THE
UNIVERSITY
OF RHODE ISLAND

University of Rhode Island
DigitalCommons@URI

Chemistry Faculty Publications

Chemistry

2002

Stationary Tempering and the Complex Quadrature Problem

Dubravko Sabo

J. D. Doll

See next page for additional authors

Follow this and additional works at: https://digitalcommons.uri.edu/chm_facpubs

Terms of Use
All rights reserved under copyright.

Citation/Publisher Attribution

Sabo, D., Doll, J. D., & Freeman, D. L. (2002). Stationary tempering and the complex quadrature problem. *Journal of Chemical Physics*, *116*(9), 3509-3520. doi: 10.1063/1.1446431
Available at: http://dx.doi.org/10.1063/1.1446431

This Article is brought to you for free and open access by the Chemistry at DigitalCommons@URI. It has been accepted for inclusion in Chemistry Faculty Publications by an authorized administrator of DigitalCommons@URI. For more information, please contact digitalcommons@etal.uri.edu.

Authors
Dubravko Sabo, J. D. Doll, and David L. Freeman

This article is available at DigitalCommons@URI: https://digitalcommons.uri.edu/chm_facpubs/18

JOURNAL OF CHEMICAL PHYSICS VOLUME 116, NUMBER 9 1 MARCH 2002

ARTICLES

Stationary tempering and the complex quadrature problem

Dubravko Sabo and J. D. Doll
Department of Chemistry, Brown University, Providence, Rhode Island 02912

David L. Freeman[a]
Department of Chemistry, University of Rhode Island, Kingston, Rhode Island 02881

(Received 24 September 2001; accepted 4 December 2001)

In the present paper we describe a stochastic quadrature method that is designed for the evaluation of generalized, complex averages. Motivated by recent advances in sparse sampling techniques, this method is based on a combination of parallel tempering and stationary phase filtering methods. Numerical applications of the resulting ''stationary tempering'' approach are presented. We also examine inherent structure decomposition from a probabilistic clustering perspective. © *2002 American Institute of Physics.* [DOI: 10.1063/1.1446431]

I. INTRODUCTION

Since their formal introduction approximately five decades ago,[1,2] Monte Carlo methods have emerged as unusually robust and powerful tools for the study many-dimensional problems.[3] These methods have two features that are of particular note. First, because they tend to treat all problems on a common footing, Monte Carlo methods effectively ''uncouple'' the physical and computational complexity of a given application. Second, as observed from the beginning,[1] these methods have an inherent ''replica'' character that makes them unusually well-suited for implementation in parallel computing environments.

In the present developments we are concerned principally with Monte Carlo approaches as they relate to the evaluation of general statistical averages. To focus the discussion more tightly, we wish to consider the problem of constructing moment generating functions of the form

$$\phi(\boldsymbol{\eta}) = \frac{\int e^{-S(x)} e^{i\boldsymbol{\eta}\cdot\mathbf{x}} d\mathbf{x}}{\int e^{-S(\mathbf{x})} d\mathbf{x}}. \tag{1.1}$$

As will be more fully discussed in Sec. II, the construction of moment generating functions can be considered the prototypical averaging problem. In Eq. (1.1) $\boldsymbol{x}$ represents the natural variables of an ''action,'' $S(\boldsymbol{x})$. We assume for the moment that this action is real-valued and known and that the associated probability distribution, $e^{-S(\boldsymbol{x})}$, is integrable. In any particular application, the identification of $S(\boldsymbol{x})$ is an important, but separate task.

Stochastic quadrature methods estimate averages such as those in Eq. (1.1) by discrete sums of the form[2–5]

$$\phi_N(\boldsymbol{\eta}) = \frac{1}{N}\sum_{n=1}^{N} e^{i\boldsymbol{\eta}\cdot\boldsymbol{x}_n}. \tag{1.2}$$

Provided that the configurations appearing in the sum, $\{\boldsymbol{x}_n\}$, are appropriately sampled from the probability distribution, $e^{-S(x)}$, $\phi_N(\boldsymbol{\eta})$ converges to $\phi(\boldsymbol{\eta})$ as $N\to\infty$. To see more clearly the nature of the sampling problem that is involved, it is useful to recast Eq. (1.1) in inherent structure form.[6,7] As discussed in Appendix A, Eq. (1.1) can be written as

$$\phi(\boldsymbol{\eta}) = \Sigma_\alpha \Gamma_\alpha \langle e^{i\boldsymbol{\eta}\cdot\mathbf{x}}\rangle_\alpha, \tag{1.3}$$

where the summation is over the inherent structures of the probability distribution, $e^{-S(\mathbf{x})}$. It is important to note that the *relevant* number of inherent structures in the summation can be vastly smaller than the *total* number of such structures.[8] In Eq. (1.3) the bracketed quantity denotes a local, inherent structure average while Γ_α is the corresponding statistical weight [cf. Eqs. (A4) and (A5), respectively].

Equation (1.3) emphasizes the two distinct computational problems we face when constructing a stochastic quadrature average. These problems are:

(i) performing a local statistical average within a specified inherent structure; and
(ii) including properly the contributions of all relevant inherent structures.

While typically not treated as separate, identifiable tasks, these steps are implicitly present in the evaluation of any generalized average.

Performing a local average over a well-defined probability distribution is the quintessential Monte Carlo problem and is soluble with well-established methods.[2,3,5] Assuring a proper accounting of the contributions of all relevant inherent structures, however, is a potentially more involved issue whose complexity depends sensitively on the details of the probability distribution in question. When the inherent structures of the distribution are not simply connected (i.e., the distribution is ''sparse''), for example, specialized techniques are generally required in order to achieve a proper sampling.[4] The technical difficulties of assuring a proper sampling notwithstanding, the general methodology for constructing averages of the type in Eq. (1.1) is well-developed and numerically robust when $S(\boldsymbol{x})$ is a real-valued function.

Our ability to treat generalized averages such as those in Eq. (1.1) drops significantly when $S(\boldsymbol{x})$ becomes complex.

---

[a] Electronic mail: freeman@chm.uri.edu

This drop is understandable since a key element in the development of general numerical tools for treating the real-valued limit, the probabilistic interpretation of the distribution $e^{-S(\boldsymbol{x})}$, is lost under such circumstances and the direct use of stochastic quadrature methods becomes problematic. While limited progress can be made using approximate saddle point methods,[9] general, arbitrarily refinable numerical methods remain elusive. This lack of broadly applicable numerical methods is unfortunate in view of the many problems that can be expressed in terms of such generalized, complex-valued averages.

In the present paper, we discuss a general method for the treatment of complex averages that combines stationary phase filtering[10] and sparse sampling techniques.[11–14] Stationary phase Monte Carlo methods provide a means for creating self-adaptive numerical filters that suppress troublesome phase oscillations and provide a natural means for locating the stationary phase regions of the problem. Moreover, these filters produce a natural importance function that can serve as the basis of a practical Monte Carlo procedure. The importance functions so generated, however, are often sparse. Recent developments involving replica and tempering approaches have greatly expanded our general ability to treat sparse sampling problems.[13–16] Because the topics are closely linked, advances in sparse sampling methods have important implications for the complex Monte Carlo problem. These implications, in large measure, are both the motivation for and the major emphasis of the present developments.

The remainder of this paper is organized as follows: Section II outlines the central components of our proposed method and presents a simple numerical example designed to illustrate key elements of its application. Section II also includes a brief review of stationary phase filtering methods. Section III contains sample numerical applications, including an example designed to illustrate the applicability of the present approach to real-time, path-integral dynamical problems.

II. FORMAL DEVELOPMENTS

In the present section we present a generalization of stochastic quadrature methods designed to treat complex averages. This method, described in Sec. II B, combines features adapted from stationary phase filtering techniques[10] and from sparse sampling developments.[11–14] Section II A presents a brief summary of stationary phase filtering methods in order to highlight key features of the method and to establish a consistent notation for the discussion.

A. Stationary phase Monte Carlo methods

In what follows we focus attention on evaluating the generalized moment generating function, $\phi(\boldsymbol{\eta})$, defined by

$$\phi(\boldsymbol{\eta})=\langle e^{i\boldsymbol{\eta}\cdot\boldsymbol{x}}\rangle_S . \tag{2.1}$$

The bracketed quantity on the right-hand side of Eq. (2.1) denotes the average defined by Eq. (1.1). In Eq. (2.1) and the remainder of the present discussion, however, we no longer assume that the average is over a positive probability density, but instead that it involves an exponential of a complex action, $S(\boldsymbol{x})$.

