The Journal of Chemical Physics

RESEARCH ARTICLE | JULY 07 2023

Family of Gaussian wavepacket dynamics methods from the perspective of a nonlinear Schrödinger equation

Jiří J. L. Vaníček

J. Chem. Phys. 159, 014114 (2023)
https://doi.org/10.1063/5.0146680

The Journal of Chemical Physics ARTICLE pubs.aip.org/aip/jcp

Family of Gaussian wavepacket dynamics methods from the perspective of a nonlinear Schrödinger equation

Cite as: J. Chem. Phys. 159, 014114 (2023); doi: 10.1063/5.0146680
Submitted: 15 February 2023 • Accepted: 8 June 2023 •
Published Online: 7 July 2023

Jiří J. L. Vaníček[a)]

AFFILIATIONS

Laboratory of Theoretical Physical Chemistry, Institut des Sciences et Ingénierie Chimiques, Ecole Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland

[a)]Author to whom correspondence should be addressed: jiri.vanicek@epfl.ch

ABSTRACT

Many approximate solutions of the time-dependent Schrödinger equation can be formulated as exact solutions of a nonlinear Schrödinger equation with an effective Hamiltonian operator depending on the state of the system. We show that Heller's thawed Gaussian approximation, Coalson and Karplus's variational Gaussian approximation, and other Gaussian wavepacket dynamics methods fit into this framework if the effective potential is a quadratic polynomial with state-dependent coefficients. We study such a nonlinear Schrödinger equation in full generality: we derive general equations of motion for the Gaussian's parameters, demonstrate time reversibility and norm conservation, and analyze conservation of energy, effective energy, and symplectic structure. We also describe efficient, high-order geometric integrators for the numerical solution of this nonlinear Schrödinger equation. The general theory is illustrated by examples of this family of Gaussian wavepacket dynamics, including the variational and nonvariational thawed and frozen Gaussian approximations and their special limits based on the global harmonic, local harmonic, single-Hessian, local cubic, and local quartic approximations for the potential energy. We also propose a new method by augmenting the local cubic approximation with a single fourth derivative. Without substantially increasing the cost, the proposed "single-quartic" variational Gaussian approximation improves the accuracy over the local cubic approximation and, at the same time, conserves both the effective energy and symplectic structure, unlike the much more expensive local quartic approximation. Most results are presented in both Heller's and Hagedorn's parametrizations of the Gaussian wavepacket.

I. INTRODUCTION

Semiclassical trajectory-based methods for solving the time-dependent Schrödinger equation (TDSE) avoid the exponential scaling of the exact quantum solution and, in contrast to classical methods, can capture various quantum effects at least qualitatively. Semiclassical methods have been successfully applied to the calculation of vibrational and electronic spectra, fluorescence and internal conversion rates, diffusion constants, and rate constants of chemical reactions.[1–5]

Multi-trajectory semiclassical methods, such as the initial value representation,[3] frozen Gaussian approximation,[6] Herman–Kluk propagator,[7] phase averaging,[2] hybrid dynamics,[8] or multiple-spawning,[9] employ an ensemble of trajectories and, therefore, account for both the nonlinear spreading and interference between various parts of the wavepacket. Trajectory ensembles were even used to capture relativistic effects.[10] Converging such methods numerically, however, often requires many trajectories, which can become computationally prohibitive if the potential energy surfaces on which the trajectories evolve are expensive. Unfortunately, this happens in the most interesting modern applications, where the potential energy surfaces are evaluated with *ab initio* electronic structure codes.[11–14]

Single-trajectory semiclassical approximations, although obviously much cruder, provide a practical alternative because they avoid the issue of convergence over the ensemble and permit using accurate potential energy surfaces. In addition, they provide a simpler physical interpretation and preserve more geometric

properties of the exact solution. Among the earliest such methods, Heller's "thawed" Gaussian approximation[15–17] propagates a Gaussian wavepacket along a classical trajectory and lets its width evolve using the local harmonic approximation for the potential. It is very efficient, permits an on-the-fly *ab initio* implementation,[18–21] and includes some anharmonic effects that are completely missing in the simpler global harmonic models.

Although this local harmonic thawed Gaussian wavepacket dynamics (TGWD) conserves the norm of the wavepacket, this method conserves neither the exact nor the local harmonic effective energy. Heller appreciated[22] very early the capability of the time-dependent variational principle[23–25] to improve the accuracy of semiclassical dynamics. Applying the variational principle to the thawed Gaussian ansatz, Heather and Metiu[93] and Coalson and Karplus[26] obtained elegant equations for propagating the Gaussian wavefunction optimally. This variational Gaussian approximation is symplectic,[25,27,28] conserves the energy exactly, and may even capture shallow tunneling, but it does so at a much higher cost because it requires the expectation values of the potential energy, gradient, and Hessian. Noting that the local harmonic TGWD can also be obtained by applying the variational principle to the local harmonic approximation of the potential, Pattanayak and Schieve improved the accuracy of Heller's method by including the third derivative of the potential in their "extended" semiclassical wavepacket dynamics.[29,30] Ohsawa and Leok[31] pointed out that this method is—similar to the variational TGWD, but unlike the local harmonic TGWD—exactly symplectic, so they called it "symplectic semiclassical wavepacket dynamics."[31–34]

To accelerate on-the-fly *ab initio* applications of the local harmonic TGWD, Begušić, Cordova, and Vaníček proposed[35] the single-Hessian approximation,[21,36] in which the trajectory evolves according to the original potential but the width of the wavepacket feels a constant curvature. Remarkably, this simplified approximation preserves the effective energy exactly and has a Hamiltonian structure in an augmented phase space.[35]

All of the approximations mentioned in the three preceding paragraphs propagate a Gaussian wavepacket in a time-dependent quadratic potential whose parameters depend on the instantaneous state of the system. Therefore, each of these approximations is also an exact solution of a certain nonlinear TDSE, i.e., a Schrödinger equation whose Hamiltonian depends on the quantum state. The most famous nonlinear Schrödinger equation is probably the Gross–Pitaevskii equation,[37–39] which describes approximately the Bose–Einstein condensates, but many other approximations can be stated as exact solutions of a TDSE with a state-dependent Hamiltonian operator.[35] The time-dependent variational principle[23–25] seeks optimal solutions in a nonlinear manifold of possible solutions and, thus, yields many examples of nonlinear TDSEs, including the time-dependent Hartree,[23,40,41] time-dependent Hartree–Fock,[23,42] or multi-configurational time-dependent Hartree method.[43,44] Most nonlinear TDSEs, however, do not rely on the variational principle—representative examples[45,46] are Heller's thawed Gaussian approximation[15] and the local control theory,[47–49] which seeks a state-dependent electric field that increases or decreases an observable of interest.

This paper explores Gaussian wavepacket dynamics from the perspective of a nonlinear TDSE with an effective potential that is a quadratic polynomial of coordinates with state-dependent coefficients. The remainder of the paper is organized as follows: Sec. II reviews the basic properties of the linear TDSE in order to highlight the differences from the nonlinear TDSE, presented in Sec. III. The formalism developed in Sec. III is completely general and also applies to nonlinear TDSEs that do not result from the variational principle and to non-Gaussian wavepackets. In contrast to the variational or symplectic approaches,[27,31] the formalism is elementary and does not rely on the very elegant, but advanced symplectic formulation of Hamiltonian quantum dynamics.

The main Sec. IV defines Gaussian wavepacket dynamics as an example of a nonlinear TDSE. We derive equations of motion for the parameters of the Gaussian in terms of the coefficients of the effective quadratic potential and analyze time reversibility and the conservation of norm, energy, effective energy, and symplectic structure. Although the focus on Gaussian wavepackets makes it possible to obtain more detailed results, the formalism remains more general than the variational and symplectic approaches because the nonlinear TDSE analyzed in Sec. IV still does not have to arise from the variational principle and the equations of motion do not have to be Hamiltonian for a non-canonical symplectic structure. Examples of wavepacket dynamics employing Gaussians with a flexible width are presented in Sec. VI and Gaussians with a fixed width in Sec. VII. In Sec. VI G, we propose a new, single-quartic TGWD, which—in contrast to the similarly accurate but much more expensive local quartic approximation—conserves both the symplectic structure and effective energy and, at the same time, improves the accuracy over the local cubic approximation without increasing the cost. The proposed method allows for tunneling, but, unlike Coalson and Karplus's variational TGWD, makes it possible to evaluate analytically the expectation values of the potential, gradient, and Hessian; it is, therefore, a natural extension of Heller's thawed Gaussian approximation, which uses classical trajectories and cannot describe tunneling. Geometric integrators for the general Gaussian wavepacket dynamics are described in Sec. VIII. Section IX translates all results from Heller's to Hagedorn's parametrization of the Gaussian wavepacket. Finally, Sec. X discusses the relationship between the three approaches to the TGWD and concludes the paper. For reference, the nonstandard notation used in this paper is summarized in Table I.

TABLE I. Summary of nonstandard notation.

Quantity and equation, where defined	Symbol and definition
Shifted position (27)	$x := q - q_t$
Shifted position operator (31)	$\hat{x} := \hat{q} - q_t$
Scaled and shifted position (34)	$\xi := A_t \cdot x + p_t$
Position covariance (56)	$\Sigma_t := \mathrm{Cov}(\hat{q})$
Real part of the width matrix (103)	$\mathcal{A} := \mathrm{Re}\, A$
Imaginary part of the width matrix (104)	$\mathcal{B} := \mathrm{Im}\, A$

II. LINEAR SCHRÖDINGER EQUATION

Let us briefly review the properties of the *linear* time-dependent Schrödinger equation (TDSE)

$$i\hbar|\dot{\psi}(t)\rangle = \hat{H}|\psi(t)\rangle, \tag{1}$$

in which the wavepacket $\psi(t)$ is driven by the time-independent Hermitian linear Hamiltonian operator $\hat{H}$. The Hamiltonian is said to be linear because it is independent of the state ψ and not because of a linear dependence on coordinates. Indeed, $\hat{H}$ can be a nonlinear function of coordinates and still be a linear operator.

The state $\psi(t)$ at time t can be obtained from the initial state $\psi(0)$ formally as $|\psi(t)\rangle = \hat{U}(t)|\psi(0)\rangle$, where $\hat{U}(t) = \exp(-it\hat{H}/\hbar)$ is the time evolution operator. Because $\hat{H}$ is a linear operator, so is $\hat{U}(t)$. The evolution is time-reversible because $\hat{U}(-t)\hat{U}(t)|\psi(0)\rangle = |\psi(0)\rangle$.

The exact quantum evolution with a time-independent Hermitian linear Hamiltonian $\hat{H}$ conserves both the norm $\|\psi(t)\|$ of the quantum state and its energy

$$E := \langle \hat{H} \rangle, \tag{2}$$

where $\langle \hat{A} \rangle \equiv \langle \hat{A} \rangle_{\psi(t)} := \langle \psi(t)|\hat{A}|\psi(t)\rangle$ denotes the expectation value of operator $\hat{A}$ in the state $\psi(t)$. Both conservation properties follow from a general expression

$$d\langle \hat{A} \rangle/dt = (i\hbar)^{-1}\langle [\hat{A}, \hat{H}] \rangle \tag{3}$$

for the time dependence of $\langle \hat{A} \rangle$, applied to the identity operator ($\hat{A} = \hat{1}$) or to the Hamiltonian ($\hat{A} = \hat{H}$). The linear time evolution also conserves the inner product $\langle \psi(t)|\phi(t)\rangle$ of two different states.

Here, we will usually assume that the Hamiltonian is separable into a sum

$$\hat{H} = \hat{T} + \hat{V} = T(\hat{p}) + V(\hat{q}) \tag{4}$$

of a kinetic energy term $\hat{T} \equiv T(\hat{p})$, depending only on momentum p, and a potential energy term $\hat{V} \equiv V(\hat{q})$, depending only on position q. Both q and p are D-dimensional vectors. We call Hamiltonians described by Eq. (4) "separable," without requiring that the potential energy $V(q)$ itself be separable into a sum $V_1(q_1) + \cdots + V_D(q_D)$ of D functions, each depending on a single degree of freedom; beware that many authors require this property in the definition of separability. While the potential energy function $V(q)$ can be an arbitrary real-valued function, for the kinetic energy we shall assume the quadratic form

$$T(p) = p^T \cdot m^{-1} \cdot p/2, \tag{5}$$

where m is the (not necessarily diagonal) positive-definite real symmetric $D \times D$ mass matrix. In a linear TDSE, neither $\hat{T}$ nor $\hat{V}$ depends on the state ψ.

III. NONLINEAR SCHRÖDINGER EQUATION

When Eq. (1) is solved approximately, its approximate solution can often[25,45,46] be interpreted as the *exact* solution of the *nonlinear* Schrödinger differential equation

$$i\hbar|\dot{\psi}(t)\rangle = \hat{H}_{\text{eff}}[\psi(t)]|\psi(t)\rangle \tag{6}$$

with an effective Hamiltonian operator $\hat{H}_{\text{eff}}(\psi)$ depending on the state ψ (see Fig. 1). Although one may envision a more general nonlinearity where $\hat{H}_{\text{eff}}(\psi)|\psi\rangle$ in the right-hand side of Eq. (6) would be replaced by an arbitrary functional of ψ, the slightly less general case described by the "quasi-linear" form of Eq. (6) is much more interesting because it preserves some features of the linear Schrödinger equation. Notation $\hat{H}_{\text{eff}}(\psi)|\psi\rangle$ reflects our assumption that whereas the mapping $\hat{H}_{\text{eff}} : |\psi\rangle \mapsto \hat{H}_{\text{eff}}(\psi)|\psi\rangle$ is nonlinear, the mapping $\hat{H}_{\text{eff}}(\psi) : |\phi\rangle \mapsto \hat{H}_{\text{eff}}(\psi)|\phi\rangle$ is linear for all ψ. In addition, we shall assume that the expectation value

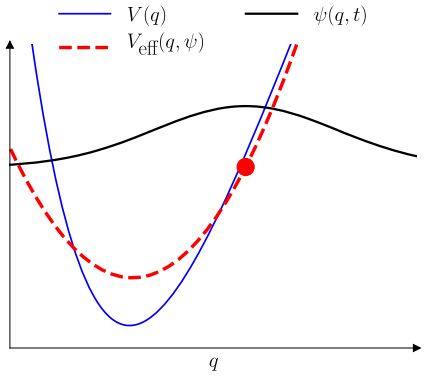

FIG. 1. Example of a wavepacket $\psi(q,t)$ that solves exactly a nonlinear Schrödinger equation (6) with a separable, state-dependent effective Hamiltonian $\hat{H}_{\text{eff}}(\psi) = T(\hat{p}) + V_{\text{eff}}(\hat{q};\psi)$. Here, the wavepacket is a Gaussian (26), and the effective potential $V_{\text{eff}}(q;\psi)$ is a quadratic polynomial (30) whose coefficients are given by the single quartic variational approximation [Eqs. (96)–(99)] to the original, Morse potential $V(q)$.

$$\langle \hat{H}_{\text{eff}}(\psi) \rangle_\phi := \langle \phi|\hat{H}_{\text{eff}}(\psi)\phi\rangle, \tag{7}$$

which generalizes the energy, is real for any ψ and ϕ. This condition implies that $\hat{H}_{\text{eff}}(\psi)$, considered as a linear operator (i.e., with fixed ψ), is Hermitian:

$$\langle \phi|\hat{H}_{\text{eff}}(\psi)\theta\rangle = \langle \hat{H}_{\text{eff}}(\psi)\phi|\theta\rangle \text{ for all } \psi, \phi, \theta. \tag{8}$$

The state at time $t \geq t_0$ can be obtained from the initial state at time t_0 as $|\psi(t)\rangle = \hat{U}(t, t_0; \psi)|\psi(t_0)\rangle$, i.e., by the multiplication with the evolution operator,

$$\hat{U}(t, t_0; \psi) = \mathcal{T} \exp\left[-\frac{i}{\hbar}\int_{t_0}^{t} \hat{H}_{\text{eff}}[\psi(t')]dt'\right], \tag{9}$$

where $\mathcal{T}$ indicates time ordering. The evolution operator is nonlinear because it depends on the propagated state. Yet, the evolution guided by the nonlinear Schrödinger equation (6) remains time-reversible,

$$\hat{U}(t_0, t; \psi)\hat{U}(t, t_0; \psi)|\psi(t_0)\rangle = |\psi(t_0)\rangle, \tag{10}$$

because, for $t_0 \leq t$,

$$\begin{aligned}\hat{U}(t_0, t; \psi) &= \bar{\mathcal{T}} \exp\left[-\frac{i}{\hbar}\int_{t}^{t_0} \hat{H}_{\text{eff}}[\psi(t')]dt'\right] \\ &= \bar{\mathcal{T}} \exp\left[\frac{i}{\hbar}\int_{t_0}^{t} \hat{H}_{\text{eff}}[\psi(t')]dt'\right] = \hat{U}(t, t_0; \psi)^{-1},\end{aligned} \tag{11}$$

where $\bar{\mathcal{T}}$ denotes reverse time ordering. Because the Hamiltonian is state-dependent, the scalar product generally depends on time:

$$\frac{d}{dt}\langle\psi|\phi\rangle = \langle\dot{\psi}|\phi\rangle + \langle\psi|\dot{\phi}\rangle = (i\hbar)^{-1}\langle\psi|[-\hat{H}_{\text{eff}}(\psi) + \hat{H}_{\text{eff}}(\phi)]\phi\rangle \neq 0, \tag{12}$$

where we used relation (A1) from Appendix A 1 for $\langle\dot{\psi}|\phi\rangle$.

Let us inspect the conservation of norm and energy by the nonlinear Schrödinger equation (6). These and other useful properties are obtained by generalizing Eq. (3) to the state-dependent operators $\hat{H}_{\text{eff}}(\psi)$ and $\hat{A}(\psi)$. In Appendix A 2, we prove that the expectation value $\langle\hat{A}(\psi)\rangle := \langle\psi|\hat{A}(\psi)\psi\rangle$ evolves in time according to the equation

$$\frac{d\langle\hat{A}(\psi)\rangle}{dt} = \left\langle\frac{d}{dt}\hat{A}(\psi)\right\rangle + \frac{1}{i\hbar}\langle[\hat{A}(\psi), \hat{H}_{\text{eff}}(\psi)]\rangle. \tag{13}$$

For a linear, state-independent operator $\hat{A}$, the general equation (13) simplifies to

$$d\langle\hat{A}\rangle/dt = (i\hbar)^{-1}\langle[\hat{A}, \hat{H}_{\text{eff}}(\psi)]\rangle. \tag{14}$$

Applying this expression, which is analogous to Eq. (3) for the linear TDSE, to the identity operator demonstrates that the nonlinear TDSE (6) conserves the norm,

$$\frac{d}{dt}\|\psi(t)\|^2 = \frac{d}{dt}\langle\hat{1}\rangle = \frac{1}{i\hbar}\langle[\hat{1}, \hat{H}_{\text{eff}}(\psi)]\rangle = 0, \tag{15}$$

while its application to the exact Hamiltonian yields the time dependence of energy,

$$\dot{E} = d\langle\hat{H}\rangle/dt = (i\hbar)^{-1}\langle[\hat{H}, \hat{H}_{\text{eff}}(\psi)]\rangle. \tag{16}$$

As a result, the energy E may not be conserved under evolution with the effective Hamiltonian $\hat{H}_{\text{eff}}(\psi)$.

