Quantum statistical calculations and symplectic corrector algorithms

Siu A. Chin
Department of Physics, Texas A&M University, College Station, Texas 77843, USA
(Received 1 December 2003; published 30 April 2004)

The quantum partition function at finite temperature requires computing the trace of the imaginary time propagator. For numerical and Monte Carlo calculations, the propagator is usually split into its kinetic and potential parts. A higher-order splitting will result in a higher-order convergent algorithm. At imaginary time, the kinetic energy propagator is usually the diffusion Green's function. Since diffusion cannot be simulated backward in time, the splitting must maintain the positivity of all intermediate time steps. However, since the trace is invariant under similarity transformations of the propagator, one can use this freedom to "correct" the split propagator to higher order. This use of similarity transforms classically gives rise to symplectic corrector algorithms. The split propagator is the symplectic kernel and the similarity transformation is the corrector. This work proves a generalization of the Sheng-Suzuki theorem: no positive time step propagators with only kinetic and potential operators can be corrected beyond second order. Second-order forward propagators can have fourth-order traces only with the inclusion of an additional commutator. We give detailed derivations of four forward correctable second-order propagators and their minimal correctors.

DOI: 10.1103/PhysRevE.69.046118 PACS number(s): 05.30.−d, 02.70.Ss, 05.10.−a, 95.10.Ce

I. INTRODUCTION

The quantum partition function requires computing the trace,

$$Z = \mathrm{Tr}(\rho) = \mathrm{Tr}(\mathrm{e}^{-\beta H}), \tag{1.1}$$

where ρ is imaginary time propagator, $\beta=1/(k_B T)$ is the inverse temperature, and $H=T+V$ is the usual Hamiltonian operator. Although specific forms of the kinetic and potential energy operators will not be used in the following, it is useful to keep in mind the many-body case where $T=(-\hbar^2/2m)\Sigma_{i=1}\nabla_i^2$ and $V=\Sigma_{i<j}v(r_{ij})$. In numerical or Monte Carlo calculations, the imaginary time propagator is first discretized as

$$e^{-\beta(T+V)} = [e^{\varepsilon(T+V)}]^n, \tag{1.2}$$

where $\varepsilon=-\Delta\beta=-\beta/n$, and the short-time propagator $e^{\varepsilon(T+V)}$ is then approximated in various ways. One systematic method is to decompose, or split, the short-time propagator into the product form

$$e^{\varepsilon(T+V)} \approx \prod_{i=1}^{N} e^{t_i\varepsilon T} e^{v_i\varepsilon V}, \tag{1.3}$$

with coefficients $\{t_i, v_i\}$ determined by the required order of accuracy. For quantum statistical calculations, since $\langle \mathbf{r}'|e^{t_i\varepsilon T}|\mathbf{r}\rangle \propto e^{-(\mathbf{r}'-\mathbf{r})^2/(2t_i\Delta\beta)}$ is the diffusion kernel, the coefficient t_i must be positive in order for it to be simulated or integrated. If t_i were negative, the kernel is unbounded and unnormalizable, and no probabilistic based (Monte Carlo) simulation is possible. However, as first proved by Sheng [1], and later by Suzuki [2], beyond second order, any factorization of the form (1.3) *must* contain some negative coefficients in the set $\{t_i, v_i\}$. Goldman and Kaper [3] further proved that any factorization of the form (1.3) must contain at least one negative coefficient for *both* operators. Thus, despite myriad of factorization schemes of the form (1.3) proposed in the classical symplectic integrator literature [4–7], none can be used for doing quantum statistical calculations beyond second order. It is only recently that fourth order, all positive-coefficient factorization schemes have been found [8,9] and applied to time-irreversible problems containing the diffusion kernel [10–14]. In order to bypass the Sheng-Suzuki's theorem, one must include other operators in the factorization (1.3), such as the double commutator $[V,[T,V]]$, where $[A,B]\equiv AB-BA$.

In computing the quantum partition function Z, only the trace of $\rho=e^{-\beta H}$ is required. Since the trace is invariant under the similarity transformation

$$\tilde{\rho} = S\rho S^{-1}, \tag{1.4}$$

one is free to use any such $\tilde{\rho}$ to compute Z. This is immaterial if ρ is known exactly. However, if the short-time propagator is only known approximately, then one may use a clever choice of S to further improve the approximation. This is a well-known idea in many areas of physics. For example, to calculate the exact quantum many-body ground state using the diffusion Monte Carlo algorithm, one can choose $S=\phi_0$, where ϕ_0 is a known trial function close to the exact ground state. This is the idea of "importance sampling" as introduced by Kalos *et al.* [15]. Its operator formulation as described above has been implemented by Chin [16] some time ago. Similar ideas have been used to improve path integrals, as detailed by Kleinert [17]. If the short-time propagator is approximated by the product form (1.3), the error terms can be calculated explicitly and eliminated by S. When implemented classically, these are known as symplectic "corrector," or "process" algorithms [18–23]. In this context the propagator ρ is the kernel algorithm and S is the corrector. Since S disappears in the calculation of Z, there is no restriction on the form of S. If S were also expanded in the product form (1.3), there is no restriction on the sign of its coefficients. This suggests that there may exist a product form (1.3) of ρ with only positive coefficients such that its trace is correct to higher order. This would not be precluded by the existing Sheng-Suzuki theorem.

