \frac{1}{2}\int_0^\infty p\,dp\,J_0(\sqrt{\epsilon}p)\exp(-\tfrac{1}{4}p^2), \tag{C2}$$

which is a standard form (AS 11.4.29):

$$I(\epsilon;\infty) = e^{-\epsilon}. \tag{C3}$$

Normalization (B29) follows from this as a special case.

Result (C3) is not in obvious agreement with the exact result that $P_0=0$ for $\bar{\varphi}>1$ ($\epsilon>N$). However, upon rewriting Eq. (C3) in terms of $P_0=2\bar{\varphi}NI$, one has

$$\lim_{N\to\infty} P_0(\bar{\varphi}) = \lim_{N\to\infty} 2\bar{\varphi}Ne^{-N\bar{\epsilon}} \tag{C4a}$$

$$= 4\bar{\varphi}\delta(\bar{\epsilon}) \tag{C4b}$$

$$= 2\delta(\bar{\varphi}) \tag{C4c}$$

[consistently normalized as $\int_0^\infty d\bar{\varphi}P_0(\bar{\varphi})=1$].

Since $\epsilon=N\bar{\epsilon}$, result (C3) suggests plotting $J(\epsilon;N)\doteq -\ln I/N$. That is done in Fig. 18 for $N=5$ to 10. Form (C3) is seen to be a good approximation for $\bar{\epsilon}\lesssim 0.5$. For $\bar{\epsilon}\lesssim 1$, a noticeable departure from Eq. (C3) is seen; the upward curvature with $\lim_{\bar{\epsilon}\to 1} J=\infty$ is required in order to satisfy $I=0$ for $\bar{\epsilon}>1$. Except for an overall N-dependent height, the curvature near $\bar{\epsilon}=1$ is seen to approach a limiting form. In Sec. C 6, we will use a saddle-point method to reproduce the large-$\bar{\epsilon}$ behavior quite well.

2. Standard polar representation of $I(\epsilon;N)$

For later use it is convenient to revert to a double integral by recalling that, for any function $F(p)$, where $p\doteq|\mathbf{p}|$,

$$\frac{1}{2\pi}\int d\mathbf{p}e^{i\mathbf{p}\cdot\boldsymbol{\varphi}}F(p) = \int_0^\infty p\,dp\,J_0(|\boldsymbol{\varphi}|p)F(p); \tag{C5}$$

we identify

$$\epsilon \doteq |\varphi|^2. \tag{C6}$$

[Of course, this was from where form (B28) originally came.] Thus,

$$I(\epsilon;N) = \frac{1}{4\pi}\int d\mathbf{p}e^{i\mathbf{p}\cdot\boldsymbol{\varphi}}J_0^N(p/\sqrt{N}), \tag{C7}$$

with $I(\epsilon;\infty)$ following from the use of Eq. (C1). If we introduce $\bar{\varphi}$ and q by

$$\varphi \doteq \sqrt{N}\bar{\varphi}, \quad p \doteq \sqrt{N}q, \tag{C8}$$

we obtain from Eq. (C7) the standard form

$$I(\epsilon;N) = \frac{N}{2\pi}\mathrm{Re}\int_{-\pi/2}^{\pi/2} d\theta \int_0^\infty q\,dq\,\exp[N\Phi(q;\theta)], \tag{C9}$$

where

$$\Phi(q;\theta) \doteq i\bar{\varphi}_q(\theta)q + \ln J_0(q), \tag{C10a}$$

$$\bar{\varphi}_q(\theta) \doteq \bar{\varphi}\cos\theta \tag{C10b}$$

(the q subscript on $\bar{\varphi}_q$ reminds one that $\bar{\varphi}_q$ is the projection of $\boldsymbol{\varphi}$ onto the $\mathbf{q}$ vector, not onto a fixed Cartesian $\hat{\mathbf{x}}$). We took advantage of the symmetry in θ to restrict the integral to $(-\frac{1}{2}\pi,\frac{1}{2}\pi)$ at the expense of taking the real part. [Because $\bar{\varphi}_x(\theta)$ is even, I is also twice the integral from 0 to $\frac{1}{2}\pi$; however, for later use we prefer the symmetrical integration domain.] For $N\to\infty$, one may replace

$$\ln J_0(q) \to -\tfrac{1}{4}q^2. \tag{C11}$$

3. Cumulant expansions and small ϵ

A variety of techniques are available to treat the limit $\bar{\epsilon}\ll 1$. Those include cumulant expansions [8] of the positive-definite function $J_0^N(p/\sqrt{N})$ and asymptotic expansion of the integral representation of formula (C7) arising from Laplace transformation. The details are somewhat lengthy and tedious, so are omitted here because of length constraints;[8] one finds

$$\bar{I}(\bar{\epsilon};N) \approx \exp[-N(\bar{\epsilon}+\tfrac{1}{4}\bar{\epsilon}^2)] \quad (\bar{\epsilon}\ll 1). \tag{C12}$$

This result can also be recovered from the saddle-point analysis given in Sec. C 6.

[8] An earlier and lengthier version of this paper that includes additional figures and details of the small-$\bar{\epsilon}$ calculations is available as Ref. [24].