For effective Hamiltonians, one can also study the time dependence of the effective energy

$$E_{\text{eff}} := \langle\hat{H}_{\text{eff}}(\psi)\rangle. \tag{17}$$

Because the exact energy (2) is conserved under the exact time evolution with the Hamiltonian $\hat{H}$, one might expect that the effective energy (17) would be conserved under the evolution with the effective Hamiltonian $\hat{H}_{\text{eff}}$. This is not true in general; taking $\hat{A} = \hat{H}_{\text{eff}}(\psi)$ in Eq. (13) gives

$$\dot{E}_{\text{eff}} = \langle d\hat{H}_{\text{eff}}[\psi(t)]/dt\rangle. \tag{18}$$

A. Nonlinear TDSE with a separable Hamiltonian

In what follows, we only consider separable Hamiltonians (4) and related separable effective Hamiltonians

$$\hat{H}_{\text{eff}}(\psi) = \hat{T} + \hat{V}_{\text{eff}}(\psi) = T(\hat{p}) + V_{\text{eff}}(\hat{q};\psi). \tag{19}$$

Expressed in position representation, the nonlinear Schrödinger equation (6) becomes

$$i\hbar\partial_t\psi(q,t) = -\frac{\hbar^2}{2}\nabla^T\cdot m^{-1}\cdot\nabla\psi(q,t) + V_{\text{eff}}(q;\psi)\psi(q,t). \tag{20}$$

With the obvious exception of energy conservation, many of the following results can be easily generalized to explicitly time-dependent potentials $\hat{V}(t)$. However, for the sake of brevity, we shall continue assuming that the original potential energy $\hat{V}$ is independent of time and that the effective potential: $\hat{V}_{\text{eff}} \equiv \hat{V}_{\text{eff}}(\psi)$ depends on time only implicitly, via the dependence on the state $\psi(t)$.

For separable Hamiltonians, general expressions (16) and (18) for the time dependence of energy and effective energy reduce to (see Appendix A 3 for proof)

$$\dot{E} = \text{Re}\langle\hat{p}^T\cdot m^{-1}\cdot(\hat{V}' - \hat{V}'_{\text{eff}})\rangle \tag{21}$$

and

$$\dot{E}_{\text{eff}} = \langle d\hat{V}_{\text{eff}}[\psi(t)]/dt\rangle, \tag{22}$$

where the gradient of the effective potential is defined as

$$\hat{V}'_{\text{eff}} := \partial V_{\text{eff}}(q;\psi)/\partial q|_{q=\hat{q}}. \tag{23}$$

Finally, applying Eq. (14) to the position and momentum operators shows that the Ehrenfest theorem continues to hold for the nonlinear Schrödinger equation with the separable Hamiltonian (19), namely,

$$d\langle\hat{q}\rangle/dt = m^{-1}\cdot\langle\hat{p}\rangle, \tag{24}$$

$$d\langle\hat{p}\rangle/dt = -\langle\hat{V}'_{\text{eff}}\rangle. \tag{25}$$

These two equations follow from the explicit expressions (A6) and (A7) in Appendix A 3 for the commutators of position and momentum with the effective Hamiltonian.

IV. NONLINEAR TIME-DEPENDENT SCHRÖDINGER EQUATION FOR A GAUSSIAN WAVEPACKET

A. Gaussian wavepacket

Let us consider approximate solutions of the linear Schrödinger equation (1), which are exact solutions of the nonlinear Schrödinger equation (6), in which the wavepacket $\psi(t)$ has a Gaussian form at all times. I.e., we will consider wavepackets written in position representation as

$$\psi(q,t) = \exp\left[\frac{i}{\hbar}\left(\frac{1}{2}x^T\cdot A_t\cdot x + p_t^T\cdot x + \gamma_t\right)\right], \tag{26}$$

where the shifted position vector

$$x := q - q_t, \tag{27}$$

was introduced to simplify notation, and q_t, p_t, A_t, and γ_t are Heller's[15,50,51] parameters determining the time dependence of $\psi(t)$. The parameters q_t and p_t are real D-dimensional vectors equal to the expectation values of position and momentum, A_t is a complex symmetric $D\times D$ matrix, and γ_t is a complex scalar. The positive definite imaginary part of A_t determines the width of the Gaussian, while its real part introduces a spatial chirp; the real part of γ_t gives a time-dependent phase, while its imaginary part controls the norm of $\psi(t)$, given by

$$\|\psi(t)\| = [\det(\text{Im } A_t/\pi\hbar)]^{-1/4}e^{-\text{Im}\,\gamma_t/\hbar}. \tag{28}$$

As shown above, in Eq. (15), this norm is conserved by the solutions of Eq. (6). Choosing γ_0 so that

$$\exp\left(-\mathrm{Im}\,\gamma_0/\hbar\right) = \left[\det\left(\mathrm{Im}\,A_0/\pi\hbar\right)\right]^{1/4} \tag{29}$$

ensures unit normalization at all times.

B. Nonlinear time-dependent Schrödinger equation in terms of parameters of the Gaussian wavepacket

Now we shall show that the nonlinear Schrödinger equation (20) is solved exactly by a Gaussian wavepacket (26) if and only if the effective potential energy operator $\hat{V}_{\mathrm{eff}}(\psi)$ is a quadratic polynomial

$$\begin{aligned}\hat{V}_{\mathrm{eff}}(\psi) &\equiv V_{\mathrm{eff}}(\hat{q};\psi) \equiv V_{\mathrm{eff}}(\hat{x};\psi)\\ &= V_0 + V_1^T\cdot\hat{x} + \hat{x}^T\cdot V_2\cdot\hat{x}/2\\ &= V_0 + V_1^T\cdot\hat{x} + \mathrm{Tr}\left(V_2\cdot\hat{x}\otimes\hat{x}^T\right)/2\end{aligned} \tag{30}$$

in the shifted position operator

$$\hat{x} := \hat{q} - q_t \tag{31}$$

with coefficients $V_j \equiv V_j(\psi)$ $(j = 0, 1, 2)$ that may depend on the state ψ. Note that $\hat{V}_{\mathrm{eff}}(\psi)$ depends on ψ not only through the coefficients $V_j(\psi)$ but also through $\hat{x} = \hat{q} - q_t$ since $q_t = \langle\hat{q}\rangle_\psi$. See Fig. 1 for an example.

Let us rewrite the Schrödinger equation (20) in terms of ordinary differential equations for the parameters of the Gaussian (26). Using the chain rule and Eqs. (B7) and (B8) from Appendix B, we find that the time derivative of the Gaussian wavepacket (26) is

$$\begin{aligned}\partial\psi/\partial t = &\left(\partial\psi/\partial q_t\right)^T\cdot\dot{q}_t + \left(\partial\psi/\partial p_t\right)^T\cdot\dot{p}_t\\ &+ \mathrm{Tr}\left[\left(\partial\psi/\partial A_t\right)^T\cdot\dot{A}_t\right] + \left(\partial\psi/\partial\gamma_t\right)\dot{\gamma}_t.\end{aligned} \tag{32}$$

Employing expressions (B10)–(B13) for the partial derivatives of ψ with respect to various parameters, we can rewrite the time derivative (32) as

$$i\hbar\frac{\partial\psi}{\partial t} = \left(\xi^T\cdot\dot{q}_t - x^T\cdot\dot{p}_t - \frac{1}{2}x^T\cdot\dot{A}_t\cdot x + \dot{\gamma}_t\right)\psi, \tag{33}$$

where we introduced a complex vector

$$\xi := A_t\cdot x + p_t = A_t\cdot\left(q - q_t\right) + p_t \tag{34}$$

to simplify notation. The kinetic energy acting on ψ requires differentiating ψ twice with respect to x,

$$\begin{aligned}\langle q|\hat{T}|\psi\rangle &= \left(-\hbar^2/2\right)\nabla^T\cdot m^{-1}\cdot\nabla\psi(q)\\ &= \frac{1}{2}\left[\xi^T\cdot m^{-1}\cdot\xi - i\hbar\,\mathrm{Tr}\left(m^{-1}\cdot A_t\right)\right]\psi(q),\end{aligned} \tag{35}$$

where we used Eq. (B15) from Appendix B.

Schrödinger's equation (6) can also be written as

$$0 = i\hbar|\dot{\psi}(t)\rangle - \hat{H}_{\mathrm{eff}}|\psi(t)\rangle \quad \text{and,} \tag{36}$$

using Eqs. (33) and (35), in position representation as

$$0 = \left[f(x) - V_{\mathrm{eff}}(x)\right]\psi, \tag{37}$$

where

$$f(x) = C_0 + C_1^T\cdot x + x^T\cdot C_2\cdot x/2 \tag{38}$$

is a quadratic polynomial with coefficients

$$\begin{aligned}C_0(\psi) := &\,p_t^T\cdot\dot{q}_t - \dot{\gamma}_t + \left(i\hbar/2\right)\mathrm{Tr}\left(m^{-1}\cdot A_t\right)\\ &- p_t^T\cdot m^{-1}\cdot p_t/2,\end{aligned} \tag{39}$$

$$C_1(\psi) := A_t\cdot\dot{q}_t - \dot{p}_t - A_t\cdot m^{-1}\cdot p_t, \tag{40}$$

$$C_2(\psi) := -\dot{A}_t - A_t\cdot m^{-1}\cdot A_t. \tag{41}$$

Because $f(x)$ is a quadratic polynomial, Eq. (37) is satisfied at all x if and only if $V_{\mathrm{eff}}(x;\,\psi)$ is also a quadratic polynomial in x in the form of Eq. (30) and, in addition, $C_j(\psi) = V_j(\psi)$ for $j = 0, 1, 2$. Let us summarize this in

Proposition 1 (Gaussian wavepacket in a linear or nonlinear TDSE). Gaussian wavepacket (26) solves the nonlinear TDSE (20) with a possibly state-dependent effective potential $\hat{V}_{\mathrm{eff}}(\psi)$ if and only if V_{eff} is a quadratic potential (30) and $C_j(\psi) = V_j(\psi)$ for $j = 0, 1, 2$. In particular, the Gaussian wavepacket solves the linear TDSE (1) if and only if the linear (i.e., independent of ψ) operator $\hat{V}$ is a quadratic polynomial of coordinates.

The system of equations $C_j(\psi) = V_j(\psi)$ (for $j = 0, 1, 2$) seems rather complicated to be useful in practice because Eqs. (39)–(41) couple the time derivatives of the Gaussian parameters. However, it is easy to invert this system:

Proposition 2. Let $V_0(\psi)$, $V_1(\psi)$, and $V_2(\psi)$ be, respectively, some prescribed real scalar, vector, and symmetric matrix functions of the state ψ. Then the Gaussian wavepacket (26) solves the nonlinear TDSE (20) with the effective potential (30) if and only if the parameters of the Gaussian solve the following system of ordinary differential equations:

$$\dot{q}_t = m^{-1}\cdot p_t, \tag{42}$$

$$\dot{p}_t = -V_1, \tag{43}$$

$$\dot{A}_t = -A_t\cdot m^{-1}\cdot A_t - V_2, \tag{44}$$

$$\dot{\gamma}_t = T(p_t) - V_0 + \left(i\hbar/2\right)\mathrm{Tr}\left(m^{-1}\cdot A_t\right). \tag{45}$$

Proof. Proposition 1 implies that we can replace C_j with V_j, $j = 0, 1, 2$, in Eqs. (39)–(41). Equation (44) of motion for $\dot{A}_t$ follows immediately by inverting Eq. (41) for the C_2 matrix. Equation (42) for $\dot{q}_t$ is obtained from the imaginary part of Eq. (40) for the C_1 vector,

$$0 = \mathrm{Im}\,A_t\cdot\dot{q}_t - \mathrm{Im}\,A_t\cdot m^{-1}\cdot p_t. \tag{46}$$

Because the evolution of $\psi(t)$ conserves the norm (28) and the initial state $\psi(0)$ is normalized, $\mathrm{Im}\,A_t$ must be invertible. Multiplying the

last equation on the left with $(\mathrm{Im}\, A_t)^{-1}$ yields Eq. (42) for $\dot{q}_t$. Substituting the equation of motion (42) for $\dot{q}_t$ into Eq. (40) for C_1 and Eq. (39) for C_0 yields Eq. (43) for $\dot{p}_t$ and Eq. (45) for $\dot{\gamma}_t$. The opposite implication is proved similarly: In particular, Eq. (39) follows from Eqs. (42) and (45), whereas Eqs. (42) and (43) imply Eq. (40), and Eq. (44) implies Eq. (41).

Next, after discussing various properties of the effective potential (30), we will show that Eqs. (42) and (43) of motion for position and momentum also follow, more generally and directly, from the Ehrenfest theorem.

C. Properties of the quadratic effective potential

Let us list several useful properties of the effective potential V_{eff} of Eq. (30). Since $x := q - q_t$, the gradient vector and Hessian matrix of V_{eff} are

$$V'_{\mathrm{eff}}(x) = V_1 + V_2 \cdot x, \tag{47}$$

$$V''_{\mathrm{eff}}(x) = V_2. \tag{48}$$

Here, we have used and will use short-hand notations

$$g' := \mathrm{grad}\, g := \nabla g, \tag{49}$$

$$g'' := \mathrm{Hess}\, g := \nabla \otimes \nabla^T g \tag{50}$$

for the gradient and Hessian of function $g(x)$. Expected values of the effective potential energy (30), its gradient, and its Hessian in the Gaussian wavepacket (26) are

$$\begin{aligned}\langle \hat{V}_{\mathrm{eff}} \rangle &= V_0 + V_1^T \cdot \langle \hat{x} \rangle + \mathrm{Tr}\left[V_2 \cdot \langle \hat{x} \otimes \hat{x}^T \rangle\right]/2 \\ &= V_0 + \mathrm{Tr}\left(V_2 \cdot \Sigma_t\right)/2,\end{aligned} \tag{51}$$

$$\langle \hat{V}'_{\mathrm{eff}} \rangle = V_1 + V_2 \cdot \langle \hat{x} \rangle = V_1, \tag{52}$$

$$\langle \hat{V}''_{\mathrm{eff}} \rangle = V_2, \tag{53}$$

where we have invoked relations

$$\langle \hat{x} \rangle = \langle \hat{q} - q_t \rangle = \langle \hat{q} \rangle - q_t = 0, \tag{54}$$

$$\langle \hat{x} \otimes \hat{x}^T \rangle = \langle (\hat{q} - q_t) \otimes (\hat{q} - q_t)^T \rangle = \mathrm{Cov}(\hat{q}) \tag{55}$$

for the mean and covariance of position and introduced the shorthand notation

$$\Sigma_t := \mathrm{Cov}(\hat{q}). \tag{56}$$

In a Gaussian wavepacket, $q_t = \langle \hat{q} \rangle$, $p_t = \langle \hat{p} \rangle$ and, therefore, the equations of motion (42) and (43) for position and momentum also follow immediately from Eqs. (24) and (25) of the Ehrenfest theorem for the general nonlinear TDSE (20) and from Eq. (52) for $\langle \hat{V}'_{\mathrm{eff}} \rangle$. Ehrenfest theorem was also used by Pattanayak and Schieve to derive the equations of semiquantal dynamics.[29,30] Finally, note that for a fixed state ψ, $\hat{V}_{\mathrm{eff}}(\psi)$ is a Hermitian operator because

$$\langle \hat{V}_{\mathrm{eff}}(\psi) \rangle_\phi = V_0(\psi) + \mathrm{Tr}\left[V_2(\psi) \cdot \mathrm{Cov}_\phi(\hat{q})\right]/2 \tag{57}$$

is real for any ψ and ϕ since $V_0(\psi)$, $V_2(\psi)$, and the covariance $\mathrm{Cov}_\phi(\hat{q})$ of position in the state ϕ are all real. This justifies our assumption (7) of the hermiticity of $\hat{H}_{\mathrm{eff}}$ used in the general analysis in Sec. III.

V. GEOMETRIC PROPERTIES OF GAUSSIAN WAVEPACKET DYNAMICS

As mentioned in Sec. II, the exact solution of the linear TDSE (1) with a time-independent Hamiltonian has several "geometric" properties: the time evolution is linear, unitary, norm-conserving, energy-conserving, symplectic, and time-reversible. Symplecticity means that the time evolution conserves the symplectic structure—a symplectic 2-form $\omega(\psi, \phi) := -2\hbar\, \mathrm{Im}\, \langle \psi, \phi \rangle$, defined on the Hilbert space as the imaginary part of the scalar product.[25] The loss of linearity implies that the nonlinear TDSE fails to conserve the inner product. As a result, the conservation of neither the norm nor the symplectic structure is guaranteed. Let us discuss the time reversibility and conservation of norm, energy, effective energy, and symplectic structure by the Gaussian wavepacket dynamics—the nonlinear TDSE (20) with the effective potential (30).

A. Norm conservation

As already shown in Secs. III and IV A, the norm of a Gaussian wavepacket is always conserved, although a scalar product between two different initial states is not.

B. Exact and effective energies

Quantum-mechanical energy in a state ψ driven by the separable Hamiltonian (4) is given by the sum of expectation values of kinetic and potential energies,

$$E = \langle \hat{H} \rangle = \langle \hat{T} \rangle + \langle \hat{V} \rangle. \tag{58}$$

In general, the expected value $\langle \hat{V} \rangle$ of the potential energy cannot be simplified. Because the kinetic energy has the quadratic form (5), its expected value in a Gaussian is

$$\begin{aligned}\langle \hat{T} \rangle &= \langle \hat{p}^T \cdot m^{-1} \cdot \hat{p} \rangle/2 = \mathrm{Tr}\left(m^{-1} \cdot \langle \hat{p} \otimes \hat{p}^T \rangle\right)/2 \\ &= T(p_t) + \mathrm{Tr}\left[m^{-1} \cdot \mathrm{Cov}(\hat{p})\right]/2,\end{aligned} \tag{59}$$

where the first term is the classical kinetic energy at the wavepacket's center and the second term reflects the finite width of the wavepacket; $\mathrm{Cov}(\hat{p})$ is the momentum covariance matrix (B24). In the third step of the derivation, we used Eq. (B29) from Appendix B 4.

Unlike the energy, the effective energy (17),

$$E_{\mathrm{eff}} = \langle \hat{H}_{\mathrm{eff}} \rangle = \langle \hat{T} \rangle + \langle \hat{V}_{\mathrm{eff}} \rangle, \tag{60}$$

can be evaluated fully analytically since, for quadratic effective potentials (30), $\langle \hat{V}_{\mathrm{eff}} \rangle$ is given by Eq. (51).

Because the effective potential is different from the exact potential, generally $\langle \hat{V}_{\rm eff}\rangle \neq \langle \hat{V}\rangle$ and, therefore, $E_{\rm eff} \neq E$. Below, we shall see that in the special case of the variational Gaussian wavepacket dynamics, a beautiful cancellation results in the equality $E_{\rm eff} = E$.

C. Time dependence or conservation of energy

As follows from a more universal analysis in Sec. III A, the evolution of a Gaussian wavepacket with an approximate effective Hamiltonian (19) may not conserve energy. It is a remarkable fact that energy *is* conserved in the special cases of variational thawed and frozen Gaussian wavepacket dynamics (FGWD), discussed below in Secs. VI A and VII A. More generally, energy is conserved along the solutions satisfying the Dirac–Frenkel variational principle for any, not necessarily Gaussian, wavefunction ansatz compatible with the principle (see Appendix C).[25,52–54] No other example of Gaussian wavepacket dynamics among those presented in Secs. VI and VII conserves energy.