In this work, we show that this is not possible. If ρ is approximated by the product form (1.3) with positive coefficients $\{t_i\}$, then $\tilde{\rho}$ cannot be corrected by S to higher than second order. The proof of this generalizes the Sheng-Suzuki theorem. The corrected propagator $\tilde{\rho}$ can be fourth order only if additional operators, such as $[V,[T,V]]$, are used in the splitting of ρ. By understanding the "correctability" requirement, we can systematically deduce the four fundamental correctable second-order propagators and their correctors.

In the following section, we recall some basic results of similarity transforms. Beyond second order, only a special class of approximate ρ satisfying the correctability condition can be corrected to higher order. In Sec. III, we compute the explicit form of the error coefficients required by the correctability criterion. In Sec. IV, we show that this requirement cannot be satisfied for propagators of the product form (1.3) with only positive $\{t_i\}$ coefficients. In Sec. V, based on our understanding of the correctability restriction, we deduce all four second-order correctable propagators and their minimal correctors. Some conclusions are given in Sec. VI.

II. SIMILARITY TRANSFORMS AND THE CORRECTABILITY CRITERION

Similarity transforms on approximate propagators of the product form (1.3) have been studied extensively in the context of symplectic correctors [18–22]. However, not all use the language of operators and some are specific to celestial mechanics. Here, we recall some elementary results and establish the fundamental correctability requirement in the context of quantum statistical physics.

Since

$$S\rho S^{-1}=[Se^{\varepsilon(T+V)}S^{-1}]^n, \tag{2.1}$$

it is sufficient to study the similarity transformation of the approximate short-time propagator ρ_A. Let ρ_A approximates $e^{\varepsilon(T+V)}$ in the product form such that

$$\rho_A=\prod_{i=1}^{N}e^{t_i\varepsilon T}e^{v_i\varepsilon V}=e^{\varepsilon H_A}, \tag{2.2}$$

where H_A is the approximate Hamiltonian

$$H_A=T+V+\varepsilon(e_{TV}[T,V])+\varepsilon^2(e_{TTV}[T,[T,V]] + e_{VTV}[V,[T,V]])+O(\varepsilon^3) \tag{2.3}$$

with error coefficients e_{TV}, e_{TTV}, and e_{VTV} determined by factorization coefficients $\{t_i,v_i\}$. The transformed propagator is

$$\tilde{\rho}_A=S\rho_A S^{-1}=Se^{\varepsilon H_A}S^{-1}=e^{\varepsilon(SH_AS^{-1})}=e^{\varepsilon\tilde{H}_A}, \tag{2.4}$$

where the last equality defines the transformed approximate Hamiltonian $\tilde{H}_A$. If now we take

$$S=\exp[\varepsilon C], \tag{2.5}$$

where C is the to-be-determined corrector, then we have the fundamental result

$$\begin{aligned}\tilde{H}_A&=e^{\varepsilon C}H_Ae^{-\varepsilon C}\\ &=H_A+\varepsilon[C,H_A]+\frac{1}{2}\varepsilon^2[C,[C,H_A]]+\frac{1}{3!}\varepsilon^3[C,[C,[C,H_A]]]\\ &\quad+\cdots.\end{aligned} \tag{2.6}$$

Let us first consider the case where the product form (2.2) for H_A is left-right symmetric, i.e., either $t_1=0$ and $v_i=v_{N-i+1}$, $t_{i+1}=t_{N-i+1}$ or $v_N=0$ and $v_i=v_{N-i}$, $t_i=t_{N-i+1}$. In this case, the propagator is reversible, $\rho_A(\varepsilon)\rho_A(-\varepsilon)=1$, and $H_A(\varepsilon)$ is an even function of ε with $e_{TV}=0$. In this case,

$$\begin{aligned}\tilde{H}_A&=H_A+\varepsilon[C,H_A]+\cdots\\ &=T+V+\varepsilon^2(e_{TTV}[T,[T,V]]+e_{VTV}[V,[T,V]])\\ &\quad+\varepsilon[C,T+V]+\cdots,\end{aligned} \tag{2.7}$$

and one immediately sees that the choice $C=\varepsilon C_1$ with $C_1\equiv c_{TV}[T,V]$ would eliminate either second order error term with $c_{TV}=e_{TTV}$ or $c_{TV}=e_{VTV}$. So, if H_A is constructed such that

$$e_{TTV}=e_{VTV}, \tag{2.8}$$

then both can be simultaneously eliminated by the corrector. This is the fundamental correctability requirement for correcting a second order ρ_A to fourth order. This observation can be generalized to higher order. At higher orders, H_A will have error terms of the form $[T,Q_i]$ and $[V,Q_i]$ where Q_i are some higher-order commutator generated by T and V. If H_A is of order $2n$ in ε, then $\tilde{H}_A$ can be of order $2n+2$ only if H_A's error coefficients for $[T,Q_i]$ and $[V,Q_i]$ are *equal* for all Q_i's. This fundamental corrector insight is often obscured by the more general case where odd order errors are allowed.

