Local-in-time error in variational quantum dynamics

Rocco Martinazzo[1,2,*], Irene Burghardt[3]
[1] *Department of Chemistry, Università degli Studi di Milano, Via Golgi 19, 20133 Milano, Italy**
[2] *Istituto di Scienze e Tecnologie Molecolari, CNR, via Golgi 19, 20133 Milano, Italy and*
[3] *Institute of Physical and Theoretical Chemistry, Goethe University Frankfurt, Max-von-Laue-Str. 7, D-60438 Frankfurt/Main, Germany*

The McLachlan "minimum-distance" principle for optimizing approximate solutions of the time-dependent Schrödinger equation is revisited, with a focus on the local-in-time error accompanying the variational solutions. Simple, exact expressions are provided for this error, which are then evaluated in illustrative cases, notably the widely used mean-field approach and the adiabatic quantum molecular dynamics. These findings pave the way for the rigorous development of adaptive schemes that re-size *on-the-fly* the underlying variational manifold and thus optimize the overall computational cost of a quantum dynamical simulation.

arXiv:1907.00841v1 [quant-ph] 1 Jul 2019

Introduction. Variational principles play a major role in quantum dynamics since they allow to devise general strategies to evolve wavefunctions on parametrized manifolds, in such a way to mimic as much as possible the exact quantum mechanical evolution. There exist at least three different time-dependent variational principles, namely the McLachlan[1] variational principle (MVP), the Time-Dependent Variational Principle[2] (TDVP) and the Dirac-Frenkel[3, 4] variational principle (DFVP), which are known to be equivalent to each other under mild conditions[5], usually satisfied in practice. However, these three variational principles have different origins and limitations and, indeed, only the first one represents a well-founded, general optimization scheme. The reason is that the DFVP

$$\langle\delta\Psi|(i\hbar\partial_t - H)|\Psi\rangle = 0 \tag{1}$$

is *not*, strictly speaking, a variational principle, since it is not a functional variation - in the sense that it does not refer to an action functional - but just a condition which defines an optimization problem. It closely resembles, but is stronger than, the condition

$$\Re\,\langle\delta\Psi|(i\hbar\partial_t - H)|\Psi\rangle = 0 \tag{2}$$

that results from the TDVP, which is indeed a stationary-action principle, $\delta S = \delta\int_{t_i}^{t_f} L[\Psi_t]dt = 0$, with the real Lagrangian (here for normalized wavefunctions)

$$L[\Psi_t] = \frac{i\hbar}{2}\left(\langle\Psi_t|\dot\Psi_t\rangle - \langle\dot\Psi_t|\Psi_t\rangle\right) - \langle\Psi_t|H|\Psi_t\rangle$$

This is rather appealing because of its formal resemblance with the *classical* stationary-action principle (and the ensuing possibility of a Hamiltonian dynamics of the variational parameters[2]) but it seems flawed due to the double ended boundary condition $|\delta\Psi_{t_f}\rangle = |\delta\Psi_{t_i}\rangle = 0$ which is incongruous with a *first* order equation in time (the time-dependent Schrödinger equation) which it is meant to replace (see *e.g.* Ref. [6]). A similar stationarity condition,

$$\Im\,\langle\delta\dot\Psi|(i\hbar\partial_t - H)|\Psi\rangle = 0 \tag{3}$$

defines the MVP which, contrary to the above two, is firmly rooted in purely geometrical ideas. Despite this, McLachlan's principle is perhaps the least popular of the three, firstly because the presence of the time-derivative of the wavefunction variation ($\delta\dot\Psi$) makes it less intuitive, and secondly, because the above mentioned equivalence of the three principles led researchers to focus on the DFVP and the TDVP which admit an immediate physical interpretation. In this Letter we revisit the MVP "geometrical" principle and exploit some basic, hitherto unexplored, consequences. Specifically, we will consider the local-in-time error associated with the MVP and consider its implications for variational propagation schemes.

The McLachlan minimum-distance principle. Let us first introduce some notation. In the following it is assumed that the wavefunctions we deal with lie on a manifold $\mathcal{M}\subseteq\mathcal{H}$ (the "variational manifold") that admits a smooth parametrization, *i.e.,* $|\Psi\rangle \equiv |\Psi(\mathbf{x})\rangle$ where $\mathbf{x}\in\Omega\subseteq\mathbb{R}^n$ and $\partial|\Psi\rangle/\partial x_i$'s, $\partial^2|\Psi\rangle/\partial x_i\partial x_j$'s are well-defined vectors of the Hilbert space $\mathcal{H}$ of the system. For simplicity, we assume that $\mathcal{M}$ contains its rays, in order to allow normalization of the wavefunction. The directional derivative along $\mathbf{u}\in\mathbb{R}^n$ in $\mathbf{x}_0$ is given by

$$|\delta_{\mathbf{u}}\Psi_0\rangle = \left.\frac{d\,|\Psi(\mathbf{x}_0+s\mathbf{u})\rangle}{ds}\right|_{s=0} = \sum_{i=1}^n u_i\left.\frac{\partial\,|\Psi(\mathbf{x})\rangle}{\partial x_i}\right|_{\mathbf{x}=\mathbf{x}_0}$$

and defines a generic "variation" of $|\Psi_0\rangle = |\Psi(\mathbf{x}_0)\rangle$ (*i.e.*, along $\mathbf{u}$). The vectors $|\delta_i\Psi_0\rangle \equiv \partial|\Psi\rangle/\partial x_i|_{\mathbf{x}=\mathbf{x}_0}$ ($i = 1,..n$) span a linear space of dimension n, denoted as $\mathrm{T}_0\mathcal{M}$, which is the space tangent to $\mathcal{M}$ in $|\Psi_0\rangle$. This linear space is *real,* as long as the manifold coordinates are real parameters, which is the most general case. Occasionally, one may make use of complex (analytic) parametrizations, and in that case $\mathrm{T}_0\mathcal{M}$ becomes a *complex* linear space, a sufficient condition for the equivalence of the above variational principles[5]. More generally, we say that the variation $|\delta\Psi_0\rangle\in\mathrm{T}_0\mathcal{M}$ is *complex* whenever the vector $i\,|\delta\Psi_0\rangle$ is a permitted variation, too[7], *i.e.*, $i\,|\delta\Psi_0\rangle\in\mathrm{T}_0\mathcal{M}$.

Suppose we are given $|\Psi_0\rangle\in\mathcal{M}$ as an initial state for

a short-time dynamics of time dt. The best choice for $|\Psi_0(dt)\rangle \in \mathcal{M}$, the time-evolved state, should minimize the error, that is the distance from the exact solution $|\Psi_0^{\text{exact}}(dt)\rangle$, $\varepsilon dt = ||\Psi_0(dt) - \Psi_0^{\text{exact}}(dt)||$ (here written in terms of error per unit time ε) or, equivalently,

$$\hbar\varepsilon = ||i\hbar\dot{\Psi}_0 - H\Psi_0||$$

Stationarity with respect to variations of $|\dot{\Psi}_0\rangle$ gives the McLachlan condition, Eq. 3, for $|\Psi_0\rangle$

$$\Im\langle\delta\dot{\Psi}_0|\,(i\hbar\partial_t - H)\,|\Psi_0\rangle = 0 \tag{4}$$

where $|\delta\dot{\Psi}_0\rangle$ can be thought of as a limiting difference between the tangent vectors of two neighboring paths. The invariance under scalar multiplication directly leads to norm conservation, since for $|\delta\dot{\Psi}_0\rangle = \delta\dot{\lambda}\,|\Psi_0\rangle$ (with $\delta\dot{\lambda}$ arbitrary complex) it gives

$$i\hbar\,\langle\Psi_0|\dot{\Psi}_0\rangle = \langle\Psi_0|H|\Psi_0\rangle \tag{5}$$

which implies $2\Re\langle\Psi_0|\dot{\Psi}_0\rangle = d\,\langle\Psi_0|\Psi_0\rangle/dt = 0$. At the same time, the *gauge* is fixed to $\hbar\Im\langle\Psi_0|\dot{\Psi}_0\rangle = -\langle\Psi_0|H|\Psi_0\rangle$, that is, precisely that of the exact solution, $i\hbar\,|\dot{\Psi}_0^{\text{exact}}\rangle = H\,|\Psi_0\rangle$. The same conclusions follow by taking $\mathcal{M}$ a manifold of *normalized* wavefunctions, but with a free phase factor that is then optimized[8].