To see when energy may be conserved, let us derive a universal expression for the time dependence of energy in a system propagated in a general quadratic effective potential (30). Substituting the gradient (47) of the effective potential into Eq. (21) gives

$$\begin{aligned}\dot{E} &= \mathrm{Re}\langle \hat{p}^{\,T}\cdot m^{-1}\cdot(\hat{V}' - V_1 - V_2\cdot\hat{x})\rangle \\ &= \mathrm{Tr}\,[m^{-1}\cdot \mathrm{Re}\langle(\hat{V}' - V_1 - V_2\cdot\hat{x})\otimes\hat{p}^{\,T}\rangle].\end{aligned} \tag{61}$$

Substitution of expression (B35) from Appendix B 5 for the expected value in Eq. (61) yields

$$\begin{aligned}\dot{E} &= p_t^T\cdot m^{-1}\cdot(\langle\hat{V}'\rangle - V_1) \\ &\quad + \mathrm{Tr}\,[m^{-1}\cdot(\langle\hat{V}''\rangle - V_2)\cdot \mathrm{Cov}_R(\hat{q},\hat{p})],\end{aligned} \tag{62}$$

where $\mathrm{Cov}_R(\hat{q},\hat{p}) = (\hbar/2)(\mathrm{Im}\,A_t)^{-1}\cdot \mathrm{Re}\,A_t$ is the real covariance (B26). Relation (62) helps determine when the energy is conserved exactly. First, consider a Gaussian with a purely imaginary width matrix A_t. Then $\mathrm{Cov}_R(\hat{q},\hat{p}) = 0$ and, in order for $\dot{E}$ to be zero for arbitrary m and p_t, we must have $V_1 = \langle\hat{V}'\rangle$. If A_t is purely imaginary, Eq. (44) implies that both $\mathrm{Re}\,A_t$ and $\mathrm{Cov}_R(\hat{q},\hat{p})$ will become nonzero unless $\dot{A}_t = 0$, i.e., we have a "frozen" Gaussian with constant width matrix $A_t = A_0$ and require that $V_2 = -A_0\cdot m^{-1}\cdot A_0$. If A_t is not purely imaginary, $\mathrm{Cov}_R(\hat{q},\hat{p}) \neq 0$ and we must have both $V_1 = \langle\hat{V}'\rangle$ and $V_2 = \langle\hat{V}''\rangle$. To sum up, there are two general ways to guarantee the conservation of energy: Either

$$V_1 = \langle\hat{V}'\rangle \quad \text{and} \quad V_2 = \langle\hat{V}''\rangle \tag{63}$$

for a Gaussian wavepacket with a flexible width, or

$$V_1 = \langle\hat{V}'\rangle, \quad V_2 = -A_0\cdot m^{-1}\cdot A_0, \quad \text{and} \quad \mathrm{Re}\,A_0 = 0 \tag{64}$$

for a Gaussian wavepacket with fixed width. As we shall see, these two cases occur, respectively, in the variational thawed and frozen Gaussian wavepacket dynamics.

D. Time dependence of the effective energy

In Sec. III A, we have also seen that the effective energy of a nonlinear TDSE is not always conserved. To find the time dependence of the effective energy (60) for the effective potential (30), we need the time derivative

$$d\hat{V}_{\rm eff}/dt = \dot{V}_0 + \dot{V}_1^T\cdot\hat{x} - V_1^T\cdot\dot{q}_t - \dot{q}_t^T\cdot V_2\cdot\hat{x} + \hat{x}^T\cdot\dot{V}_2\cdot\hat{x}/2. \tag{65}$$

Because the second and fourth terms of Eq. (65) vanish under the expectation value, the substitution of Eq. (65) into Eq. (22) for the time derivative of $E_{\rm eff}$ yields

$$\dot{E}_{\rm eff} = \dot{V}_0 - V_1^T\cdot\dot{q}_t + \mathrm{Tr}\,(\dot{V}_2\cdot\Sigma_t)/2. \tag{66}$$

The effective energy of the Gaussian generally depends on time. Yet, we will see that in many examples of Gaussian wavepacket dynamics, the effective energy is conserved due to the cancellation of various terms in Eq. (66). The effective energy is conserved, e.g., if Eqs. (42)–(45) of motion for q_t, p_t, A_t, and γ_t coincide with Hamilton's equations for the Hamiltonian $E_{\rm eff}$ on a symplectic manifold of Gaussian wavepackets.[31,55]

E. Time reversibility

If we denote by $\Lambda := (q,p,A,\gamma)$ the collection of parameters of the Gaussian ψ, the time evolution of $\psi(t)$ can be expressed in terms of the time evolution $\Lambda_t = \Phi(\Lambda_0,t)$ of the parameters by a flow $\Phi(\Lambda,t)$. Time reversibility (10) of the TGWD, equivalent to the condition

$$\Lambda_0 = \Phi(\Lambda_t,-t), \tag{67}$$

follows from the reversibility (10) of general nonlinear TDSE. In Sec. VIII, we will provide more explicit proof based on condition (67) and the fact that the symmetric composition of reversible flows is reversible.

F. Symplecticity

The family of Gaussian wavepackets (26) parameterized by q_t, p_t, A_t, and γ_t forms a finite-dimensional symplectic submanifold of the Hilbert space and is equipped with a certain noncanonical symplectic structure.[31] Ohsawa and Leok showed that symplectic reduction associated with norm conservation leads to a simpler symplectic form

$$\omega = \sum_{j=1}^{D} dq_j\wedge dp_j + (\hbar/4)\sum_{j,k=1}^{D} d(\mathcal{B}^{-1})_{jk}\wedge d\mathcal{A}_{kj} \tag{68}$$

on a manifold with coordinates $\Lambda := (q,p,\mathcal{A}\equiv \mathrm{Re}\,A, \mathcal{B}\equiv \mathrm{Im}\,A)$; this symplectic structure is conserved, e.g., by the variational Gaussian approximation but not by the original thawed Gaussian approximation.[31] If the effective Hamiltonian is defined as the expectation value of the exact or some approximate Hamiltonian, i.e., $H_{\rm eff} := \langle\hat{H}\rangle_\psi$ or $H_{\rm eff} := \langle\hat{H}_{\rm appr}\rangle_\psi$, and Eqs. (42)–(45) of motion are replaced with noncanonical Hamilton's equations for $H_{\rm eff}$, then the conservation of both symplectic structure and effective energy are guaranteed automatically.[31] This is because the function $H_{\rm eff}(q,p,\mathcal{A},\mathcal{B})$ provides a Hamiltonian function on a manifold with coordinates Λ and symplectic structure (68) and because every

Hamiltonian flow conserves its energy and symplectic structure. Although most examples in the following Secs. VI and VII satisfy $H_{\text{eff}} = \langle \hat{H}_{\text{appr}} \rangle_\psi$, we do not assume the validity of this relation. Therefore, the more general Eqs. (42)–(45) obtained from the perspective of the nonlinear TDSE differ from the equations of motion obtained from the Hamiltonian approach, and conservation of neither the symplectic form nor the effective energy is guaranteed.

The analysis of the symplectic structure of Gaussian wavepacket dynamics with a general effective potential (30) can be done elegantly using the formalism of symplectic geometry, as was done by Ohsawa and Leok[31] for the effective potential obtained as the expected value of some approximate potential ($V_{\text{eff}} := \langle \hat{V}_{\text{appr}} \rangle$). Because this analysis relies on non-elementary concepts of differential geometry, it will be presented elsewhere.[55]

VI. THAWED GAUSSIAN WAVEPACKET DYNAMICS

The reader may ask whether there exist any interesting effective quadratic Hamiltonians for which the preceding general analysis is useful. Indeed, there are many; five such Hamiltonians are hidden behind the *variational Gaussian approximation*,[22,26] Heller's original *thawed Gaussian approximation*,[15] the *single-Hessian thawed Gaussian approximation*,[35] the *local cubic variational TGWD* (also known as the *extended semiclassical dynamics*[30] or *symplectic semiclassical wavepacket dynamics*[32,33]), and—of course—the *global harmonic approximation*, of which the last one leads to a linear TDSE, while the first four give rise to genuine nonlinear TDSEs. Below, we also propose the *single-quartic variational TGWD*, which improves the accuracy of the local cubic variational TGWD without increasing its cost or sacrificing its geometric properties.

We now list the expansion coefficients V_0, V_1, and V_2 of the effective potential (30) for each of these approximations. Equations of motion for parameters q_t, p_t, A_t, and γ_t are in each case obtained by substituting specific expressions for V_0, V_1, and V_2 into the general Eqs. (42)–(45).

In this section, we will discuss methods employing a "thawed" Gaussian—a Gaussian wavepacket with a flexible width matrix, while in Sec. VII, we shall provide examples of methods using a "frozen" Gaussian—a Gaussian wavepacket with a time-independent width.

A. Variational TGWD

As shown in Refs. 25, 26, and 52 and here in Appendix D, the optimal solution (in the sense of the Dirac–Frenkel variational principle[23–25]) of the TDSE (1) with a Gaussian ansatz (26) is the *variational TGWD*, or variational Gaussian approximation,[22,25,26,52,56] which corresponds to an effective potential (30) with coefficients

$$V_0 = \langle \hat{V} \rangle - \text{Tr}\left[\langle \hat{V}'' \rangle \cdot \Sigma_t\right]/2, \quad V_1 = \langle \hat{V}' \rangle, \quad V_2 = \langle \hat{V}'' \rangle. \tag{69}$$

Inserting V_j from Eq. (69) into general Eqs. (51)–(53) shows that in the variational TGWD the exact and effective potentials have the same expectation values of the potential energy, gradient, and Hessian,

$$\langle \hat{V}_{\text{eff}} \rangle = \langle \hat{V} \rangle, \quad \langle \hat{V}'_{\text{eff}} \rangle = \langle \hat{V}' \rangle, \quad \langle \hat{V}''_{\text{eff}} \rangle = \langle \hat{V}'' \rangle. \tag{70}$$

The first equality results from a beautiful cancellation of two terms in the expression

$$\begin{aligned} \langle \hat{V}_{\text{eff}} \rangle &= \langle \hat{V} \rangle - \text{Tr}\left[\langle \hat{V}'' \rangle \cdot \Sigma_t\right]/2 + \text{Tr}\left[\langle \hat{V}'' \rangle \cdot \Sigma_t\right]/2 \\ &= \langle \hat{V} \rangle \end{aligned} \tag{71}$$

and implies that the effective energy is exact ($E_{\text{eff}} = E$) for variational TGWD, even though the propagation itself may be far from exact. Equations of motion that follow from the effective potential (69) were originally derived (differently) by Coalson and Karplus[26] and are equivalent to those of Theorem 3.2 by Ohsawa and Leok[31] and Theorem 3.11 by Lasser and Lubich.[52] Poirier derived these equations using quantum trajectories.[57]

The variational TGWD is symplectic.[27,31] Because any solution derived from the Dirac–Frenkel variational principle conserves energy (see Appendix C)[25,52] and because the exact and effective energies are equal in the variational TGWD, this approximation conserves the effective energy, too. The conservation of the exact and effective energies by the variational TGWD also follows directly from the general expressions (62) for dE/dt and (66) for dE_{eff}/dt. See Appendix E for this more "pedestrian" proof of $\dot{E}_{\text{eff}} = 0$.

The variational TGWD has been extended from real-time to imaginary-time quantum dynamics in order to describe the equilibrium properties of van der Waals clusters[58] and the time-correlation functions of liquid *para*-hydrogen.[59]

B. Local harmonic TGWD

In his original thawed Gaussian approximation,[1,15] Heller did not invoke the variational principle and avoided the expensive evaluation of expectation values needed in Eq. (69) by making the *local harmonic approximation*, in which the effective potential in Eq. (30) depends on ψ only via q_t and its coefficients,

$$V_0 = V(q_t), \quad V_1 = V'(q_t), \quad V_2 = V''(q_t), \tag{72}$$

are the coefficients of the truncated, second-order Taylor expansion of $V(\hat{q})$ about q_t. Heller's local harmonic TGWD (72) can also be obtained from the variational TGWD (69) if the local harmonic approximation is used to evaluate the expectation values $\langle \hat{V}^{(j)} \rangle$. In Sec. VI E, we prove this statement for an arbitrary "local quadratic" approximation for V.

If we substitute expressions for V_0 and V_1 from Eq. (72) into the general equation (51), we find that

$$\langle \hat{V}_{\text{eff}} \rangle = V(q_t) + \text{Tr}\left[V''(q_t) \cdot \Sigma_t\right]/2 \neq \langle \hat{V} \rangle, \tag{73}$$

and, therefore, $E_{\text{eff}} \neq E$. The local harmonic TGWD conserves neither exact nor effective energy. Whereas the nonconservation of the exact energy was proven, in general, in the discussion following Eq. (62) in Sec. V C, the nonconservation of the effective energy follows from Eq. (66) because

$$\begin{aligned} \dot{E}_{\text{eff}} &= V'(q_t)^T \cdot \dot{q}_t - V'(q_t)^T \cdot \dot{q}_t + \text{Tr}\left(B_t \cdot \Sigma_t\right)/2 \\ &= \text{Tr}\left(B_t \cdot \Sigma_t\right)/2, \end{aligned} \tag{74}$$

where B_t is the matrix obtained from the contraction of vector $\dot{q}_t$ with the symmetric rank-3 tensor $V'''(q_t)$. Using Einstein's convention for a sum over repeated indices,

$$(B_t)_{jk} := dV''(q_t)_{jk}/dt = (\dot{q}_t)_i V'''(q_t)_{ijk} = (p_t)_l (m^{-1})_{li} V'''(q_t)_{ijk}. \quad (75)$$

Lauvergnat *et al.*[60] derived equations of motion for the local harmonic TGWD in generalized coordinates.

C. Single-Hessian TGWD

The most expensive part of a higher-dimensional calculation using the local harmonic TGWD is, of course, the evaluation of the Hessian matrix $V''(q_t)$. In the *single-Hessian approximation*,[35,61,62] the Hessian is computed only once, at a reference geometry q_{ref}, but the energies and gradients are still computed at each point q_t along the trajectory. Within the *single-Hessian TGWD*, the effective potential (30) again depends on ψ only via q_t:

$$V_0 = V(q_t), \quad V_1 = V'(q_t), \quad V_2 = V''(q_{\text{ref}}). \quad (76)$$

Although the single-Hessian TGWD does not conserve energy, it is symplectic, and Eq. (66) implies the conservation of effective energy,[35]

$$\dot{E}_{\text{eff}} = V'(q_t)^T \cdot \dot{q}_t - V'(q_t)^T \cdot \dot{q}_t + \frac{1}{2} \text{Tr}\left[\frac{dV''(q_{\text{ref}})}{dt} \cdot \Sigma_t\right] = 0, \quad (77)$$

because $dV''(q_{\text{ref}})/dt = 0$ as q_{ref} is constant. Because of its efficiency and improved geometric properties, the single-Hessian TGWD was implemented in the electronic structure package Turbomole.[21]

D. Global harmonic TGWD

Among all thawed Gaussian approximations, the least expensive but crudest one is the *global harmonic TGWD*, in which the effective potential is the second-order Taylor expansion of V about a fixed reference geometry q_{ref},

$$V_{\text{eff}}(q) = V(q_{\text{ref}}) + V'(q_{\text{ref}})^T \cdot (q - q_{\text{ref}}) + (q - q_{\text{ref}})^T \cdot V''(q_{\text{ref}}) \cdot (q - q_{\text{ref}})/2. \quad (78)$$

This equation is not in the standard form (30), which requires an expansion about the current center q_t of the wavepacket; the coefficients V_j of the standard form (30) are obtained by evaluating derivatives $V_{\text{eff}}^{(j)}$ at q_t:

$$\begin{aligned} V_0 &= V_{\text{eff}}(q_t), \\ V_1 = V'_{\text{eff}}(q_t) &= V'(q_{\text{ref}}) + V''(q_{\text{ref}}) \cdot (q_t - q_{\text{ref}}), \\ V_2 &= V''_{\text{eff}}(q_t) = V''(q_{\text{ref}}). \end{aligned} \quad (79)$$

Because the coefficients V_0 and V_1 depend on ψ via q_t, one might think that V_{eff} is a nonlinear operator. Yet, in contrast to the previously mentioned approximations, in the global harmonic TGWD, V_{eff} is a *linear* operator; this follows clearly from Eq. (78), where V_{eff} depends on ψ via neither q_t nor any other parameter of the Gaussian.

Although the global harmonic TGWD does not conserve energy, it obviously conserves both the symplectic structure and effective energy E_{eff} because $\hat{H}_{\text{eff}} = \hat{T} + \hat{V}_{\text{eff}}$ is a linear, time-independent Hamiltonian operator. An alternative proof follows from Eq. (66):

$$\dot{E}_{\text{eff}} = [V'(q_{\text{ref}}) + V''(q_{\text{ref}}) \cdot (q_t - q_{\text{ref}})]^T \cdot \dot{q}_t - [V'(q_{\text{ref}}) + V''(q_{\text{ref}}) \cdot (q_t - q_{\text{ref}})]^T \cdot \dot{q}_t = 0. \quad (80)$$

E. Variational Gaussian approximation applied to any local quadratic approximation for V

Effective potentials used in the local harmonic, single-Hessian, and global harmonic TGWD are all quadratic functions of nuclear coordinates and, as a result, can be obtained either directly (as suggested above) or by an alternative procedure consisting of two steps: First, approximate the exact potential V with a state-dependent approximation V_{appr}. Then, apply the variational TGWD to V_{appr} instead of V. To see this, note that if

$$V_{\text{appr}}(x) = v_0 + v_1^T \cdot x + x^T \cdot v_2 \cdot x/2 \quad (81)$$

is a quadratic polynomial of x, inserting V_{appr} instead of V into the variational expressions (69) for V_j yields

$$V_0 = \langle \hat{V}_{\text{appr}} \rangle - \text{Tr}\left[\langle \hat{V}''_{\text{appr}} \rangle \cdot \Sigma_t\right]/2 = v_0 + \text{Tr}\,(v_2 \cdot \Sigma_t)/2 - \text{Tr}\,(v_2 \cdot \Sigma_t)/2 = v_0, \quad (82)$$

$$V_1 = \langle \hat{V}'_{\text{appr}} \rangle = v_1, \quad (83)$$

$$V_2 = \langle \hat{V}''_{\text{appr}} \rangle = v_2, \quad (84)$$

where the expected values were evaluated by applying Eqs. (51)–(53) to V_{appr} instead of V_{eff}. The effective potential is equal to the original approximate potential,

$$V_{\text{eff}}(x) = V_{\text{appr}}(x). \quad (85)$$

The expectation values of the effective potential, its gradient, and its Hessian are, therefore, equal to the corresponding properties of the approximate potential:

$$\langle \hat{V}_{\text{eff}} \rangle = \langle \hat{V}_{\text{appr}} \rangle = v_0 + \text{Tr}\,(v_2 \cdot \Sigma_t)/2, \quad (86)$$

$$\langle \hat{V}'_{\text{eff}} \rangle = \langle \hat{V}'_{\text{appr}} \rangle = v_1, \quad (87)$$

$$\langle \hat{V}''_{\text{eff}} \rangle = \langle \hat{V}''_{\text{appr}} \rangle = v_2. \quad (88)$$

The variational principle is not needed if the effective potential is *defined* by Eq. (85). The equality of the effective and exact energies requires that $v_0 = \langle \hat{V} \rangle - \text{Tr}\,(v_2 \cdot \Sigma_t)/2$. This equality holds in the variational TGWD, where $v_2 = \langle \hat{V}'' \rangle$.