As is the case with their real-valued counterparts, knowledge of the moment generating functions defined by Eq. (2.1) is sufficient to permit the construction of more general averages. If, for example, an arbitrary function of N dimensions, $f(\boldsymbol{x})$, can be written as

$$f(\boldsymbol{x})=\int_{-\infty}^{\infty}\frac{d\boldsymbol{k}}{(2\pi)^N}\hat{F}(\boldsymbol{k})e^{i\boldsymbol{k}\cdot\boldsymbol{x}}, \tag{2.2}$$

then the average of $f(\boldsymbol{x})$ over the distribution $e^{-S(\boldsymbol{x})}$ is given by

$$\langle f(\boldsymbol{x})\rangle_S=\int_{-\infty}^{\infty}\frac{d\boldsymbol{k}}{(2\pi)^N}\hat{F}(\boldsymbol{k})\phi(\boldsymbol{k}). \tag{2.3}$$

The construction of the moment generating function, Eq. (2.1), thus represents the prototypical complex averaging problem.

Direct Monte Carlo methods are not generally useful for the evaluation of complex averages. While one can, in principle, reformulate the problem using a modulus-phase decomposition of $e^{-S(\boldsymbol{x})}$,

$$e^{-S(\boldsymbol{x})}=|e^{-S(\boldsymbol{x})}|e^{i\theta(\boldsymbol{x})}, \tag{2.4}$$

the resulting expressions are numerically ill-posed. Using Eqs. (2.1) and (2.4), for example, the moment generating function becomes

$$\phi(\boldsymbol{\eta})=\frac{\langle e^{i\boldsymbol{\eta}\cdot\boldsymbol{x}}e^{i\theta(\boldsymbol{x})}\rangle_{|e^{-S(\boldsymbol{x})}|}}{\langle e^{i\theta(\boldsymbol{x})}\rangle_{|e^{-S(\boldsymbol{x})}|}}, \tag{2.5}$$

where the bracketed quantities in Eq. (2.5) represent averages over the distribution, $|e^{-S(\boldsymbol{x})}|$. While formally correct, Eq. (2.5) is of little practical value. If the integrands are highly oscillatory (the case of primary interest), then the integrations in Eq. (2.5) are governed by the stationary phase regions of the complex action.[9] Since these regions typically represent a vanishingly small fraction of the volume for which the modulus $|e^{-S(\boldsymbol{x})}|$ is significant, the numerator and denominator in Eq. (2.5) tend to vanish separately leading to an ill-conditioned result. Worse, the severity of these numerical difficulties tends to be an exponentially increasing function of the dimensionality of the averages involved.

The immediate problems associated with complex averages (severe phase oscillations and the lack of a natural importance function) can be addressed using stationary phase Monte Carlo techniques.[10] These methods are based on the invariance of a broad class of integrals to a group of "averaging" operations. Specifically, if the integral defined by

$$I=\int d\boldsymbol{x}f(\boldsymbol{x}), \tag{2.6}$$

involves an infinite domain or if it involves a finite domain over which the integrand is periodic, then the integral is invariant to the replacement of the integrand, $f(\boldsymbol{x})$, by its "preaveraged" value. That is, the original integral is rigorously equivalent to

$$I=\int d\boldsymbol{x}\langle f(\boldsymbol{x})\rangle_\varepsilon, \tag{2.7}$$

where $\langle f(\boldsymbol{x})\rangle_\varepsilon$ is defined by

$$\langle f(\boldsymbol{x})\rangle_\varepsilon=\frac{\int d\boldsymbol{y}P_\varepsilon(\boldsymbol{y})f(\boldsymbol{x}+\boldsymbol{y})}{\int d\boldsymbol{y}P_\varepsilon(\boldsymbol{y})}. \tag{2.8}$$

To demonstrate the equality of Eqs. (2.6) and (2.7), one needs merely to substitute Eq. (2.8) into Eq. (2.7), and to invert the orders of the integration. This process reveals that Eq. (2.7) represents many, superimposed copies of the original integral defined by Eq. (2.6). This demonstration also indicates that the equality of Eqs. (2.6) and (2.7) holds for any $P_\varepsilon(\boldsymbol{x})$ that is integrable and for which the inversion of the order of integrations is valid. Thus, while it will often be convenient to choose $P_\varepsilon(\boldsymbol{x})$ to correspond to a probability density, it is important to recognize that this choice is overly restrictive. In practical terms it is necessary only to choose $P_\varepsilon(\boldsymbol{x})$ in such a manner that the convolutions defined by Eq. (2.8) converge and can be readily evaluated.

Equations (2.6)–(2.8) define a group of operations for which a broad class of integrals remain invariant. By allowing us to remould troublesome integrands without altering the values of the associated integrals, this set of operations provides us with a useful tool for the reformulation of the complex Monte Carlo average, Eq. (2.1). Using Eq. (2.8), for example, we can rewrite Eq. (2.1) as

$$\phi(\boldsymbol{\eta})=\frac{\int d\boldsymbol{x}\langle e^{-S(\boldsymbol{x})}e^{i\boldsymbol{\eta}\cdot\boldsymbol{x}}\rangle_\varepsilon}{\int d\boldsymbol{x}\langle e^{-S(\boldsymbol{x})}\rangle_\varepsilon}. \tag{2.9}$$

We can, in fact, go further and utilize different preaveraging distributions in the numerator and denominator of Eq. (2.9). For simplicity, however, we consider here only the simpler result in which a common distribution is utilized in both terms.

A key feature of Eq. (2.9) is that we have introduced a controllable set of length scales into the problem. The preaverages in Eq. (2.9) tend to destroy the regions of nonstationary phase, thus making the modified integrands ‘‘simpler’’ than their original counterparts. If, for example, S is a rapidly varying function (on the length scale of P_ε) in a particular region, then preaveraging tends to destroy $e^{-S(\boldsymbol{x})}$ in that region. If, on the other hand, S is slowly varying, then preaveraging tends to leave $e^{-S(\boldsymbol{x})}$ relatively unchanged. By construction, therefore, the modified integrands in Eq. (2.9) tend to be oscillatory where they are unimportant and important where they are nonoscillatory. The degree to which the preaveraging suppresses the phase oscillations in the original integrands is set by the (controllable) length scales of the preaveraing distribution, P_ε.

In addition to suppressing phase oscillations, stationary phase filtering produces a natural importance function for the complex quadrature problem. The selection of this importance function is a matter of key practical significance in implementing the method. As emphasized by Kalos and Whitlock,[3] importance sampling is a valuable, general tool in Monte Carlo approaches. By reducing the integrand’s fluctuations, a suitably chosen importance function can produce a large, often infinite, reduction in the variance associated with statistical estimates of the associated integral. To identify the most appropriate importance function for the complex quadrature problem, we rewrite Eq. (2.9) for the moment generating function as

$$\phi(\boldsymbol{\eta})=\frac{\int d\boldsymbol{x}W(\boldsymbol{x})\left\{\frac{\langle e^{-S(\boldsymbol{x})}e^{i\boldsymbol{\eta}\cdot\boldsymbol{x}}\rangle_\varepsilon}{W(\boldsymbol{x})}\right\}}{\int d\boldsymbol{x}W(\boldsymbol{x})\left\{\frac{\langle e^{-S(\boldsymbol{x})}\rangle_\varepsilon}{W(\boldsymbol{x})}\right\}}, \tag{2.10}$$

or, more compactly as

$$\phi(\boldsymbol{\eta})=\frac{\left\langle\frac{\langle e^{-S(\boldsymbol{x})}e^{i\boldsymbol{\eta}\cdot\boldsymbol{x}}\rangle_\varepsilon}{W(\boldsymbol{x})}\right\rangle_W}{\left\langle\frac{\langle e^{-S(\boldsymbol{x})}\rangle_\varepsilon}{W(\boldsymbol{x})}\right\rangle_W}. \tag{2.11a}$$

The bracket notation in Eq. (2.11a) denotes an average over an (as yet to be identified) importance function, $W(\boldsymbol{x})$. The generalization to include the use of different preaveraging distributions and importance functions in the numerator and denominator of Eq. (2.11a) is again straightforward. For completeness, if we were to utilize the preaveraging distributions $P_{\varepsilon_N}(\boldsymbol{x})$ and $P_{\varepsilon_D}(\boldsymbol{x})$ and importance functions $W_N(\boldsymbol{x})$ and $W_D(\boldsymbol{x})$ in the numerator and denominator of Eq. (2.9), for example, then the expression for the moment generating function would become

$$\phi(\boldsymbol{\eta})=\frac{\left\langle\frac{\langle e^{-S(\boldsymbol{x})}e^{i\boldsymbol{\eta}\cdot\boldsymbol{x}}\rangle_{\varepsilon_N}}{W_N(\boldsymbol{x})}\right\rangle_{W_N}}{\left\langle\frac{\langle e^{-S(\boldsymbol{x})}\rangle_{\varepsilon_D}}{W_D(\boldsymbol{x})}\right\rangle_{W_D}}\frac{\int W_N(\boldsymbol{x})d\boldsymbol{x}}{\int W_D(\boldsymbol{x})d\boldsymbol{x}}. \tag{2.11b}$$

The bracketed terms in Eq. (2.11b) represent averages over the associated distributions and can be evaluated using conventional Monte Carlo methods. The ratio of the integrals of the two importance functions in Eq. (2.11b) is analogous to equilibrium partition function ratios and can be estimated using methods similar to those designed to compute thermodynamic free-energy differences.[4,5,17] Equation (2.11b) is useful in situations where the integration regions that are ‘‘important’’ to the numerator and to the denominator differ significantly.