Next, we consider the optimization of the path. When the time-dependence in $|\Psi_t\rangle$ comes *only* from variational parameters, $|\delta\dot{\Psi}_0\rangle$ is nothing else that an arbitrary element of $\mathrm{T}_0\mathcal{M}$. In other words, in this case holds

$$\Im\langle\delta\Psi_0|\,(i\hbar\partial_t - H)\,|\Psi_0\rangle = 0 \tag{6}$$

since $\mathrm{T}_0\mathcal{M}$ is a linear space and its elements are just the wavefunction variations. Eq. 6 is only apparently similar to Eq. 2 (though they both reduce to the Dirac-Frenkel condition, Eq. 1, for complex variations). This becomes clear when evaluating it for $|\delta\Psi_0\rangle = |\dot{\Psi}_0\rangle$, the time derivative of the variational solution which is a legitimate element of $\mathrm{T}_0\mathcal{M}$, since Eq. 6 gives

$$\hbar\,\langle\dot{\Psi}_0|\dot{\Psi}_0\rangle = \Im\langle\dot{\Psi}_0|H|\Psi_0\rangle \tag{7}$$

which is a genuine consequence of the McLachlan principle. The same manipulation in the TDVP gives a different (though rather important) condition, namely energy conservation, $\Re\langle\dot{\Psi}_0|H|\Psi_0\rangle = \frac{1}{2}\frac{d}{dt}\langle\Psi_0|H|\Psi_0\rangle = 0$. Eq. 7 gives immediately a "boundedness theorem"

$$\hbar||\dot{\Psi}_0|| \leq ||H\Psi_0|| \tag{8}$$

but it is actually more powerful, as is shown in the following.

Local-in-time error. The value of the distance at the variational minimum, denoted as $\varepsilon_{\mathcal{M}}$,

$$\varepsilon_{\mathcal{M}}[\Psi_0] = \hbar^{-1}\min_{u\in\mathrm{T}_0\mathcal{M}}||i\hbar u - H\Psi_0||$$

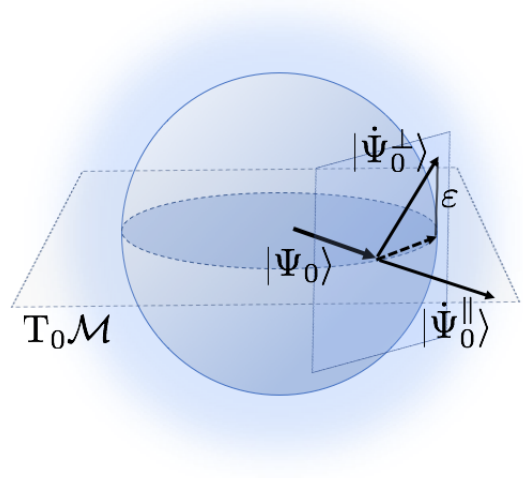

Figure 1. Schematics illustrating the local-in-time error ε when $\mathcal{M}$ contains its rays and $\mathrm{T}_0\mathcal{M}$ is complex-linear. Here, the sphere represents the unit sphere of normalized vectors, and $|\dot{\Psi}_0^{\parallel}\rangle$ and $|\dot{\Psi}_0^{\perp}\rangle \equiv |\dot{\Psi}_0^{+,\text{exact}}\rangle$ are the "irrelevant" and "relevant" components of the exact time derivative, given by $(i\hbar)^{-1}\bar{E}_0\,|\Psi_0\rangle$ and $(i\hbar\Delta E_0)^{-1}(H - \bar{E}_0)\,|\Psi_0\rangle$, respectively. Note that the graphics cannot describe the fact that both components preserve the norm.

is a functional of $|\Psi_0\rangle$, depending on the chosen manifold $\mathcal{M}$. It represents the distance of the manifold $\mathcal{M}$ from the exact solution in $|\Psi_0\rangle$, *i.e.,* a local-in-time measure of the performance of the variational method associated to $\mathcal{M}$. Figuratively, it gives a "skin" of finite thickness to the manifold $\mathcal{M}$ that locally measures the accuracy of the variational method associated to $\mathcal{M}$, for the given dynamical problem. Importantly, it also sets an *a posteriori* upper bound to the wavefunction error[9]

$$||\Psi_0(t) - \Psi_0^{\text{exact}}(t)|| \leq \int_0^t \varepsilon_{\mathcal{M}}[\Psi_0(\tau)]d\tau \tag{9}$$

and can thus be used confidently to minimize the error over time when acting on $\mathcal{M}$ (see Supplemental Material, SM). Using Eq. 7 one easily finds

$$\varepsilon^2_{\mathcal{M}}[\Psi_0] = \frac{1}{\hbar^2}\left(||H\Psi_0||^2 - \hbar^2||\dot{\Psi}_0||^2\right) \tag{10}$$

which is a simple, exact expression for the local-in-time error. When $\mathrm{T}_0\mathcal{M}$ is complex-linear, this is a simple consequence of the fact that the variational condition can be recast as an orthogonal projection[9], namely $i\hbar\,|\dot{\Psi}_0\rangle = \mathcal{P}_0 H\,|\Psi_0\rangle$ where $\mathcal{P}_0$ is the projector onto $\mathrm{T}_0\mathcal{M}$; however, this condition is not necessary for Eq. 10 to hold, when the MVP is used. In the following, we show how $\varepsilon^2_{\mathcal{M}}$ can be used in practice to assess quantitatively the quality of a variational approximation and how to improve it when necessary.

We first rewrite Eq. 10 in a more appealing form, since it is invariant under a shift of the Hamiltonian ($H \rightarrow$

$H^\epsilon = H - \epsilon$) provided, of course, the *gauge* is modified accordingly ($|\Psi_0\rangle \to |\Psi_0^\epsilon\rangle = \exp(+\frac{i}{\hbar}\epsilon t)\,|\Psi_0\rangle$). Hence, it is convenient to choose as reference energy the average energy of the state $|\Psi_0\rangle$, denoted here and in the following as $\bar{E}_0$, resulting in the corresponding "standard" *gauge* $|\Psi_0^\perp\rangle := |\Psi_0^{\bar{E}_0}\rangle$. With this *gauge*, Eq. 10 takes the form

$$\varepsilon^2_{\mathcal{M}}[\Psi_0] = \frac{1}{\hbar^2}\left(\Delta E_0^2 - \hbar^2||\dot{\Psi}_0^+||^2\right) \tag{11}$$

where $\Delta E_0^2 = \langle (H-\bar{E}_0)^2\rangle_0$ is the energy variance and $|\dot{\Psi}_0^+\rangle$ satisfies $\langle\Psi_0|\dot{\Psi}_0^+\rangle = 0$. Again, this admits a simple interpretation since the action of H on a given vector $|\Psi_0\rangle$ can always be split into a component along $|\Psi_0\rangle$ and one orthogonal to it, $|\Psi_0^\perp\rangle$, namely[10]

$$\begin{aligned} H\,|\Psi_0\rangle &= \bar{E}_0\,|\Psi_0\rangle + \Delta E_0\,|\Psi_0^\perp\rangle \\ &= i\hbar\,|\dot{\Psi}_0^\parallel\rangle + i\hbar\,|\dot{\Psi}_0^\perp\rangle \end{aligned}$$

where $|\Psi_0^\perp\rangle = (H-\bar{E}_0)\,|\Psi_0\rangle\,/\Delta E_0$ is a normalized vector orthogonal to $|\Psi_0\rangle$. The two components $|\dot{\Psi}_0^\parallel\rangle$ and $|\dot{\Psi}_0^\perp\rangle$ are, respectively, the "irrelevant" and "relevant" components of the exact time-derivative (see Fig. 1). The latter reduces to the time-derivative of the exact wavefunction in the standard *gauge*, $i\hbar\,|\dot{\Psi}_0^{+,\text{exact}}\rangle = i\hbar\,|\dot{\Psi}_0^\perp\rangle = \Delta E_0\,|\Psi_0^\perp\rangle$, and thus ΔE_0 determines the "intrinsic" length of this derivative. We note that the decomposition of Eq. 11 is different from the approach of Ref. [9] where the error is written in terms of the deviation of the tangent space projection from the exact solution.

Interestingly, when the equations of motion can be recast in the form $i\hbar\,|\dot{\Psi}_0^+\rangle = H_v\,|\Psi_0\rangle$, where H_v is a "variational" (self-adjoint) Hamiltonian operator, the error becomes a measure of the ability of $\mathcal{M}$ to account for the energy fluctuations,

$$\varepsilon^2_{\mathcal{M}}[\Psi_0] = \frac{1}{\hbar^2}\left(\Delta E_0^2 - \Delta E_{v,0}^2\right)$$

where $\Delta E_{v,0}^2 = \langle\Psi_0|H_v^2|\Psi_0\rangle$ is the variance of the "effective" energy[11]. This variational energy variance is bounded, $\Delta E_{v,0}^2 \le \Delta E_0^2$, and attains its maximum value for the exact solution.