Sheng [1] and Suzuki [2] independently proved that no ρ_A of the form (2.2) can have positive coefficients t_i beyond second order. More precisely, if ρ_A is of the product form (2.2) with positive t_i's such that $e_{TV}=0$, then both e_{TTV} and e_{VTV} cannot be zero. We will prove a more general theorem that the product form (2.2) with positive t_i's such that $e_{TV}=0$ cannot be *corrected* beyond second order, i.e., e_{TTV} can never equal to e_{VTV}. From this perspective, the Sheng-Suzuki theorem is a special case where the common value for both coefficients is zero.

In the general case where $e_{TV}\neq 0$, we have

$$\begin{aligned}\tilde{H}_A&=T+V+\varepsilon(e_{TV}[T,V])+\varepsilon^2(e_{TTV}[T,[T,V]]\\ &\quad+e_{VTV}[V,[T,V]])+\varepsilon[C,T+V]+\varepsilon^2e_{TV}[C,[T,V]]\\ &\quad+\tfrac{1}{2}\varepsilon^2[C,[C,T+V]]+O(\varepsilon^3).\end{aligned} \tag{2.9}$$

Since $[c_TT+c_VV,T+V]=(c_T-c_V)[T,V]$, the linear term in ε can be eliminated if we choose $C=C_0\equiv c_TT+c_VV$ such that

$$(c_T - c_V) = -e_{TV}. \tag{2.10}$$

This is the first-order correctability condition. This means that with a suitable choice of c_T and c_V, a first-order propagator can always be corrected to second order. Hence, the trace of any first-order propagator is always second order. For example, the trace $\mathrm{Tr}(e^{\varepsilon T}e^{\varepsilon V})$ is second-order despite its appearance.

With the first-order correctability condition satisfied, the remaining commutators in Eq. (2.9) are either $[T,[T,V]]$ or $[V,[T,V]]$, and can again be corrected by adding to C the term $\varepsilon C_1 = \varepsilon c_{TV}[T,V]$. Thus with

$$C = C_0 + \varepsilon C_1 = c_T T + c_V V + \varepsilon c_{TV}[T,V] \tag{2.11}$$

such that $(c_T - c_V) = -e_{TV}$, we have

$$\begin{aligned}\widetilde{H}_A &= T + V + \varepsilon^2(e_{TTV}[T,[T,V]] + e_{VTV}[V,[T,V]]) \\ &\quad + \varepsilon^2[C_1, T+V] + \varepsilon^2 e_{TV}[C_0,[T,V]] \\ &\quad + \tfrac{1}{2}\varepsilon^2[C_0,[C_0,T+V]] + O(\varepsilon^3), \\ &= T + V + \varepsilon^2\left(e_{TTV} - c_{TV} + \tfrac{1}{2}c_T e_{TV}\right)[T,[T,V]] \\ &\quad + \varepsilon^2\left(e_{VTV} - c_{TV} + \tfrac{1}{2}c_V e_{TV}\right)[V,[T,V]] + O(\varepsilon^3).\end{aligned} \tag{2.12}$$

If we now choose $c_{TV} = e_{TTV} + \frac{1}{2}c_T e_{TV}$ to eliminate the error term $[T,[T,V]]$, then the error term $[V,[T,V]]$ can vanish only if

$$e_{TTV} = e_{VTV} + \tfrac{1}{2}(e_{TV})^2. \tag{2.13}$$

This is the general second-order correctability requirement for correcting any first-order propagator beyond second order. The major result of this work is to show that this condition cannot be satisfied for product decomposition of the form (2.2) with only positive t_i coefficients.

III. DETERMINING THE ERROR COEFFICIENTS

To check whether the correctability requirement, Eq. (2.13), can ever be satisfied by an approximate propagator of the product form (2.2), we need to determine e_{TV}, e_{TTV}, and e_{VTV} in terms of $\{t_i, v_i\}$. From the assumed equality

$$\prod_{i=1}^{N} e^{t_i\varepsilon T}e^{v_i\varepsilon V} = e^{\varepsilon H_A}, \tag{3.1}$$

with H_A given by Eq. (2.2), we can expand both sides and compare terms order by order in powers of ε. The left-hand side of Eq. (3.1) can be expanded as

$$\begin{aligned}&e^{\varepsilon t_1 T}e^{\varepsilon v_1 V}e^{\varepsilon t_2 T}e^{\varepsilon v_2 V}\cdots e^{\varepsilon t_N T}e^{\varepsilon v_N V} \\ &\quad = 1 + \varepsilon\left(\sum_{i=1}^{N} t_i\right)T + \varepsilon\left(\sum_{i=1}^{N} v_i\right)V + \cdots,\end{aligned} \tag{3.2}$$

and the right-hand side as

$$\begin{aligned}e^{\varepsilon H_A} &= 1 + \varepsilon(T+V) + \varepsilon^2 e_{TV}[T,V] + \varepsilon^3 e_{TTV}[T,[T,V]] \\ &\quad + \varepsilon^3 e_{VTV}[V,[T,V]] + \frac{1}{2}\varepsilon^2(T+V)^2 + \frac{1}{2}\varepsilon^3 e_{TV}\{(T+V) \\ &\quad \times [T,V] + [T,V](T+V)\} + \frac{1}{3!}\varepsilon^3(T+V)^3 + \cdots.\end{aligned} \tag{3.3}$$