If v_0 only depends on q_t, Eq. (66) implies that

$$\dot{E}_{\text{eff}} = [v'_0(q_t) - v_1]^T \cdot \dot{q}_t + \text{Tr}\,(\dot{v}_2 \cdot \Sigma_t)/2.$$

This shows that if the variational principle is applied to an approximate instead of an exact Hamiltonian, neither the exact nor effective energy is conserved, in general. The case of local harmonic TGWD demonstrates that the symplectic structure is generally not conserved either.

F. Local cubic variational TGWD

Compared with Heller's original TGWD, the variational TGWD is hard to evaluate in practice because expectation values $\langle \hat{V}^{(j)} \rangle$ typically cannot be obtained analytically. However, as we have just shown, to go beyond Heller's method, the variational TGWD must be combined with a more accurate approximation for V than the local harmonic one. Then the effective quadratic potential V_{eff} will, obviously, differ from the approximate potential V_{appr}. An obvious and simplest possible choice for V_{appr} that still permits evaluating the expectation values analytically is the *local cubic approximation,*

$$\begin{aligned} V_{\text{appr}}(q) := V(q_t) + V'(q_t)^T \cdot x + x^T \cdot V''(q_t) \cdot x/2 \\ + V'''(q_t)_{ijk}\, x_i x_j x_k/3!, \end{aligned} \tag{89}$$

which was also used for evaluating the matrix elements of the potential in the variational multi-configurational Gaussian method.[63] Expectation values for V_{appr} are

$$\langle \hat{V}''_{\text{appr}} \rangle = V''(q_t), \tag{90}$$

$$\langle \hat{V}'_{\text{appr}} \rangle_i = V'(q_t)_i + V'''(q_t)_{ijk}\, \Sigma_{t,jk}/2, \tag{91}$$

$$\langle \hat{V}_{\text{appr}} \rangle = V(q_t) + \text{Tr}\left[V''(q_t) \cdot \Sigma_t\right]/2. \tag{92}$$

Substitution of these into the variational expressions (69) for V_j yields the effective potential coefficients,

$$V_0 = V(q_t), \quad V_{1,i} = V'(q_t)_i + V'''(q_t)_{ijk}\, \Sigma_{t,jk}/2, \quad V_2 = V''(q_t) \tag{93}$$

of the *local cubic variational TGWD*. The equations of motion obtained from the effective potential (93) are equivalent to those of "symplectic semiclassical wavepacket dynamics" [Eq. (36) in Ref. 31, Eq. (19) in Ref. 32, and Proposition 4.4 in Ref. 33] and, indirectly, to those of "extended semiclassical dynamics" of Pattanayak and Schieve.[30] The method is symplectic,[31] and only the equation for $\dot{p}_t$ differs from the local harmonic TGWD due to a nonclassical term in the force, given by $-V_1$.

Equation (66) implies that the local cubic variational TGWD conserves the effective energy because

$$\begin{aligned} \dot{E}_{\text{eff}} &= \dot{V}_0 - V_1^T \cdot \dot{q}_t + \text{Tr}\left(\dot{V}_2 \cdot \Sigma_t\right)/2 \\ &= \dot{V}(q_t) - V'(q_t)^T \cdot \dot{q}_t - V'''(q_t)_{ijk}\, \Sigma_{t,jk} \dot{q}_{t,i}/2 \\ &\quad + \text{Tr}\left\{\left[dV''(q_t)/dt\right] \cdot \Sigma_t\right\}/2 = 0, \end{aligned} \tag{94}$$

where, in the last step, the second and fourth terms cancel the first and third terms [see Eq. (75)].

G. Single-quartic variational TGWD (see Fig. 1)

To increase the accuracy over the local cubic approximation, the "obvious" logical step is to include the local fourth derivative of V. However, evaluating this *local quartic approximation* is expensive, and, similar to the local harmonic TGWD, the *local quartic variational TGWD* conserves neither the effective energy nor the symplectic structure. Instead, in analogy to the single-Hessian TGWD, a much more effective approach is to apply the variational TGWD to the *single-quartic approximation,*

$$\begin{aligned} V_{\text{appr}}(q) := V(q_t) + V'(q_t)^T \cdot x + x^T \cdot V''(q_t) \cdot x/2 \\ + V'''(q_t)_{ijk}\, x_i x_j x_k/3! \\ + V^{(4)}(q_{\text{ref}})_{ijkl}\, x_i x_j x_k x_l/4!, \end{aligned} \tag{95}$$

which augments the local cubic approximation (89) with the evaluation of a single fourth derivative at a reference geometry q_{ref}. The coefficients of the effective potential of this *single-quartic variational TGWD* are obtained by applying the variational formulas (69) to the expectation values of the potential, gradient, and Hessian of V_{appr}:

$$V_{2,ij} = \langle \hat{V}''_{\text{appr}} \rangle_{ij} = V''(q_t)_{ij} + V^{(4)}(q_{\text{ref}})_{ijkl}\, \Sigma_{t,kl}/2, \tag{96}$$

$$V_{1,i} = \langle \hat{V}'_{\text{appr}} \rangle_i = V'(q_t)_i + V'''(q_t)_{ijk}\, \Sigma_{t,jk}/2, \tag{97}$$

$$\begin{aligned} \langle \hat{V}_{\text{appr}} \rangle = V(q_t) + \text{Tr}\left[V''(q_t) \cdot \Sigma_t\right]/2 \\ + V^{(4)}(q_{\text{ref}})_{ijkl}\, \Sigma_{t,ij} \Sigma_{t,kl}/8, \end{aligned} \tag{98}$$

$$\begin{aligned} V_0 &= \langle \hat{V}_{\text{appr}} \rangle - \text{Tr}\left[\langle \hat{V}''_{\text{appr}} \rangle \cdot \Sigma_t\right]/2 \\ &= V(q_t) - V^{(4)}(q_{\text{ref}})_{ijkl}\, \Sigma_{t,ij} \Sigma_{t,kl}/8. \end{aligned} \tag{99}$$

Here, in deriving expressions for $\langle \hat{V}'_{\text{appr}} \rangle$ and $\langle \hat{V}_{\text{appr}} \rangle$, we used the identities $\langle x_i x_j x_k \rangle = 0$ and

$$\langle x_i x_j x_k x_l \rangle = \Sigma_{t,ij} \Sigma_{t,kl} + \Sigma_{t,ik} \Sigma_{t,jl} + \Sigma_{t,il} \Sigma_{t,jk}; \tag{100}$$

since $V^{(4)}(q_{\text{ref}})_{ijkl}$ is a totally symmetric tensor,

$$V^{(4)}(q_{\text{ref}})_{ijkl} \langle x_i x_j x_k x_l \rangle = 3 V^{(4)}(q_{\text{ref}})_{ijkl}\, \Sigma_{t,ij} \Sigma_{t,kl}.$$

In the single-quartic variational TGWD, equations for $\dot{q}_t$ and $\dot{p}_t$ remain the same as in the local cubic variational TGWD, but equations for $\dot{A}_t$ and $\dot{\gamma}_t$ change due to the changes in V_0 and V_2. Because only a single fourth derivative is needed, the computational cost is only slightly higher than the cost of the local cubic variational TGWD. Assuming that the fourth derivative must be evaluated by finite difference, the increase in cost is negligible in typical simulations, where the number of time steps is much larger than the number of degrees of freedom.

Remarkably, Eq. (66) implies that the single-quartic variational TGWD conserves the effective energy because

$$\begin{aligned}\dot{E}_{\text{eff}} &= \dot{V}_0 - V_1^T \cdot \dot{q}_t + \text{Tr}\left(\dot{V}_2 \cdot \Sigma_t\right)/2 \\ &= \dot{V}(q_t) - V^{(4)}(q_{\text{ref}})_{ijkl}\dot{\Sigma}_{t,ij}\Sigma_{t,kl}/4 \\ &\quad - V'(q_t)^T \cdot \dot{q}_t - V'''(q_t)_{ijk}\Sigma_{t,jk}\dot{q}_{t,i}/2 \\ &\quad + \frac{1}{2}\,\text{Tr}\left[\frac{dV''(q_t)}{dt} \cdot \Sigma_t\right] + \frac{1}{4}V^{(4)}(q_{\text{ref}})_{ijkl}\dot{\Sigma}_{t,ij}\Sigma_{t,kl} \\ &= 0, \end{aligned} \tag{101}$$

where, in the last step, the first term cancels the third one, the second term cancels the sixth, and the fourth term cancels the fifth [see Eq. (75)]. Because the single-quartic variational TGWD does not increase the cost over the local cubic variational TGWD and because it, in contrast to the local quartic TGWD, conserves the symplectic structure and effective energy, this method appears to be the logical and promising choice in calculations whose goal is to improve geometric properties and accuracy beyond Heller's original local harmonic TGWD.

VII. FROZEN GAUSSIAN WAVEPACKET DYNAMICS

In more restrictive approximations, the Gaussian wavepacket has a constant width matrix $A_t = A_0$. Clearly, all such "frozen Gaussian" approximations[6] require that $\dot{A}_t = 0$ and Eq. (44) implies that the coefficient V_2 in the effective quadratic potential (30) must satisfy

$$V_2 = -A_t \cdot m^{-1} \cdot A_t = -A_0 \cdot m^{-1} \cdot A_0. \tag{102}$$

One can freely choose only the coefficients V_0 and V_1, but not V_2. The frozen Gaussian approximation is typically used in multi-trajectory methods, which can describe wavepacket spreading and distortion without requiring a flexible width; the description of the nonlinear spreading of a wavepacket was, indeed, Heller's motivation in proposing the frozen Gaussian approximation.[6] Other methods employing frozen Gaussians are the semiclassical Herman–Kluk propagator[3,7,52,56] and its extensions,[64–66] which assign each trajectory a weight factor depending on the stability matrix; imaginary-time frozen Gaussian dynamics,[67] which treats the quantum Boltzmann operator instead of the time evolution operator; and Gaussian basis methods,[68–72] which allow coupling between the trajectories.

Here, we focus only on *single-trajectory* methods employing a Gaussian with a fixed width. These methods form a family of *frozen Gaussian wavepacket dynamics* (FGWD), which has an interesting relation to the single-Hessian approximation (76) of Sec. VI C. The FGWD is on one hand, a special case of the single-Hessian TGWD because one can think of V_2 from Eq. (102) as a reference Hessian of a harmonic potential whose ground state is the initial state and, on the other hand, a generalization because one is still free to choose V_0 and V_1.

Let

$$\mathcal{A} := \text{Re}\, A_0 = \text{Re}\, A_t, \tag{103}$$

$$\mathcal{B} := \text{Im}\, A_0 = \text{Im}\, A_t \tag{104}$$

denote the real and imaginary parts of the constant width matrix. It follows from Eq. (102) that

$$\text{Re}\, V_2 = -\mathcal{A} \cdot m^{-1} \cdot \mathcal{A} + \mathcal{B} \cdot m^{-1} \cdot \mathcal{B}, \tag{105}$$

$$\text{Im}\, V_2 = -\mathcal{A} \cdot m^{-1} \cdot \mathcal{B} - \mathcal{B} \cdot m^{-1} \cdot \mathcal{A}. \tag{106}$$

Recall that in order for $\hat{V}_{\text{eff}}(\psi)$ to be a Hermitian operator with real expectation values, coefficient $V_2(\psi)$ must be a real matrix. In Appendix F, it is shown that setting $\text{Im}\, V_2 = 0$ in Eq. (106) implies that the initial width matrix A_0 of the frozen Gaussian is purely imaginary,

$$\mathcal{A} = 0. \tag{107}$$

As a consequence,

$$V_2 = \text{Re}\, V_2 = \mathcal{B} \cdot m^{-1} \cdot \mathcal{B}. \tag{108}$$

Because $\mathcal{A} = 0$ in the FGWD, $\text{Cov}_R(\hat{q}, \hat{p}) = 0$ and the equations of motion (42)–(45) become

$$\dot{q}_t = m^{-1} \cdot p_t, \tag{109}$$

$$\dot{p}_t = -V_1, \tag{110}$$

$$A_t = A_0 = \text{const.} = i\mathcal{B}, \tag{111}$$

$$\dot{\gamma}_t = T(p_t) - V_0 - (\hbar/2)\,\text{Tr}\left(m^{-1} \cdot \mathcal{B}\right). \tag{112}$$

Equations (62) and (66) for the time dependence of exact and effective energies reduce to

$$\dot{E} = p_t^T \cdot m^{-1} \cdot (\langle \hat{V}' \rangle - V_1), \tag{113}$$

$$\begin{aligned}\dot{E}_{\text{eff}} &= \dot{V}_0 - V_1^T \cdot \dot{q}_t = \dot{V}_0 + \dot{p}_t^T \cdot m^{-1} \cdot p_t \\ &= \dot{V}_0 + \dot{T}(p_t). \end{aligned} \tag{114}$$

Now, let us examine examples of effective potentials giving rise to FGWD; V_2 is always given by Eq. (108).

A. Variational FGWD

Application of the Dirac–Frenkel variational principle to the frozen Gaussian ansatz is similar to the derivation of the variational TGWD in Appendix D except that we lose parameter A_t and, therefore, Eqs. (D8) and (D13). The remaining equations are either $0 = \langle g \rangle = \langle xg \rangle$ or the equivalent $0 = \langle g \rangle = \langle \nabla g \rangle$, which give

$$\begin{aligned} V_0 &= \langle \hat{V} \rangle - \text{Tr}\left[V_2 \cdot \text{Cov}(\hat{q})\right]/2 \\ &= \langle \hat{V} \rangle + (i\hbar/4)\,\text{Tr}\left(A_0 \cdot m^{-1} \cdot A_0 \cdot A_0^{-1}\right) \\ &= \langle \hat{V} \rangle - (\hbar/4)\,\text{Tr}\left(m^{-1} \cdot \mathcal{B}\right), \\ V_1 &= \langle \hat{V}' \rangle. \end{aligned} \tag{115}$$

Equation (45) for $\dot{\gamma}_t$ becomes

$$\begin{aligned}\dot{\gamma}_t &= T(p_t) - \langle \hat{V} \rangle - (\hbar/4)\,\text{Tr}\left(m^{-1} \cdot \mathcal{B}\right) \\ &= T(p_t) - \langle \hat{V} \rangle - \text{Tr}\left[m^{-1} \cdot \text{Cov}(\hat{p})\right]/2 \\ &= 2T(p_t) - \langle \hat{V} \rangle - \langle \hat{T} \rangle = 2T(p_t) - \langle \hat{H} \rangle. \end{aligned} \tag{116}$$

Since $V_1 = \langle \hat{V}' \rangle$ and $E_{\text{eff}} = E$ in the variational FGWD, Eq. (113) implies that the exact and effective energies are conserved: $\dot{E} = \dot{E}_{\text{eff}} = 0$.

B. Variational FGWD with classical trajectories

In Heller's original frozen Gaussian approximation,[6] the Gaussians move along classical trajectories. Such an approximation can be obtained by applying the Dirac–Frenkel variational principle to the Gaussian ansatz, in which q_t and p_t are required to follow classical Hamilton's equations of motion and γ_t is the only variationally optimized parameter. (Strictly speaking, this is a generalization of the Dirac–Frenkel variational principle because the approximation manifold is time-dependent.) The derivation is analogous to that of the variational TGWD except that now we only have Eq. (D9) related to the parameter γ_t: $0 = \langle g \rangle$. The resulting method corresponds to an effective potential with coefficients

$$V_0 = \langle \hat{V} \rangle - (\hbar/4)\,\mathrm{Tr}\,(m^{-1} \cdot \mathcal{B}) \quad \text{and} \quad V_1 = V'(q_t). \tag{117}$$

As in the previous case, the exact and effective energies are equal, $E = E_{\mathrm{eff}}$; however, neither energy is conserved, and their common time dependence is

$$\dot{E} = \dot{E}_{\mathrm{eff}} = p_t^T \cdot m^{-1} \cdot [\langle \hat{V}' \rangle - V'(q_t)]. \tag{118}$$

C. Variational frozen Gaussian approximation applied to any local quadratic approximation for V

In analogy to the TGWD, one may apply the variational frozen Gaussian approximation (115) to an approximate, state-dependent potential V_{appr} from Eq. (81) instead of V. The only difference is that for the frozen Gaussian approximation, the coefficient v_2 of V_{appr} is fixed and always equal to V_2 from Eq. (108). Derivation from Sec. VI E combined with the variational approximation (115) yields an effective potential (30) with

$$V_0 = v_0 \quad \text{and} \quad V_1 = v_1. \tag{119}$$

Neither the exact nor the effective energy is conserved in general; their time dependencies are

$$\dot{E} = p_t^T \cdot m^{-1} \cdot (\langle \hat{V}' \rangle - v_1) \tag{120}$$

and, if the coefficients v_0 and v_1 only depend on q_t,

$$\dot{E}_{\mathrm{eff}} = \dot{v}_0 - v_1^T \cdot \dot{q}_t = \left[v_0'(q_t) - v_1(q_t)\right]^T \cdot \dot{q}_t. \tag{121}$$

D. Local harmonic FGWD = single-Hessian FGWD

Since the coefficient v_2 of V_{appr} is always equal to V_2 from Eq. (108), the local harmonic and single-Hessian approximations applied to the frozen Gaussian are equivalent; the effective potential coefficients from Eq. (119) are

$$V_0 = v_0 = V(q_t) \quad \text{and} \quad V_1 = v_1 = V'(q_t). \tag{122}$$

The same method is obtained regardless of whether the FGWD assumes classical or variational trajectories. Whereas the exact energy generally depends on time,

$$\dot{E} = p_t^T \cdot m^{-1} \cdot [\langle \hat{V}' \rangle - V'(q_t)], \tag{123}$$

Eq. (121) implies that the effective energy is conserved,

$$\dot{E}_{\mathrm{eff}} = [V'(q_t) - V'(q_t)]^T \cdot \dot{q}_t = 0. \tag{124}$$

E. Global harmonic FGWD

Likewise, combining the frozen Gaussian ansatz with the global harmonic approximation yields an approximation equivalent to V_{eff} with coefficients

$$V_0 = v_0 = V_{\mathrm{appr}}(q_t) \quad \text{and} \quad V_1 = v_1 = V'_{\mathrm{appr}}(q_t). \tag{125}$$

This result follows because the Hessian of the global harmonic approximation is again constrained by Eq. (108) and $V''_{\mathrm{appr}}(q_t) = v_2 = V_2$. Whereas the exact energy, in general, depends on time,

$$\dot{E} = p_t^T \cdot m^{-1} \cdot [\langle \hat{V}' \rangle - V'_{\mathrm{appr}}(q_t)], \tag{126}$$

Eq. (121) implies conservation of the effective energy,

$$\dot{E}_{\mathrm{eff}} = [V'_{\mathrm{appr}}(q_t) - V'_{\mathrm{appr}}(q_t)]^T \cdot \dot{q}_t = 0. \tag{127}$$

VIII. GEOMETRIC INTEGRATORS

As mentioned in Sec. V, exact solutions of the nonlinear TDSE (20) have certain geometric properties, such as norm conservation and time reversibility. For some effective potentials, the exact solutions also conserve the symplectic structure, energy, or effective energy. The numerical solution of the nonlinear equation, however, requires further approximations, including time discretization. "Geometric integrators" are numerical algorithms[28,73] that preserve some or all of the geometric properties of the exact solution, regardless of the discretization time step.