Now, upon factoring out ΔE_0^2, which is common to any manifold containing $|\Psi_0\rangle$, we write

$$\varepsilon^2_{\mathcal{M}}[\Psi_0] = \frac{\Delta E_0^2}{\hbar^2}\left(1 - r^2_{\mathcal{M}}[\Psi_0]\right) \;\text{ with }\; r^2_{\mathcal{M}}[\Psi_0] := \frac{\hbar^2||\dot{\Psi}_0^+||^2}{\Delta E_0^2}$$

where we have introduced the dimensionless index $r_{\mathcal{M}}[\Psi_0] \in [0,1]$ (see Eq. 8) with the properties

$$\mathcal{M}' \supseteq \mathcal{M} \implies r_{\mathcal{M}'}[\Psi_0] \ge r_{\mathcal{M}}[\Psi_0]$$

$$r_{\mathcal{M}}[\Psi_0] = 1 \iff \dot{\Psi}_0 = \dot{\Psi}_0^{\text{exact}}$$

We thus see that the ratio $r_{\mathcal{M}}[\Psi_0]$ is a convenient measure of the performance of a variational method for the given dynamical problem.

The above result can be generalized to the case in which the manifold $\mathcal{M}$ is time-dependent, $\mathcal{M} = \mathcal{M}(t)$, and the time-derivative of the wavefunction contains both a variational ($|\dot{\Psi}_v\rangle \in \mathrm{T}_0\mathcal{M}(0)$) and a non-variational ($|\dot{\Psi}_n\rangle$) contribution, *i.e.*, $|\dot{\Psi}_0\rangle = |\dot{\Psi}_v\rangle + |\dot{\Psi}_n\rangle$. In this case energy is not conserved

$$\frac{dE_0}{dt} = 2\Re\,\langle\dot{\Psi}_0|H|\Psi_0\rangle = 2\Re\left(\langle\dot{\Psi}_n|(H - i\hbar\partial_t)\Psi_0\rangle\right)$$

but the error takes yet a simple form

$$\varepsilon^2_{\mathcal{M}}[\Psi_0] = \frac{1}{\hbar^2}\left(||H\Psi_0 - i\hbar\dot{\Psi}_n||^2 - \hbar^2||\dot{\Psi}_v||^2\right)$$

see SM for details.

Examples. As a first example, we consider a simple one-dimensional system whose wavefunction $|\Psi_0\rangle$ is constrained to have a Bargmann form[12, 13], $|\Psi_0\rangle = C\exp\left(za^\dagger\right)|0\rangle$, where the phonon annhilation operator a reads as $a = \frac{\hat{q}}{2\Delta q} + i\frac{\hat{p}}{2\Delta p}$, $\hat{q}$ and $\hat{p}$ being the usual coordinate and momentum operators and Δq, Δp being two parameters satisfying $\Delta q\Delta p = \hbar/2$, and representing, respectively, the coordinate and momentum width of the state. Finally, $|0\rangle$ is the vacuum state ($a\,|0\rangle = 0$) and $C, z \in \mathbb{C}$ parametrize the vector. This is a semiclassical approximation to the dynamics, also known as Frozen Gaussian approximation (FGA)[14], since the variational equations of motion reduce to evolution laws for the average position and momentum of the wavepacket, $q_0 = 2\Delta q\Re z$ and $p_0 = 2\Delta p\Im z$, respectively. A straightforward calculation gives the equation of motion for z (see SM for details), $\dot{z} = i\hbar^{-1}\,\langle\Psi_0|[H,a]|\Psi_0\rangle\,/\,\langle\Psi_0|\Psi_0\rangle$, and the time-derivative of the wavefunction in the standard *gauge*, $|\dot{\Psi}_0^+\rangle = \dot{z}(a^\dagger - z^*)\,|\Psi_0\rangle$. Thus, the error due to the FGA to the dynamics follows as $\hbar^2\varepsilon^2 = \Delta E_0^2 - \hbar^2|\dot{z}|^2$, where (for $H = \frac{p^2}{2m} + V$) the second term on the r.h.s. is just the variance of the following variational Hamiltonian

$$H_v = \frac{p_0}{m}\delta\hat{p} + \langle V'\rangle\,\delta\hat{q}$$

where $\delta\hat{p} = \hat{p} - p_0$, $\delta\hat{q} = \hat{q} - q_0$ and $V' = \frac{dV(\hat{q})}{dq}$. The error is easily seen to vanish when H takes a harmonic form, *i.e.*, $H = H_{\text{HO}} = \hbar\omega a^\dagger a + \lambda a^\dagger + \lambda^* a$ ($\omega \in \mathbb{R}$, $\lambda \in \mathbb{C}$), and in general it reads as, to lowest order in Δq,

$$\hbar\varepsilon \approx \Delta q^2\sqrt{\frac{m^2\Delta^4}{2} + \left(\frac{|V_0^{(3)}|^2}{6} + \frac{m\Delta^2}{2}V_0^{(4)}\right)\Delta q^2}$$

where $V_0^{(n)}$ is the nth derivative of the potential in q_0, $m\Delta^2 = V_0^{(2)} - m\omega^2$, and $\omega = \hbar/2m\Delta q^2$ (see SM). In locally harmonic potentials ($V^{(2)} > 0$), one may set Δq to make the first term on the r.h.s. vanishing and obtain

$\hbar\varepsilon \approx \hbar^3|V_0^{(3)}|/8\sqrt{6}[mV_0^{(2)}]^{3/2}$, although this condition only holds at $t=0$ if Δq is kept frozen.

As a second example, let us consider the general $N-$particle Hamiltonian $H=\sum_{i=1}^N h_i + V$ (where h_i are one-particle operators and V is a many-body interaction potential) and the mean-field *ansatz* of the time-dependent Hartree method, $|\Psi_0\rangle = \Pi_{i=1}^N |\phi_i\rangle$ where the ϕ_i's are variational single-particle functions (spf's), subject only to the normalization condition $\langle\phi_i|\phi_i\rangle = 1$. Application of the DF condition, Eq. 6, gives the equations of motion of the spf's in the form (SM)

$$i\hbar\,|\dot{\phi}_i\rangle = \left(H_i + g_i - \bar{E}_0\right)|\phi_i\rangle$$

where $H_i = \langle\Psi^i|H|\Psi^i\rangle$ is the mean-field Hamiltonian for the ith degree of freedom ($|\Psi^i\rangle = \Pi_{j\neq i}|\phi_j\rangle$ is the ith single-hole wavefunction) and $g_i = i\hbar\,\langle\phi_i|\dot{\phi}_i\rangle \in \mathbb{R}$ are arbitrary *gauge* terms that enforce the normalization conditions. As shown in SM, the total time-derivative of the state vector in the standard *gauge* follows as

$$i\hbar\,|\dot{\Psi}_0^+\rangle = H^0_{\text{mf}}|\Psi_0\rangle \quad H^0_{\text{mf}} = \sum_{i=1}^N \left(H_i - \bar{E}_0\right), \quad \langle H_{\text{mf}}\rangle = 0$$

(here H^0_{mf} is the appropriate variational Hamiltonian for the problem) and thus it holds $\hbar^2||\dot{\Psi}_0^+||^2 = \Delta E^2_{\text{mf},0}$, where $\Delta E^2_{\text{mf},0} = \sum_{i=1}^N \Delta E^2_{i,0}$ and $\Delta E^2_{i,0} = \langle\left(H_i - \bar{E}_0\right)^2\rangle_0$ are the one-particle energy fluctuations. Furthermore, since $H - \bar{E}_0 = H_{\text{mf}} + \Delta V$, where $\Delta V = V + (N-1)\langle V\rangle - \sum_{i=1}^N v_i$ is the zero-mean fluctuating potential, the energy variance can be given in a simple form (here $\Delta V_0^2 = \langle\Delta V^2\rangle_0$)

$$\Delta E_0^2 = \Delta E^2_{\text{mf},0} + \Delta V_0^2 + 2\sum_{i=1}^N \Re\,\langle H_i\Delta V\rangle_0$$

and $r^2_{\text{mf}}[\Psi_0] \equiv \Delta E^2_{\text{mf},0}/\Delta E_0^2$. The above expression clearly shows the key role played by the potential energy fluctuations in limiting the reliability of the mean-field approach and indicates that

$$\varepsilon_{\text{mf}} = \hbar^{-1}\left[\Delta V_0^2 + 2\sum_{i=1}^N \Re\,\langle H_i\Delta V\rangle_0\right]^{1/2}$$

is the appropriate expression for the correlation error intrinsic in the TDH method. Notice that from the inequality $\sum_{i=1}^N \Re\,\langle H_i\Delta V\rangle_0 = \Re\,\langle(H_{\text{mf}}\Delta V\rangle_0 \leq \Delta E_{\text{mf},0}\Delta V_0$ follows a simple lower bound for the $r-$index, namely $r_{\text{mf}}[\Psi_0] \geq \Delta E_{\text{mf},0}/(\Delta E_{\text{mf},0} + \Delta V_0)$.