Why is Eq. (2.11) any better than Eq. (2.5), a result we previously dismissed as unacceptable? The essential advantage of Eq. (2.11) over Eq. (2.5) is that the important regions of the integration are no longer automatically a vanishingly small fraction of the total volume. By choosing the importance function wisely, we can restrict attention to the integration regions that dominate the results, thereby avoiding the ill-posed nature of Eq. (2.5).

Generalizing slightly the arguments given by Kalos and Whitlock,[3] it is not hard to show that a reasonable choice for the importance function in Eq. (2.11a) is the modulus of the *preaveraged* complex exponential. That is, it is sensible to select

$$W_\varepsilon(\boldsymbol{x})=|\langle e^{-S(\boldsymbol{x})}\rangle_\varepsilon|. \tag{2.12}$$

This choice amounts to the modulus-phase decomposition of the *preaveraged* exponential

$$\langle e^{-S(\boldsymbol{x})}\rangle_\varepsilon = |\langle e^{-S(\boldsymbol{x})}\rangle_\varepsilon| e^{i\chi(\boldsymbol{x})}. \tag{2.13}$$

Unlike the corresponding result for the "bare" exponential, Eq. (2.4), the modulus-phase decomposition in Eq. (2.13) leads to a computationally viable result. In particular, the modulus, $|\langle e^{-S(x)}\rangle_\varepsilon|$, tends to be small in nonstationary phase regions and the phase, $\chi(\boldsymbol{x})$, weakly oscillatory where the importance function is significant.

The choice of importance functions given in Eq. (2.12) smoothly bridges steepest descents and stationary phase situations.[9] If, for example, the imaginary portion of the action is small, then the important regions of $W_\varepsilon(\boldsymbol{x})$ correspond to minima of the real part of the action. If, on the other hand, the imaginary portion of the action becomes more dominant, $W_\varepsilon(\boldsymbol{x})$ increasingly reflects the stationary phase regions of the problem. Because the stationary phase regions of the problem are no longer a vanishingly small fraction of the integration volume, the numerator and denominator of Eq. (2.11), unlike those of Eq. (2.5), do not automatically tend to vanish.

It is convenient to express Eq. (2.11a) in terms of the inherent structures of the stationary phase importance function, $W_\varepsilon(\boldsymbol{x})$, given by Eq. (2.12). Using the methods of Appendix A [c.f. Eqs. (A3)–(A6)], we can express the moment generating function for this distribution as

$$\phi(\boldsymbol{\eta}) = \frac{\sum_\alpha \Gamma_\alpha \left\langle \frac{\langle e^{-S(\boldsymbol{x})} e^{i\boldsymbol{\eta}\cdot\boldsymbol{x}}\rangle_\varepsilon}{W_\varepsilon(\boldsymbol{x})} \right\rangle_\alpha}{\sum_\alpha \Gamma_\alpha \left\langle \frac{\langle e^{-S(\boldsymbol{x})}\rangle_\varepsilon}{W_\varepsilon(\boldsymbol{x})} \right\rangle_\alpha}, \tag{2.14}$$

where

$$\Gamma_\alpha = \frac{\int_\alpha d\boldsymbol{x} W_\varepsilon(\boldsymbol{x})}{\int d\boldsymbol{x} W_\varepsilon(\boldsymbol{x})}. \tag{2.15}$$

The ratio Γ_α represents the fractional statistical weight of inherent structure α. Since Eq. (2.14) is a ratio, we can further factor out a common inherent structure weight from both the numerator and denominator. Denoting this common statistical weight as Γ_1, we have

$$\phi(\boldsymbol{\eta}) = \frac{\sum_\alpha \gamma_\alpha \left\langle \frac{\langle e^{-S(\boldsymbol{x})} e^{i\boldsymbol{\eta}\cdot\boldsymbol{x}}\rangle_\varepsilon}{W_\varepsilon(\boldsymbol{x})} \right\rangle_{W_\alpha}}{\sum_\alpha \gamma_\alpha \left\langle \frac{\langle e^{-S(\boldsymbol{x})}\rangle_\varepsilon}{W_\varepsilon(\boldsymbol{x})} \right\rangle_{W_\alpha}}, \tag{2.16}$$

where γ_α is given by

$$\gamma_\alpha = \frac{\Gamma_\alpha}{\Gamma_1}. \tag{2.17}$$

We now turn to the problem of developing a viable procedure for evaluating the preaverages that appear in Eq. (2.8). We assume, for simplicity, that these involve a normalized Gaussian distribution of the form

$$P_\varepsilon(y) = \frac{\exp\left(-\frac{1}{2}\boldsymbol{y}^T\cdot(\boldsymbol{\varepsilon}^2)^{-1}\cdot\boldsymbol{y}\right)}{\sqrt{(2\pi)^N \mathrm{Det}(\boldsymbol{\varepsilon}^2)}}. \tag{2.18}$$

Here, $\boldsymbol{\varepsilon}^2$ is the positive definite covariance matrix that defines the Gaussian distribution and whose linear dimension is equal to the number of integration variables in Eq. (2.8). Although this matrix will often be taken to be diagonal in practical applications, it is both straightforward and useful to retain the generalized quadratic form in Eq. (2.18). Since it is not generally possible to evaluate the necessary Gaussian integrals analytically, it is convenient to consider an implementation based on gradient approximations to the various preaverages. Methods designed to correct for the effects of these gradient approximations have been considered elsewhere.[18] Expanding the complex action through second-order, and performing the associated integrations[19] we obtain

$$\langle e^{-S(\boldsymbol{x})}\rangle_\varepsilon = e^{-S(\boldsymbol{x})} \frac{\exp\left(\frac{1}{2}\boldsymbol{B}^T\cdot(\boldsymbol{1}+\boldsymbol{\varepsilon}^T\boldsymbol{S}''\boldsymbol{\varepsilon})^{-1}\cdot\boldsymbol{B}\right)}{\sqrt{\mathrm{Det}(\boldsymbol{1}+\boldsymbol{\varepsilon}^T\boldsymbol{S}''\boldsymbol{\varepsilon})}}, \tag{2.19}$$

where the matrix $\boldsymbol{\varepsilon}$ is formally the "square-root" of the matrix $\boldsymbol{\varepsilon}^2$ [i.e., $(\boldsymbol{\varepsilon})(\boldsymbol{\varepsilon}) = (\boldsymbol{\varepsilon}^2)$]. The vector $\boldsymbol{B}$ in Eq. (2.19) is given by matrix product expression

$$\boldsymbol{B} = \boldsymbol{\varepsilon}\cdot\boldsymbol{S}', \tag{2.20}$$

and the matrices of first and second derivatives are given by

$$(\boldsymbol{S}')_n = \frac{\partial S}{\partial x_n} \tag{2.21}$$

and

$$(\boldsymbol{S}'')_{n,m} = \frac{\partial^2 S}{\partial x_n \partial x_m}, \tag{2.22}$$

respectively. Combining Eqs. (2.12) and (2.19)–(2.22) we obtain an approximation to the stationary phase importance function, $W_\varepsilon(\boldsymbol{x})$, in terms of the action and its first and second derivatives. Using methods adapted from semiclassical collision theory,[20] one can assign the proper branch of the square root appearing in Eq. (2.19) from the phases of the eigenvalues of the complex matrix that appears in the determinant of Eq. (2.19). Examining Eq. (2.19) in detail, we see that preaveraging process acts as a generalized "band-pass filter" that modifies the original integrand only slightly in near stationary phase regions, but effectively damps it to zero elsewhere.

Conventional stationary phase expressions can be recovered from Eq. (2.14) by expanding the gradient approximation to the stationary phase importance function [obtained by combining Eqs. (2.12) and (2.19)] to second-order about the associated stationary phase point and performing the resulting integrations analytically. More generally, however, the stationary phase importance function serves as the starting point for refinable, stochastic approach to the complex quadrature problem.

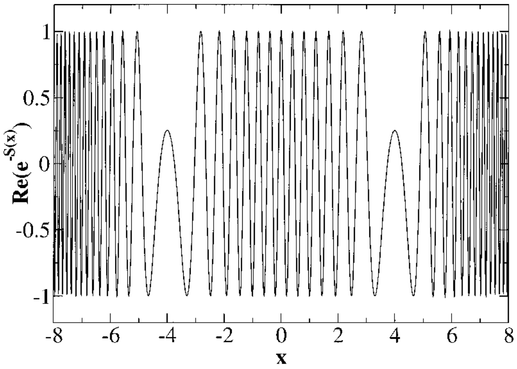

FIG. 1. Plot of the real part of $e^{-S(x,t)}$ for the Airy action, Eq. (2.23), ($t = -16$).