Let us describe geometric integrators for the nonlinear Schrödinger equation (20) with a general quadratic effective potential (30). Because the effective Hamiltonian (19) is separable into a kinetic energy term depending only on $\hat{p}$ and a potential term depending only on $\hat{q}$, we can, under a rather weak additional assumption on V_{eff} (see Sec. VIII B), employ the explicit splitting method. In this method, equations of motion are solved analytically for both the kinetic and potential propagation steps, in which $\hat{H}_{\mathrm{eff}} := T(\hat{p})$ and $\hat{H}_{\mathrm{eff}} := V_{\mathrm{eff}}(\hat{q}; \psi)$, respectively.

By composing exactly solved kinetic and potential propagations with the same time step Δt, one obtains—depending on the ordering of composition—either the "VT" or "TV" algorithm, which approximates the evolution driven by $\hat{H}_{\mathrm{eff}} = T(\hat{p}) + V_{\mathrm{eff}}(\hat{q}; \psi)$ with first-order accuracy in the time step Δt. Composing, in turn, the VT with TV algorithm, both with the time step $\Delta t/2$, yields, depending on the order of composition, either the "VTV" or "TVT" second-order algorithm, which are analogs of the Verlet algorithm[74] for classical molecular dynamics and of the split-operator algorithm[75] for quantum dynamics. They also generalize Faou and Lubich's algorithm[27] for the variational TGWD to the TGWD with a general effective potential (30). Both VTV and TVT algorithms are symmetric and, therefore, time-reversible.

Explicit geometric integrators of arbitrary even orders in Δt are then obtained by applying recursive[76,77] or nonrecursive[78,79] symmetric composition schemes to the second-order algorithms.[28,73]

For details on how high-order integrators are generated from the elementary algorithms for kinetic and potential propagations, see, e.g., Refs. 46, 80, and 81. Because this procedure is general, one can use for the splitting and composition the same algorithm[46,80,81] as for any other classical or quantum dynamics; below, we only need to describe exact solutions for the kinetic and potential propagations.

The algorithm presented below suggests how to write a single computer program that can evaluate all methods from Secs. VI and VII, regardless of whether they were derived using the variational principle, the Hamiltonian approach, or, more generally, the perspective of the nonlinear TDSE. One simply invokes the potential propagation with different coefficients, V_0, V_1, and V_2.

A. Kinetic propagation

Equations of motion for the kinetic propagation are obtained by considering only the kinetic term in the effective Hamiltonian, i.e., by setting $\hat{H}_{\text{eff}} = T(\hat{p})$, which results in a problem equivalent to solving the propagation of a free-particle Gaussian wavepacket. Setting $V_0 = V_1 = V_2 = 0$ in Eqs. (42)–(45) yields the system

$$\dot{q}_t = m^{-1} \cdot p_t, \tag{128}$$

$$\dot{p}_t = 0, \tag{129}$$

$$\dot{A}_t = -A_t \cdot m^{-1} \cdot A_t, \tag{130}$$

$$\dot{\gamma}_t = T(p_t) + (i\hbar/2)\,\mathrm{Tr}\left(m^{-1} \cdot A_t\right) \tag{131}$$

of ordinary differential equations for parameters $\Lambda_t := (q_t, p_t, A_t, \gamma_t)$ whose exact solution is the flow $\Lambda_t = \Phi_T(\Lambda_0, t)$ given explicitly by

$$q_t = q_0 + tm^{-1} \cdot p_0, \tag{132}$$

$$p_t = p_0, \tag{133}$$

$$\begin{aligned} A_t &= \left(A_0^{-1} + tm^{-1}\right)^{-1} = A_0 \cdot \left(\mathrm{Id}_D + tm^{-1} \cdot A_0\right)^{-1} \\ &= \left(\mathrm{Id}_D + tA_0 \cdot m^{-1}\right)^{-1} \cdot A_0, \end{aligned} \tag{134}$$

$$\gamma_t = \gamma_0 + tT(p_0) + (i\hbar/2)\ln\det\left(\mathrm{Id}_D + tm^{-1} \cdot A_0\right). \tag{135}$$

In the last equation, the continuity of γ_t and, therefore, continuity of $\psi(t)$, is guaranteed for sufficiently small time steps t if one takes the principal branch of the logarithm—the branch on which the imaginary part of the logarithm lies in the interval $(-\pi, \pi)$.[21] The first of the three alternatives for evaluating A_t behaves better numerically despite requiring two instead of one matrix inverse at each step (m^{-1} can be precomputed). In the FGWD, Eq. (130) is replaced with $\dot{A}_t = 0$, Eq. (134) with $A_t = A_0$, and Eq. (135) with

$$\gamma_t = \gamma_0 + t[T(p_0) - (\hbar/2)\,\mathrm{Tr}\left(m^{-1} \cdot \mathcal{B}\right)]. \tag{136}$$

Because $T(\hat{p})$ is a special case of a linear Hamiltonian operator $\hat{H}$, kinetic propagation conserves the scalar product, norm, and symplectic structure. Kinetic propagation is also time-reversible because $\Lambda_0 = \Phi_T(\Lambda_t, -t)$, which follows by inverting Eqs. (132)–(135) explicitly:

$$q_0 = q_t - tm^{-1} \cdot p_t, \tag{137}$$

$$p_0 = p_t, \tag{138}$$

$$A_0 = \left(A_t^{-1} - tm^{-1}\right)^{-1}, \tag{139}$$

$$\gamma_0 = \gamma_t - tT(p_t) + (i\hbar/2)\ln\det\left(\mathrm{Id}_D - tm^{-1} \cdot A_t\right). \tag{140}$$

In FGWD, Eqs. (139) and (140) become $A_0 = A_t$ and

$$\gamma_0 = \gamma_t + t[T(p_t) - (\hbar/2)\,\mathrm{Tr}\left(m^{-1} \cdot \mathcal{B}\right)]. \tag{141}$$

Equations (132)–(135) and (140) for the forward and backward propagations are derived in Appendix G.

B. Potential propagation

Equations of motion for the potential propagation are obtained by considering only the potential energy term in the effective Hamiltonian, i.e., $\hat{H}_{\text{eff}}(\psi) = V_{\text{eff}}(\hat{q}; \psi)$. Setting $T(p) = 0$ is equivalent to taking the limit $m \to \infty$ (or $m^{-1} \to 0$) in Eqs. (42)–(45), which yields the system

$$\dot{q}_t = 0, \tag{142}$$

$$\dot{p}_t = -V_1, \tag{143}$$

$$\dot{A}_t = -V_2, \tag{144}$$

$$\dot{\gamma}_t = -V_0. \tag{145}$$

This system can be solved analytically if the coefficients V_0, V_1, and V_2 depend on the state $\psi(t)$ only via q_t and $\mathrm{Im}\,A_t$ but are independent of p_t, $\mathrm{Re}\,A_t$, and γ_t. This assumption, which holds for all approximations from Secs. VI and VII, results in the trivial solution

$$q_t = q_0, \tag{146}$$

$$p_t = p_0 - tV_1(q_0, \mathrm{Im}\,A_0), \tag{147}$$

$$A_t = A_0 - tV_2(q_0, \mathrm{Im}\,A_0), \tag{148}$$

$$\gamma_t = \gamma_0 - tV_0(q_0, \mathrm{Im}\,A_0). \tag{149}$$

In FGWD, Eqs. (144) and (148) are replaced with $\dot{A}_t = 0$ and $A_t = A_0$.

Equation (146) follows immediately from Eq. (142). Because V_2 is real, Eq. (144) implies that $d\mathrm{Im}\,A_t/dt = 0$ and, therefore, $\mathrm{Im}\,A_t = \mathrm{Im}\,A_0$. As a consequence, if the coefficients $V_j \equiv V_j(q_t, \mathrm{Im}\,A_t)$ depend only on q_t and $\mathrm{Im}\,A_t$, then each V_j remains unchanged during the potential propagation, and Eqs. (143)–(145) can be solved separately to yield Eqs. (147)–(149), respectively.

The assumption is obviously satisfied in the global harmonic, local harmonic, and single-Hessian TGWD, for which the coefficients of V_{eff} depend on $\psi(t)$ only via q_t. In the variational methods, coefficients of the effective potential depend on the expectation values

$$\langle \hat{V}^{(j)} \rangle = \int V^{(j)}(q_t + x)\rho(x)d^D x, \tag{150}$$

which depend on both q_t and $\mathrm{Im}\,A_t$ but on no other parameters of $\psi(t)$ because the density $\rho(x) = |\psi(x)|^2$ depends on $\psi(t)$ only via $\mathrm{Im}\,A_t$ [see Eq. (B16)].

Potential propagation conserves the norm of $\psi(t)$ because $V_{\mathrm{eff}}(\hat{q};\psi)$ is a special case of $\hat{H}_{\mathrm{eff}}$. Potential propagation is also time-reversible because $\Lambda_0 = \Phi_{V_{\mathrm{eff}}}(\Lambda_t, -t)$, which, in turn, follows by inverting Eqs. (146)–(149):

$$q_0 = q_t, \tag{151}$$

$$p_0 = p_t + tV_1(q_t, \mathrm{Im}\,A_t), \tag{152}$$

$$A_0 = A_t + tV_2(q_t, \mathrm{Im}\,A_t), \tag{153}$$

$$\gamma_0 = \gamma_t + tV_0(q_t, \mathrm{Im}\,A_t). \tag{154}$$

In FGWD, Eq. (153) is replaced with $A_0 = A_t$.

In general, in particular, propagation with V_{eff} does not conserve symplectic structure; in particular, potential propagation in neither the local harmonic nor the local quartic variational TGWD is symplectic.[31] In contrast, potential propagations in all other presented examples of the TGWD, i.e., the variational, single-Hessian, global harmonic, local cubic variational, and single-quartic variational TGWD are symplectic.

C. Geometric properties of integrators

If the exact solution of the nonlinear TDSE has a certain geometric property for any Hamiltonian $\hat{H}_{\mathrm{eff}} = T(\hat{p}) + V_{\mathrm{eff}}(\hat{q};\psi)$, then both the kinetic and potential steps share the same property because they can be thought of as exact solutions of nonlinear TDSEs with effective Hamiltonians $\hat{H}_{\mathrm{eff}} = T(\hat{p})$ and $\hat{H}_{\mathrm{eff}}(\psi) = V_{\mathrm{eff}}(\hat{q};\psi)$. This implication justifies the norm conservation by kinetic and potential propagations. Although the implication is also true for the exact and effective energies, it is "useless" because the definitions of these energies are different for the three effective Hamiltonians: $\hat{T}$, $\hat{V}_{\mathrm{eff}}$, and $\hat{T} + \hat{V}_{\mathrm{eff}}$.

If a geometric property preserved by kinetic and potential steps is also preserved under the composition of flows, it is preserved by the TV and VT integrators as well as by their arbitrary compositions. This is again true for the norm and useless for the exact and effective energies. It is also true for symplecticity; as a consequence, if the elementary propagation with V_{eff} is symplectic, then so is an arbitrary integrator based on composing VT and TV steps. Time reversibility requires that the composition of time-reversible maps be symmetric. VT and TV integrators are not reversible, whereas TVT and VTV integrators and their symmetric compositions are reversible.

In summary, all integrators obtained by symmetric compositions of the VTV and TVT algorithms are norm-conserving and time-reversible. They are also symplectic *if* the exact solution of the nonlinear TDSE itself is symplectic. Due to the splitting, however, for a given finite time step Δt, they conserve neither the exact nor the effective energy, even if the exact solution of the nonlinear TDSE does. [This happens for the exact energy, conserved by the variational TGWD and variational FGWD. The numerical integrators do conserve the exact energy but only approximately, with an error $O(\Delta t^M)$, where the order M is greater or equal to the order of the method.] General proofs of these statements can be found in Refs. 28, 45, 73, 80, and 81.

D. Geometric properties of Gaussian wavepacket dynamics

Typically, one first demonstrates a geometric property of an approximation, before analyzing the preservation of this property by a numerical integrator. This is how we have treated norm conservation. However, sometimes it is easier to go "the other way." If an integrator is "consistent," i.e., at least first-order accurate in the time step Δt, and preserves a given geometric property during both kinetic or potential propagations, then considering the limit $\Delta t \to 0$ shows that the exact solution of the nonlinear TDSE has the same geometric property. The application of this idea to the time-reversible second-order TVT algorithm proves the time reversibility of general Gaussian wavepacket dynamics. The application of this idea to symplecticity shows that if the potential propagation step with V_{eff} is symplectic, then so is the Gaussian wavepacket dynamics with this effective potential. Whereas the symplecticity of the variational and local cubic variational TGWD was demonstrated in Refs. 27 and 31, a detailed analysis of the symplecticity of the TGWD with a general effective potential (30) will be presented elsewhere because this analysis relies on non-elementary tools of symplectic geometry.[55]

IX. HAGEDORN PARAMETRIZATION

Hagedorn[82] proposed an alternative parametrization of the Gaussian wavepacket, in which the equations of motion and other properties become simpler. Below, we translate the preceding results from Heller's parametrization (q, p, A, γ) to Hagedorn's parametrization (q, p, Q, P, S).[25,52,83] For derivations, see Appendix B 7.

A. Gaussian wavepacket

As shown in Appendix B 7, in Hagedorn's parametrization, the Gaussian wavepacket (26) can be written as

$$\psi(q,t) = (\pi\hbar)^{-D/4}(\det Q_t)^{-1/2} \times \exp\left[\frac{i}{\hbar}\left(\frac{1}{2}x^T \cdot P_t \cdot Q_t^{-1} \cdot x + p_t^T \cdot x + S_t\right)\right], \tag{155}$$

where q_t and p_t are, as before, real D-vectors of position and momentum, Q_t and P_t are complex $D \times D$ matrices, related to the width matrix A_t, and S_t is a real scalar generalizing classical action and related to Heller's parameter γ_t. Matrices Q_t and P_t have several remarkable properties, listed in Appendix B 7. In particular, the often needed expressions for the position and momentum covariances assume symmetric and decoupled forms

$$\mathrm{Cov}(\hat{q}) = (\hbar/2) Q_t \cdot Q_t^\dagger \quad \text{and} \quad \mathrm{Cov}(\hat{p}) = (\hbar/2) P_t \cdot P_t^\dagger. \tag{156}$$

B. Equations of motion

In Appendix B 7, the nonlinear TDSE (20) is shown to be equivalent to a system of ordinary differential equations for parameters q_t, p_t, Q_t, P_t, and S_t. Whereas Eqs. (42)–(43) for $\dot{q}_t$ and $\dot{p}_t$ remain unchanged, Eqs. (44) and (45) for $\dot{A}_t$ and $\dot{\gamma}_t$ are replaced with three equations,

$$\dot{Q}_t = m^{-1} \cdot P_t, \tag{157}$$

$$\dot{P}_t = -V_2 \cdot Q_t, \tag{158}$$

$$\dot{S}_t = T(p_t) - V_0. \tag{159}$$

A family of high-order geometric integrators for the numerical propagation can be obtained, as in Heller's parametrization, by combining the concepts of splitting into the sequence of kinetic and potential propagations and of the symmetric composition of the symmetric second-order TVT or VTV algorithm. We, therefore, only need to derive expressions for the elementary kinetic and potential propagations, which are generalizations of Faou and Lubich's algorithm[25,27] for the variational TGWD to the generalized thawed Gaussian wavepacket dynamics expressed in Eqs. (42), (43), and (157)–(159).

C. Kinetic propagation

If the Hamiltonian consists only of the kinetic energy, $\hat{H}_{\mathrm{eff}} = T(\hat{p})$, equations of motion (157)–(159) reduce to

$$\dot{Q}_t = m^{-1} \cdot P_t, \tag{160}$$

$$\dot{P}_t = 0, \tag{161}$$

$$\dot{S}_t = T(p_t). \tag{162}$$

Because the momentum p_t is constant during the kinetic step, this system has the analytical solution

$$Q_t = Q_0 + t m^{-1} \cdot P_0, \tag{163}$$

$$P_t = P_0, \tag{164}$$

$$S_t = S_0 + t T(p_0), \tag{165}$$

which is time-reversible [i.e., $\Lambda_0 = \Phi_T(\Lambda_t, -t)$ if $\Lambda := (q, p, Q, P, S)$ and $\Lambda_t = \Phi_T(\Lambda_0, t)$] since the inversion of Eqs. (163)–(165) gives

$$Q_0 = Q_t - t m^{-1} \cdot P_t, \tag{166}$$

$$P_0 = P_t, \tag{167}$$

$$S_0 = S_t - t T(p_t). \tag{168}$$

D. Potential propagation

If $\hat{H}_{\mathrm{eff}}(\psi) = V_{\mathrm{eff}}(\hat{q}; \psi)$, Eqs. (157)–(159) reduce to

$$\dot{Q}_t = 0, \tag{169}$$

$$\dot{P}_t = -V_2 \cdot Q_t, \tag{170}$$

$$\dot{S}_t = -V_0. \tag{171}$$

Since the position q_t remains constant during the potential step, under the assumption that the coefficients V_0, V_1, and V_2 only depend on q_t and Q_t but not on other Hagedorn parameters, this system has the exact solution

$$Q_t = Q_0, \tag{172}$$

$$P_t = P_0 - t V_2(q_0, Q_0) \cdot Q_0, \tag{173}$$

$$S_t = S_0 - t V_0(q_0, Q_0). \tag{174}$$

The assumption holds for all examples from Secs. VI and VII. In the global harmonic, local harmonic, and single-Hessian TGWD, coefficients V_j depend only on q_t; in the variational methods, expected values $\langle \hat{V}^{(j)} \rangle$ and, hence, coefficients V_j depend only on q_t and Q_t because the density $\rho(x)$ depends only on position covariance, which depends only on Q_t [see Eqs. (150), (B16), and (156)]. Potential propagation is time-reversible, i.e., $\Lambda_0 = \Phi_{V_{\mathrm{eff}}}(\Lambda_t, -t)$ if $\Lambda_t = \Phi_{V_{\mathrm{eff}}}(\Lambda_0, t)$, since the inversion of Eqs. (172)–(174) yields

$$Q_0 = Q_t, \tag{175}$$

$$P_0 = P_t + t V_2(q_t, Q_t) \cdot Q_t, \tag{176}$$

$$S_0 = S_t + t V_0(q_t, Q_t). \tag{177}$$

X. DISCUSSION AND CONCLUSION

In conclusion, we have discussed Gaussian wavepacket dynamics from the perspective of a nonlinear Schrödinger equation, which is complementary to the variational[25,27] and symplectic[27,31–33] perspectives. The more general state-dependent quadratic potential

appearing in the nonlinear TDSE describing Gaussian wavepacket dynamics guarantees norm conservation and time reversibility but not always the conservation of energy, effective energy, or symplectic structure. Depending on the choice of the coefficients of this potential, one obtains a large family of both well-known and new Gaussian wavepacket dynamics methods. Among the latter, the single-quartic variational TGWD is promising because it is symplectic, conserves effective energy, and increases accuracy over the local cubic variational TGWD without substantially increasing its cost.