Finally, as a last example we consider the error intrinsic to the adiabatic (Born-Oppenheimer) dynamics, a common strategy to tackle molecular problems where the electronic degrees of freedom are averaged out with the well-known *ansatz*

$$|\Psi_0\rangle = \int d\mathbf{X}\psi(\mathbf{X})\,|\Phi_n(\mathbf{X})\rangle\,|\mathbf{X}\rangle$$

Here $\mathbf{X}$ represents the nuclear degrees of freedom, and $|\Phi_n(\mathbf{X})\rangle$ is the nth eigenstate of the electronic Hamiltonian with clamped nuclei at $\mathbf{X}$, *i.e.*, the electronic operator $h_{\text{el}}(\mathbf{X})$ defined by $\langle\mathbf{X}|H-T|\mathbf{X}'\rangle = h_{\text{el}}(\mathbf{X})\delta(\mathbf{X}-\mathbf{X}')$, H being the total Hamiltonian and T the kinetic energy of the nuclei. Application of the variational principle gives the equation of motion for the "nuclear wavefunction" $\psi(\mathbf{X})$ in the nth electronic state

$$H_n\psi = i\hbar\frac{\partial\psi}{\partial t}, \qquad H_n = \langle T\rangle_n + E_n(\mathbf{X})$$

where $E_n(\mathbf{X})$ is the electronic energy and

$$\begin{aligned}\langle T\rangle_n =& T - i\sum_i \frac{\hbar}{M_i}\left\langle\Phi_n|\frac{\partial\Phi_n}{\partial R_i}\right\rangle_{\text{el}} P_i + \\ &- \sum_i \frac{\hbar^2}{2M_i}\left\langle\Phi_n|\frac{\partial^2\Phi_n}{\partial R_i^2}\right\rangle_{\text{el}}\end{aligned}$$

is a self-adjoint operator, the nuclear kinetic energy operator averaged over the electronic state[15]. This gives the rate of variation of the wavefunction in the standard *gauge* as

$$\hbar^2||\dot{\Psi}_0^+||^2 = \int d\mathbf{X}\psi^*(\mathbf{X})\left[\langle T\rangle_n + (E_n(\mathbf{R}) - \bar{E}_0)\right]^2\psi(\mathbf{X})$$

while the energy variance reads as

$$\Delta E_0^2 = \int d\mathbf{X}\psi^*(\mathbf{X})\,\langle\left[T + (h_{\text{el}}(\mathbf{X}) - \bar{E}_0)\right]^2\rangle_n\,\psi(\mathbf{X})$$

Hence, the local-in-time error in the adiabatic approximation takes the form of a nuclear kinetic energy fluctuation term

$$\varepsilon^2[\Psi_0] = \frac{1}{\hbar^2}\int d\mathbf{X}\psi^*(\mathbf{X})\left[\langle T^2\rangle_n - \langle T\rangle_n^2\right]\psi(\mathbf{X})$$

This can also be put in a form that makes explicit the contributions of electronic transitions, that is, upon introducing $\phi_{m\leftarrow n}(\mathbf{X}) = \langle\Phi_m|T|\Phi_n\rangle_{\text{el}}\,\psi(\mathbf{X})$,

$$\varepsilon^2[\Psi_0] = \frac{1}{\hbar^2}\sum_{m\neq n}\int d\mathbf{X}|\phi_{m\leftarrow n}(\mathbf{X})|^2$$

Here, the amplitudes read explicitly as

$$\begin{aligned}\phi_{m\leftarrow n}(\mathbf{X}) = -\sum_{i,\alpha}\frac{\hbar^2}{2M_i}&\left[\frac{F^{i,\alpha}_{mn}(\mathbf{X})}{\Delta E_{mn}(\mathbf{X})}\frac{\partial\psi(\mathbf{X})}{\partial\mathbf{X}_{i,\alpha}}\right.\\ &\left.+B^{i,\alpha}_{mn}(\mathbf{X})\psi(\mathbf{X})\right]\end{aligned}$$

where $\Delta E_{mn} = E_m - E_n$, i and α label the nuclei and their coordinates, respectively, $F^{i,\alpha}_{mn} = \langle\Phi_m|F^{i,\alpha}|\Phi_n\rangle$ where $F^{i,\alpha}$ is the operator for the α component of the force acting on the nucleus i, and $B^{i,\alpha}_{mn} = \left\langle\Phi_m|\frac{\partial^2\Phi_n}{\partial\mathbf{X}^2_{i,\alpha}}\right\rangle$.

Adaptive propagation schemes. Eq. 11 represents a rigorous criterion to optimize *on-the-fly* the computational cost of a quantum dynamical simulation, as it

can be used to re-size the underlying variational manifold in order to keep the error below a specified "tolerable" value (see also Eq. 9). We sketch here its application to a rather popular and quite efficient variational method for high-dimensional systems, the multiconfiguration time-dependent Hartree (MCTDH) method [16–19]. In this method the wavefunction takes the form $|\Psi_0\rangle = \sum_I C_I |\Phi_I\rangle$ where C_I's are complex coefficients, $I = (i_1, i_2, ..i_N)$ is a multi-index and $|\Phi_I\rangle = |\phi_{i_1}\phi_{i_2}..\phi_{i_N}\rangle$ (where $i_k = 1, ..n_K$) are configurations of fully flexible spf's. Of interest here is the possibility of changing *on-the-fly* the number of spfs, which means varying both the size of the secular problem for the amplitude coefficients and the number of spfs to be optimized. Notice that this would solve from the outset the problem of regularizing solutions that contain configurations with vanishing weight. We focus on the "spawning" process[20], *i.e.* the generation of new spfs and related configurations, which becomes necessary when, in the course of the dynamics, the local error ε exceeds some given threshold, thereby signaling the need for a more flexible manifold. If the main correction comes from single excitations of the "occupied" configurations $|\Phi_I\rangle$, the "best" spf $|\eta\rangle$ to add to the kth degree of freedom is the one the maximizes the expectation value of a certain reduced, self-adjoint "rate" operator $\Gamma^{(k)}$ for the kth mode (see SM), among those single-particle states that lie in the orthogonal complement of both the occupied spfs for the kth mode ($|\phi_{i_k}\rangle$, $i_k = 1, n_k$) and their time-derivatives. The reduced operator reads as

$$\Gamma^{(k)} = \sum_{I(k)} \langle \Phi_{I(k)}|H|\Psi_0\rangle \langle \Psi_0|H|\Phi_{I(k)}\rangle$$

where $\Phi_{I(k)}$ is a kth hole configuration and the scalar products are taken over all modes except the kth. Then, the reduction of the local-in-time (squared) error when adding such spf is given by $\langle \eta|\Gamma^{(k)}|\eta\rangle / \hbar^2$ (see SM for details).

Conclusions. Variational solutions of the time-dependent Schrödinger equation have an intrinsic measure of their reliability, a local-in-time error that measures the departure from the instantaneous exact solution. Simple expressions have been provided for this error in some relevant cases, with the aim of showing how the error helps to assess quantitatively the reliability of the variational method for a given dynamical problem. Future applications involve the development of adaptive propagation schemes that re–size *on-the-fly* the variational manifold, and optimize the computational cost for a target accuracy.

* rocco.martinazzo@unimi.it

[1] A. McLachlan, Molecular Physics **8**, 39 (1964).

[2] P. Kramer and M. Saraceno, *Geometry of the time-dependent variational principle in quantum mechanics* (Springer-Verlag, 1981) p. 98.

[3] P. A. M. Dirac and P. A. M., Mathematical Proceedings of the Cambridge Philosophical Society **26**, 376 (1930).

[4] J. Frenkel, *Wave Mechanics, Advanced General Theory* (Clarendon Press, Oxford, 1934) p. 524.

[5] J. Broeckhove, L. Lathouwers, E. Kesteloot, and P. Van Leuven, Chemical Physics Letters **149**, 547 (1988).

[6] G. Vignale, Physical Review A **77**, 062511 (2008).