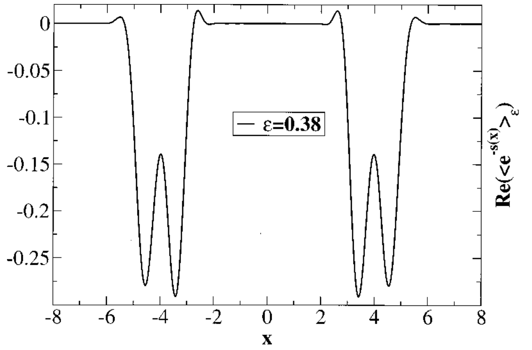

FIG. 3. Plot of the real part of the second-order gradient approximation to $\langle e^{-S(x,t)}\rangle_\varepsilon$ for the Airy action. Results are shown for $t=-16$, $\varepsilon=0.38$.

It is useful to illustrate the application of the stationary phase Monte Carlo method with a simple, one-dimensional example. For simplicity, we assume that the ‘‘action’’ is given by

$$S(x,t)=-i\left(\frac{x^3}{3}+tx\right), \tag{2.23}$$

where t corresponds to a (real) controllable parameter. Using standard special function definitions, the moment generating function for this action specified by Eq. (2.1) is given analytically by

$$\phi(\eta)=\frac{Ai(t+\eta)}{Ai(t)}, \tag{2.24}$$

where $Ai(t)$ is the familiar Airy function.[21]

The stationary phase preaverages and importance functions required for a numerical treatment based on gradient-level stationary phase filtering methods are obtained by combining Eqs. (2.19)–(2.23) and (2.12). Figure 1 displays the real portion of the ‘‘bare’’ complex exponential, Re$[e^{-S(\mathbf{x},\mathbf{t})}]$, for the case where $t=-16$. For negative values of the parameter t, there are two real-valued stationary phase points, $x=\pm\sqrt{-t}$. Rapid oscillations, evident in Fig. 1 except in the vicinity of the stationary phase regions, make the direct numerical integration of Eq. (2.1) difficult. Second-order gradient approximations to the importance function, shown in Fig. 2 as a function of the preaveraging length scale, ε, display the characteristic band-pass nature of the importance function. As the parameter t increases toward zero and ultimately becomes positive, the stationary phase points coalesce and eventually become complex. Reflecting this behavior, the corresponding inherent structures of $W_\varepsilon(x)$ merge, and, as t becomes positive, are strongly attenuated. As will be discussed more fully in Sec. II B, the sharpness of the inherent structures of $W_\varepsilon(\mathbf{x})$ is a function of the preaveraging length scale, ε. In general terms, maximum compression of the importance function is achieved when the filter length scale is comparable to the natural width of the corresponding stationary phase region.[18] In the case of Fig. 2, this corresponds to a choice of $\varepsilon=0.38$.

Figure 3 shows the real portion of the preaveraged complex exponential, Re$[\langle e^{-S(\mathbf{x},\mathbf{t})}\rangle_\varepsilon]$, again for the case where $t=-16$. Comparing Figs. 1 and 3, we see that stationary phase filtering has suppressed the troublesome oscillations in the problem, thereby dramatically simplifying the original integrand. Figure 4 compares the moment generating functions obtained using various preaveraging length scales with the exact analytic result obtained from Eq. (2.24). The preaveraged integrands [cf. Eq. (2.19)] are sufficiently well behaved that the results in Fig. 4 can be (and were) produced by simply integrating Eq. (2.11a) with a standard numerical

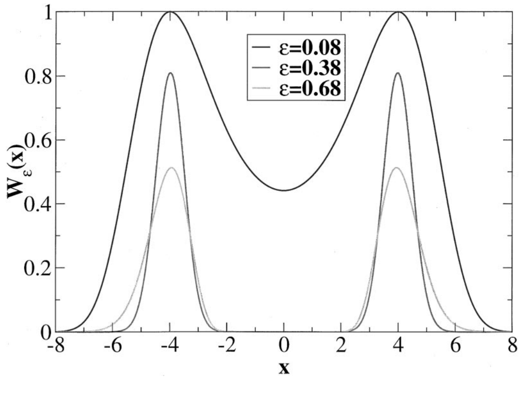

FIG. 2. Second-order gradient result for the stationary phase importance function, $W_\varepsilon(x,t)$, for the Airy action. Plots correspond to $\varepsilon=(0.08,0.38,0.68)$. All results correspond to $t=-16$. Note the variation of the dispersion about the stationary phase regions with filter length scale.

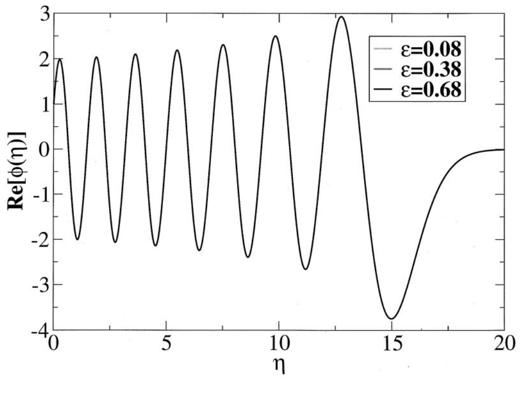

FIG. 4. Plot of the real part of the second-order gradient approximation to the moment generating function, $\phi(\eta,t=-16)$, for the Airy action. Results for $\varepsilon=(0.08,0.38,0.68)$ were obtained by integrating Eq. (2.9) using direct quadrature methods. Only one curve is visible because, as discussed in the text, the second-order results are exact for this example.

package such as Mathematica. More generally, as discussed in Sec. III, such results will be obtained from Monte Carlo estimates of the integrals involved. We see from Fig. 4 that the second-order gradient results are quite good. They are, in fact, *exact* for this example. This somewhat surprising result is a consequence of the cubic nature of the Airy action [Eq. (2.23)]. While the second-order gradient treatments of the preaverages in the numerator and denominator of Eq. (2.11a) are separately approximate, each term differs from the corresponding exact value by a common, multiplicative factor. This cancellation of errors does not hold for Eq. (2.11b) where different preaveraging distributions for numerator and denominator are involved.

B. Stationary tempering

As summarized in the preceding section, the stationary phase Monte Carlo method addresses two of the central issues associated with the application of stochastic quadrature techniques to complex averages. The method suppresses troublesome phase oscillations, and, in the process, provides a natural importance function. There is, however, a Faustian element to these developments. As is apparent in Fig. 2, the importance function produced by the preaveraging process can be sparse in nature. That is, there is no guarantee that the inherent structures of the stationary phase importance function are connected. Under such circumstances specialized methods are required in order to assure a valid statistical sampling.[4]

Sparse sampling problems arise in a number of familiar contexts, the most common of which are perhaps the ‘‘rare event’’ problems associated with thermally activated processes. At low temperatures, the equilibrium configurational distributions for general physical systems typically cluster about isolated inherent structures that correspond to the various local minima of the potential energy surface of the system. Unless special care is exercised, conventional sampling methods typically fail to sample the entire distribution, instead becoming trapped in these isolated, local minima. Techniques for detecting and overcoming sparse sampling difficulties have been discussed elsewhere.[4,5,11–14]

A general strategy for dealing with sparse distributions is to create a computational ensemble that is ‘‘richer’’ in the space of trial moves than are conventional Monte Carlo sampling procedures. A way of doing this is combine results from several simulations, each corresponding to a different value of a control parameter such as the temperature. Assuming that the control parameter involved influences the connectedness of the relevant probability distribution, one then performs random walks for each of the computational elements involved, using configurations selected from the computational ensemble to supplement conventional trial moves. While the details of the various steps differ depending on the particular sparse sampling procedure involved, the essential result is that if the acceptance–rejection logic of the process is properly designed and if the computational ensemble is suitably chosen, then the approach produces a proper sampling for *all* elements in the ensemble, including those for which the associated probability densities are sparse. Both J-walking and parallel tempering techniques are based on this general strategy.[11–16]

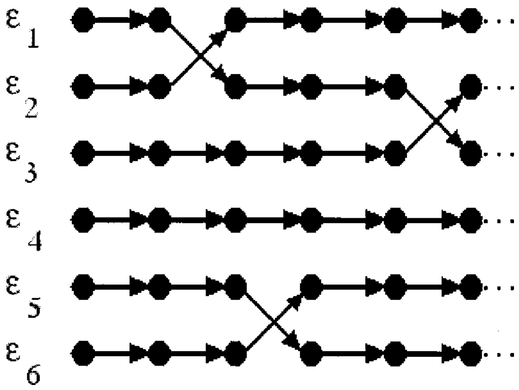

FIG. 5. Schematic illustration of the stationary tempering method. In this approach, the control parameter that governs the sparseness of the associated probability distribution is the stationary phase filter length scale, ε. In the short segment depicted in the illustration, the simulation corresponding to filter length scale ε_4 does not undergo a tempering exchange attempt.