The general form of presentation suggests how all single-trajectory Gaussian wavepacket dynamics methods can be implemented in a single, universal computer code, in which one only needs to modify the three coefficients of the effective potential to obtain any one of the specific methods. Moreover, we have described a single, universal high-order geometric integrator for Gaussian wavepacket dynamics, which generalizes Faou and Lubich's integrator for the variational TGWD.[27]

Many but not all TGWD methods can be obtained by the variational and symplectic approaches. If the variational principle is applied to the exact potential V, the resulting method is symplectic and conserves energy.[27] Neither property is guaranteed if this principle is applied to an approximate potential V_{appr}: local harmonic approximation provides a counterexample. Remarkably, the symplectic approach[31] always conserves both the symplectic structure and effective energy. If applied to the local harmonic approximation, the symplectic method conserves the local harmonic energy, although the resulting equations of motion are equivalent to those obtained by applying the variational principle to the local cubic approximation. Heller's original thawed Gaussian approximation, therefore, cannot be derived by the symplectic approach. As opposed to both variational and symplectic approaches, the more general nonlinear TDSE (20) does not require a specific form of the wavepacket; e.g., if it is applied directly to the local cubic approximation, without invoking the variational principle, an initial Gaussian wavepacket will quickly lose its Gaussian form.

Single-trajectory Gaussian wavepacket dynamics clearly cannot propagate wavefunctions of more general forms, needed in many chemical physics applications. This issue has been addressed partially in the extended thawed Gaussian approximation,[16,36,84,85] which propagates a Gaussian multiplied with a linear polynomial and, thus, can describe electronic spectra beyond the Condon approximation[84,85] or rates of internal conversion.[86] The same issue is fully resolved by Hagedorn wavepackets,[52,82,83] which can propagate arbitrary wavefunctions. Indeed, Ohsawa generalized the symplectic formulation to such non-Gaussian states.[87] The generalization of the present analysis to the dynamics of wavepackets of arbitrary shapes is, therefore, also interesting and in progress. In particular, any of the effective potentials described here will preserve the form not only of Gaussian but also of Hagedorn wavepackets. Yet, even the simple Gaussian wavepacket dynamics discussed here improves substantially electronic spectra calculations over the standard global harmonic approaches, which completely ignore the anharmonicity of the potential energy surface. Among the different methods mentioned here, the variational,[26] local cubic variational,[29,31] and single-quartic variational TGWD can even approximate tunneling and, therefore, also deserve further attention.

ACKNOWLEDGMENTS

The author thanks Roya Moghaddasi Fereidani for producing Fig. 1, Frank Grossmann, Fabian Kröninger, Caroline Lasser, Christian Lubich, Tomoki Ohsawa, and Bill Poirier for discussions, and Tomislav Begušić for suggesting the proof in Appendix F. The author acknowledges the financial support from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (Grant Agreement No. 683069—MOLEQULE) as well as from the COST Action CA21101—Confined Molecular Systems: from a new generation of materials to the stars (COSY) of the European Community.

AUTHOR DECLARATIONS

Conflict of Interest

The author has no conflicts to disclose.

Author Contributions

Jiří Vaníček: Writing – original draft (equal); Writing – review & editing (equal).

DATA AVAILABILITY

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

APPENDIX A: PROPERTIES OF THE NONLINEAR TDSE

1. Nonconservation of the inner product

In Sec. III, the nonconservation (12) of the inner product by the nonlinear TDSE follows from the relation

$$\begin{aligned}\langle\dot\psi|\phi\rangle &= \langle\phi|\dot\psi\rangle^* = -(i\hbar)^{-1}\langle\phi|\hat H_{\text{eff}}(\psi)\psi\rangle^*\\ &= i\hbar^{-1}\langle\hat H_{\text{eff}}(\psi)\psi|\phi\rangle = i\hbar^{-1}\langle\psi|\hat H_{\text{eff}}(\psi)\phi\rangle.\end{aligned}\tag{A1}$$

2. Evolution of the expected value of a nonlinear operator

To prove the relation (13), let us evaluate the time derivative of a nonlinear observable $\langle\hat A(\psi)\rangle$:

$$\begin{aligned}\frac{d}{dt}\langle\hat A(\psi)\rangle &= \frac{d}{dt}\langle\psi|\hat A(\psi)\psi\rangle\\ &= \langle\dot\psi|\hat A(\psi)\psi\rangle + \left\langle\psi\middle|\frac{d\hat A(\psi)}{dt}\psi\right\rangle + \langle\psi|\hat A(\psi)\dot\psi\rangle\\ &= i\hbar^{-1}\langle\psi|\hat H_{\text{eff}}(\psi)\hat A(\psi)\psi\rangle + \langle d\hat A(\psi)/dt\rangle\\ &\quad - i\hbar^{-1}\langle\psi|\hat A(\psi)\hat H_{\text{eff}}(\psi)\psi\rangle\\ &= \langle d\hat A(\psi)/dt\rangle - i\hbar^{-1}\langle[\hat A(\psi),\hat H_{\text{eff}}(\psi)]\rangle.\end{aligned}\tag{A2}$$

In the third step of the derivation, we used the relation

$$\langle\dot{\psi}|\hat{A}(\psi)\psi\rangle = \langle\hat{A}(\psi)\psi|\dot{\psi}\rangle^* = i\hbar^{-1}\langle\hat{A}(\psi)\psi|\hat{H}_{\text{eff}}(\psi)\psi\rangle^* = i\hbar^{-1}\langle\hat{H}_{\text{eff}}(\psi)\psi|\hat{A}(\psi)\psi\rangle = i\hbar^{-1}\langle\psi|\hat{H}_{\text{eff}}(\psi)\hat{A}(\psi)\psi\rangle, \tag{A3}$$

which follows easily from the hermiticity of $\hat{H}_{\text{eff}}(\psi)$ viewed as a linear operator for a fixed state ψ.

3. Time dependence of energy, position, and momentum in a nonlinear TDSE with a separable Hamiltonian

For separable Hamiltonians, the commutator on the right-hand side of Eq. (16) becomes

$$[\hat{H}, \hat{H}_{\text{eff}}(\psi)] = [\hat{V} - \hat{V}_{\text{eff}}(\psi), \hat{T}], \tag{A4}$$

where

$$2[\hat{V},\hat{T}] = \hat{p}^T\cdot m^{-1}\cdot[\hat{V},\hat{p}] + [\hat{V},\hat{p}^T]\cdot m^{-1}\cdot\hat{p} = i\hbar(\hat{p}^T\cdot m^{-1}\cdot\hat{V}' + \hat{V}'^T\cdot m^{-1}\cdot\hat{p}), \tag{A5}$$

and an analogous relation holds for $[\hat{V}_{\text{eff}}(\psi), \hat{T}]$. Substituting Eqs. (A4) and (A5) into the general equation (16) for the time dependence of energy yields Eq. (21).

The Ehrenfest theorem [Eqs. (24) and (25)] for the evolution of q and p follows directly from the commutators

$$2[\hat{q}_j, \hat{H}_{\text{eff}}] = [\hat{q}_j, \hat{p}^T\cdot m^{-1}\cdot\hat{p}] = [\hat{q}_j,\hat{p}^T]\cdot m^{-1}\cdot\hat{p} + \hat{p}^T\cdot m^{-1}\cdot[\hat{q}_j,\hat{p}] = 2i\hbar(m^{-1}\cdot\hat{p})_j, \tag{A6}$$

$$[\hat{p}, \hat{H}_{\text{eff}}] = -i\hbar\hat{V}'_{\text{eff}}. \tag{A7}$$

APPENDIX B: PROPERTIES OF THE GAUSSIAN WAVEPACKET

1. Derivatives of functions of a vector or matrix

We need several matrix relations, which can be found, e.g., in Ref. 88. A derivative of a scalar function $\psi(x)$ of a vector x is defined by

$$[\partial\psi(x)/\partial x]_i := \partial\psi(x)/\partial x_i. \tag{B1}$$

Likewise, a derivative of a scalar function $\psi(A)$ of a general, not necessarily symmetric, matrix A is defined by

$$[\partial\psi(A)/\partial A]_{ij} := \partial\psi(A)/\partial A_{ij}. \tag{B2}$$

With these definitions, if a is a vector, we get

$$\partial(a^T\cdot x)/\partial x = a, \tag{B3}$$

$$\partial(x^T\cdot A\cdot x)/\partial A = x\otimes x^T, \tag{B4}$$

$$\partial(x^T\cdot A\cdot x)/\partial x = (A + A^T)\cdot x. \tag{B5}$$

If $B = B^T$ is a symmetric matrix, Eq. (B5) reduces to

$$\partial(x^T\cdot B\cdot x)/\partial x = 2B\cdot x. \tag{B6}$$

If x and A depend on time t, then

$$d\psi(x)/dt = (\partial\psi/\partial x_i)\dot{x}_i = (\partial\psi/\partial x)^T\cdot\dot{x}, \tag{B7}$$

$$d\psi(A)/dt = (\partial\psi/\partial A_{ij})\dot{A}_{ij} = \text{Tr}\left[(\partial\psi/\partial A)^T\cdot\dot{A}\right]. \tag{B8}$$

The divergence of a vector-valued function $v(A,x) = A\cdot x$ is

$$\nabla^T\cdot v = \partial(A_{ij}x_j)/\partial x_i = A_{ij}\delta_{ji} = \text{Tr}\,A. \tag{B9}$$

2. Derivatives of the Gaussian amplitude

Derivatives of the Gaussian (26) with respect to the four parameters are

$$\frac{\partial\psi}{\partial q_t} = \left(\frac{\partial x}{\partial q_t}\right)^T\cdot\frac{\partial\psi}{\partial x} = -\frac{\partial\psi}{\partial x} = -\frac{i}{\hbar}\xi\psi, \tag{B10}$$

$$\partial\psi/\partial p_t = i\hbar^{-1}x\psi, \tag{B11}$$

$$\partial\psi/\partial A_t = i\hbar^{-1}(1/2)x\otimes x^T\psi, \tag{B12}$$

$$\partial\psi/\partial\gamma_t = i\hbar^{-1}\psi, \tag{B13}$$

where ξ is given in Eq. (34) and relation (B4) was used to obtain Eq. (B12). The first and second derivatives of ψ with respect to the coordinate q (or x), needed in $\langle q|\hat{T}|\psi\rangle$, are

$$\nabla\psi = \partial\psi/\partial x = i\hbar^{-1}\xi\psi, \tag{B14}$$

$$\nabla^T\cdot m^{-1}\cdot\nabla\psi = \frac{i}{\hbar}\left[\left(\nabla^T\cdot m^{-1}\cdot\xi\right)\psi + (\nabla^T\psi)\cdot m^{-1}\cdot\xi\right] = \frac{i}{\hbar}\left[\text{Tr}\left(m^{-1}\cdot A_t\right) + \frac{i}{\hbar}\xi^T\cdot m^{-1}\cdot\xi\right]\psi, \tag{B15}$$

where Eqs. (B9) and (B10) were used in the last step.

3. Derivatives of the Gaussian density

The probability density of the Gaussian (26) is

$$\rho(x) := |\psi(x)|^2 = \left[\det(2\pi\Sigma_t)\right]^{-1/2}e^{-x^T\cdot\Sigma_t^{-1}\cdot x/2}, \tag{B16}$$

where $\Sigma_t = \text{Cov}(x) = \text{Cov}(q)$ is the position covariance matrix (B23), which is a real, symmetric, positive definite, and invertible matrix because $\text{Im}\,A_t$ is.

Derivatives of the Gaussian density (B16) with respect to the coordinate q (or x) are

$$\nabla\rho = \partial\rho/\partial x = -\Sigma_t^{-1}\cdot x\rho, \tag{B17}$$

$$\nabla\otimes\nabla^T\rho = \left(\Sigma_t^{-1}\cdot x\otimes x^T\cdot\Sigma_t^{-1} - \Sigma_t^{-1}\right)\rho. \tag{B18}$$

Because Σ_t is invertible, the above equations for the gradient and Hessian of the density imply that

$$x\rho = -\Sigma_t \cdot \nabla\rho, \tag{B19}$$

$$x \otimes x^T\rho = \left(\Sigma_t \cdot \nabla \otimes \nabla^T \cdot \Sigma_t + \Sigma_t\right)\rho. \tag{B20}$$

4. Covariances

Here, we list explicit expressions for the frequently needed position and momentum covariance matrices. Recall that the (generally complex) cross-covariance matrix of Hermitian vector operators $\hat{A}$ and $\hat{B}$ is given by

$$\begin{aligned}\mathrm{Cov}(\hat{A},\hat{B}) &:= \langle(\hat{A}-\langle\hat{A}\rangle)\otimes(\hat{B}-\langle\hat{B}\rangle)^T\rangle \\ &= \langle\hat{A}\otimes\hat{B}^T\rangle - \langle\hat{A}\rangle\otimes\langle\hat{B}\rangle^T,\end{aligned} \tag{B21}$$

and that one writes $\mathrm{Cov}(\hat{A})$ instead of $\mathrm{Cov}(\hat{A},\hat{A})$ for the (always real) autocovariance of $\hat{A}$. For $\hat{A}\neq\hat{B}$, a real cross-covariance matrix is obtained by symmetrization,

$$\begin{aligned}\mathrm{Cov}_R(\hat{A},\hat{B}) &= [\mathrm{Cov}(\hat{A},\hat{B}) + \mathrm{Cov}(\hat{B},\hat{A})^T]/2 \\ &= \mathrm{Re}\,\mathrm{Cov}(\hat{A},\hat{B}).\end{aligned} \tag{B22}$$

The position, momentum, and position-momentum covariance matrices in a Gaussian (26) [or (155)] are

$$\mathrm{Cov}(\hat{q}) = (\hbar/2)(\mathrm{Im}\,A_t)^{-1} = (\hbar/2)Q_t\cdot Q_t^\dagger, \tag{B23}$$

$$\mathrm{Cov}(\hat{p}) = (\hbar/2)A_t\cdot(\mathrm{Im}\,A_t)^{-1}\cdot A_t^* = (\hbar/2)P_t\cdot P_t^\dagger, \tag{B24}$$

$$\begin{aligned}\mathrm{Cov}(\hat{q},\hat{p}) &= \mathrm{Cov}(\hat{q})\cdot A_t = (\hbar/2)(\mathrm{Im}\,A_t)^{-1}\cdot A_t \\ &= (\hbar/2)Q_t\cdot Q_t^\dagger\cdot P_t\cdot Q_t^{-1} = \hbar(i + Q_t\cdot P_t^\dagger/2),\end{aligned} \tag{B25}$$

$$\mathrm{Cov}_R(\hat{q},\hat{p}) = (\hbar/2)(\mathrm{Im}\,A_t)^{-1}\cdot\mathrm{Re}\,A_t = (\hbar/2)\mathrm{Re}(Q_t\cdot P_t^\dagger), \tag{B26}$$

where we have listed expressions in both Heller's and Hagedorn's parametrizations. The first equality in the expression for $\mathrm{Cov}(\hat{q},\hat{p})$ holds because

$$\begin{aligned}\mathrm{Cov}(\hat{q},\hat{p}) &= \langle(\hat{q}-q_t)\otimes(\hat{p}-p_t)^T\rangle \\ &= \int \psi(x)^* x\otimes(\xi-p_t)^T\psi(x)d^Dx \\ &= \langle x\otimes(A_t\cdot x)^T\rangle = \langle x\otimes x^T\rangle\cdot A_t,\end{aligned} \tag{B27}$$

whereas the last equality is obtained by the substitution $Q_t^\dagger\cdot P_t = 2iI_D + P_t^\dagger\cdot Q_t$, which follows from Eq. (B46).

We frequently need explicit expressions for the expectation value $\langle\hat{A}\otimes\hat{B}^T\rangle$. Since $\langle\hat{q}\rangle = q_t$ and $\langle\hat{p}\rangle = p_t$, the general relation (B21) implies that

$$\langle\hat{q}\otimes\hat{q}^T\rangle = q_t\otimes q_t^T + \mathrm{Cov}(\hat{q}), \tag{B28}$$

$$\langle\hat{p}\otimes\hat{p}^T\rangle = p_t\otimes p_t^T + \mathrm{Cov}(\hat{p}), \tag{B29}$$

$$\langle\hat{q}\otimes\hat{p}^T\rangle = q_t\otimes p_t^T + \mathrm{Cov}(\hat{q},\hat{p}). \tag{B30}$$

For a frozen Gaussian, $\dot{A}_t = 0$ and $A_t = A_0 = \text{const}$. If $A_t = A_0$ is purely imaginary, i.e., $A_0 = i\,\mathrm{Im}\,A_0 = i\mathcal{B}$, covariance expressions (B23)–(B26) reduce to

$$\mathrm{Cov}(\hat{q}) = (i\hbar/2)A_0^{-1} = (\hbar/2)\mathcal{B}^{-1}, \tag{B31}$$

$$\mathrm{Cov}(\hat{p}) = -(i\hbar/2)A_0 = (\hbar/2)\mathcal{B}, \tag{B32}$$

$$\mathrm{Cov}(\hat{q},\hat{p}) = i\hbar/2, \tag{B33}$$

$$\mathrm{Cov}_R(\hat{q},\hat{p}) = 0. \tag{B34}$$

Because the position covariance appears frequently in the text, we use the shorthand notation $\Sigma_t := \mathrm{Cov}(\hat{q})$.

5. Time dependence of energy in the TGWD

To gain further insight into the time dependence of energy, let us express the expected value in Eq. (61) for $\dot{E}$ in terms of differences $\langle\hat{V}'\rangle - V_1$ and $\langle\hat{V}''\rangle - V_2$; to simplify the derivation, we employ notation $\hat{y} := \hat{p} - p_t$ for the displaced momentum operator:

$$\begin{aligned}&\langle(\hat{V}' - V_1 - V_2\cdot\hat{x})\otimes(p_t+\hat{y})^T\rangle \\ &\quad= \langle(\hat{V}'-V_1)\otimes(p_t+\hat{y})^T\rangle - V_2\cdot\langle\hat{x}\otimes(p_t+\hat{y})^T\rangle \\ &\quad= (\langle\hat{V}'\rangle - V_1)\otimes p_t^T + \langle\hat{V}'\otimes\hat{y}^T\rangle - V_2\cdot\langle\hat{x}\otimes\hat{y}^T\rangle \\ &\quad= (\langle\hat{V}'\rangle - V_1)\otimes p_t^T + \langle\hat{V}''\rangle\cdot\mathrm{Cov}(\hat{q},\hat{p}) - V_2\cdot\mathrm{Cov}(\hat{q},\hat{p}) \\ &\quad= (\langle\hat{V}'\rangle - V_1)\otimes p_t^T + (\langle\hat{V}''\rangle - V_2)\cdot\mathrm{Cov}(\hat{q},\hat{p}).\end{aligned} \tag{B35}$$

In the second step, we used $\langle\hat{x}\rangle = \langle\hat{y}\rangle = 0$ and, in the third step, invoked the definition (B25) of the position-momentum covariance $\mathrm{Cov}(\hat{q},\hat{p})$ and the relation

$$\langle\hat{V}'\otimes(\hat{p}-p_t)^T\rangle = \langle\hat{V}''\rangle\cdot\mathrm{Cov}(\hat{q},\hat{p}), \tag{B36}$$

proven in Appendix B 6. Substitution of Eq. (B35) in Eq. (61) for $\dot{E}$ yields the final expression (62) for the time dependence of energy.