Matching the first-order terms in ε gives the primary constraints

$$\sum_{i=1}^{N} t_i = 1 \quad \text{and} \quad \sum_{i=1}^{N} v_i = 1. \tag{3.4}$$

To determine the error coefficients, we "tag" a particular operator in Eq. (3.3) whose coefficient contains e_{TV}, e_{TTV}, or e_{VTV} and match the same operator's coefficients in the expansion of Eq. (3.2). For example, in the ε^2 terms of Eq. (3.3), the coefficient of the operator TV is $\left(\frac{1}{2} + e_{TV}\right)$. Equating this to the coefficients of TV from Eq. (3.2) gives

$$\frac{1}{2} + e_{TV} = \sum_{i=1}^{N} s_i v_i, \tag{3.5}$$

where we have introduced the variable

$$s_i = \sum_{j=1}^{i} t_j. \tag{3.6}$$

This way of computing TV from Eq. (3.2) corresponds to first picking out a V operator from among all the v_i terms, and then combine all the t_i terms to its left in the exponential to generate a T operator. Alternatively, the same coefficient can also be expressed as

$$\frac{1}{2} + e_{VT} = \sum_{i=1}^{N} t_i u_i, \tag{3.7}$$

where

$$u_i = \sum_{j=i}^{N} v_j. \tag{3.8}$$

This way of computing TV corresponds to first picking out a T operator from among all the t_i terms, and then combine all the v_i terms to its right in the exponential to generate a V operator. To demonstrate how these variables are to be used, we can directly prove the equality of Eqs. (3.5) and (3.7). First, note that $s_N = 1$ and $u_1 = 1$. Second, since $t_i = s_i - s_{i-1}$, at $i=1$ we must consistently set $s_0 = 0$. Similarly, since $v_i = u_i - u_{i+1}$, we must set $u_{N+1} = 0$. Therefore we have

$$\sum_{i=1}^{N} s_i v_i = \sum_{i=1}^{N} s_i(u_i - u_{i+1}) = \sum_{i=1}^{N}(s_i - s_{i-1})u_i = \sum_{i=1}^{N} t_i u_i. \tag{3.9}$$

The determination of error coefficients is simplified if we pick operators whose expansion coefficients are easy to calculate. Matching the coefficients of operators TTV and TVV (note, *not* the operator VTV) yields

$$\frac{1}{6}+\frac{1}{2}e_{TV}+e_{TTV}=\frac{1}{2}\sum_{i=1}^{N}s_i^2 v_i=\frac{1}{2}\sum_{i=1}^{N}(s_i^2-s_{i-1}^2)u_i, \tag{3.10}$$

$$\frac{1}{6}+\frac{1}{2}e_{TV}-e_{TVT}=\frac{1}{2}\sum_{i=1}^{N}t_i u_i^2. \tag{3.11}$$

IV. PROVING THE MAIN RESULT

Using the expression for e_{TVT} from Eq. (3.11), the correctability requirement (2.13) reads

$$\frac{1}{2}\sum_{i=1}^{N}t_i u_i^2=a, \tag{4.1}$$

with

$$a=\tfrac{1}{2}\left(\tfrac{1}{2}+e_{TV}\right)^2+\tfrac{1}{24}-e_{TTV} \tag{4.2}$$

and e_{TV}, e_{TTV} given by Eqs. (3.7) and (3.10), respectively. In Suzuki's proof [2], he recognizes that in terms of the variable $\sqrt{t_i}u_i$, Eq. (4.1) is a hypersphere and Eqs. (3.7) and (3.10) are hyperplanes. His proof is based on a geometric demonstration that his hyperplane cannot intersect his hypersphere. While this geometric language is very appealing, it is cumbersome when dealing with more than one hyperplane. We will use a different strategy.

If t_i are all positive, then the left-hand side of Eq. (4.1) is a positive-definite quadratic form in u_i. There would be no real solutions for u_i if the minimum of the quadratic form is greater than a. Our strategy is therefore to minimize the quadratic form subject to constraints (3.7) and (3.10),

$$\sum_{i=1}^{N}t_i u_i=b, \tag{4.3}$$

$$\sum_{i=1}^{N}t_i(s_i+s_{i-1})u_i=c, \tag{4.4}$$

with $b=\frac{1}{2}+e_{VT}$, $c=\frac{1}{3}+e_{TV}+2e_{TTV}$, and show that the resulting minimum is always greater than a. [The primary constraints (3.4) are just $s_N=1$ and $u_1=1$.]

For constrained minimization, one can use the method of Lagrange multiplier. Minimizing

$$F=\frac{1}{2}\sum_{i=1}^{N}t_i u_i^2-\lambda_1\left(\sum_{i=1}^{N}t_i u_i-b\right)-\lambda_2\left(\sum_{i=1}^{N}t_i(s_i+s_{i-1})u_i-c\right) \tag{4.5}$$

gives

$$u_i=\lambda_1+\lambda_2(s_i+s_{i-1}). \tag{4.6}$$

Substituting this back to satisfy constraints (4.3) and (4.4) determines λ_1 and λ_2:

$$\lambda_1+\lambda_2=b, \tag{4.7}$$

$$\lambda_1+\lambda_2+g\lambda_2=c. \tag{4.8}$$

The only nontrivial evaluation is $\sum_{i=1}^{N}t_i(s_i+s_{i-1})^2=1+g$, where

$$g=\sum_{i=1}^{N}(s_i^2 s_{i-1}-s_i s_{i-1}^2). \tag{4.9}$$

The minimum of the quadratic form is therefore

$$\begin{aligned}F&=\frac{1}{2}\sum_{i=1}^{N}t_i\left[\lambda_1+\lambda_2(s_i+s_{i-1})\right]^2\\&=\frac{1}{2}\left[(\lambda_1+\lambda_2)^2+g\lambda_2^2\right]\\&=\frac{1}{2}\left[b^2+\frac{1}{g}(c-b)^2\right].\end{aligned} \tag{4.10}$$