Sparse sampling methods offer a practical way to address the sampling difficulties associated with the stationary phase Monte Carlo method. As suggested by Fig. 2, a natural control parameter in such applications is the length-scale of the filtering process rather than the physical temperature. In Fig. 5 we present a schematic outline of a ‘‘stationary tempering’’ approach based on a combination of parallel tempering and stationary phase Monte Carlo techniques. In this method, parallel tempering simulations are performed on an ensemble of importance distributions corresponding to various stationary phase filter length-scales. As discussed in Sec. II A, if the filtering process involved is implemented exactly, then the results calculated via stationary phase Monte Carlo are formally independent of choice of filter length scale. The necessary statistical errors in the calculated results, on the other hand, vary with the sparse sampling control parameter. We are thus free to select from our sparse sampling computational ensemble an ‘‘optimal’’ filter length scale based on a variance minimization criteria. The required statistical errors can be estimated in practice by examining the numerical scatter in various elements within the stationary tempering ensemble. In situations where the filtering methods involved are implemented in an approximate gradient manner (i.e., without ‘‘corrections’’), the calculated results are no longer guaranteed to be exact. In such circumstances, the most reliable results will be those for which the underlying gradient approximation is the most accurate (i.e., results obtained using the smallest, computationally viable filter length scales). A practical indicator of the adequacy of gradient-based methods is thus the degree to which there exists a length-scale independent ‘‘plateau value’’ in the computed results at smaller filter length scales. Numerical applications of the stationary tempering approach will be discussed in Sec. III.

We close this section by making two points concerning the nature of the importance function, $W_\varepsilon(\boldsymbol{x})$, defined by Eq. (2.12). First, it will prove useful to consider the auxiliary moment generating function associated with $W_\varepsilon(\boldsymbol{x})$, $\phi_W(\boldsymbol{\eta})$, by

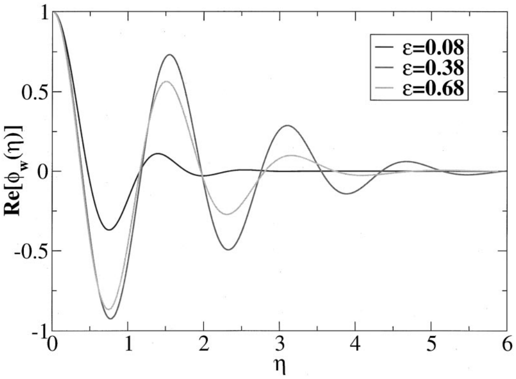

FIG. 6. Plots of the real part of the auxiliary moment generating function, $\phi_W(\eta,t=-16)$, defined by Eq. (2.25) for the Airy action for $\varepsilon=(0.08,0.38,0.68)$. As discussed in the text, this function provides a probe of the structure of the stationary phase importance function, $W_\varepsilon(x,t)$.

$$\phi_W(\boldsymbol{\eta})=\frac{\int W_\varepsilon(\boldsymbol{x})e^{i\boldsymbol{\eta}\cdot\boldsymbol{x}}d\boldsymbol{x}}{\int W_\varepsilon(\boldsymbol{x})d\boldsymbol{x}}. \tag{2.25}$$

This moment generating function contains information that is useful in optimizing the performance of stationary tempering techniques. Specifically, the "sharpness" of the stationary phase filters in the stationary tempering ensemble can be quantified by examining the large argument behavior of the corresponding auxiliary moment generating function, $\phi_W(\boldsymbol{\eta})$. The narrower the filter is about a particular stationary phase region, the longer-lived will be that inherent structure's contribution to the associated moment generating function, $\phi_W(\boldsymbol{\eta})$. Comparison of Figs. 2, 6, and 7 illustrates this correlation for the Airy example. The asymptotic behavior of ϕ_W thus provides us with a second means for monitoring the numerical performance of the stationary phase tempering method, the first being the length-scale dependence of the statistical variance in the computed results.

The second point we wish to make concerns the inherent structure analysis of the stationary phase importance function, $W_\varepsilon(\boldsymbol{x})$. In the Stillinger–Weber approach,[6,7] inherent structures of a given distribution are generated by direct-space quench procedures. The integration domain is decomposed into cells by assigning each point in the volume to the nearest local maxima of the distribution. As discussed in Appendix A, it is useful to consider the inherent structure problem from an alternative point of view. In applied mathematics, a common problem involves the reconstruction of a probability distribution from a proper statistical sampling of that distribution. In this "probabilistic clustering" problem,[22] one is attempting to identify within the large statistical data set originally provided a few, meaningful parameters that characterize the underlying probability distribution. So phrased, this data compression task is a (slightly) modified version of the inherent structure problem. Stationary tempering methods, in conjunction with probabilistic clustering techniques, thus provide an alternate way to probe the inherent structures of the stationary phase importance function, $W_\varepsilon(\boldsymbol{x})$.

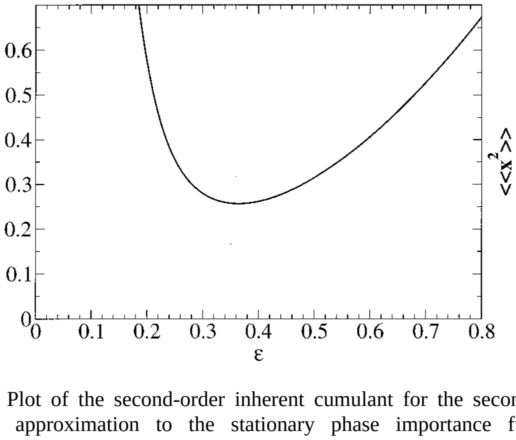

FIG. 7. Plot of the second-order inherent cumulant for the second-order gradient approximation to the stationary phase importance function, $W_\varepsilon(x,t)$, for the Airy action as a function of filter length-scale, ε. The minimum, which corresponds to the most highly compressed importance function, corresponds to $\varepsilon=0.38$.

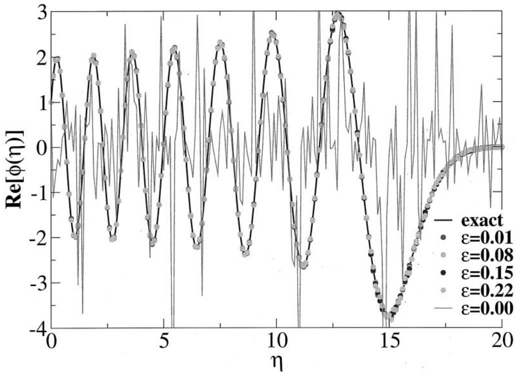

FIG. 8. Real portion of the moment generating function for the Airy action $(t=-16)$. Solid line (smooth) is exact result [Eq. (2.24)]. Plotting symbols correspond to results of second-order gradient stationary tempering calculations for variety of filter length scales. Results for the different ε values agree because, as discussed in the text, the second-order results for this particular example are exact. Stationary tempering calculations were performed using of the order of 10^7 Monte Carlo points. The erratic solid line corresponds to a conventional Monte Carlo simulation $(\varepsilon=0)$ with 10^6 points. Error bars for all tempering calculations are smaller than the corresponding plotting symbols.

III. NUMERICAL EXAMPLES

In the present section, we wish to illustrate the formal developments of Sec. II with a number of applications. These include both simple, pedagogical results and results from more physically relevant, dynamical path-integration applications. Unless noted otherwise, all results utilize an uncorrected, second-order implementation of the stationary tempering approach.

We begin by returning to the simple, Airy action example discussed in Sec. II A. As illustrated by Figs. 1–4, the stationary phase filtering approach dramatically simplifies the construction of the generalized moment generating function and produces excellent (in this case exact) results. As was emphasized in Sec. II A, however, the results displayed in Fig. 4 were generated by directly integrating Eq. (2.5) with standard quadrature methods, not by statistical means. It is thus important to demonstrate that stationary tempering methods can reliably produce these results. Figure 8 compares $\phi(\eta)$ calculated with the stationary tempering approach described in Sec. II and the exact results of Eq.

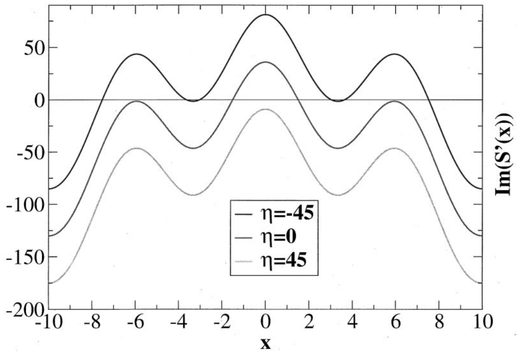

FIG. 9. Plots of $\partial(S(x,t)-i\eta x)/\partial x$ for modified Airy action ($t=-36$, $\eta=-45,0,45$).