6. Proof of Eq. (B36): $\langle\hat{V}'\otimes(\hat{p}-p_t)^T\rangle = \langle\hat{V}''\rangle\cdot\mathrm{Cov}(\hat{q},\hat{p})$

Equation (B36) follows immediately from the relations

$$\langle\hat{V}'\otimes(\hat{p}-p_t)^T\rangle = \langle\hat{V}'\otimes\hat{x}^T\rangle\cdot A_t, \tag{B37}$$

$$\langle\hat{V}'\otimes x^T\rangle = \langle\hat{V}''\rangle\cdot\mathrm{Cov}(\hat{q}), \tag{B38}$$

and Eq. (B25) for the position-momentum covariance. To prove Eq. (B37), we note that

$$\begin{aligned}&\langle\hat{V}'\otimes(\hat{p}-p_t)^T\rangle \\ &\quad= \int\psi(q,t)^*V'(q)\otimes(-i\hbar\nabla - p_t)^T\psi(q,t)d^Dq \\ &\quad= \int\psi(x,t)^*V'(q_t+x)\otimes(A_t\cdot x)^T\psi(x,t)d^Dx \\ &\quad= \langle\hat{V}'\otimes\hat{x}^T\rangle\cdot A_t.\end{aligned} \tag{B39}$$

Equation (B38) follows from Eq. (B19) and by-parts integration,

$$\begin{aligned}\langle \hat{V}' \otimes \hat{x}^T \rangle &= \int V'(q_t + x) \otimes x^T \rho(x) d^D x \\ &= -\int \nabla V(q_t + x) \otimes \nabla^T \rho(x) d^D x \cdot \mathrm{Cov}(\hat{q}) \\ &= \int \rho(x) \nabla \otimes \nabla^T V(q_t + x) d^D x \cdot \mathrm{Cov}(\hat{q}) \\ &= \langle \nabla \otimes \nabla^T V \rangle \cdot \mathrm{Cov}(\hat{q}). \end{aligned} \tag{B40}$$

7. Derivation of the Gaussian wavepacket's form and of the equations of motion in Hagedorn's parametrization

Here, we justify Hagedorn's form (155) of the Gaussian wavepacket as well as Eqs. (157)–(159) of motion for Hagedorn's parameters Q_t, P_t, and S_t.

Every complex symmetric matrix A with a positive definite imaginary part can be factorized as[52]

$$A = P \cdot Q^{-1}, \tag{B41}$$

where Q and P are complex invertible matrices such that the real $2D \times 2D$ matrix

$$Y := \begin{pmatrix} \mathrm{Re}\, Q & \mathrm{Im}\, Q \\ \mathrm{Re}\, P & \mathrm{Im}\, P \end{pmatrix} \tag{B42}$$

is symplectic, i.e.,

$$Y^T \cdot J \cdot Y = J, \tag{B43}$$

where

$$J = \begin{pmatrix} 0 & -\mathrm{Id}_D \\ \mathrm{Id}_D & 0 \end{pmatrix} \tag{B44}$$

is the standard symplectic matrix. This condition is equivalent[52] to the requirement that the matrices Q and P satisfy the relations

$$Q^T \cdot P - P^T \cdot Q = 0, \tag{B45}$$

$$Q^\dagger \cdot P - P^\dagger \cdot Q = 2i\, \mathrm{Id}_D. \tag{B46}$$

The imaginary part of A can be computed from Q as[52]

$$\mathrm{Im}\, A = (Q \cdot Q^\dagger)^{-1}. \tag{B47}$$

If one imposes the equation

$$\dot{Q}_t = m^{-1} \cdot P_t \tag{B48}$$

of motion for Q_t, then P_t must satisfy the equation

$$\begin{aligned}\dot{P}_t = \dot{A}_t \cdot Q_t + A_t \cdot \dot{Q}_t = -A_t \cdot m^{-1} \cdot A_t \cdot Q_t \\ - V_2 \cdot Q_t + A_t \cdot m^{-1} \cdot P_t = -V_2 \cdot Q_t. \end{aligned} \tag{B49}$$

In the derivation, we used the Leibniz rule, Eq. (44) for $\dot{A}_t$, and factorization (B41) of A_t. It can be shown[52] that if parameters Q_t and P_t are propagated with Eqs. (B48) and (B49) and satisfy relations (B45)–(B47) at time zero, they satisfy those relations for all times.

Equation (45) for $\dot{\gamma}_t$ can be rewritten as

$$\dot{\gamma}_t = \dot{S}_t + \frac{i\hbar}{2} \frac{d}{dt} \ln \det Q_t, \tag{B50}$$

where we have defined a generalized real action S_t by

$$\dot{S}_t = T(p_t) - V_0. \tag{B51}$$

The second term in Eq. (B50) follows from Eq. (45) since

$$\begin{aligned}\mathrm{Tr}\left(m^{-1} \cdot A_t\right) = \mathrm{Tr}\left(m^{-1} \cdot P_t \cdot Q_t^{-1}\right) = \mathrm{Tr}\left(\dot{Q}_t \cdot Q_t^{-1}\right) \\ = (\det Q_t)^{-1} \cdot d(\det Q_t)/dt. \end{aligned} \tag{B52}$$

Here, we used the factorization (B41) of A_t, the equation of motion (B48) for Q_t, and the formula for the derivative of a determinant. By integrating Eq. (B50) for $\dot{\gamma}_t$, we can rewrite the wavepacket (26) at time t as

$$\begin{aligned}\psi(q,t) = (\det Q_0 / \det Q_t)^{1/2} \\ \times \exp\left[\frac{i}{\hbar}\left(\frac{1}{2} x^T \cdot P_t \cdot Q_t^{-1} \cdot x + p_t^T \cdot x + S_t - S_0 + \gamma_0\right)\right]. \end{aligned} \tag{B53}$$

Recall that the initial wavepacket $\psi(q,0)$ is normalized if γ_0 satisfies Eq. (29). Setting the initial value of γ_t to $\gamma_0 = S_0 + i\, \mathrm{Im}\, \gamma_0$, where

$$\begin{aligned}\mathrm{Im}\, \gamma_0 = -\hbar \ln \det (\mathrm{Im}\, A_0 / \pi\hbar)^{1/4} = \hbar \ln \det (\pi\hbar Q_0 \cdot Q_0^\dagger)^{1/4} \\ = \hbar \ln \left[(\pi\hbar)^D |\det Q_0|^2\right]^{1/4} = \frac{\hbar}{2} \ln \left[(\pi\hbar)^{D/2} |\det Q_0|\right], \end{aligned}$$

we obtain a simple formula for the wavepacket parameterized by Q_t, P_t, and S_t instead of A_t and γ_t,

$$\begin{aligned}\psi(x,t) = (\pi\hbar)^{-D/4} \left(\frac{\det Q_0}{\det Q_t |\det Q_0|}\right)^{1/2} \\ \times \exp\left[\frac{i}{\hbar}\left(\frac{1}{2} x^T \cdot P_t \cdot Q_t^{-1} \cdot x + p_t^T \cdot x + S_t\right)\right]. \end{aligned} \tag{B54}$$

The prefactor can be further simplified with another choice of the initial value of γ_t, namely,

$$\gamma_0 = S_0 + \hbar\varphi_0 + i\mathrm{Im}\, \gamma_0, \tag{B55}$$

where $\varphi_0 = -\frac{1}{2} \arg \det Q_0$, for which the wavepacket assumes a simple Hagedorn form (155), valid at all times.

APPENDIX C: DIRAC–FRENKEL VARIATIONAL PRINCIPLE, NONLINEAR TDSE, AND CONSERVATION OF ENERGY AND NORM

One often seeks an approximate solution of the TDSE only within a certain subset M of the full Hilbert space $\mathcal{H}$. If one seeks

a solution among states $\phi(t;\theta_1,\ldots,\theta_N) \in \mathcal{H}$ that depend on N parameters $\theta_j, j = 1,\ldots,N$, the Dirac–Frenkel variational principle states that the optimal solution satisfies the equation

$$\langle\delta\phi|[i\hbar(d/dt) - \hat{H}]|\phi(t)\rangle = 0, \tag{C1}$$

where $\delta\phi$ is an arbitrary variation of the solution, i.e., an infinitesimal change of ϕ such that $\phi + \delta\phi$ is still in the approximation manifold M. [More precisely, $\delta\phi$ is an arbitrary "tangent vector" to the manifold M at the point $\phi(t)$.] The variation $\delta\phi$ can be expressed in terms of the variations of its N parameters as

$$\delta\phi = \sum_{j=1}^{N} \frac{\partial\phi}{\partial\theta_j}\delta\theta_j. \tag{C2}$$

Because the variations of parameters are independent, the Dirac–Frenkel principle (C1) requires the following equation to be satisfied for each of the parameters:

$$0 = \left\langle \frac{\partial\phi}{\partial\theta_j} \middle| \left(i\hbar\frac{d}{dt} - \hat{H}\right) \middle| \phi \right\rangle \text{ for } j = 1,\ldots,N. \tag{C3}$$

This variational principle provides a rich class of nonlinear TDSEs because Eq. (C1) is equivalent[52] to

$$i\hbar|\dot{\phi}(t)\rangle = \hat{P}(\phi(t))\hat{H}|\phi(t)\rangle, \tag{C4}$$

where $\hat{P}(\phi)$ is the projection on the tangent space of M at the state ϕ. In other words, the solution $\phi(t)$ satisfies the nonlinear TDSE (6) with the effective Hamiltonian

$$\hat{H}_{\text{eff}}(\phi) := \hat{P}(\phi(t))\hat{H}. \tag{C5}$$

Remarkably, solutions of the Dirac–Frenkel equation (C1) preserve several properties of the exact solution of the TDSE (1). In particular, the energy is conserved along the variational solutions $\phi(t)$ satisfying Eq. (C1), and the conservation of the norm requires only a weak additional assumption. The energy is conserved because[25]

$$\begin{aligned}\frac{d}{dt}\langle\hat{H}\rangle &= \frac{d}{dt}\langle\phi(t)|\hat{H}|\phi(t)\rangle = \langle\dot{\phi}|\hat{H}|\phi\rangle + \langle\phi|\hat{H}|\dot{\phi}\rangle \\ &= 2\,\mathrm{Re}\langle\dot{\phi}|\hat{H}|\phi\rangle = 2\mathrm{Re}\langle\dot{\phi}|i\hbar\dot{\phi}\rangle \\ &= 2\,\mathrm{Re}(i\hbar\|\dot{\phi}\|^2) = 0,\end{aligned} \tag{C6}$$

where, in the fourth step, we invoked the variational principle (C1) with $\delta\phi \propto \dot{\phi}$. If the manifold has the ray property ($\lambda\phi \in M$ for each complex number λ and each $\phi \in M$), then the norm remains constant because[25]

$$\begin{aligned}\frac{d}{dt}\|\phi(t)\|^2 &= \frac{d}{dt}\langle\phi(t)|\phi(t)\rangle = \langle\dot{\phi}|\phi\rangle + \langle\phi|\dot{\phi}\rangle = 2\mathrm{Re}\langle\phi|\dot{\phi}\rangle \\ &= 2\hbar^{-1}\mathrm{Im}\langle\phi|i\hbar\dot{\phi}\rangle = 2\hbar^{-1}\mathrm{Im}\langle\phi|\hat{H}\phi\rangle \\ &= 2\hbar^{-1}\mathrm{Im}\langle\hat{H}\rangle = 0,\end{aligned} \tag{C7}$$

where, in the fifth step, we invoked the variational principle (C1) with $\delta\phi \propto \phi$. ϕ is in the tangent space because of the ray property. If the ray property does not hold with given parameters, it will hold if we augment[89] the parameter set by a prefactor λ of the state ϕ.

Note that the conservation of energy requires only the real part of Eq. (C1), which is sometimes referred to as the *Lagrangian variational principle* (or *time-dependent variational principle*),[90] whereas the conservation of norm requires only the imaginary part of Eq. (C1), which is sometimes referred to as the *McLachlan variational principle*.[40] For Gaussian wavepackets, the three forms of the variational principle are equivalent.[25,52,91]

APPENDIX D: VARIATIONAL GAUSSIAN APPROXIMATION

The variational Gaussian approximation (or variational TGWD) follows from the Dirac–Frenkel principle:

Proposition (Variational TGWD). The Dirac–Frenkel variational principle (C1) for the TDSE (1) applied to the Gaussian ansatz (26) yields the following equations of motion for the Gaussian's parameters:

$$\dot{q}_t = m^{-1}\cdot p_t, \tag{D1}$$

$$\dot{p}_t = -\langle\hat{V}'\rangle, \tag{D2}$$

$$\dot{A}_t = -A_t\cdot m^{-1}\cdot A_t - \langle\hat{V}''\rangle, \tag{D3}$$

$$\begin{aligned}\dot{\gamma}_t = {}& T(p_t) - \langle\hat{V}\rangle + (\hbar/4)\,\mathrm{Tr}\left[\langle\hat{V}''\rangle\cdot(\mathrm{Im}\,A_t)^{-1}\right] \\ &+ (i\hbar/2)\,\mathrm{Tr}\left(m^{-1}\cdot A_t\right).\end{aligned} \tag{D4}$$

These are equivalent to Eqs. (42)–(45) for the TGWD satisfying the nonlinear TDSE (20) with the effective potential (30) and coefficients (69).

Proof. Since the manifold of Gaussian wavepackets has the ray property, the variational solution conserves the norm of the wavefunction (see Appendix C).[25] For the Gaussian ansatz (26), this implies [see Eq. (28)] that

$$\exp\left(-\mathrm{Im}\,\gamma_t/\hbar\right) = \left[\det\left(\mathrm{Im}\,A_t/\pi\hbar\right)\right]^{1/4} \tag{D5}$$

for all times t if the initial norm is 1. For Gaussian (26) with density $\rho(x)$ [Eq. (B16)], variational equations (C3) for parameters q_t, p_t, A_t, and γ_t are, respectively,

$$0 = \int (A_t\cdot x + p_t)g(x)\rho(x)d^Dx, \tag{D6}$$

$$0 = \int x g(x)\rho(x)d^Dx, \tag{D7}$$

$$0 = \int x\otimes x^T g(x)\rho(x)d^Dx, \tag{D8}$$

$$0 = \int g(x)\rho(x)d^Dx, \tag{D9}$$

where

$$g(x) := f(x) - V(q_t + x) \tag{D10}$$

is the difference between the quadratic polynomial (38) and the potential energy $V(q)$. Because ρ is normalized, the variational equations (D6)–(D9) can be expressed as

$$0 = \langle (A_t \cdot x + p_t) g(x) \rangle, \tag{D11}$$

$$0 = \langle x g(x) \rangle, \tag{D12}$$

$$0 = \left\langle x \otimes x^T g(x) \right\rangle, \tag{D13}$$

$$0 = \langle g(x) \rangle. \tag{D14}$$

The first equation, which follows from the second and fourth equations, is redundant. The last three equations, which are independent, are equivalent to the system

$$0 = \langle g(x) \rangle, \tag{D15}$$

$$0 = \langle \nabla g(x) \rangle, \tag{D16}$$

$$0 = \left\langle \nabla \otimes \nabla^T g(x) \right\rangle. \tag{D17}$$

Equation (D16) is equivalent to Eq. (D12) because

$$\begin{aligned} \langle \nabla g(x) \rangle &= \int \rho \nabla g \, d^D x = - \int g \nabla \rho \, d^D x \\ &= \int g \Sigma_t^{-1} \cdot x \rho \, d^D x = \Sigma_t^{-1} \cdot \langle x g(x) \rangle. \end{aligned} \tag{D18}$$

Here, we integrated by parts and used relation (B17). Likewise, Eqs. (D13) and (D14) are equivalent to Eqs. (D15) and (D17) since

$$\begin{aligned} \left\langle \nabla \otimes \nabla^T g(x) \right\rangle &= \int \rho \nabla \otimes \nabla^T g \, d^D x = \int g \nabla \otimes \nabla^T \rho \, d^D x \\ &= \int g \left(\Sigma_t^{-1} \cdot x \otimes x^T \cdot \Sigma_t^{-1} - \Sigma_t^{-1} \right) \rho \, d^D x \\ &= \Sigma_t^{-1} \cdot \langle x \otimes x^T g(x) \rangle \cdot \Sigma_t^{-1} - \Sigma_t^{-1} \langle g(x) \rangle, \end{aligned}$$

where we integrated twice by parts and used Eq. (B18).

Recalling the definition (D10) of $g(x)$, we find that g, its gradient, and its Hessian can be written explicitly as

$$g(x) = x^T \cdot C_2 \cdot x/2 + x^T \cdot C_1 + C_0 - V(q_t + x), \tag{D19}$$

$$g'(x) = C_2 \cdot x + C_1 - V'(q_t + x), \tag{D20}$$

$$g''(x) = C_2 - V''(q_t + x). \tag{D21}$$

Substituting expressions (D19)–(D21) for g, g', and g'' into Eqs. (D15)–(D17) for the expectation values yields

$$\begin{aligned} 0 &= \langle x^T \cdot C_2 \cdot x \rangle/2 + C_1^T \cdot \langle x \rangle + C_0 - \langle \hat{V} \rangle \\ &= \mathrm{Tr}\,(C_2 \cdot \Sigma_t)/2 + C_0 - \langle \hat{V} \rangle, \end{aligned} \tag{D22}$$

$$0 = C_2 \cdot \langle x \rangle + C_1 - \langle \hat{V}' \rangle = C_1 - \langle \hat{V}' \rangle, \tag{D23}$$

$$0 = C_2 - \langle \hat{V}'' \rangle, \tag{D24}$$

where we have used relations $\langle x \rangle = 0$ for the position and

$$\langle x^T \cdot C_2 \cdot x \rangle = \mathrm{Tr}\,(C_2 \cdot \langle x \otimes x^T \rangle) = \mathrm{Tr}\,(C_2 \cdot \Sigma_t) \tag{D25}$$

for the expectation value of a quadratic form.

The system of Eqs. (D22)–(D24) has the solution

$$C_0 = \langle \hat{V} \rangle - \mathrm{Tr}\,[\langle \hat{V}'' \rangle \cdot \Sigma_t]/2, \tag{D26}$$

$$C_1 = \langle \hat{V}' \rangle, \tag{D27}$$

$$C_2 = \langle \hat{V}'' \rangle. \tag{D28}$$

In view of Propositions 1 and 2 from Sec. IV B, we find that the variational TGWD is equivalent to the nonlinear TDSE (20) with an effective potential (30) whose parameters V_0, V_1, and V_2 satisfy conditions (69) and are all real, as required in Proposition 2. Altogether, the Dirac–Frenkel principle applied to the Gaussian (26) is equivalent to the system (42)–(45) with coefficients V_j given by Eq. (30), i.e., to the system (D1)–(D4).