[7] This condition is satisfied by *any* variation when $\mathrm{T}_0\mathcal{M}$ happens to be complex-linear. The converse is also true, that is if $i \times \mathrm{T}_0\mathcal{M} = \mathrm{T}_0\mathcal{M}$ then $\mathrm{T}_0\mathcal{M}$ is complex-linear.

[8] It is worth emphasizing that, under such circumstances, the above defined "differential" distance depends on both the manifold (*i.e.,* the shape of the trial wavefunction) *and* the *gauge*. This becomes obvious when considering the *gauge* transformation $|\Psi_t\rangle = e^{i\Theta t}|\bar{\Psi}_t\rangle$ (where $|\bar{\Psi}_{t=0}\rangle = |\Psi_0\rangle$) and computing the time-derivative at $t = 0$, $|\dot{\Psi}_0\rangle = i\Theta|\Psi_0\rangle + |\dot{\bar{\Psi}}_0\rangle$. Optimization of the *gauge* can be achieved by considering the appropriate variation $|\delta\dot{\Psi}_0\rangle = i\delta\Theta|\Psi_0\rangle$ in Eq. 4, and results in the condition $-\hbar\Im\langle\Psi_0|\dot{\Psi}_0\rangle = \langle\Psi_0|H|\Psi_0\rangle$ that can be combined with norm conservation to give Eq. 5.

[9] C. Lubich, *From Quantum to Classical Molecular Dynamics: Reduced Models and Numerical Analysis (Zurich Lectures in Advanced Mathematics)* (European Mathematical Society, 2008) p. 156.

[10] E. Pollak, Journal of Chemical Theory and Computation **15**, 1498 (2019).

[11] This happens, for instance, when $\mathrm{T}_0\mathcal{M}$ is complex linear and $\mathcal{M}$ contains its rays. In that case, $H_v \equiv \mathcal{P}_0(H - \bar{E}_0)\mathcal{P}_0$ and the error can also be written in the form $\hbar^2\varepsilon^2_{\mathcal{M}}[\Psi_0] = \langle\Psi_0|(H - \bar{E}_0)\mathcal{Q}_0(H - \bar{E}_0)|\Psi_0\rangle$ where $\mathcal{Q}_0 = 1 - \mathcal{P}_0$.

[12] V. Bargmann, Communications on Pure and Applied Mathematics **14**, 187 (1961).

[13] C. W. C. W. Gardiner and P. P. Zoller, *Quantum noise : a handbook of Markovian and non-Markovian quantum stochastic methods with applications to quantum optics* (Springer, 2004) p. 449.

[14] E. J. Heller, The Journal of Chemical Physics **75**, 2923 (1981).

[15] Here, for the second term on the r.h.s. it holds $\left\langle \Phi_n | \frac{\partial \Phi_n}{\partial R_i} \right\rangle_{\text{el}} \equiv \Im \left\langle \Phi_n | \frac{\partial \Phi_n}{\partial R_i} \right\rangle_{\text{el}}$ because of norm conservation. This term vanishes in the presence of time-reversal invariance (*i.e.*, when the electronic wavefunctions can be chosen globally real).

[16] H.-D. Meyer, U. Manthe, and L. S. Cederbaum, Chem. Phys. Lett. **165**, 73 (1990).

[17] M. H. Beck, A. Jäckle, G. A. Worth, and H.-D. Meyer, Phys. Rep. **324**, 1 (2000).

[18] H.-D. Meyer, F. Gatti, and G. A. Worth, eds., *Multidimensional Quantum Dynamics: MCTDH Theory and Applications* (Wiley-VCH, Weinheim, 2009).

[19] H.-D. Meyer, Wiley Interdisciplinary Reviews: Computational Molecular Science **2**, 351 (2012).

[20] D. Mendive-Tapia, T. Firmino, H.-D. Meyer, and F. Gatti, Chem. Phys. **482**, 113 (2017).

SUPPLEMENTAL MATERIAL

A posteriori error bound

Following Ref. [9], let $|\Psi(t)\rangle$ and $|\tilde{\Psi}(t)\rangle$ be, respectively, an approximate and the exact solution of the TDSE with the same initial state, $|\Psi(0)\rangle = |\tilde{\Psi}(0)\rangle \equiv |\Psi_0\rangle$ and $|\Delta\Psi\rangle = |\Psi(t)\rangle - |\tilde{\Psi}(t)\rangle$. From the identity

$$i\hbar\,|\Delta\dot{\Psi}\rangle - H\,|\Delta\Psi\rangle = i\hbar\,|\dot{\Psi}\rangle - H\,|\Psi\rangle$$

it follows

$$\hbar\Re\,\langle\Delta\Psi|\Delta\dot{\Psi}\rangle = \Im\,\langle\Delta\Psi|i\hbar\dot{\Psi} - H\Psi\rangle$$

Here, $\Re\,\langle\Delta\Psi|\Delta\dot{\Psi}\rangle = \frac{1}{2}\frac{d}{dt}||\Delta\Psi||^2 = ||\Delta\Psi||\frac{d||\Delta\Psi||}{dt}$ and thus

$$\hbar||\Delta\Psi||\frac{d||\Delta\Psi||}{dt} = \Im\,\langle\Delta\Psi|i\hbar\dot{\Psi} - H\Psi\rangle \leq ||\Delta\Psi||\;||i\hbar\dot{\Psi} - H\Psi||$$

i.e.,

$$\frac{d}{dt}||\Delta\Psi|| \leq \frac{1}{\hbar}||i\hbar\dot{\Psi} - H\Psi||$$

which integrated gives

$$||\Delta\Psi|| \leq \frac{1}{\hbar}\int_0^t ||i\hbar\dot{\Psi}(\tau) - H\Psi(\tau)||d\tau$$

When $|\Psi(t)\rangle \in \mathcal{M}$ is a variational solution the integrand on the r.h.s. takes at any time its minimum value and it is just the local-in-time error $\varepsilon_{\mathcal{M}}[\Psi(t)]$ defined in the main text, Eq. 10.

The above bound also contraints the error in autocorrelation functions (here and below $||\Psi_0|| = 1$)

$$|\,\langle\Psi(t)|\Psi_0\rangle - \langle\tilde{\Psi}(t)|\Psi_0\rangle\,| = |\,\langle\Delta\Psi|\Psi_0\rangle\,| \leq ||\Delta\Psi||$$

and in the average values of any bounded observable,

$$\begin{aligned}|\,\langle\Psi(t)|A|\Psi(t)\rangle - \langle\tilde{\Psi}(t)|A|\tilde{\Psi}(t)\rangle\,| &= \\ = |\,\langle\Delta\Psi|A|\Psi(t)\rangle + \langle\tilde{\Psi}(t)|A|\Delta\Psi\rangle\,| \\ \leq ||\Delta\Psi||\left(||A\Psi|| + ||A\tilde{\Psi}||\right) \\ \leq 2||A||_\infty||\Delta\Psi||\end{aligned}$$

where $||A||_\infty$ is the operator norm.

Error and energy drift with time-dependent manifolds

We address here in some detail the situation where the manifold $\mathcal{M}$ is time-dependent and the time-derivative of the wavefunction contains both a variational and a non-variational contribution

$$|\dot{\Psi}_0\rangle = |\dot{\Psi}_0^T\rangle + |\dot{\Phi}_0\rangle$$

(here the superscript T reminds us that $|\dot{\Psi}_0^T\rangle \in \mathrm{T}_0\mathcal{M}$, the space tangent to $\mathcal{M}(t)$ at $t = 0$) . This may happen, for instance, when the manifold is described by a set of variational parameters $x_1, x_2, ..x_N$ and a number of additional time-dependent parameters $y_1, y_1, ..y_M$ which, for computational efficiency, are evolved according to some physically sound law ("guided" parameters), simpler than the variational equations of motion. In such circumstances, the (partial) variational condition $\Im\,\langle\dot{\Psi}_0^T|(i\hbar\partial_t - H)\Psi_0\rangle = 0$ leads to

$$\hbar\,\langle\dot{\Psi}_0^T|\dot{\Psi}_0^T\rangle = \Im\,\langle\dot{\Psi}_0^T|H\Psi_0 - i\hbar\dot{\Phi}_0\rangle$$

which generalizes Eq. 7. Hence, for the error it follows

$$\langle\Big(i\hbar\dot{\Phi}_0 - H\Psi_0\Big) + i\hbar\dot{\Psi}_0^T|\,\Big(i\hbar\dot{\Phi}_0 - H\Psi_0\Big) + i\hbar\dot{\Psi}_0^T\rangle =$$
$$||i\hbar\dot{\Phi}_0 - H\Psi_0||^2 + \hbar^2||\dot{\Psi}_0^T||^2 - 2\hbar\Im\,\langle\dot{\Psi}_0^T|i\hbar\dot{\Phi}_0 - H\Psi_0\rangle$$

and thus

$$\varepsilon^2_{\mathcal{M}}[\Psi_0] = \frac{1}{\hbar^2}\left(||H\Psi_0 - i\hbar\dot{\Phi}_0||^2 - \hbar^2||\dot{\Psi}_0^T||^2\right)$$

(cfr. Eq. 10) and the inequality

$$\hbar||\dot{\Psi}_0^T|| \leq ||H\Psi_0 - i\hbar\dot{\Phi}_0||$$

that can be considered a generalization of the boundedness theorem above to the case in which the manifold is time-dependent. Here, the appearance of $i\hbar\dot{\Phi}_0$ on the r.h.s. of the inequality can be understood in the limiting case where the non-variational time-derivative comes from an effective Hamiltonian, *i.e.* $i\hbar\,|\dot{\Phi}_0\rangle = H^{\mathrm{eff}}\,|\Psi_0\rangle$, since in such case the above inequality reduces to

$$\hbar||\dot{\Psi}_0^T|| \leq ||(H - H^{\mathrm{eff}})\Psi_0||$$

a rather reasonable result.