(2.20). The numerical calculations were performed using on the order of 10^7 Monte Carlo points. The tempering ensemble utilized in these calculations involved seven filter parameters distributed uniformly over the closed interval [0.01, 0.43]. As can be seen in Fig. 2, the tempering ensemble thus includes importance functions ranging from sharply defined to diffuse. The optimal filter parameter, determined on the basis of the sharpness of the importance function, $W_\varepsilon(x)$, about the stationary phase regions (cf. Fig. 7) and on the rate of decay of auxiliary moment generating function (cf. Fig. 6), $\phi_W(\eta)$, is approximately $\varepsilon=0.38$. Trial moves that involve the exchange of configurations between a single, randomly selected pair of tempering distributions with different filter parameters were attempted after every ten Monte Carlo steps. We see from Fig. 8 that the results of stationary tempering are excellent. Of particular note is the way in which the method ''uniformly'' treats the transition from oscillatory to damped behavior (i.e., the coalescence of the real-valued stationary phase points). We also see the dramatic increase in noise as the stationary phase filtering is turned off, a reminder of how ill-suited conventional Monte Carlo methods are for the complex quadrature problem.

It is important to demonstrate that the stationary tempering approach is not limited to applications where the stationary phase structure of the problem is simple. To this end, we consider a ''modified'' Airy action of the form

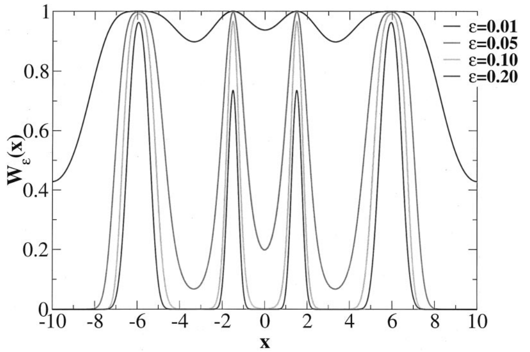

FIG. 10. Plots of the second-order gradient stationary phase importance function, $W_\varepsilon(x,t)$, for modified Airy action; $t=-36$, $\eta=0$, $\varepsilon=0.01$, 0.05, 0.10, 0.20. Note the sparseness of the distributions at larger ε values.

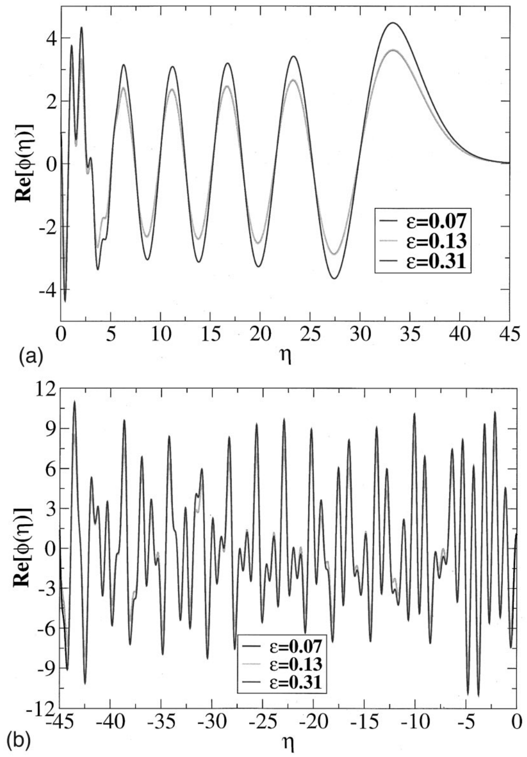

FIG. 11. Real part of second-order gradient approximation to the moment generating function for modified Airy action. The plot is broken into two panels in order to expand the plotting scale. Results shown were computed for various filter length scales by direct quadrature integration of Eq. (2.9). Note that the results are independent of filter length scale as ε tends to zero. The number of stationary phase regions involved for the various values of η can be determined from Fig. 9.

$$S(x,t)=-i\left(\frac{x^3}{3}+t\sin(x)\right). \tag{3.1}$$

Figure 9 plots the imaginary part of the derivative of the action as a function of position for various values of η and t. The number of stationary phase points (zeros of the derivative) varies depending on the parameter values. These stationary phase regions are visible in Fig. 10 as the inherent structures of the importance function for this modified action.

Moment generating functions for the modified Airy action, produced by directly integrating Eq. (2.9) for various filter length scales using Mathematica, are displayed in Fig. 11. Unlike the gradient results for the Airy action, the calculated values for the modified Airy example are not exact and a dependence on the filter parameter can be seen at larger length scales. As illustrated in Fig. 11, however, the results for smaller filter length scales approach a common, limiting value, an indication that the corresponding gradient results are reliable. We note that for larger length scales, where the uncorrected gradient results are less accurate, the errors appear to be only in the amplitude of the calculated results. The frequency dependence, on the other hand, appears to be faithfully reproduced. Further work will be required to establish the degree of generality of this behavior.

Figure 12 compares moment generating functions for the modified Airy action obtained using stationary tempering

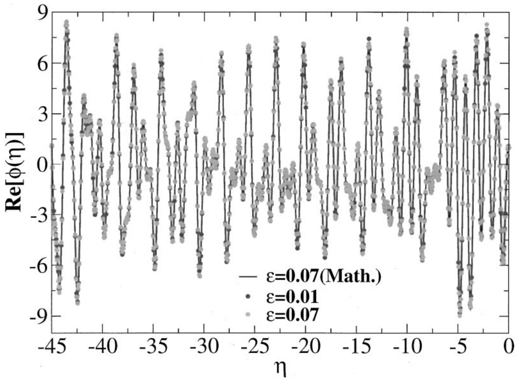

FIG. 12. Comparison of stationary tempering results with corresponding results of Fig. 11. See Fig. 13 for an illustration of the importance of sparse sampling. Error bars for both tempering results presented average less than 0.2 over the range of η values displayed.

methods to the quadrature results of Fig. 11. The tempering ensemble consisted of six filter length scales distributed evenly in the range of [0.01, 0.31]. Of the order of 10^7 Monte Carlo points were utilized in all calculations. Random, pairwise tempering exchanges were attempted every 10 Monte Carlo moves. The agreement between the stationary tempering and quadrature results is excellent. In Fig. 13 we present moment generating function results produced using stationary phase filtering *without tempering*. While correct (for the present example) at smaller length scales, nontempering results are generally unreliable. In particular, we see in Fig. 13 that the results computed at larger filter length scales are qualitatively incorrect. The reason for this breakdown is apparent in Fig. 10. At smaller filter length scales, the stationary phase importance function remains sufficiently connected that conventional sampling is, for this example, sufficient. As the stationary phase filtering becomes more aggressive, however, the stationary phase importance function becomes more disconnected. Under such conditions conventional techniques essentially become ‘‘locked’’ in one of the many inherent structures and the numerical results produced become unreliable. In the present example, this breakdown of conventional methods is evident. In general, however, this is not the case. Moreover, without sparse sampling methods it becomes extremely difficult to separate the effects of sampling from those of the gradient implementation of the stationary phase Monte Carlo approach.

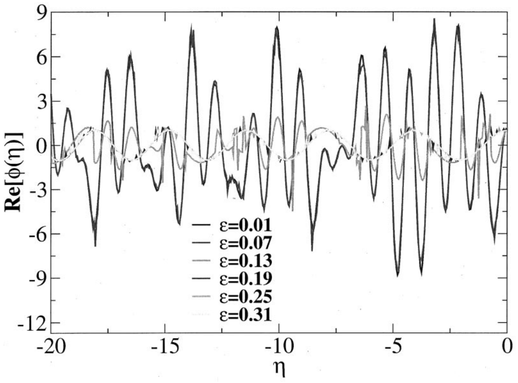

FIG. 13. As in Fig. 12, but *without* parallel tempering. Note the breakdown of the Monte Carlo results at larger filter length scales where the stationary phase importance function, $W_\varepsilon(x,t)$, is sparse (cf. Fig. 10). Results in both Figs. 12 and 13 utilized the same number (10^7) of Monte Carlo points.

Finally, we wish to illustrate the application of the stationary tempering method to a problem that is representative of those that arise in the context of dynamical path integration. To this end, we examine the calculation of the moment generating function associated with the complex action for a particle of mass m moving in a one-dimensional Lennard-Jones cage.[23] In the Fourier path integral language,[10] the relevant portion of the action for the problem is

$$S(\boldsymbol{a},\beta_c)=\Sigma_{k=1}^{\infty}a_k^2/2\sigma_k^2+\beta_c\overline{V}, \tag{3.2}$$

where β_c is the complex temperature

$$\beta_c=\frac{\beta}{2}+i\,\frac{t}{\hbar}, \tag{3.3}$$

β is the (reciprocal) temperature

$$\beta=\frac{1}{k_BT}, \tag{3.4}$$

t is the physical time, and σ_k^2 is given by

$$\sigma_k^2=\frac{2\beta_c\hbar^2}{m\pi^2k^2}. \tag{3.5}$$

In terms of the Fourier coefficients, $\{a_k\}$, the path, $x(u)$ is given by

$$x(u)=x+(x'-x)u+\sum_{k=1}^{k_{\max}}a_k\sin(k\pi u), \tag{3.6}$$

where x and x' are the values of the path at $u=0$ and $u=1$, respectively. In Eq. (3.6) $\overline{V}$ represents the average of the potential energy along the path, $x(u)$

$$\overline{V}=\int_0^1 du\,V(x(u)). \tag{3.7}$$

The degree of detail in the path is controlled by the number of path variables, $k_{\max}$, utilized in the calculation, a number that will vary depending on the system as well as on the time and temperature.