4. The limit $\bar{\epsilon} \gg 1$

First, we will reestablish result (B21) that

$$\bar{I}(\bar{\epsilon};N) \equiv 0 \quad (\bar{\epsilon} \gg 1). \tag{C13}$$

Now $\bar{I}$ is proportional to the integral

$$F(\bar{\varphi};\theta) = \int_0^\infty q dq e^{N\Phi(q,\theta)}. \tag{C14}$$

For real θ the integrand is analytic everywhere in the finite complex q plane. An application of Cauchy's theorem then leads one to $2i\,\mathrm{Im}\,F = -C$, where C is the integral around a large semicircle in the upper half of the q plane. If the integral vanishes on that semicircle as its radius approaches infinity, one may then deduce that $\mathrm{Im}\,F(\theta) \equiv 0$ for all θ between $-\frac{1}{2}\pi$ and $\frac{1}{2}\pi$. Since F is an analytic function of θ, F itself must vanish. Now for sufficiently large $|q|$, $J_0(q) \sim (2/\pi q)^{1/2}\cos(q-\frac{1}{4}\pi)$. With $q = x+iy$, one has $|\cos q| = (\cos^2 x \cosh^2 y + \sin^2 x \sinh^2 y)^{1/2}$, which can be reduced with the aid of various identities to $|\cos q| = (\cosh^2 y - \sin^2 x)^{1/2} \leqslant \cosh|y|$. Convergence is then controlled by

$$\mathrm{Re}[\,i\bar{\varphi}q + \ln J_0(q)\,] = -\bar{\varphi}y + \ln|J_0(q)| \tag{C15a}$$

$$= -\bar{\varphi}y + \ln|\cos(q-\tfrac{1}{4}\pi)| + \frac{1}{2}\ln(2/\pi|q|) \tag{C15b}$$

$$< -\bar{\varphi}y + \ln\cosh y. \tag{C15c}$$

The least-convergent case is for large y, where $\cosh y \sim \frac{1}{2}e^y$; hence $\mathrm{Re}\,\varphi < -\bar{\varphi}y + y = (1-\bar{\varphi})y$. Therefore, for $\bar{\varphi} > 1$ convergence is assured and result (C13) follows. It is easy to see that for sufficiently large N it holds also for $\bar{\varphi} = 1$.

5. Large N

It is noteworthy that result (C3) is valid for all ϵ in spite of the rapid oscillations of $J_0(\sqrt{\epsilon}p)$ for large ϵ; no asymptotics are involved. This is unfortunate, in a way, because the case of large but finite N is not exactly solvable and some asymptotic methods will be required. It is therefore useful to give an alternate derivation of result (C3) in order to explain why the rapid oscillations for large ϵ do not lead to a simplifying asymptotics and to motivate later work. We will show that result (C3) can be considered to follow from a steepest-descent calculation that is exact for the present case.

We may choose either a Cartesian (p_x, p_y) or polar (p, θ) representation. Each has certain advantages. In general, one expects the polar representation to be superior because it makes explicit the symmetry that the integral depends on only $\varphi \equiv |\boldsymbol{\varphi}|$, not φ_x and φ_y separately. However, the p integral runs from 0 to ∞, leading to concerns about contributions from the vicinity of the origin. In the Cartesian representation, the contours run from $-\infty$ to ∞ and the p_x and p_y integrals behave quite symmetrically; however, that representation obscures the dependence on φ alone.

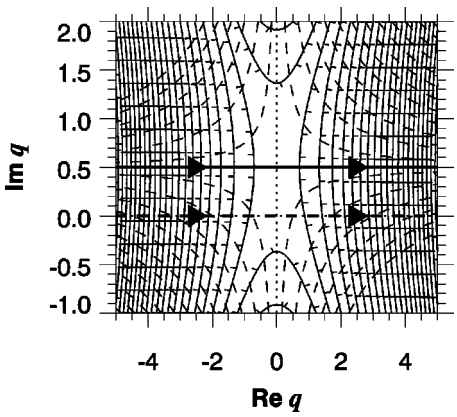

FIG. 19. The simple saddle point for the Cartesian representation with $N = \infty$. Dashed line, original contour; solid line, path of steepest descent. Note that the contour of steepest descent is parallel to the abscissa for a simple quadratic saddle.

a. Cartesian formulation

We begin with the Cartesian formulation. Then,

$$I(\epsilon;\infty) = X(\varphi_x)X(\varphi_y), \tag{C16}$$

where

$$X(\varphi_x) \doteq \frac{1}{2\sqrt{\pi}} \int_{-\infty}^{\infty} dp_x e^{ip_x\varphi_x} \exp(-\tfrac{1}{4}p_x^2) \tag{C17a}$$

$$= \frac{1}{2\sqrt{\pi}} e^{-\varphi_x^2} \int_{-\infty}^{\infty} dp_x \exp[-\tfrac{1}{4}(p_x - 2i\varphi_x)^2]. \tag{C17b}$$

Upon introducing $\bar{p}_x \doteq p_x - 2i\varphi_x$, one finds

$$X(\varphi_x) = \frac{1}{2\sqrt{\pi}} e^{-\varphi_x^2} \int_{\bar{\mathcal{C}}} d\bar{p}_x \exp(-\tfrac{1}{4}\bar{p}_x^2), \tag{C18}$$

where the contour $\bar{\mathcal{C}}$ is a horizontal line at a distance of $2\varphi_x$ below the real axis. However, since the $\bar{p}_x$ integral converges within the 90° cones centered on the real axis, one may use Cauchy's theorem to deform $\bar{\mathcal{C}}$ to the real axis. The resulting integral, of a Gaussian with variance $\sigma^2 = 2$, is standard; one finds $X(\varphi_x) = e^{-\varphi_x^2}$. Then,

$$I(\epsilon;\infty) = e^{-\varphi_x^2} e^{-\varphi_y^2} = e^{-|\varphi|^2} = e^{-\epsilon}. \tag{C19}$$

The oscillations were transformed away by the contour deformation. That is, the original contour $\mathcal{C}$ may be deformed to the path of steepest descent that passes over the simple saddle centered at $p_x = 2i\varphi_x$ (see Fig. 19). Although the location of this saddle moves to $i\infty$ as $\varphi_x \sim \sqrt{\epsilon} \to \infty$, the contributions to the integral along the path of steepest descent are sensibly independent of that location, coming from a region of $O(1)$ centered on the saddle.