To minimize F, one must maximize g. Solving $\partial g/\partial s_i=0$ gives $s_i=(s_{i+1}+s_{i-1})/2$, which means that s_i is linear in i. The normalization $s_N=1$ fixes $s_i=i/N$, giving

$$g_{max}=\frac{1}{3}\left(1-\frac{1}{N^2}\right). \tag{4.11}$$

This is indeed a maximum since one can directly verify that $\partial^2 g/\partial s_i^2=-2(s_{i+1}-s_{i-1})<0$. Hence, at any finite N,

$$\begin{aligned}F>\tfrac{1}{2}[b^2+3(c-b)^2]&=\tfrac{1}{2}\left(\tfrac{1}{2}+e_{TV}\right)^2+\tfrac{3}{2}\left(2e_{TTV}-\tfrac{1}{6}\right)^2\\&=a+6e_{TTV}^2.\end{aligned} \tag{4.12}$$

Thus the minimum of the quadratic form is always higher than the value required by the correctability condition. Hence, no real solutions for u_i are possible if t_i are all positive.

We note that the above proof is independent of e_{TV}. For $e_{TV}=0$, the correctability condition is just $e_{TTV}=e_{VTV}$. Hence for symmetric decompositions with positive t_i's, where $e_{TV}=0$ is automatic, we have as a corollary that e_{TTV} can never equal to e_{VTV}.

V. CORRECTABLE FORWARD PROPAGATORS AND THEIR CORRECTORS

The last section is the main result of this work. Here, we show how the correctability criterion can be applied systematically to deduce forward correctable second-order propagators and their minimal correctors.

The proof of noncorrectability is limited to the conventional product form (2.2), which factorizes the propagator only in terms of operators T and V. As shown in the last section, symmetrically decomposed positive-time-step propagators cannot be corrected beyond second order because e_{TTV} cannot be made equal to e_{VTV}. For example, the second-order propagator

$$\exp\left(\tfrac{1}{2}\varepsilon T\right)\exp(\varepsilon V)\exp\left(\tfrac{1}{2}\varepsilon T\right) \tag{5.1}$$

has $t_1=t_2=1/2$, $v_1=u_1=1$, $s_1=1/2$, and $e_{TV}=0$. From Eqs. (3.10) and (3.11), we can determine indeed that the two error coefficients are not equal:

$$e_{TTV}=\tfrac{1}{2}\left(\tfrac{1}{2}\right)^2 1-\tfrac{1}{6}=-\tfrac{1}{24},$$

$$e_{TVT}=\tfrac{1}{6}-\tfrac{1}{2}\left(\tfrac{1}{2}\right)1=-\tfrac{1}{12}. \tag{5.2}$$

A simple way to force them equal is to directly incorporate either operator $[T,[T,V]]$ or $[V,[T,V]]$ in the factorization process. Since $[V,[T,V]]=(\hbar^2/m)\Sigma_i|\nabla_i\Sigma_{j\neq i}v(r_{ij})|^2$ is just another potential function, Suzuki [24] suggested that one should keep the operator $[V,[T,V]]$. If now we add $\frac{1}{24}\varepsilon^3[V,[T,V]]$ to εV in Eq. (5.1), we can change the coefficient e_{VTV} from $-1/12$ to $-1/24$, matching that of e_{TTV}. The result is still only a second-order propagator

$$\rho_{TI}=\exp\left(\tfrac{1}{2}\varepsilon T\right)\exp\left(\varepsilon V+\tfrac{1}{24}\varepsilon^3[V,[T,V]]\right)\exp\left(\tfrac{1}{2}\varepsilon T\right), \tag{5.3}$$

but now has a fourth-order trace. This propagator was first obtained by Takahashi and Imada [25,26] by directly computing the trace. It is a remarkable find given how little they had to work with. This derivation explains, without doing any trace calculation, why the propagator worked.

The alternative of keeping $[T,[T,V]]$ would require adding $-\frac{1}{24}\varepsilon^3[T,[T,V]]$ to make e_{TTV} equal to e_{VTV}'s value of $-1/12$. This operator is too complicated for practical use, but in the case of the harmonic oscillator, it can be combined with the kinetic energy operator:

$$\rho'_{2B}=\exp\left(\tfrac{1}{2}\varepsilon T-\tfrac{1}{48}\varepsilon^3[T,[T,V]]\right)\exp(\varepsilon V)\times\exp\left(\tfrac{1}{2}\varepsilon T-\tfrac{1}{48}\varepsilon^3[T,[T,V]]\right). \tag{5.4}$$