Figure 14 displays moment generating functions for the cage problem calculated with the (second-order gradient) stationary tempering method. These results correspond to Lennard-Jones interaction parameters representative of molecular hydrogen[24] $(\varepsilon_{\mathrm{LJ}}/k_{\mathrm{B}}=34.2\ \mathrm{K},\sigma_{\mathrm{LJ}}=2.96\ \mathrm{\AA})$. Four Fourier path variables are included in all calculations and the one-dimensional path averages of the potential energy, Eq. (3.7), are evaluated using 32 point Gauss quadrature. Results in the panels (a)–(c) correspond the physical times of 0, 100, and 200 fs, respectively. The temperature in all cases is 300 K. These simulations utilize filter parameters based on results suggested by Fourier path integral analysis of the simple harmonic oscillator. That is, the stationary phase filter parameter for the kth Fourier path variable is assumed to be proportional to the natural width of the corresponding harmonic path variable. Specifically, ε_k, is taken to be

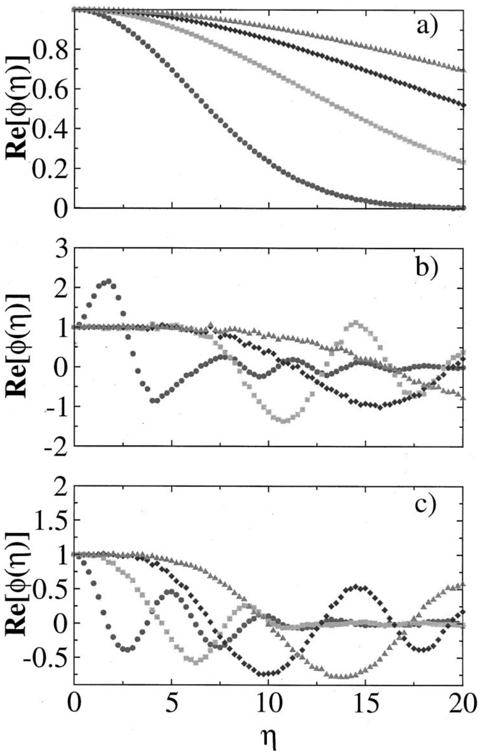

FIG. 14. Real part of moment generating function for LJ cage problem. System parameters correspond to those of molecular hydrogen ($\varepsilon_{\mathrm{LJ}}/k_B = 34.2\,\mathrm{K}, \sigma_{\mathrm{LJ}} = 2.96\,\text{Å}$) and $T = 300\,\mathrm{K}$. Plots, in the order of increasing length scale of their variation, correspond to $\boldsymbol{\eta}$ being oriented along the Fourier components a_1, a_2, a_3, a_4, respectively. Four Fourier components were utilized in the paths for all calculations. Panels (a)–(c) correspond to $t = (0-200)$ fs, respectively. Results plotted correspond to $\varepsilon_0 = 0.4$ [cf. Eq. (3.8)]. While not shown (for reasons of clarity), the calculated results obtained using $\varepsilon_0 = 0.2$ are statistically indistinguishable from those shown in Fig. 14.

$$\varepsilon_k = \varepsilon_0 \left(|\sigma_k^2| \Bigg/ \left[1 + \left(\frac{|\beta_c| \hbar \omega_0}{k\pi} \right)^2 \right] \right)^{1/2}, \tag{3.8}$$

where ε_0 is an overall scale parameter, σ_k^2 is the free-particle Gaussian variance [Eq. (3.5), and ω_0 is an adjustable constant, taken here to be $\beta\hbar\omega_0 = 1$]. In the simulations for both $t = 0$ and 100 fs, the tempering ensemble utilized three values of ε_0 (0.2, 0.4, 0.8) while the results for $t = 200$ fs utilized four filter length scales (0.2, 0.4, 0.8, 1.6). Of the order of 2×10^6 Monte Carlo points were utilized in all calculations. The frequency of pairwise tempering exchange attempts was one in every ten Monte Carlo steps.

The dynamics of the cage problem are reflected in the time evolution of the moment generating function seen in Fig. 14. At shorter times [panel (a)], the moment generating function viewed in various directions is essentially gaussian in nature. At longer times, however, damped oscillatory behavior is evident. Using the stationary tempering method, we have been able to extend calculations such as those in Fig. 14 to more than 1 ps for this system at 300 K, times for which the ratio of the physical to thermal time, $t/(\beta\hbar/2)$, exceeds 80. It should be noted that while the complex-temperature matrix elements of one-dimensional cage problem are readily calculable by a variety of means, the moment generating functions in Fig. 14 are not. Thus, unlike previous examples in this section, the calculated results in Fig. 14 cannot be readily checked by comparison with separately available analytic or exact numerical values. As discussed in Sec. II, however, a practical indicator of the adequacy of gradient-based methods is the degree to which there exists a length-scale independent ''plateau value'' in the computed results at smaller filter length scales. In the present calculations, numerical results obtained using $\varepsilon_0 = 0.2$ and $\varepsilon_0 = 0.4$ are statistically indistinguishable, an indication that the gradient-based results shown in Fig. 14 are reliable.

Finally, as discussed in Appendix A, the temporal evolution of the moment generating function reflects the low-order, dynamical inherent structures of the problem. Probabilistic clustering methods applied to the stationary tempering results thus provide a means for both revealing and characterizing the dynamical inherent structures that are associated with the quantum dynamics. Such analyses and more extensive dynamical path-integral applications will be considered elsewhere.

IV. DISCUSSION AND SUMMARY

We have examined numerical methods for evaluating complex generalizations of conventional equilibrium averages. Such complex averages arise naturally in a number of contexts, including semiclassical and path-integral dynamical applications. Unfortunately, such problems typically lack the natural ''importance function'' structure of their equilibrium counterparts. Instead of a positive Boltzmann distribution, the ''weights'' in such generalized averages are complex-valued objects that prevent the direct application of conventional stochastic quadrature techniques.

Previous efforts have shown that stationary phase filtering techniques[10] are useful in confronting the immediate problems associated with the complex quadrature problem. Based on the invariance of the underlying complex average to a general class of ''preaveraging'' operations, these methods suppress troublesome phase oscillations and produce a natural stationary-phase importance function by means of a numerical ''filtering'' process. Such self-adaptive filtering techniques permit the rigorous reformulation of the original complex average in Monte Carlo form. In the stationary phase importance functions so produced, however, the statistically important regions of the integration volume are generally not connected in a simple manner. Stationary phase filtering methods, by themselves, thus exchange one set of difficulties (oscillations and lack of importance function) for another (sparse sampling).

In the present paper we have introduced a ''stationary tempering'' technique that combines stationary phase filtering[10] and parallel tempering methods.[11–16] The result is an approach that is designed to cope with the twin issues of phase oscillations and sparse sampling that arise in the treatment of generalized complex averages. Here the length-scale of the stationary-phase filtering process (as opposed to the system temperature) serves as the control parameter for the creation of a ensemble in which the computational elements exhibit varying degrees of connectedness.

We have examined a number of numerical applications of the parallel tempering approach to explore the basic approach and to establish its utility. Our results for illustrative examples, including prototypical dynamical path integral applications, have indicated that the stationary tempering method provides an effective means for addressing the complex quadrature problem. Our results have also suggested practical methods for optimizing the numerical efficiency of the stationary tempering method. These include methods based on the minimization of the statistical variance of the calculation with respect to the filter length-scale and on the asymptotic behavior of the moment generating function associated with the stationary-phase importance function.

ACKNOWLEDGMENTS

The authors gratefully acknowledge support from the National Science Foundation through awards CHE-9714970, CDA-9724347, and CHE-0095053. J.D.D. wishes to express his special thanks to Professor Hannes Jonsson and to the University of Iceland for their hospitality and support during a visit in which portions of the present research were completed.

APPENDIX A: INHERENT STRUCTURE, CUMULANTS AND PROBABILISTIC CLUSTERING

Given a moment generating function of a specified probability density, $W(\boldsymbol{x})$

$$\phi_W(\boldsymbol{\eta})=\frac{\int W(\boldsymbol{x})e^{i\boldsymbol{\eta}\cdot\boldsymbol{x}}d\boldsymbol{x}}{\int W(\boldsymbol{x})d\boldsymbol{x}}, \tag{A1}$$

conventional cumulants are defined by[25]

$$\phi_W(\boldsymbol{\eta})=\exp\left\{\sum_{m=1}\frac{(i)^m}{m!}\langle\langle(\boldsymbol{\eta}\cdot\boldsymbol{x})^m\rangle\rangle\right\}. \tag{A2}$$

While convenient, such cumulants are ''global'' properties that do not directly reflect the ''inherent structure'' of the underlying probability distribution.