As is well known, this exact result for integration along a contour of steepest descent over a simple saddle is identical to the result of the standard algorithm that writes

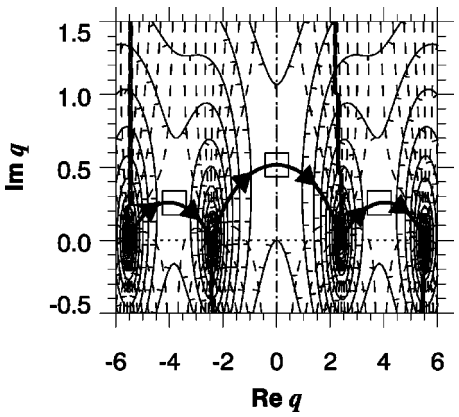

FIG. 20. Contour plot of $\varphi(q)$ for $\overline{\varphi}=0.25$. Saddle points are at the centers of the squares. Solid lines, $\mathrm{Re}\,\Phi=\mathrm{const}$; dashed lines, $\mathrm{Im}\,\Phi=\mathrm{const}$. Medium-thickness lines, branch cuts; heavy curve, path of steepest descent.

$$\int_{\mathcal{C}} dz\,\exp[N\Phi(z)] \approx e^{N\Phi(z_0)} \int_{\mathcal{C}} dz\,\exp[\tfrac{1}{2}N\Phi''(z_0)(z-z_0)^2] \tag{C20a}$$

$$= \left(\frac{2\pi}{N|\Phi''|}\right)^{1/2} e^{N\Phi(z_0)}, \tag{C20b}$$

where $\varphi'(z_0)=0$ and we assume for this example that $\Phi''(z_0)$ is real and negative. This suggests that in the limit of large but finite N a steepest-descent calculation will still be useful, though no longer exact.

In the limit $N\to\infty$, one thus observes that the integral factors in a Cartesian representation and may be evaluated by performing two independent contour integrations. Unfortunately, for finite N the presence of higher-order terms $[O(p^4)]$ in the expansion of $J_0^N(p/\sqrt{N})$ prevents such a factorization. It is therefore desirable to consider the polar representation in some detail.

b. Polar formulation

The form of the q integral in Eq. (C9) suggests the use of the method of steepest descent, although we will see that there are complications in the present case because the q integral begins at $q=0$, not $q=-\infty$. Define

$$I_{\pm} \doteq \frac{N}{2\pi} \int_{-\pi/2}^{\pi/2} d\theta \int_0^{\pm\infty} q\,dq\,e^{N\Phi}, \tag{C21}$$

so that $I=\mathrm{Re}\,I_+$. Since the integrand is an analytic function of q, one may use Cauchy's theorem to prove that

$$I_+ = I_- + S, \tag{C22}$$

where S is the contribution from the path $\mathcal{S}$ of steepest descent. A representative contour plot of $\Phi(q)$ is shown in Fig. 20. It reveals the existence of saddle points located near the zeros of $J_1(q)$, and of sinks located at the zeros of $J_0(q)$. The path of steepest descent is also shown in Fig. 20. Further discussion of the saddle points is given in Sec. C 6.

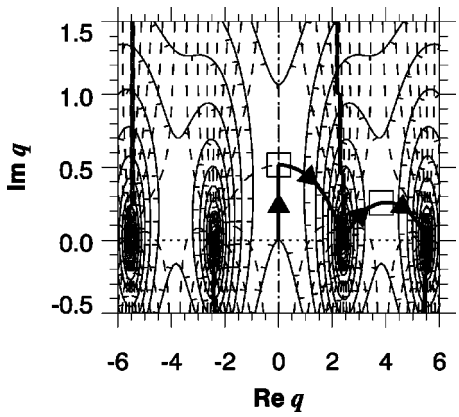

FIG. 21. Alternate integration contour: $\mathcal{U}$ (vertical line) $+\mathcal{S}'$ (curved line).

Unfortunately, by replacing $q\to -q$ in I_- and noting that $J_0(q)$ is an even function, one can easily prove that $I_- = I_+^*$, so $\mathrm{Re}\,I_- = \mathrm{Re}\,I_+$. Upon taking the real part of Eq. (C22), one therefore proves that $\mathrm{Re}\,S=0$ and the steepest-descent contribution to $I=\mathrm{Re}\,I_+$ vanishes when q is integrated over the entire path $\mathcal{S}$ from $-\infty$ to ∞ (and when θ is integrated over the entire contour running from $-\frac{1}{2}\pi$ to $\frac{1}{2}\pi$). Therefore introducing I_- is not helpful. It is more productive to note that

$$I_+ = U + S', \tag{C23}$$

where U is the contribution from the (uphill) path $\mathcal{U}$ along the imaginary axis beginning at $q=0$ and ending at the central saddle point at $q=i\hat{q}$ ($\hat{q}$ being real), and S' is the integral over the right-hand half of $\mathcal{S}$ (see Fig. 21). Because the horizontal symmetry has been broken, it is not true that S' is purely imaginary; one has

$$I = \mathrm{Re}\,U + \mathrm{Re}\,S'. \tag{C24}$$

Before attempting a general analysis, it is useful to explicitly verify Eq. (C23) in the special case (C11), for which all integrals can be performed exactly and one can easily understand the sizes of the various contributions. We thus consider

$$I_+ = \frac{N}{2\pi} \int_{-\pi/2}^{\pi/2} d\theta \int_0^{\infty} q\,dq\,e^{N\Phi}, \tag{C25}$$

where

$$\Phi \doteq iq\overline{\varphi}_q - \frac{1}{4}q^2. \tag{C26}$$

With $q=x+iy$, the path of steepest descent is readily seen to be horizontal, $y=\hat{q}$ with

$$\hat{q} = 2\overline{\varphi}_q. \tag{C27}$$

Upon introducing the normalized horizontal distance p from the saddle point by

$$q = i\hat{q} + (2/N)^{1/2} p, \tag{C28}$$

one has

$$S' = \frac{1}{\pi}\int_{-\pi/2}^{\pi/2} d\theta e^{-\epsilon_q}\int_0^\infty dp(p+i\sqrt{2\epsilon_p})e^{-p^2/2} \tag{C29a}$$

$$= S'_r + iS'_i\,, \tag{C29b}$$

where

$$S'_r \doteq \frac{1}{\pi}\int_{-\pi/2}^{\pi/2} d\theta e^{-\epsilon_q}, \tag{C30a}$$

$$S'_i \doteq \frac{1}{\sqrt{\pi}}\int_{-\pi/2}^{\pi/2} d\theta e^{-\epsilon_q}\sqrt{\epsilon_q}. \tag{C30b}$$