APPENDIX E: CONSERVATION OF E_{eff} BY THE VARIATIONAL TGWD

Although $E_{\mathrm{eff}} = E = \mathrm{const}$ for any solution of the Dirac–Frenkel variational principle (see Appendix C), it is instructive to demonstrate the conservation of the effective energy explicitly for the variational TGWD, regarded as a solution of the nonlinear TDSE (6). Applying the general expression (66) for $\dot{E}_{\mathrm{eff}}$ to the variational coefficients (69) gives

$$\begin{aligned} \dot{E}_{\mathrm{eff}} &= \frac{d}{dt}[\langle \hat{V} \rangle - \mathrm{Tr}\,(V_2 \cdot \Sigma_t)/2] - V_1^T \cdot \dot{q}_t + \mathrm{Tr}\,(\dot{V}_2 \cdot \Sigma_t)/2 \\ &= \frac{d\langle \hat{V} \rangle}{dt} - \frac{1}{2} \mathrm{Tr}\,(V_2 \cdot \dot{\Sigma}_t) - \langle \hat{V}' \rangle^T \cdot m^{-1} \cdot p_t, \end{aligned} \tag{E1}$$

where two terms $\mathrm{Tr}\,(\dot{V}_2 \cdot \Sigma_t)/2$ of opposite signs cancel each other. The time derivative of E_{eff} becomes

$$\begin{aligned} \dot{E}_{\mathrm{eff}} = \mathrm{Tr}\,(m^{-1} \cdot \mathrm{Re}\langle \hat{V}' \otimes \hat{p}^T \rangle) - \mathrm{Tr}\,[V_2 \cdot \mathrm{Cov}_R(\hat{q}, \hat{p}) \cdot m^{-1}] \\ - \mathrm{Tr}\,(m^{-1} \cdot \langle \hat{V}' \rangle \otimes p_t^T) \end{aligned} \tag{E2}$$

because

$$\frac{d\langle \hat{V} \rangle}{dt} = \mathrm{Re}\langle \hat{p}^T \cdot m^{-1} \cdot \hat{V}' \rangle = \mathrm{Tr}\,(m^{-1} \cdot \mathrm{Re}\langle \hat{V}' \otimes \hat{p}^T \rangle),$$

which can be derived in the same way as Eq. (21), and

$$\begin{aligned} \mathrm{Tr}\,(V_2 \cdot \dot{\Sigma}_t) &= (\hbar/2) \,\mathrm{Tr}\,[V_2 \cdot (m^{-1} \cdot \mathcal{A} \cdot \mathcal{B}^{-1} + \mathcal{B}^{-1} \cdot \mathcal{A} \cdot m^{-1})] \\ &= 2 \,\mathrm{Tr}\,[V_2 \cdot \mathrm{Cov}_R(\hat{q}, \hat{p}) \cdot m^{-1}]. \end{aligned} \tag{E3}$$

To find expression (E3) for $\mathrm{Tr}\,(V_2 \cdot \dot{\Sigma}_t)$, we used the notation $\mathcal{A} := \mathrm{Re}\,A_t$ and $\mathcal{B} := \mathrm{Im}\,B_t$ and the relation

$$\begin{aligned}\dot{\Sigma}_t &= -(\hbar/2)\mathcal{B}^{-1}\cdot\dot{\mathcal{B}}\cdot\mathcal{B}^{-1}\\ &= (\hbar/2)(m^{-1}\cdot\mathcal{A}\cdot\mathcal{B}^{-1}+\mathcal{B}^{-1}\cdot\mathcal{A}\cdot m^{-1}),\end{aligned}\tag{E4}$$

which follows from Eqs. (B23) and (44). Equation (E3) then follows from the symmetry of matrices $\mathcal{A}$, $\mathcal{B}^{-1}$, m^{-1}, and V_2, and from Eq. (B26) for $\mathrm{Cov}_R(\hat{q},\hat{p})$ because

$$\begin{aligned}\mathrm{Tr}\left(V_2\cdot m^{-1}\cdot\mathcal{A}\cdot\mathcal{B}^{-1}\right) &= \mathrm{Tr}\left[\left(V_2\cdot m^{-1}\cdot\mathcal{A}\cdot\mathcal{B}^{-1}\right)^T\right]\\ &= \mathrm{Tr}\left(\mathcal{B}^{-1}\cdot\mathcal{A}\cdot m^{-1}\cdot V_2\right)\\ &= \mathrm{Tr}\left(V_2\cdot\mathcal{B}^{-1}\cdot\mathcal{A}\cdot m^{-1}\right).\end{aligned}$$

Using identity (B36) and definition (B22) of $\mathrm{Cov}_R(\hat{q},\hat{p})$ shows that the effective energy is conserved:

$$\begin{aligned}\dot{E}_{\mathrm{eff}} &= \mathrm{Tr}\left\{m^{-1}\cdot\left[\mathrm{Re}\langle\hat{V}'\otimes(\hat{p}-p_t)^T\rangle-\langle\hat{V}''\rangle\cdot\mathrm{Cov}_R(\hat{q},\hat{p})\right]\right\}\\ &= 0.\end{aligned}\tag{E5}$$

APPENDIX F: PROOF THAT $\mathcal{A} = \mathrm{Re}\,A = 0$ IN FGWD

Here, we prove that setting $\mathrm{Im}\,V_2 = 0$ in Eq. (106) for the FGWD implies that the width matrix A of the frozen Gaussian is purely imaginary. In fact, we will show more generally that if m^{-1} and $\mathcal{B}$ are positive-definite real symmetric $D\times D$ matrices and $\mathcal{A}$ a real (but not necessarily symmetric) $D\times D$ matrix, then the equation

$$\mathcal{A}\cdot m^{-1}\cdot\mathcal{B}+\mathcal{B}\cdot m^{-1}\cdot\mathcal{A}=0\tag{F1}$$

implies that $\mathcal{A}=0$. First, note that matrix m^{-1} has a unique positive-definite real symmetric square root $m^{-1/2}$.[92] Multiplication of Eq. (F1) by $m^{-1/2}$ both from the left and from the right yields

$$\mathcal{A}'\cdot\mathcal{B}'+\mathcal{B}'\cdot\mathcal{A}'=0\tag{F2}$$

where $\mathcal{A}' = m^{-1/2}\cdot\mathcal{A}\cdot m^{-1/2}$ is a real matrix and $\mathcal{B}' = m^{-1/2}\cdot\mathcal{B}\cdot m^{-1/2}$ is a positive-definite real symmetric matrix,[92] whose eigenvectors v_j form a basis of $\mathbb{R}^D$. In this basis, matrix elements of Eq. (F2) are

$$0 = v_j^T\cdot(\mathcal{A}'\cdot\mathcal{B}'+\mathcal{B}'\cdot\mathcal{A}')\cdot v_k = (\lambda_j+\lambda_k)v_j^T\cdot\mathcal{A}'\cdot v_k.$$

Since all eigenvalues λ_j of $\mathcal{B}'$ are strictly positive, $v_j\cdot\mathcal{A}'\cdot v_k = 0$ for all j and k. Therefore, $\mathcal{A}'=0$. Because $m^{-1/2}$ is positive-definite, it has an inverse $m^{1/2}$ and

$$\mathcal{A} = m^{1/2}\cdot\mathcal{A}'\cdot m^{1/2} = 0.\tag{F3}$$

APPENDIX G: KINETIC PROPAGATION

Here we derive analytical solutions (132)–(135) for the kinetic propagation. Equations (132) and (133) for q_t and p_t follow from Eqs. (128) and (129) because $p_t = \mathrm{const}$. To solve the differential Eq. (130) for A_t, we use the relation for the derivative of a matrix inverse,[88]

$$d(A_t^{-1})/dt = -A_t^{-1}\cdot\dot{A}_t\cdot A_t^{-1} = m^{-1},\tag{G1}$$

where Eq. (130) was used in the second step. The differential equation (G1) has the solution

$$\begin{aligned}A_t^{-1} &= A_0^{-1}+tm^{-1} = A_0^{-1}\cdot(\mathrm{Id}_D+tA_0\cdot m^{-1})\\ &= (\mathrm{Id}_D+tm^{-1}\cdot A_0)\cdot A_0^{-1},\end{aligned}\tag{G2}$$

and taking the inverse of the three alternative expressions for A_t^{-1} gives the three formulas for A_t in Eq. (134). To find γ_t, we substitute the expression for $m^{-1}\cdot A_t$ from Eq. (130) into Eq. (131) and use the formula for the derivative of a logarithm of a determinant,

$$\begin{aligned}\dot{\gamma}_t &= T(p_t)-(i\hbar/2)\,\mathrm{Tr}\,(A_t^{-1}\cdot\dot{A}_t)\\ &= T(p_t)-(i\hbar/2)\frac{d}{dt}\ln\det A_t.\end{aligned}\tag{G3}$$

Since $p_t = p_0$, the solution of this differential equation is

$$\begin{aligned}\gamma_t &= \gamma_0+tT(p_0)-(i\hbar/2)\ln\left[(\det A_t)/(\det A_0)\right]\\ &= \gamma_0+tT(p_0)-(i\hbar/2)\ln\det(A_0^{-1}\cdot A_t).\end{aligned}\tag{G4}$$

Inserting Eq. (134) for A_t gives Eq. (135) for γ_t.

Among Eqs. (132)–(135), the only one that is nontrivial to invert is Eq. (135) for γ_t, whose inversion (140) follows from Eqs. (132) and (G2) because

$$\begin{aligned}\mathrm{Id}_D+tm^{-1}\cdot A_0 = A_t^{-1}\cdot A_0 &= (A_0^{-1}\cdot A_t)^{-1} = \left[(A_t^{-1}-tm^{-1})\cdot A_t\right]^{-1}\\ &= \left(\mathrm{Id}_D-tm^{-1}\cdot A_t\right)^{-1}.\end{aligned}\tag{G5}$$

REFERENCES

[1] E. J. Heller, Acc. Chem. Res. **14**, 368 (1981).
[2] S. Mukamel, *Principles of Nonlinear Optical Spectroscopy*, 1st ed. (Oxford University Press, New York, 1999).
[3] W. H. Miller, J. Phys. Chem. A **105**, 2942 (2001).
[4] D. J. Tannor, *Introduction to Quantum Mechanics: A Time-Dependent Perspective* (University Science Books, Sausalito, 2007).
[5] E. J. Heller, *The Semiclassical Way to Dynamics and Spectroscopy* (Princeton University Press, Princeton, NJ, 2018).
[6] E. J. Heller, J. Chem. Phys. **75**, 2923 (1981).
[7] M. F. Herman and E. Kluk, Chem. Phys. **91**, 27 (1984).
[8] F. Grossmann, J. Chem. Phys. **125**, 014111 (2006).
[9] M. Ben-Nun and T. J. Martínez, J. Chem. Phys. **108**, 7244 (1998).
[10] H.-M. Tsai and B. Poirier, J. Phys.: Conf. Ser. **701**, 012013 (2016).
[11] M. Ben-Nun, J. Quenneville, and T. J. Martínez, J. Phys. Chem. A **104**, 5161 (2000).
[12] M. Ceotto, S. Atahan, S. Shim, G. F. Tantardini, and A. Aspuru-Guzik, Phys. Chem. Chem. Phys. **11**, 3861 (2009).
[13] J. Tatchen and E. Pollak, J. Chem. Phys. **130**, 041103 (2009).
[14] B. F. E. Curchod and T. J. Martínez, Chem. Rev. **118**, 3305 (2018).
[15] E. J. Heller, J. Chem. Phys. **62**, 1544 (1975).
[16] S. Y. Lee and E. J. Heller, J. Chem. Phys. **76**, 3035 (1982).
[17] D. J. Tannor and E. J. Heller, J. Chem. Phys. **77**, 202 (1982).
[18] M. Wehrle, M. Šulc, and J. Vaníček, J. Chem. Phys. **140**, 244114 (2014).
[19] M. Wehrle, S. Oberli, and J. Vaníček, J. Phys. Chem. A **119**, 5685 (2015).
[20] T. Begušić, J. Roulet, and J. Vaníček, J. Chem. Phys. **149**, 244115 (2018).
[21] T. Begušić, E. Tapavicza, and J. Vaníček, J. Chem. Theory Comput. **18**, 3065 (2022).
[22] E. J. Heller, J. Chem. Phys. **64**, 63 (1976).

[23]P. A. M. Dirac, Math. Proc. Cambridge Philos. Soc. **26**, 376 (1930).
[24]J. Frenkel, *Wave Mechanics* (Clarendon Press, Oxford, 1934).
[25]C. Lubich, *From Quantum to Classical Molecular Dynamics: Reduced Models and Numerical Analysis*, 12th ed. (European Mathematical Society, Zürich, 2008).
[26]R. D. Coalson and M. Karplus, J. Chem. Phys. **93**, 3919 (1990).
[27]E. Faou and C. Lubich, Comput. Visualization Sci. **9**, 45 (2006).
[28]E. Hairer, C. Lubich, and G. Wanner, *Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations* (Springer, Berlin, Heidelberg; New York, 2006).
[29]A. K. Pattanayak and W. C. Schieve, Phys. Rev. Lett. **72**, 2855 (1994).
[30]A. K. Pattanayak and W. C. Schieve, Phys. Rev. E **50**, 3601 (1994).
[31]T. Ohsawa and M. Leok, J. Phys. A: Math. Theor. **46**, 405201 (2013).
[32]T. Ohsawa, J. Math. Phys. **56**, 032103 (2015).
[33]T. Ohsawa, Lett. Math. Phys. **105**, 1301 (2015).
[34]T. Ohsawa and C. Tronci, J. Math. Phys. **58**, 092105 (2017).
[35]T. Begušić, M. Cordova, and J. Vaníček, J. Chem. Phys. **150**, 154117 (2019).
[36]T. Begušić and J. Vaníček, J. Chem. Phys. **153**, 024105 (2020).
[37]E. P. Gross, Nuovo Cimento **20**, 454 (1961).
[38]L. P. Pitaevskii, Sov. Phys. JETP **13**, 451 (1961).
[39]R. Carles, Ann. Henri Poincaré **3**, 757 (2002).
[40]A. D. McLachlan, Mol. Phys. **8**, 39 (1964).
[41]P. Jungwirth and R. B. Gerber, J. Chem. Phys. **102**, 6046 (1995).
[42]A. D. McLachlan and M. A. Ball, Rev. Mod. Phys. **36**, 844 (1964).
[43]H.-D. Meyer, U. Manthe, and L. S. Cederbaum, Chem. Phys. Lett. **165**, 73 (1990).
[44]M. Beck, A. Jäckle, G. Worth, and H.-D. Meyer, Phys. Rep. **324**, 1 (2000).
[45]J. Roulet and J. Vaníček, J. Chem. Phys. **154**, 154106 (2021).
[46]J. Roulet and J. Vaníček, J. Chem. Phys. **155**, 204109 (2021).
[47]R. Kosloff, S. A. Rice, P. Gaspard, S. Tersigni, and D. J. Tannor, Chem. Phys. **139**, 201 (1989).
[48]P. Marquetand and V. Engel, J. Chem. Phys. **127**, 084115 (2007).
[49]V. Engel, C. Meier, and D. J. Tannor, Adv. Chem. Phys. **141**, 29 (2009).
[50]E. J. Heller, J. Chem. Phys. **65**, 4979 (1976).
[51]E. J. Heller, Acc. Chem. Res. **39**, 127 (2006).
[52]C. Lasser and C. Lubich, Acta Numer. **29**, 229 (2020).
[53]L. Hackl, T. Guaita, T. Shi, J. Haegeman, E. Demler, and J. I. Cirac, SciPost Phys. **9**, 048 (2020).
[54]C. Lasser and C. Su, J. Math. Phys. **63**, 072107 (2022).
[55]J. Vaníček, "Symplectic and Hamiltonian structures of Gaussian wavepacket dynamics" (unpublished).
[56]M. Werther, S. L. Choudhury, and F. Grossmann, Int. Rev. Phys. Chem. **40**, 81 (2021).
[57]B. Poirier, "Dirac–Frenkel variational principle as applied to quantum trajectories" (unpublished).
[58]P. A. Frantsuzov and V. A. Mandelshtam, J. Chem. Phys. **121**, 9247 (2004).
[59]I. Georgescu, J. Deckman, L. J. Fredrickson, and V. A. Mandelshtam, J. Chem. Phys. **134**, 174109 (2011).
[60]D. Lauvergnat, E. Baloïtcha, G. Dive, and M. Desouter-Lecomte, Chem. Phys. **326**, 500 (2006).
[61]T. Begušić and J. Vaníček, J. Phys. Chem. Lett. **12**, 2997 (2021).
[62]T. Begušić and J. Vaníček, Chimia **75**, 261 (2021).
[63]M. Bonfanti, J. Petersen, P. Eisenbrandt, I. Burghardt, and E. Pollak, J. Chem. Theory Comput. **14**, 5310 (2018).
[64]A. L. Kaledin and W. H. Miller, J. Chem. Phys. **118**, 7174 (2003).
[65]M. Ceotto, G. Di Liberto, and R. Conte, Phys. Rev. Lett. **119**, 010401 (2017).
[66]M. S. Church, S. V. Antipov, and N. Ananth, J. Chem. Phys. **146**, 234104 (2017).
[67]H. Cartarius and E. Pollak, J. Chem. Phys. **134**, 044107 (2011).
[68]T. J. Martínez and R. D. Levine, J. Chem. Soc., Faraday Trans. **93**, 941 (1997).
[69]M. S. Child and D. V. Shalashilin, J. Chem. Phys. **118**, 2061 (2003).
[70]I. Burghardt, K. Giri, and G. A. Worth, J. Chem. Phys. **129**, 174104 (2008).
[71]M. Šulc, H. Hernández, T. J. Martínez, and J. Vaníček, J. Chem. Phys. **139**, 034112 (2013).
[72]B. Gu and S. Garashchuk, J. Phys. Chem. A **120**, 3023 (2016).
[73]B. Leimkuhler and S. Reich, *Simulating Hamiltonian Dynamics* (Cambridge University Press, 2004).
[74]L. Verlet, Phys. Rev. **159**, 98 (1967).
[75]M. D. Feit, J. A. Fleck, Jr., and A. Steiger, J. Comput. Phys. **47**, 412 (1982).
[76]H. Yoshida, Phys. Lett. A **150**, 262 (1990).
[77]M. Suzuki, Phys. Lett. A **146**, 319 (1990).
[78]W. Kahan and R.-C. Li, Math. Comput. **66**, 1089 (1997).
[79]M. Sofroniou and G. Spaletta, Optim. Method Software **20**, 597 (2005).
[80]S. Choi and J. Vaníček, J. Chem. Phys. **150**, 204112 (2019).
[81]J. Roulet, S. Choi, and J. Vaníček, J. Chem. Phys. **150**, 204113 (2019).
[82]G. A. Hagedorn, Commun. Math. Phys. **71**, 77 (1980).
[83]G. A. Hagedorn, Ann. Phys. **269**, 77 (1998).
[84]A. Patoz, T. Begušić, and J. Vaníček, J. Phys. Chem. Lett. **9**, 2367 (2018).
[85]A. Prlj, T. Begušić, Z. T. Zhang, G. C. Fish, M. Wehrle, T. Zimmermann, S. Choi, J. Roulet, J.-E. Moser, and J. Vaníček, J. Chem. Theory Comput. **16**, 2617 (2020).
[86]M. Wenzel and R. Mitric, J. Chem. Phys. **158**, 034105 (2023).
[87]T. Ohsawa, Nonlinearity **31**, 1807 (2018).
[88]K. B. Petersen and M. S. Pedersen, *The Matrix Cookbook* (Technical University of Denmark, 2012), available at https://www2.imm.dtu.dk/pubdb/pubs/3274-full.html.
[89]L. Joubert-Doriol and A. F. Izmaylov, J. Chem. Phys. **148**, 114102 (2018).
[90]P. Kramer and M. Saraceno, *Geometry of the Time-Dependent Variational Principle in Quantum Mechanics*, Lecture Notes in Physics (Springer-Verlag, Berlin, 1981), Vol. 140.
[91]J. Broeckhove, L. Lathouwers, E. Kesteloot, and P. V. Leuven, Chem. Phys. Lett. **149**, 547 (1988).
[92]P. R. Halmos, *Finite Dimensional Vector Spaces* (Princeton University Press, 1942).
[93]R. Heather and H. Metiu, Chem. Phys. Lett. **118**, 558 (1985).