It is instructive at this point to consider these results in view of the energy conservation since when the wavefunction contains "guided" parameters energy is no longer conserved. Thus in the following we assume that three variational principles are equivalent to each other on $\mathcal{M}$ and consider the energy change per unit time

$$\begin{aligned}W_0 = \frac{dE_0}{dt} &= 2\Re\,\langle\dot{\Psi}_0|H|\Psi_0\rangle \\ &= 2\Re\,\langle\dot{\Psi}_0^T|H|\Psi_0\rangle + 2\Re\,\langle\dot{\Phi}_0|H|\Psi_0\rangle \\ &= 2\Re\left(\langle\dot{\Phi}_0|(H - i\hbar\partial_t)\Psi_0\rangle\right)\end{aligned}$$

where the last equality follow from the Dirac-Frenkel condition $\langle\dot{\Psi}_0^T|(i\hbar\partial_t - H)\Psi_0\rangle = 0$, namely from

$$\langle\dot{\Psi}_0^T|H|\Psi_0\rangle = i\hbar\,\langle\dot{\Psi}_0^T|\dot{\Psi}_0\rangle = i\hbar\,\langle\dot{\Psi}_0|\dot{\Psi}_0\rangle - i\hbar\,\langle\dot{\Phi}_0|\dot{\Psi}_0\rangle$$

When optimizing also w.r.t. $|\delta\dot{\Phi}_0\rangle$, the above equation shows that the (magnitude of the) energy drift is stationary at the variational minimum

$$\delta W_0 = -2\Re\left(\langle\delta\dot{\Phi}_0|(i\hbar\partial_t - H)\Psi_0\rangle\right) = 0$$

a trivial result because we already known that $|W_0|$ is actually at its minimum under such circumstances ($|W_0| = 0$), but, in general, it shows that optimizing the guide (under given constraints) minimizes the energy dritft. In this context it is worth noticing that for a variational solution it must hold

$$|W_0| \leq 2||\dot{\Phi}_0|| \; ||(i\hbar\partial_t - H)\Psi_0|| = 2\hbar\varepsilon_{\mathcal{M}}[\Psi_0]||\dot{\Phi}_0||$$

that can be converted into a lower bound on the variational solution in terms of energy drift,

$$\varepsilon_{\mathcal{M}}[\Psi_0] \geq \frac{|W_0|}{2\hbar||\dot{\Phi}_0||}$$

Thus, optimization of the guide (minimization of $|W_0|$) effectively lowers the bound by reducing the error contribution due to the non-conservation of the energy.

Mean-field approximation

Let us consider the general $N-$particle Hamiltonian $H = \sum_{i=1}^{N} h_i + V$, where h_i are one-particle operators and V is a many-body interaction potential, and the mean-field *ansatz* of the time-dependent Hartree method,

$$|\Psi_0\rangle = \Pi_{i=1}^{N} |\phi_i\rangle$$

where ϕ_i's are variational single-particle functions, subjected only to the normalization condition $\langle\phi_i|\phi_i\rangle = 1$ that is enforced through the *guage* terms $i\hbar\langle\phi_i|\dot{\phi}_i\rangle = g_i \in \mathbb{R}$. Application of the DF condition, Eq. 6, gives the equations of motion of the spf's. To this end, it is worth noticing that it suffices to consider only the special (complex) spf's variations satisfying $\langle\delta\phi_i|\phi_i\rangle = 0$ (*i.e.* $|\delta\phi_i\rangle \in V_i' \equiv \{|\phi_i\rangle\}^{\perp}$) along with the Dirac-Frenkel condition (Eq. 6) since the general stationary condition adds nothing (this is evident upon introducing the projector $P_i = |\phi_i\rangle\langle\phi_i|$ and noticing that $\Re\langle\delta_i\Psi_0|P_i\,(i\hbar\partial_t - H)\,|\Psi_0\rangle \equiv 0$ when $\Re\langle\delta\phi_i|\phi_i\rangle = 0$ and $i\hbar\langle\phi_k|\dot{\phi}_k\rangle = g_k \in \mathbb{R}$).

Thus, the requirement $\left(i\hbar|\dot{\phi}_i\rangle - H_i|\phi_i\rangle\right) \in V_i^{\perp} = \{|\phi_i\rangle\}^{\perp\perp}$ gives $i\hbar|\dot{\phi}_i\rangle - H_i|\phi_i\rangle = \alpha|\phi_i\rangle$, where α is easily found to be $\alpha = i\hbar\langle\phi_i|\dot{\phi}_i\rangle - \bar{E}_0 \equiv g_i - \bar{E}_0$, and the equations of motion take the form

$$i\hbar|\dot{\phi}_i\rangle = \left(H_i + g_i - \bar{E}_0\right)|\phi_i\rangle$$

where $H_i = \langle\Psi^i|H|\Psi^i\rangle$ is the mean-field Hamiltonian for the ith degree of freedom ($|\Psi^i\rangle = \Pi_{j\neq i}|\phi_j\rangle$ is the ith single-hole wavefunction) and $E = \langle\Psi|H|\Psi\rangle \equiv \langle H_i\rangle$. It follows that the total time-derivative of the state vector satisfies

$$i\hbar|\dot{\Psi}\rangle = H_{\rm mf}|\Psi\rangle$$

where the mean-field (total) Hamiltonian $H_{\rm mf}$ reads as

$$H_{\rm mf} = \sum_{i=1}^{N}(H_i + g_i - \bar{E}_0)$$

The optimal *gauge* condition on the total wavefunction requires $\sum_i g_i = 0$ and thus, introducing now the initial time $t = 0$,

$$i\hbar|\dot{\Psi_0^+}\rangle = H_{\rm mf}^0|\Psi_0\rangle \quad H_{\rm mf}^0 = \sum_{i=1}^{N}\left(H_i - \bar{E}_0\right), \quad \langle H_{\rm mf}\rangle = 0$$

Now, the mean-field Hamiltonians H_i read as $H_i = h_i + \sum_j \epsilon_j - \epsilon_i + v_i$ (where $\epsilon_i = \langle\phi_i|h_i|\phi_i\rangle$ is the average one-particle energy on the ith degree and $v_i = \langle\Psi^i|V|\Psi^i\rangle$ is the ith mean-field potential) hence it is easy to check that it holds $H - \bar{E}_0 = H_{\rm mf} + \Delta V$ where

$$\Delta V = V + (N-1)\bar{V} - \sum_{i=1}^{N} v_i$$

is the zero-mean fluctuating potential ($\bar{V} = \langle V\rangle \equiv \langle v_i\rangle$ for any i). Thus,

$$\Delta E_0^2 = \Delta E_{\rm mf,0}^2 + \langle\Delta V^2\rangle_0 + 2\sum_{i=1}^{N}\Re\langle H_i\Delta V\rangle_0$$

and

$$r_{\rm mf}^2[\Psi_0] = \frac{\Delta E_{\rm mf,0}^2}{\Delta E_{\rm mf,0}^2 + \Delta V_0^2 + 2\sum_{i=1}^{N}\Re\langle H_i\Delta V\rangle_0}$$

where $\Delta V_0^2 = \langle\Delta V^2\rangle_0$ and $\Delta E_{\rm mf,0}^2 \equiv \sum_{i=1}^{N}\Delta E_{i,0}^2$, being $\Delta E_{i,0}^2 = \langle\left(H_i - \bar{E}_0\right)^2\rangle_0$ the one-particle energy fluctuations (one may further notice that they consist of both a "kinetic" and a "potential" term, since $H_i - E = (h_i - \epsilon_i) + \left(v_i - \bar{V}\right)$).