This can also be written in the form of

$$\rho_{2B}=\exp\left(\tfrac{1}{2}\varepsilon V\right)\exp\left(\varepsilon T-\tfrac{1}{24}\varepsilon^3[T,[T,V]]\right)\exp\left(\tfrac{1}{2}\varepsilon V\right). \tag{5.5}$$

In this case $\exp\left(\frac{1}{2}\varepsilon V\right)\exp(\varepsilon T)\exp\left(\frac{1}{2}\varepsilon V\right)$ has $e_{TTV}=1/12$ and $e_{VTV}=1/24$, and propagator ρ_{2B} corresponds to changing e_{TTV}'s value to match that of e_{VTV}. The Takahashi-Imada propagator (5.3) can also be written as

$$\rho'_{TI}=\exp\left(\tfrac{1}{2}\varepsilon V+\tfrac{1}{48}\varepsilon^3[V,[T,V]]\right)\exp(\varepsilon T)\times\exp\left(\tfrac{1}{2}\varepsilon V+\tfrac{1}{48}\varepsilon^3[V,[T,V]]\right), \tag{5.6}$$

corresponding to changing e_{VTV}'s value to match that of e_{TTV}. These are the four fundamental correctable second-order propagators with a fourth-order trace.

For the computation of the trace, it is unnecessary to know the corrector explicitly. In other cases, such as symplectic corrector algorithms, one may wish to apply the corrector occasionally to see the working of the corrected fourth-order propagator $\tilde{\rho}$. We will give a detailed derivation of correctors for propagators (5.3)–(5.6), cumulating in a set of four minimal correctors. These minimal correctors with analytical coefficients have not been previously described in the literature [18–23].

For the Takahashi-Imada propagator, we have $e_{TTV}=e_{VTV}=e_2$ with $e_2=-1/24$. From Eq. (2.7), we see that a possible corrector is $C=e_2\varepsilon[T,V]$. This can be constructed in a straightforward manner as suggested by Wisdom *et al.* [18]. Since

$$\begin{aligned}B(v_1,t_1)&\equiv\exp(\varepsilon v_1 V)\exp(\varepsilon t_1 T)\exp(-\varepsilon v_1 V)\exp(-\varepsilon t_1 T)\\&=\exp\Big(-v_1t_1\varepsilon^2[T,V]-\tfrac{1}{2}t_1^2v_1\varepsilon^3[T,[T,V]]\\&\qquad-\tfrac{1}{2}t_1v_1^2\varepsilon^3[V,[T,V]]+O(\varepsilon^4)\Big),\end{aligned} \tag{5.7}$$

by setting $v_1t_1=(1/48)$, the following product is a workable corrector

$$B(v_1,t_1)B(-v_1,-t_1)=\exp\left(-\tfrac{1}{24}\varepsilon^2[T,V]+O(\varepsilon^4)\right). \tag{5.8}$$

Note that it is important to have the operator V before T to generate a negative e_2 coefficient. However, without fully determining both v_1 and t_1, this corrector clearly underutilizes $B(v_1,t_1)$. It requires eight operators, which is far from optimal. We will show below that four is sufficient.

Let $H=T+V$ and $G=[T,V]$. Since $H_A=H+e_2\varepsilon^2[H,G]$, we can see from Eq. (2.7) that adding a term c_0H to C will not affect the corrector term $\varepsilon[C,T+V]$, but such a term will generate unwanted third-order terms $c_0e_2\varepsilon^3[H,[H,G]]$ from $\varepsilon[C,H_A]$ and $\frac{1}{2}c_0e_2\varepsilon^3[H,[G,H]]$ from $\frac{1}{2}\varepsilon[C,[C,H_A]]$. To cancel them, we must add another term $c_2\varepsilon^2[H,G]$ to the corrector such that $c_2=\frac{1}{2}c_0e_2$. Thus the corrector can have the more general form

$$\exp(\varepsilon C)=\exp\left(c_0\varepsilon H+e_2\varepsilon^2G+\tfrac{1}{2}c_0e_2\varepsilon^3[H,G]\right)+O(\varepsilon^4) \tag{5.9}$$

$$=\exp(c_0\varepsilon H)\exp(e_2\varepsilon^2G)+O(\varepsilon^4), \tag{5.10}$$

where the second line follows from the fundamental Baker-Campbell-Hausdorff formula, $\exp(A)\exp(B)=\exp\{A+B+(1/2)[A,B]\cdots\}$. To exploit the use of the free parameter c_0, we can approximate $\exp(c_0\varepsilon H)$ by

$$\begin{aligned}&\exp\left(\varepsilon\frac{c_0}{2}V\right)\exp(\varepsilon c_0T)\exp\left(\varepsilon\frac{c_0}{2}V\right)\\&\quad=\exp\left(c_0\varepsilon H+\frac{1}{12}c_0^3\varepsilon^3[T,[T,V]]+\frac{1}{24}c_0^3\varepsilon^3[V,[T,V]]\right)\\&\qquad+O(\varepsilon^5),\end{aligned} \tag{5.11}$$

and the term $\exp(e_2\varepsilon^2G)$ by $B(v_1,t_1)$. We can now choose c_0,v_1,t_1 such that $v_1t_1=1/24$ and the third-order terms in Eq. (5.11) exactly cancel the third-order terms in Eq. (5.7): $\frac{1}{2}t_1^2v_1=\frac{1}{12}c_0^3$, $\frac{1}{2}t_1v_1^2=\frac{1}{24}c_0^3$. This gives $c_0=1/(2\times3^{1/6})$, $v_1=1/(4\sqrt{3})$, and $t_1=1/(2\sqrt{3})$. The result is a corrector with six operators:

$$S=\exp\left(\varepsilon\frac{c_0}{2}V\right)\exp(\varepsilon c_0T)\exp\left[\varepsilon\left(\frac{c_0}{2}+v_1\right)V\right]\times\exp(\varepsilon t_1T)\exp(-\varepsilon v_1V)\exp(-\varepsilon t_1T). \tag{5.12}$$