Following Stillinger and Weber,[6,7] we can decompose the integration domain in Eq. (A1) into nonoverlapping domains by assigning each point in the integration volume to one of the local maxima of the probability distribution through suitably designed quench procedures.

Alternatively, we can perform a Wigner–Seitz decomposition based on the local maxima of the probability distribution. However, we choose to define them, the average in Eq. (A1) can be written in terms of the ''inherent structures'' of the probability distribution as

$$\phi_W(\boldsymbol{\eta})=\sum_\alpha \Gamma_\alpha\langle e^{i\boldsymbol{\eta}\cdot\boldsymbol{x}}\rangle_\alpha . \tag{A3}$$

Here $\langle e^{i\boldsymbol{\eta}\cdot\boldsymbol{x}}\rangle_\alpha$ denotes a ''local'' average of $e^{i\boldsymbol{\eta}\cdot\boldsymbol{x}}$ over the integration domain corresponding to the αth inherent structure

$$\langle e^{i\eta\cdot x}\rangle_\alpha=\frac{\int_\alpha d\boldsymbol{x}W(\boldsymbol{x})e^{i\boldsymbol{\eta}\cdot\boldsymbol{x}}}{\int_\alpha d\boldsymbol{x}W(\boldsymbol{x})}, \tag{A4}$$

whose corresponding fractional weight, Γ_α, is given by

$$\Gamma_\alpha=\frac{\int_\alpha d\boldsymbol{x}W(\boldsymbol{x})}{\int d\boldsymbol{x}W(\boldsymbol{x})}. \tag{A5}$$

Expanding the *local* averages in Eq. (A3) [as opposed to the *global* average in Eq. (A1)] in cumulant form gives

$$\phi_W(\boldsymbol{\eta})=\sum_\alpha \Gamma_\alpha\exp\left\{\sum_{m=1}\frac{(i)^m}{m!}\langle\langle(\boldsymbol{\eta}\cdot\boldsymbol{x})^m\rangle\rangle_\alpha\right\}, \tag{A6}$$

where $\langle\langle(\boldsymbol{\eta}\cdot\boldsymbol{x})^m\rangle\rangle_\alpha$ represents the mth-order ''inherent cumulant'' (i.e., the mth-order cumulant associated with the αth inherent structure) of the distribution, $W(\boldsymbol{x})$. Unlike the analogous quantities in Eq. (A2), the inherent cumulants defined by Eq. (A6) are ''local'' quantities that reflect directly the properties of the inherent structures of the distribution, $W(\boldsymbol{x})$.

An important problem in applied mathematics is the reconstruction of an unknown probability distribution, $W(\boldsymbol{x})$, from a statistical sampling of configurations, $\{\boldsymbol{x}_n\}$. The probabilistic clustering problem, is effectively the inverse of the sparse sampling problem. Rather than the focus being to produce a statistical sampling of a given distribution, in clustering applications we seek instead to reconstruct an unknown distribution from a known statistical sampling. This problem is of broad significance in information technology and has received considerable attention.[22]

Our point in emphasizing the links between sparse sampling, probabilistic clustering and inherent structure methods is that advances in one topic contain implications for the others. In particular, our increased ability to sample sparse distributions appears to offer an alternative approach to the inherent structure decomposition problem. Equation (A6), for example, expresses the moment generating function in a crystallographic-like manner in terms of parameters that characterize the strength, position, dispersion, etc. of the inherent structures of the probability density. This structural analogy is relevant because we have a practical way of ''measuring'' $\phi_W(\boldsymbol{\eta})$. Although computational as opposed to experimental, we can estimate Eq. (A6) with Monte Carlo methods. Such an estimate, $\phi_{W,N}(\boldsymbol{\eta})$, defined by

$$\phi_{W,N}(\boldsymbol{\eta})=\frac{1}{N}\sum_{n=1}^{N} e^{i\boldsymbol{\eta}\cdot\boldsymbol{x}_n}, \tag{A7}$$

provides us with a refinable means of evaluating $\phi_W(\boldsymbol{\eta})$ numerically in terms of configurations, $\{\boldsymbol{x}_n\}$, which can, in turn, be produced using the sparse sampling methods described in Sec. II. Viewed from this perspective, performing an inherent structure decomposition of the probability distribution $W(\boldsymbol{x})$ amounts to forcing the ansatz for $\phi_W(\boldsymbol{\eta})$, Eq. (A6), and the ''measured'' result, defined by Eq. (A7), to ''agree.'' This sparse sampling approach to inherent structure analysis will be discussed in greater detail elsewhere.

[1] N. Metropolis and S. Ulam, J. Am. Stat. Assoc. **247**, 335 (1949).
[2] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. M. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).
[3] M. Kalos and P. Whitlock, *Monte Carlo Methods* (Wiley-Interscience, New York, 1986).
[4] J. P. Valleau and G. M. Torrie, in *Modern Theoretical Chemistry*, edited by B. J. Berne (Plenum, New York, 1977).
[5] J. P. Valleau and S. G. Whittington, in *Modern Theoretical Chemistry*, edited by B. J. Berne (Plenum, New York, 1977).
[6] F. H. Stillinger and T. A. Weber, Phys. Rev. A **28**, 2408 (1983).
[7] F. H. Stillinger and T. A. Weber, Science **225**, 983 (1984).

[8] F. H. Stillinger, Phys. Rev. E **59**, 48 (1999).
[9] G. F. Carrier, M. Krook, and C. E. Pearson, *Functions of a Complex Variable: Theory and Technique* (McGraw-Hill, New York, 1966).
[10] V. S. Filinov, Nucl. Phys. B **271**, 717 (1986); J. D. Doll and D. L. Freeman, Adv. Chem. Phys. **73**, 289 (1988); N. Makri and W. H. Miller, Chem. Phys. Lett. **139**, 10 (1987); N. Makri and W. H. Miller, J. Chem. Phys. **89**, 2170 (1988); For a recent review, see J. D. Doll, M. Eleftherios, S. A. Corcelli, and D. L. Freeman, in *Quantum Monte Carlo Methods in Physics and Chemistry, Vol. C 526*, edited by M. P. Nightingale and C. J. Umrigar (Kluwer, Dordrecht, 1999), pp. 213–245 and references therein.
[11] D. D. Frantz, D. L. Freeman, and J. D. Doll, J. Chem. Phys. **93**, 2769 (1990).
[12] D. D. Frantz, D. L. Freeman, and J. D. Doll, J. Chem. Phys. **97**, 5713 (1992).
[13] E. Marinari and G. Parisi, Europhys. Lett. **19**, 451 (1992).
[14] C. J. Geyer and E. A. Thompson, J. Am. Stat. Assoc. **90**, 909 (1995).
[15] M. Falcioni and M. W. Deem, J. Chem. Phys. **110**, 1754 (1999).
[16] J. P. Neirotti, F. Calvo, D. L. Freeman, and J. D. Doll, J. Chem. Phys. **112**, 10340 (2000).
[17] M. P. Allen and D. J. Tildesley, *Computer Simulations of Liquids* (Oxford Science Publications, Oxford, 1987).
[18] J. D. Doll, D. L. Freeman, and M. J. Gillan, Chem. Phys. Lett. **143**, 277 (1988).
[19] L. D. Landau and E. M. Lifshitz, *Statistical Physics*, 2nd ed. (Addison-Wesley, Reading, 1969), pp 346–350.
[20] W. H. Miller, Science **233**, 171 (1986).
[21] *Handbook of Mathematical Functions, Vol. 55*, edited by M. Abramowitz and I. A. Stegun (U.S. Department of Commerce, Washington DC, 1964).
[22] B. G. Mirkin, *Mathematical Classification and Clustering* (Kluwer, Dordrecht, 1996).
[23] D. L. Freeman, R. D. Coalson, and J. D. Doll, J. Stat. Phys. **43**, 931 (1986).
[24] J. D. Doll and D. L. Freeman, J. Chem. Phys. **111**, 7685 (1999).
[25] H. Risken, *The Fokker-Planck Equation* (Springer-Verlag, Berlin, 1984).

The Journal of Chemical Physics is copyrighted by the American Institute of Physics (AIP). Redistribution of journal material is subject to the AIP online journal license and/or AIP copyright. For more information, see http://ojps.aip.org/jcpo/jcpcr/jsp
Copyright of Journal of Chemical Physics is the property of American Institute of Physics and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.

The Journal of Chemical Physics is copyrighted by the American Institute of Physics (AIP). Redistribution of journal material is subject to the AIP online journal license and/or AIP copyright. For more information, see http://ojps.aip.org/jcpo/jcpcr/jsp