Because $\epsilon_q = O(N)$ [for $\overline{\varphi} = O(1)$], the imaginary part of S' is much larger in magnitude than the real part. Explicitly,

$$S'_r = \frac{1}{\pi}\int_{-\pi/2}^{\pi/2} d\theta e^{-\epsilon\cos^2\theta} = e^{-\epsilon/2} I_0(\tfrac{1}{2}\epsilon) = O(\epsilon^{-1/2}); \tag{C31}$$

S'_i is $O(1)$. For both S'_r and S'_i, note that for large ϵ the contributions to the integrals come from regions within a distance of the order of $\sqrt{\epsilon}$ from the end points.

To evaluate U, we introduce the normalized vertical distance p from the saddle point by

$$q = i[\hat{q} + (2/N)^{1/2}p], \tag{C32}$$

so

$$U = -\frac{1}{\pi}\int_{-\pi/2}^{\pi/2} d\theta e^{-\epsilon_q}\int_{-\sqrt{2\epsilon_q}}^{0} dp(p+\sqrt{2\epsilon_q})e^{p^2/2}. \tag{C33}$$

The $p\,dp$ integral can readily be done, yielding a contribution

$$U_1 = 1 - S'_r\,, \tag{C34}$$

where we noted definition (C30a); the second of these terms cancels the real part of Eq. (C29b). The $\sqrt{\epsilon_q}$ term gives a contribution

$$U_2 = -\frac{2}{\pi}\int_0^{\pi/2} d\theta e^{-\epsilon\cos\theta}\sqrt{2\epsilon}\cos\theta\int_0^{\sqrt{2\epsilon}\cos\theta} dp e^{p^2/2}. \tag{C35}$$

This can be evaluated exactly by introducing $y \doteq \sqrt{2\epsilon}\sin\theta$. Then,

$$U_2 = -\frac{2}{\pi}e^{-\epsilon}\int_0^{\sqrt{2\epsilon}} dy e^{y^2/2}\int_0^{\sqrt{2\epsilon-y^2}} dp e^{p^2/2}. \tag{C36}$$

The double integral can be interpreted as a 2D Cartesian integral over a quarter circle of radius $\sqrt{2\epsilon}$, so it is most conveniently done in polar coordinates:

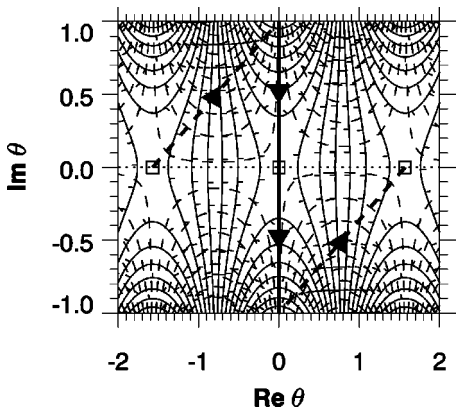

FIG. 22. Contour plot of $-\cos^2\theta$. Dashed lines, contours $\mathcal{C}$; solid line; contour $\mathcal{D}$. Contours $\mathcal{C}$ can be deformed into any curves that connect the indicated endpoints.

$$U_2 = -\frac{2}{\pi}e^{-\epsilon}\int_0^{\pi/2} d\beta\int_0^{\sqrt{2\epsilon}} \rho\, d\rho\, e^{\rho^2/2} \tag{C37a}$$

$$= e^{-\epsilon} - 1. \tag{C37b}$$

The -1 cancels against the 1 in Eq. (C34), so

$$U = e^{-\epsilon} - S'_r\,, \tag{C38}$$

and, upon adding Eqs. (C29b) and (C38),

$$I = \mathrm{Re}\,U + \mathrm{Re}\,S' = e^{-e}, \tag{C39}$$

in agreement with Eq. (C19).

Because the dominant term in S' is imaginary, so does not contribute to I, one might have hoped that $I \approx \mathrm{Re}\,U$. Unfortunately that is not the case; since according to Eq. (C30a) $S'_r = O(\epsilon^{-1/2})$, the second term of Eq. (C38) is much larger than the first, so there is a cancellation of relatively large terms between S' and U. This poses a problem in the general case (C10), where the position of the saddle point obeys an implicit equation and the θ integrals that arise cannot be done analytically. Therefore, we will consider deformations of the θ contour and show how to extract the correct answer from Eq. (C30b). Ignore for a moment the constraint that the end points of the θ integral are pinned to $\pm\frac{1}{2}\pi$. One has explicitly

$$S'_i = \left(\frac{\epsilon}{\pi}\right)^{1/2}\int_{-\pi/2}^{\pi/2} d\theta\cos\theta e^{-\epsilon\cos^2\theta}. \tag{C40}$$

The contour plot of $-\cos^2\theta$ in the complex θ plane is shown in Fig. 22. Saddle points are found at $\theta = \pm\frac{1}{2}\pi$ and at $\theta = 0$. Consider, in particular, the integral from $q = i\infty$ so $q = -i\infty$, which is along a contour $\mathcal{D}$ of constant phase. With $q = x + iy$ and $\cos\theta = \cos x\cosh y - i\sin x\sinh y$, which reduces to $\cos\theta = \cosh y$ on $\mathcal{D}$, one has

$$S'_i = \left(\frac{\epsilon}{\pi}\right)^{1/2}\int_{-\infty}^{\infty} dy\cosh y e^{-\epsilon\cosh^2 y}. \tag{C41}$$

With $z \doteq \sinh y$ and $\cosh^2 y = 1 + \sinh^2 y$, this becomes

$$S_i' = \left(\frac{\epsilon}{\pi}\right)^{1/2} e^{-\epsilon} \int_{-\infty}^{\infty} dz e^{-\epsilon z^2} = e^{-\epsilon}. \tag{C42}$$

Thus, one obtains the correct answer by integrating the dominant q-saddle-point contribution along the θ contour $\mathcal{D}$ of steepest descent. Note that although that integral can be done exactly in the present case, it can also be recovered exactly by the standard procedure of Taylor expansion near the saddle point $\theta=0$ and performing a Gaussian integral.