Since this corrector has made good use of all the parameters, it is surprising that one can find an even shorter corrector. Instead of $B(v_1,t_1)$, consider just

$$\exp(\varepsilon d_0V)\exp(\varepsilon d_0T) = \exp\left(d_0\varepsilon H-\tfrac{1}{2}d_0^2[T,V]+\tfrac{1}{12}d_0^3\varepsilon^3[T,[T,V]] - \tfrac{1}{12}d_0^3\varepsilon^3[V,[T,V]]\right)+O(\varepsilon^4). \tag{5.13}$$

The corrector

$$\begin{aligned}S_{TI}&=\exp\left(\varepsilon\frac{c_0}{2}V\right)\exp(\varepsilon c_0T)\exp\left[\varepsilon\left(\frac{c_0}{2}+d_0\right)V\right]\exp(\varepsilon d_0T)\\&=\exp\Big\{(c_0+d_0)\varepsilon H+\left(-\tfrac{1}{2}d_0^2\right)\varepsilon^2G\\&\quad+\tfrac{1}{2}\left(-\tfrac{1}{2}d_0^2\right)(c_0+d_0)[H,G]+\tfrac{1}{12}(c_0^3+4d_0^3)\varepsilon^3[T,[T,V]]\\&\quad+\tfrac{1}{24}(c_0^3+4d_0^3)\varepsilon^3[V,[T,V]]\Big\}+O(\varepsilon^4)\end{aligned} \tag{5.14}$$

will have the correct value for e_2 if we take $d_0^2/2=1/24$, fixing $d_0=(1/2\sqrt{3})$. The corrector will also be of the form (5.9) after both commutators have been eliminated by setting $c_0^3=-4d_0^3$, giving $c_0=-1/(2^{1/3}\sqrt{3})$. This is the minimal corrector for the Takahashi-Imada propagator.

The corrector of the form (5.9) is completely determined by a single number e_2. Its sign dictates the order of the T and V operators, and its value fixes their coefficients. For the alternative propagator ρ'_{2B}, Eq. (5.4), with $e_2=-1/12$, its corrector is of the same form as Eq. (5.14), but now with $d_0=1/\sqrt{6}$ and $c_0=-2^{1/6}/\sqrt{3}$.

For positive values of e_2, the corrector is of the form

$$\begin{aligned}S&=\exp\left(\varepsilon\frac{c_0}{2}T\right)\exp(\varepsilon c_0V)\exp\left[\varepsilon\left(\frac{c_0}{2}+d_0\right)T\right]\exp(\varepsilon d_0V)\\&=\exp\Big\{(c_0+d_0)\varepsilon H+\left(\tfrac{1}{2}d_0^2\right)\varepsilon^2G+\tfrac{1}{2}\left(\tfrac{1}{2}d_0^2\right)(c_0+d_0)[H,G]\\&\quad-\tfrac{1}{24}(c_0^3+4d_0^3)\varepsilon^3[T,[T,V]]-\tfrac{1}{12}(c_0^3+4d_0^3)\varepsilon^3[V,[T,V]]\Big\}\\&\quad+O(\varepsilon^4).\end{aligned} \tag{5.15}$$

Propagator ρ_{2B} is dual to the TI propagator with $e_2=1/24$. Its corrector is of the form (5.15) but with same coefficients $d_0=1/(2\sqrt{3})$ and $c_0=-1/(2^{1/3}\sqrt{3})$. The ρ'_{TI} propagator (5.6) with $e_2=1/12$ is dual to ρ'_{2B}. Its corrector is of the form (5.15) with $d_0=1/\sqrt{6}$ and $c_0=-2^{1/6}/\sqrt{3}$. These compact correctors are fitting companions to their equally compact propagators.

VI. CONCLUSIONS

In this work, we proved a fundamental result on the correctability of forward time step propagators. We show that if $\rho=e^{\varepsilon(T+V)}$ were to be approximated by the product form (2.2), then no product form with positive coefficients $\{t_i\}$ is correctable beyond second order. Whereas a conventional higher-order propagator requires its error terms to vanish, a correctable propagator only requires its error terms to satisfy the correctability condition. The latter requirement seemed far less stringent. A surprising element of this work is that this is not the case. For symmetric decomposition with positive $\{t_i\}$, the two second-order error coefficients cannot both vanish because, they can never be equal. The correctability requirement itself is stringent enough. This proof of noncorrectability generalizes the previous work of Sheng [1] and Suzuki [2].

From knowing correctability requirement, we derived systematically the four forward correctable second-order propagators and their minimal correctors. These minimal correctors follow from a more general form (5.10) of the corrector with free parameters. Much of the existing literature on symplectic corrector is rather opaque, concerned only with how to satisfy "order conditions" numerically [22,23]. This work suggests that a more analytical approach is possible.