Coherent state (or Frozen Gaussian) approximation

We detail here the case of a coherent state approximation to the dynamics by considering a situation slightly more general than the one presented in the main text, namely a system with two degrees of freedom to which we apply the mean-field approximation

$$|\Psi_0\rangle = |\phi_1\rangle|\phi_2\rangle$$

and force the single particle function of the second degree to take the form of a normalized coherent-state (CS)

$$|\phi_2\rangle \equiv |\theta, z\rangle = \exp\left(i\theta - \frac{|z|^2}{2} + za^{\dagger}\right)|0\rangle$$

This wavefunction is a "precusor" of the Ehrenfest method, with $|\phi_1\rangle$ describing an "electronic" system and

$|\phi_2\rangle$ a "semiclassical" nuclear degree of freedom. The equation of motion for $|\phi_2\rangle$ (or, better, z) can be derived either from the Dirac-Frenkel condition, Eq. 1, or from the McLachlan minimum-distance condition, Eq. 4. For illustrative purposes we follow the second route, and consider

$$\langle\phi_1\delta\dot{\phi}_2|\left[i\hbar\,|\dot{\phi}_1\phi_2\rangle + i\hbar\,|\phi_1\dot{\phi}_2\rangle - H\,|\phi_1\phi_2\rangle\right] = 0$$

which is the appropriate condition for optimizing $|\phi_2\rangle$. Notice that, though not evident from the chosen parametrization, the CS variations can be considered complex, as seen by considering the unnormalized Bargmann vectors $C\exp(za^\dagger)\,|0\rangle$ and the complex analytic parametrization $(C,z)\in\mathbb{C}^2$ (as mentioned in the main text). Some lenghty but simple algebra leads to

$$\begin{aligned}-i\delta\dot{\theta}\left\{-\hbar\dot{\theta} - \hbar\Im\left(\dot{z}z^*\right) + g - \bar{E}_0\right\} + \\ \delta\dot{z}^*\left\{i\hbar\dot{z} + \frac{z}{2}\left[-\hbar\dot{\theta} - \hbar\Im\left(\dot{z}z^*\right) + g + \bar{E}_0\right] - \langle\phi_2|aH_{\rm cl}|\phi_2\rangle\right\} + \\ \delta\dot{z}\left\{\frac{z^*}{2}\left[\hbar\dot{\theta} + \hbar\Im\left(\dot{z}z^*\right) - g + \bar{E}_0\right]\right\} = 0\end{aligned}$$

where $g = i\hbar\,\langle\phi_1|\dot{\phi}_1\rangle$, $H_{\rm cl} = \langle\phi_1|H|\phi_1\rangle$ and $\bar{E}_0 = \langle\Psi_0|H|\Psi_0\rangle$. Hence, the optimal gauge θ (a concept that only becomes meaningful in view of computing an error) is such that

$$\hbar\dot{\theta} = -\hbar\Im\left(\dot{z}z^*\right) + g - \bar{E}_0$$

and the stationary condition reduces to

$$i\hbar\dot{z} + \bar{E}_0 - \langle\phi_2|aH_{\rm cl}|\phi_2\rangle = 0$$

i.e.,

$$\dot{z} = \frac{i}{\hbar}\,\langle\phi_2|[H_{\rm cl},a]|\phi_2\rangle$$

It follows

$$i\hbar\,|\dot{\Psi}_0\rangle = \left[(H_q + g - \bar{E}_0) + (-\hbar\dot{\theta} - i\hbar\Re\left(\dot{z}z^*\right) + i\hbar\dot{z}a^\dagger)\right]|\Psi_0\rangle$$

(where $H_q = \langle\phi_2|H|\phi_2\rangle$) and thus, upon replacing $\dot{\theta}$ with its optimal value,

$$i\hbar\,|\dot{\Psi}_0\rangle \equiv \left[H_q + i\hbar\dot{z}(a^\dagger - z^*)\right]|\Psi_0\rangle$$

in such a way that it holds $i\hbar\,\langle\Psi_0|\dot{\Psi}_0\rangle = \bar{E}_0$ as required by the minimum-distance principle. Note that the *gauge* term θ is irrelevant for the parameter dynamics, and can be safely neglected when deriving the equation of motion for z from the Dirac-Frenkel condition. However, such a term is needed in order to make $|\dot{\Psi}_0\rangle$ appropriate for computing the error, and needs to be obtained separately when using the Dirac-Frenkel principle.

Finally, with the Hamiltonian referenced to E, we write the variational error using Eq. 11 where now

$$i\hbar\,|\dot{\Psi}_0^+\rangle = \left[H_q - \bar{E}_0 + i\hbar\dot{z}(a^\dagger - z^*)\right]|\Psi_0\rangle$$

leads to the simple expression

$$\hbar^2||\Psi_0^+||^2 = \langle\Psi_0|\left(H_q - \bar{E}_0\right)^2|\Psi_0\rangle + \hbar^2|\dot{z}|^2$$

By comparing this expression with the above one obtained for the general TDH case one finds that

$$\varepsilon_z^2 = \langle\Psi_0|\left(\frac{H_{\rm cl} - \bar{E}_0}{\hbar}\right)^2|\Psi_0\rangle - |\dot{z}|^2 \geq 0$$

is the genuine error due to the coherent-state approximation. The case considered in the main text can be obtained by setting $|\phi_1\rangle \equiv 1$, $g = 0$, $H_q = \bar{E}_0$ and $H_{\rm cl} = H$.

Local-in-time error in the FGA

The local-in-time error derived above, $\hbar^2\varepsilon^2 = \Delta E_0^2 - \hbar^2|\dot{z}|^2$, is easily seen to vanish when the Hamiltonian takes a harmonic oscillator (HO) form, $H = H_{\rm HO} = \hbar\omega a^\dagger a + 2\hbar\Re(\lambda a^\dagger)$. This rather general result in this context follows easily by observing that, on the one hand, it holds

$$\dot{z}_{\rm HO} = -i\left(\omega a + \lambda\right)$$

and, on the other hand,

$$\left(H_{\rm HO} - E_{\rm HO}\right)|z\rangle = \hbar\left(\omega z + \lambda\right)\left(a^\dagger - z^*\right)|z\rangle$$

In view of the above, we write $H = \frac{p^2}{2m} + V = H_{\rm HO} + W$, where W is assumed to be local, $W = W(\hat{q})$ (see below). We find, on the one hand, $\dot{z} = \dot{z}_{\rm HO} - \frac{i}{2\Delta p}\,\langle W'\rangle$ and, on the other hand,

$$\Delta E_0^2 = \Delta E_{\rm HO}^2 + \Delta W^2 + 2\Re\,\langle z|\left(W - \langle W\rangle\right)\left(H_{\rm HO} - E_{\rm HO}\right)|z\rangle$$

The last term on the r.h.s. can be rearranged into

$$\begin{aligned}\Re\,\langle z|\left(W - \langle W\rangle\right)\left(H_{\rm HO} - E_{\rm HO}\right)|z\rangle = \\ \hbar\Re\left[\left(\omega z + \lambda\right)\langle z|\left(W - \langle W\rangle\right)a^\dagger|z\rangle\right]\end{aligned}$$

where $\langle z|\left(W - \langle W\rangle\right)a^\dagger|z\rangle \equiv \frac{\hbar}{2\Delta p}\,\langle W'\rangle$. It follows

$$\Delta E_0^2 - \hbar^2|\dot{z}|^2 = \Delta W^2 - \Delta x^2\,\langle W'\rangle^2$$

Next, we choose $H_{\rm HO}$ such that $H - H_{\rm HO}$ is a purely local potential. To this end we set $\Delta q^2 = \hbar/2m\omega$ and, for $H_{\rm HO}$ in the form $H_{\rm HO} = \frac{p^2}{2m} + \frac{m\omega^2}{2}(q - \bar{q})^2$, we obtain $W = V - \frac{m\omega^2}{2}\left(q - \bar{q}\right)^2$ and then fix $\bar{q}$ by enforcing the condition $\langle W'\rangle = 0$, *i.e.*,

$$\bar{q} = q_0 - \frac{\langle V'\rangle}{m\omega^2}$$

where $q_0 = \langle z|\hat{q}|z\rangle$. This reduces the problem of finding the error to that of computing ΔW^2 . Upon using the condition above $m\omega^2(q_0 - \bar{q}) = \langle V'\rangle$, we readily find

$$W - \langle W\rangle = V - \langle V'\rangle\,\delta\hat{q} - \frac{m\omega^2}{2}\left(\delta\hat{q}^2 - \Delta q^2\right)$$

where $\delta\hat{q} = q - q_0$. Finally, expanding the potential around q_0, squaring and averaging

$$\Delta W^2 \approx \frac{m^2\Delta^4}{2}\Delta q^4 + \left(\frac{|V_0^{(3)}|^2}{6} + \frac{m\Delta^2}{2}V_0^{(4)}\right)\Delta q^6$$

where $m\Delta^2 := V''(q_0) - m\omega^2$, $V_0^{(n)}$ is a shorthand for the nth derivative of the potential evaluated in q_0 and $\langle(q-q_0)^{2n}\rangle = \Delta q^{2n}(n-1)!!$ for $n = 2$ has been used.