Of course, S_i' is integrated along the real axis $\mathcal{R}$, not along $\mathcal{D}$. However, since the integrand is analytic, one can write schematically

$$S_i' = \int_{\mathcal{R}} d\theta = \int_{\mathcal{C}} d\theta + \int_{\mathcal{D}} d\theta, \tag{C43}$$

where $\mathcal{C}$ is an arbitrary contour that connects $\theta=-\frac{1}{2}\pi$ to $\theta=i\infty$ and $\theta=-i\infty$ to $\theta=\frac{1}{2}\pi$; see Fig. 22. Since S_i' is real by definition, one has $\mathrm{Im}\int_{\mathcal{C}} = -\mathrm{Im}\int_{\mathcal{D}}$. Thus, with $\int_{\mathcal{R}} d\theta = \int_{\mathcal{C}} d\theta + \int_{\mathcal{D}} d\theta$, one has

$$I_{\mp} = \int_{\mathcal{R}} d\theta \int_U dq + \int_{\mathcal{R}} d\theta S', \tag{C44a}$$

$$I = \mathrm{Re}\Bigg(\underbrace{\int_{\mathcal{R}} d\theta \int_U dq + \int_{\mathcal{R}} d\theta S_r'}_{I} + \underbrace{\int_{\mathcal{C}} d\theta i S_i'}_{-I} \Bigg) + \mathrm{Re} \underbrace{\int_{\mathcal{D}} d\theta i S_i'}_{I}. \tag{C44b}$$

We have shown with the underbraces the values of various terms. This procedure of extracting the steepest-descent contribution to the θ integral can be viewed as adding $0=-I+I$ to the first underbraced terms of Eq. (C44b). However, that interpretation has the deficiency that one could equally well add $\lambda\times 0 = -\lambda I + \lambda I$, where λ is an arbitrary real number; that would make the numerical coefficient of the $\mathcal{D}$ contribution uncertain. That $\lambda=1$ is seen more fundamentally from decomposition (C44a) (with no real part taken) and by the demonstration that we obtain the correct answer for the special case (C11).

6. Saddle-point analysis of the limit $\overline{\epsilon} \lesssim 1$

By appealing to the results for the special $N=\infty$ case treated in the last section, we will in the general case thus approximate

$$I \approx \frac{1}{2}\left(\frac{N}{2\pi}\right) \int_{\mathcal{D}} d\theta \int_{\mathcal{S}} q dq e^{N\Phi}. \tag{C45}$$

The factor of $\frac{1}{2}$ takes into account that we are to integrate only over the right-hand half of the q saddle.

For the general structure of the complex q plane, we refer again to Fig. 20. It is easy to verify the existence and orientation of the saddle points for sufficiently small $\overline{\varphi_q}$. Stationary points are determined from the solution(s) of

$$0 = \frac{\partial}{\partial q}\Phi(q;\theta) = i\overline{\varphi_q} - \frac{J_1(q)}{J_0(q)}. \tag{C46}$$

The solutions of this transcendental equation determine functions $\hat{q}^{(i)} \equiv \hat{q}^{(i)}(\theta)$, where we label the various roots by i; again, the hat signifies dependence on θ. For small $\overline{\varphi_x}$, either J_1 must be relatively small or J_0 must be relatively large. Since the Bessel functions are bounded according to $|J_\nu(z)| \leqslant \frac{1}{2}|z|^\nu \exp(|y|)/\Gamma(\nu+1)$ for $\nu \geqslant -\frac{1}{2}$ (AS 9.1.62) and since for large z the asymptotic forms $J_\nu(z) = [2/(\pi z)]^{1/2}[\cos(z-\frac{1}{2}\nu\pi-\frac{1}{4}\pi)+e^{|y|}O(|z|^{-1})]$ (AS 9.2.1) differ merely by a phase factor, the only possibility is that $J_1(q)$ is small, i.e., that q lies near the zeros $j_{1,s}$ of J_1. Therefore, we write

$$q = j_{1,s} + i\delta \quad (s=0,\ \pm 1,\ \pm 2,\ \ldots) \tag{C47}$$

$(j_{1,0}=0)$. Then,

$$J_1(q) \approx J_1(j_{1,s}) + \tfrac{1}{2}[J_0(j_{1,s}) - J_2(j_{1,s})] i\delta \tag{C48a}$$

$$= \tfrac{1}{2}[J_0(j_{1,s}) - J_2(j_{1,s})] i\delta, \tag{C48b}$$

and Eq. (C46) reduces to

$$\delta = \frac{2\overline{\varphi_q}}{[1 - J_2(j_{1,s})/J_0(j_{1,s})]}. \tag{C49}$$

Because the zeros of the Bessel functions interlace, it is always the case that $J_2(j_{1,s})/J_0(j_{1,s}) < 0$, so δ is always positive. For $s=0$, one has $J_2(j_{1,0}) = J_2(0) = 0$, so we recover Eq. (C27), $q \approx \hat{q}^{(A)} \approx 2i\overline{\varphi_x}$.