The Takahashi-Imada type of propagators considered here are unique in that they are the only known second-order, forward-time-step propagators with a fourth-order trace. If one is willing to evaluate the potential at least twice, then with the inclusion of $[V,[T,V]]$, one can make both error coefficients e_{TTV} and e_{VTV} vanish [8,9]. The result is a whole family of positive time step fourth-order propagators [27–30] with a fourth-order trace. While this class of forward decomposition algorithms is indispensable for solving time-irreversible problems [10–14], they are less interesting from the point of view of calculating the trace. For correctable propagators, their key attraction is that one can obtain a higher-order trace without using a higher-order propagator. Methods and results of this work can be used to study ways of correcting these fourth-order propagators to higher orders.

Note added in proof. F. Casas has informed me that he and S. Blanes have also proved the result in Sec. IV using a different method; see Ref. [31]

ACKNOWLEDGMENTS

I wish to thank J. Boronat and J. Casulleras for their invitation to lecture in Barcelona which initiated this work, E. Krotscheck for his interest and hospitality at Linz, H. Forbert for discussing the correctability requirement, and G. Chen on the use of constrained minimization. This work was supported, in part, by National Science Foundation Grant No. DMS-0310580.

[1] Q. Sheng, IMA J. Numer. Anal. **9**, 199 (1989).
[2] M. Suzuki, J. Math. Phys. **32**, 400 (1991).
[3] D. Goldman and T. J. Kaper, SIAM (Soc. Ind. Appl. Math.) J. Numer. Anal. **33**, 349 (1996).
[4] R. I. McLachlan and P. Atela, Nonlinearity **5**, 542 (1991).
[5] R. I. McLachlan, SIAM J. Sci. Comput. (USA) **16**, 151 (1995).
[6] P. V. Koseleff, in *Integration Algorithms and Classical Mechanics* (American Mathmatical Society, Providence, RI, 1996), Vol. 10, p. 103.
[7] R. I. McLachlan and G. R. W. Quispel, Acta Numerica **11**, 241 (2002).
[8] M. Suzuki, in *Computer Simulation Studies in Condensed Matter Physics VIII*, edited by D. Landau, K. Mon, and H. Shuttler (Springler, Berlin, 1996).
[9] S. A. Chin, Phys. Lett. A **226**, 344 (1997).
[10] H. A. Forbert and S. A. Chin, Phys. Rev. E **63**, 016703 (2001).
[11] H. A. Forbert and S. A. Chin, Phys. Rev. B **63**, 144518 (2001).
[12] J. Auer, E. Krotscheck, and S. A. Chin, J. Chem. Phys. **115**, 6841 (2001).
[13] O. Ciftja and S. A. Chin, Phys. Rev. B **68**, 134510 (2003).
[14] S. Jang, S. Jang, and G. A. Voth, J. Chem. Phys. **115**, 7832 (2001).
[15] M. Kalos, D. Levesque, and L. Verlet, Phys. Rev. A **9**, 2178 (1974).
[16] S. A. Chin, Phys. Rev. A **42**, 6991 (1990).
[17] H. Kleinert, *Path Integrals in Quantum Mechanics, Statistics and Polymer Physics* (World Scientific, Singapore, 1990).
[18] J. Wisdom, M. Holman, and J. Touma, in *Integration Algorithms and Classical Mechanics*, edited by J. E. Marsden, G. W. Patrick, and W. F. Shadwick (American Mathematical Society, Providence, RI, 1996).
[19] R. I. McLachan, in *Integration Algorithms and Classical Mechanics*, edited by J. E. Marsden, G. W. Patrick, and W. F. Shadwick (American Mathematical Society, Providence, RI, 1996).
[20] M. A. Lopez-Marcos, J. M. Sanz-Serna, and R. D. Skeel, in *Numerical Analysis 1995*, edited by D. F. Griffiths and G. A. Watson (Longman, Harlow, U.K., 1996), pp. 107–122.
[21] M. A. Lopez-Marcos, J. M. Sanz-Serna, and R. D. Skeel, SIAM J. Sci. Comput. (USA) **18**, 223 (1997).
[22] S. Blanes, F. Casas, and J. Ros, SIAM J. Sci. Comput. (USA) **21**, 711 (1999).
[23] S. Blanes, Appl. Numer. Math. **37**, 289 (2001).
[24] M. Suzuki, Phys. Lett. A **201**, 425 (1995).
[25] M. Takahashi and M. Imada, J. Phys. Soc. Jpn. **53**, 3765 (1984).
[26] X.-P. Li and J. Q. Broughton, J. Chem. Phys. **86**, 5094 (1987).
[27] S. A. Chin and C. R. Chen, J. Chem. Phys. **117**, 1409 (2002).
[28] I. P. Omelyan, I. M. Mryglod, and R. Folk, Phys. Rev. E **66**, 026701 (2002).
[29] I. P. Omelyan, I. M. Mryglod, and R. Folk, Comput. Phys. Commun. **151**, 272 (2003).
[30] S. A. Chin and C. R. Chen, e-print astro-ph/0304223.
[31] See GIPS preprint at http:/www.focm.net./gi/gips