To understand the orientation of these saddle points, we compute

$$\Phi'' \equiv \frac{\partial^2}{\partial q^2}\Phi(q;\theta) = -R'(q), \tag{C50}$$

where

$$R(q) \doteq J_1(q)/J_0(q). \tag{C51}$$

With the aid of the identities $J_0'(z) = -J_1(z)$ and $J_1'(z) = J_0(z) - J_1(z)/z$, one finds

$$R'(q) = \frac{J_0(q)J_1'(q) - J_1(q)J_0'(q)}{J_0(q)^2} \tag{C52a}$$

$$= 1 - \frac{1}{q}\frac{J_1(q)}{J_0(q)} + \left(\frac{J_1(q)}{J_0(q)}\right)^2. \tag{C52b}$$

For the saddles, one may simplify Eq. (C52b) with the aid of Eq. (C46) to

$$\Phi_s'' = -\left[1 - \overline{\varphi}^2 + \left(\frac{y\overline{\varphi}}{|q|^2}\right) - i\left(\frac{x\overline{\varphi}}{|q|^2}\right)\right]. \tag{C53}$$

Since $\overline{\varphi} < 1$ and $y > 0$, one has $\mathrm{Re}\,\Phi_s'' < 0$. For the principal saddle at $\hat{q}^{(A)}$ $(x=0)$, the path of steepest descent is hori-

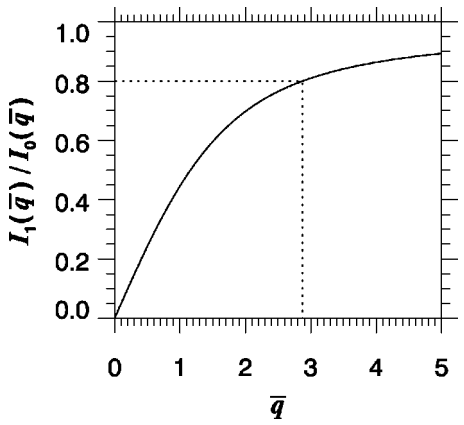

FIG. 23. Illustration of the solution of the implicit equation (C54). Solid curve, $I_1(\overline{q})/I_0(\overline{q})$; horizontal dotted line, specified $\overline{\varphi}_q$; vertical dotted line, derived $\hat{q}$. As $\overline{\varphi}\to 1$, $\hat{q}\to\infty$.

zontal. The other saddles are somewhat skewed, but the path of steepest descent is still basically from left to right.

To determine which saddle point dominates, one may formally carry out the saddle-point integrations. One can verify that the principal root $\hat{q}_x^{(A)}$ dominates for small $\overline{\varphi}_x$. One can also determine that this root continues to dominate as $\overline{\varphi}\to 1$.

For finite $\overline{\varphi}$ the principal root must be determined by solution of the transcendental equation (C46). Symmetry guarantees that $\hat{q}_x^{(A)}$ is purely imaginary for all $\overline{\varphi}$, so we introduce the real number $\overline{q}$ by $q=i\overline{q}$; thus one finds the implicit equation for $\overline{q}(\theta)$ to be

$$I_1(\hat{q})/I_0(\hat{q})=\overline{\varphi}_q(\theta), \tag{C54}$$

where I_ν is the modified Bessel function of the first kind and $\hat{q}$ is the specific $\overline{q}$ that solves Eq. (C54). [See Fig. 23 for a graphical representation of the solution of Eq. (C54).] Equation (C52b) becomes

$$R'(q)\equiv S(\overline{q})=1-\frac{I_1(\overline{q})}{\overline{q}I_0(\overline{q})}-\left(\frac{I_1(\overline{q})}{I_0(\overline{q})}\right)^2. \tag{C55}$$

This function has the properties

$$S(0)=1/2, \tag{C56a}$$

$$S(\infty)=0, \tag{C56b}$$

$$S(\overline{q})\geqslant 0; \tag{C56c}$$

it is graphed in Fig. 24. One may rewrite Eq. (C55) in a form more efficient for numerical computation by using Eq. (C54):

$$S(\hat{q}(\theta))=1-\overline{\eta}(\hat{q}(\theta))-\overline{\varphi}_q^2(\theta), \tag{C57}$$

where

$$\overline{\eta}(\overline{q}(\theta))=\overline{\varphi}_q(\theta)/\overline{q}(\theta). \tag{C58}$$

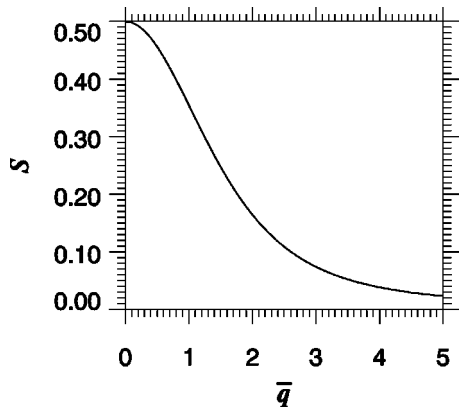

FIG. 24. The function $S(\overline{q})$ [see Eq. (C55)].

We now proceed to analyze the θ integral. Upon integrating over the principal q saddle using the standard Gaussian approximation, we find

$$\frac{1}{2}\mathcal{S}\approx\frac{N}{4\pi}\int_{\mathcal{D}}d\theta i\hat{q}(\theta)\exp[N\Psi(\theta)]\left(\frac{2\pi}{N|\Phi''|}\right)^{1/2}, \tag{C59}$$

where

$$\Psi(\theta)\doteq-\overline{\varphi}_q(\theta)\hat{q}(\theta)+\ln I_0(\hat{q}(\theta)). \tag{C60}$$

To find the stationary points, we calculate

$$\Psi'(\theta)=[-\overline{\varphi}_q(\theta)+I_1(\hat{q})/I_0(\hat{q})]\hat{q}(\theta)+\overline{\varphi}\hat{q}(\theta)\sin\theta \tag{C61a}$$

$$=\overline{\varphi}\hat{q}(\theta)\sin\theta, \tag{C61b}$$

since the term in brackets vanishes by definition of $\hat{q}$ [see Eq. (C54)]. This result has the same form seen in the special case; we are led again to $\theta=0$ as the principal root. At that point $\overline{\varphi}_q=\overline{\varphi}$ and $\Psi''(0)=\overline{\varphi}\hat{q}$, where now $\hat{q}\equiv\hat{q}(0)$. Upon completing the θ integration by integrating vertically downward on $\mathcal{D}$ using the Gaussian approximation, one finally finds

$$I(\epsilon;N)=\frac{1}{2}\left(\frac{1}{\overline{\eta}(\hat{q})S(\hat{q})}\right)^{1/2}\exp(N[-\overline{\varphi}\hat{q}+\ln I_0(\hat{q})]). \tag{C62}$$

7. Summary of the results

In summary, for any Fourier amplitude $\varphi_{\mathbf{k}}\equiv\varphi$, the natural intensity variable is

$$\overline{\epsilon}\doteq|\overline{\varphi}|^2, \tag{C63}$$

where

$$\overline{\varphi}\doteq(1+k^2)\varphi \tag{C64}$$

[see Eq. (B17)]. The fundamental probability density function $P_0(\overline{\varphi})$ is given from Eq. (B27) as $P_0(\overline{\varphi})=2N\overline{\varphi}\overline{I}(\overline{\epsilon};N)$. One has the exact result

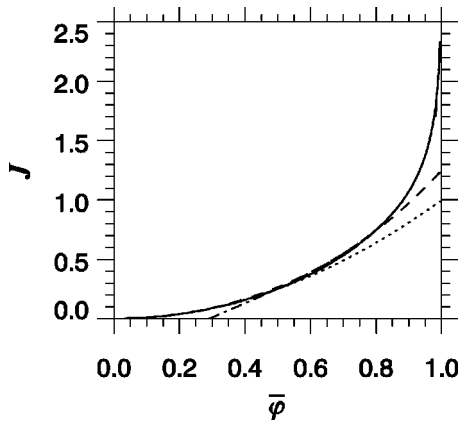

FIG. 25. Comparison of the asymptotic result (C66) with an exact numerical evaluation of $J \doteq -N^{-1}\ln\overline{I}$ for $N=17$. Solid line, exact (numerical) result; long dashed line, (C66) (overlays exact result); dotted line, y_1 (overlays exact result for $\overline{\varphi} \lesssim 0.6$); dashed line, y_2 (overlays exact result for $\overline{\varphi} \lesssim 0.8$); dash-dotted line, y_3 (overlays exact result for $\overline{\varphi} \gtrsim 0.45$). See the text for the definitions of the approximations y_i.

$$\overline{I}(\overline{\epsilon};N) \equiv 0 \quad (\overline{\epsilon} \gtrsim 1). \tag{C65}$$

For $\overline{\epsilon} < 1$ and N large, one has approximately

$$\overline{I}(\overline{\epsilon};N) \approx \left(\frac{1}{2[\,\hat{\eta}(\hat{q})S(\hat{q})\,]^{1/2}} \right) \exp[\,-N\Psi(\hat{q})\,] \tag{C66}$$

(upon introducing a minus sign into the formula to make it look more like a conventional PDF), where

$$\Psi(\overline{q}) \doteq \overline{q}\,\overline{\varphi} - \ln I_0(\overline{q}), \tag{C67a}$$

$$\overline{\eta}(\overline{q}) \doteq \overline{\varphi}/\overline{q}, \tag{C67b}$$

$$S(\overline{q}) \doteq 1 - \overline{\eta}(\overline{q}) - \overline{\varphi}^2, \tag{C67c}$$

and $\hat{q} = \hat{q}(\overline{\varphi})$ is to be determined by solution of the transcendental equation

$$I_1(\hat{q})/I_0(\hat{q}) = \overline{\varphi}. \tag{C68}$$

It is a straightforward exercise to show that for small $\overline{\varphi}$ formula (C66) reduces to

$$\overline{I}(\overline{\epsilon};n) \approx \exp[\,-N(\overline{\epsilon} + \tfrac{1}{4}\overline{\epsilon}^2)\,], \tag{C69}$$

in agreement with Eq. (C12).

The need to solve Eq. (C68), which is conventionally done by numerical iteration, means that the rigorous numerical evaluation of $\mathcal{P}_0$ will be very slow. Fortunately, we generally consider states of intensity sufficiently low that the much simpler approximation (C12) is adequate.

Numerical verification of result (C66) is virtually impossible for realistic N (e.g., $N=1597$) because of loss of precision. Nevertheless, even physically very small N may be asymptotically large. In Fig. 25, we compare a direct numerical integration of $J \doteq -N^{-1}\ln\overline{I}$ for $N=17$ with the asymptotic result (C66). The agreement is seen to be virtually perfect; the relative error (for $\overline{I}$ itself, not its logarithm) is less than 0.5% over the entire domain. Also plotted in these figures are various simpler approximations y_i, $i=1, 2, 3$. The function y is Eq. (C3), $J \approx \overline{\epsilon}$; y_2 follows from Eq. (C69), $J \approx \overline{\epsilon}(1 + \frac{1}{4}\overline{\epsilon})$. The function y_3 is the result of the asymptotic solution of Eq. (C68) for $x \doteq 1 - \overline{\varphi} \ll 1$:

$$\overline{q} \approx \left[2x\left(1 - \frac{1}{2}x \right) \right]^{-1}, \tag{C70a}$$

$$J \approx \left(\overline{q} + \frac{1}{4} \right) x - \frac{1}{2}\ln(2\pi\overline{q}) + \frac{1}{2N}\ln(4\eta S). \tag{C70b}$$

This is seen to be an excellent approximation down to about $\overline{\varphi} \approx 0.5$; it is actually used in the Monte Carlo calculations for $\overline{\varphi}$ close to 1, where the library routine that solves Eq. (C68) has difficulty in converging.

[1] D. J. Evans and G. P. Morris, Comput. Phys. Rep. **1**, 297 (1984).
[2] W. G. Hoover, Phys. Today **37**(1), 44 (1984).
[3] C. K. Birdsall and A. B. Langdon, *Plasma Physics via Computer Simulation* (McGraw-Hill, New York, 1985).
[4] W. G. Hoover, *Computational Statistical Mechanics* (Elsevier, New York, 1991).
[5] W. W. Lee, Phys. Fluids **26**, 556 (1983).
[6] J. A. Krommes, Phys. Plasmas **6**, 1477 (1999).
[7] W. W. Lee, Rev. Mod. Phys. (to be published).
[8] J. A. Krommes, Phys. Rep. **360**, 1 (2002).
[9] D. H. E. Dubin, J. A. Krommes, C. R. Oberman, and W. W. Lee, Phys. Fluids **26**, 3524 (1983).
[10] R. H. Kraichnan and D. Montgomery, Rep. Prog. Phys. **43**, 547 (1980).
[11] A. Hasegawa and K. Mima, Phys. Fluids **21**, 87 (1978).
[12] R. H. Kraichnan, Phys. Fluids **10**, 1417 (1967).
[13] D. Fyfe and D. Montgomery, Phys. Fluids **22**, 246 (1979).
[14] J. Reynders, Ph.D. thesis, Princeton University, 1992 (unpublished).
[15] J. A. Krommes, W. W. Lee, and C. Oberman, Phys. Fluids **29**, 2421 (1986).
[16] J. A. Krommes, Phys. Rev. Lett. **70**, 3067 (1993).
[17] J. A. Krommes, Phys. Fluids B **5**, 1066 (1993).
[18] J. Denavit and J. M. Walsh, Comments Plasma Phys. Controlled Fusion **6**, 209 (1981).
[19] N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller, J. Chem. Phys. **21**, 108 (1953).
[20] M. E. J. Newman and G. T. Barkema, *Monte Carlo Methods in Statistical Physics* (Clarendon, Oxford, 1999).
[21] R. C. Tolman, *The Principles of Statistical Mechanics* (Oxford University Press, Oxford, 1938).

[22] P. Terry and W. Horton, Phys. Fluids **25**, 491 (1982).
[23] J. C. Bowman, Ph.D. thesis, Princeton University 1992 (unpublished).
[24] J. A. Krommes and S. Rath, Princeton Plasma Physics Laboratory Technical Report No. PPPL-3729, 2002 (unpublished) available from www.pppl.gov/pub_report/2002/PPPL-3729.pdf.
[25] R. H. Kraichnan, J. Fluid Mech. **67**, 155 (1975).
[26] R. Kubo, *Statistical Mechanics* (North-Holland, Amsterdam, 1971).
[27] R. C. Davidson, *Methods in Nonlinear Plasma Theory* (Academic, New York, 1972).
[28] M. H. Kalos and P. A. Whitlock, *Monte Carlo Methods*, Basics Vol. 1 (Wiley, New York, 1986).
[29] G. S. Fishman, *Monte Carlo: Concepts, Algorithms, and Applications* (Springer, New York, 1996).
[30] N. G. van Kampen, *Stochastic Processes in Physics and Chemistry* (North-Holland, Amsterdam, 1981).
[31] *E. T. Jaynes: Papers on Probability, Statistics, and Statistical Physics*, edited by R. D. Rosenkrantz (Kluwer Academic, Dordrecht, 1989).
[32] G. Hu and J. A. Krommes, Phys. Plasmas **1**, 863 (1994).
[33] J. A. Krommes and G. Hu, Phys. Plasmas **1**, 3211 (1994).
[34] G. Joyce and D. Montgomery, J. Plasma Phys. **10**, 107 (1973).
[35] J. B. Taylor and B. McNamara, Phys. Fluids **14**, 1492 (1971).
[36] D. A. Schecter, D. H. E. Dubin, K. S. Fine, and C. F. Driscoll, Phys. Fluids **11**, 905 (1999).
[37] C. E. Seyler, Jr., Y. Salu, D. Montgomery, and G. Knorr, Phys. Fluids **18**, 803 (1975).
[38] C. F. Driscoll, D. Z. Jin, D. A. Schecter, and D. H. E. Dubin, Physica C **369**, 21 (2002).
[39] *Maximum Entropy in Action* edited by B. Buck and V. A. Macaulay (Clarendon, Oxford, 1991).
[40] A. E. Koniges, J. A. Crotinger, W. P. Dannevik, G. F. Carnevale, P. H. Diamond, and F. Y. Gang, Phys. Fluids B **3**, 1297 (1991).
[41] W. Horton and Y.-H. Ichikawa, *Chaos and Structures in Nonlinear Plasmas* (World Scientific, Singapore, 1996).
[42] S. Chandrasekhar, *Plasma Physics* (University of Chicago Press, Chicago, 1960).
[43] *Handbook of Mathematical Functions* edited by M. Abramowitz and I. A. Stegun, Natl. Bur. Stand. Appl. Math. Ser. No. 55 (U.S. GPO, Washington, D.C., 1964).
[44] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products* (Academic, New York, 1965).
[45] NAG FORTRAN 77 Library, Mark 19 (The Numerical Algorithms Group Ltd., Oxford, 1999).
[46] The initial account of this work was given by J. A. Krommes, Bull. Am. Phys. Soc. **37**, 1590 (1992).
[47] For some related discussion of the short-wavelength limit, see M. Ottaviani and J. A. Krommes, Phys. Rev. Lett. **69**, 2923 (1992).
[48] In the more physically reasonable *generalized Haseqawa-Mima equation*, electron response is taken to be fluidlike for $k_{\|}=0$ modes. Further references to that model and aspects of the resulting dynamics were discussed by J. A. Krommes and C.-B. Kim, Phys. Rev. E **62**, 8508 (2000).