In closing this section, we notice that from the variational equation of motion it follows

$$i\hbar\,|\dot{\Psi}_0^+\rangle = \langle[a, H]\rangle\,(a^\dagger - z^*)\,|\Psi_0\rangle$$

and thus

$$H_v = \langle[a, H]\rangle\,(a^\dagger - z^*) - \langle[a^\dagger, H]\rangle\,(a - z)$$

is the appropriate "variational" Hamiltonian. Introducing $\delta\hat{q}$ and $\delta\hat{p} = \hat{p} - p_0$, and using the expression of H above, one easily finds

$$H_v = \frac{p_o}{m}\delta\hat{p} + \langle V'\rangle\,\delta\hat{q}$$

which is a kind of Hamiltonian linearized around the average position and momentum of the wavepacket.

Adiabatic approximation

A key quantity in the adiabatic approximation is the kinetic energy operator "reduced" with respect to the electronic coordinates, $\langle T\rangle_{nm} = \langle\Phi_n(\mathbf{X})|T|\Phi_m(\mathbf{X})\rangle$. Using i to label the nuclear coordinates with mass M_i we obtain

$$\langle T\rangle_{nm} = \delta_{nm}T - i\hbar\sum_i \frac{1}{M_i}\left\langle\Phi_n(\mathbf{X})|\frac{\partial\Phi_m}{\partial X_i}\right\rangle P_i + \\ - \hbar^2\sum_i\frac{1}{2M_i}\left\langle\Phi_n(\mathbf{X})|\frac{\partial^2\Phi_m}{\partial X_i^2}\right\rangle$$

where P_i is a nuclear momentum operator. Here, for $m \neq n$ the second term can also be written in a form that makes explicit the energy differences, since it holds

$$\left\langle\Phi_n|\left[\frac{\partial}{\partial X_i}, h_{\text{el}}(\mathbf{X})\right]|\Phi_m\right\rangle \equiv \Delta E_{mn}(\mathbf{X})\left\langle\Phi_n|\frac{\partial\Phi_m}{\partial X_i}\right\rangle$$

and, on the other hand, $\left[\frac{\partial}{\partial X_i}, h_{\text{el}}(\mathbf{X})\right] = -F^i$ where F^i is the one-electron operator representing the force acting on $\mathbf{X}_i$. The operators $\langle T\rangle_{nm}$ satisfy

$$\langle T\rangle_{nm}^\dagger = \langle T\rangle_{mn}$$

as can be readily checked by either its definition or a direct calculation. In the latter case, notice that one needs the identities

$$\left\langle\Phi_n|\frac{\partial\Phi_m}{\partial X_i}\right\rangle + \left\langle\frac{\partial\Phi_n}{\partial X_i}|\Phi_m\right\rangle = 0$$

$$\left\langle\Phi_n|\frac{\partial^2\Phi_m}{\partial X_i^2}\right\rangle + 2\left\langle\frac{\partial\Phi_n}{\partial X_i}|\frac{\partial\Phi_m}{\partial X_i}\right\rangle + \left\langle\frac{\partial^2\Phi_n}{\partial X_i^2}|\Phi_m\right\rangle = 0$$

that follow from the orthonormality of the electronic states (the first make also the diagonal term $\left\langle\Phi_n(\mathbf{X})|\frac{\partial\Phi_n}{\partial X_i}\right\rangle$ vanishing in the presence of time-reversal invariance).

Finally, in the main text, we have used

$$\langle T^2\rangle_{nn} - \langle T\rangle_{nn}^2 = \sum_{m\neq n}\langle T\rangle_{nm}\langle T\rangle_{mn} \\ \equiv \sum_{m\neq n}\langle T\rangle_{mn}^\dagger\langle T\rangle_{mn}$$

to rewrite the error in terms of contributing electronic transitions $||\langle T\rangle_{mn}\psi||^2$.

Spawning in MCTDH

We sketch here a possible "spawning" algorithm in propagating high-dimensional wavepackets of the multiconfiguration time-dependent Hartree (MCTDH) type using the error expresion provided by Eq. 11. In this method the wavefunction takes the form $|\Psi_0\rangle = \sum_I C_I|\Phi_I\rangle$ where C_I's are complex coefficients, $I = (i_1, i_2, ..i_N)$ is a multi-index and $|\Phi_I\rangle = |\phi_{i_1}\phi_{i_2}..\phi_{i_N}\rangle$ (where $i_k = 1, ..n_K$) are configurations of fully flexible single-particle functions. We call $|\phi_{i_k}\rangle$ ($i_k = 1, n_k$) the "occupied" spfs for the kth mode, and $|\Phi_I\rangle$ the "occupied" configurations. The scalar product over the kth degree of freedom

$$\langle\phi_{i_k}|\Phi_J\rangle = \begin{cases} 0 & \text{if } i_k \notin J \\ |\Phi_{J_k}^{(k)}\rangle & \text{if } i_k \in J \end{cases}$$

defines the single-hole configuration $|\Phi_{J_k}^{(k)}\rangle$ (and the $N-1$ dimensional multi-index $J_k = j_1..j_{k-1}j_{k+1}..j_N$) with the spf for the kth mode removed. When "spawning" is required (*i.e.*, when the error ε exceeds some given threshold) new spfs are introduced and the set of configurations enlarged,

$$|\Psi_0\rangle \to |\Psi_0'\rangle = |\Psi_0\rangle + |\delta\Psi_0\rangle \qquad |\delta\Psi_0\rangle = \sum_J D_J|\delta\Phi_J\rangle$$

where $|\delta\Phi_J\rangle$'s have one or more occupied spfs replaced by newly generated ones. At the time of spawning, however, such an addition does *not* modify the wavefunction

($D_J \equiv 0$) but only its time-derivative

$$|\dot{\Psi}_0\rangle \to |\dot{\Psi}'_0\rangle = |\dot{\Psi}_0\rangle + |\delta\dot{\Psi}_0\rangle \qquad |\delta\dot{\Psi}_0\rangle = \sum_J \dot{D}_J\,|\delta\Phi_J\rangle$$

We consider one additional spf per mode at a time, call it $|\eta_k\rangle$ for the kth mode, and assume that the main contribution comes through single excitations, i.e.,

$$|\delta\dot{\Psi}_0\rangle \approx \sum_k |\delta\dot{\Psi}_0^k\rangle = \sum_k \sum_J \dot{D}_J^{(k)}\,|\eta_k\Phi_J^{(k)}\rangle$$

where J is now a $N-1$ dimensional index and $|\Phi_J^{(k)}\rangle$ a kth-single-hole configuration. If $|\eta_k\rangle$ is chosen to be orthogonal to both the occupied spfs ($|\phi_{i_k}\rangle$, $i_k = 1, n_k$) and their time-derivative ($|\dot{\phi}_{i_k}\rangle$, $i_k = 1, n_k$) the above time-derivative is orthogonal to

$$|\dot{\Psi}_0\rangle = \sum_I \dot{C}_I\,|\Phi_I\rangle + \sum_k \sum_I C_I\,|\dot{\phi}_{i_k}\rangle\,|\Phi_{I_k}^{(k)}\rangle$$

and thus

$$||\dot{\Psi}'_0||^2 = ||\dot{\Psi}_0||^2 + \sum_k ||\mathbf{D}^{(k)}||^2 \quad \text{where } ||\mathbf{D}^{(k)}||^2 = \sum_J |\dot{D}_J^{(k)}|^2$$

On the other hand, the amplitude coefficients of the newly introduced configurations follow from the secular problem

$$i\hbar\dot{D}_J^{(k)} = \langle\eta_k\Phi_J^{(k)}|H|\Psi_0\rangle$$

hence

$$\hbar^2||\mathbf{D}^{(k)}||^2 = \sum_J \langle\eta_k\Phi_J^{(k)}|H|\Psi_0\rangle\,\langle\Psi_0|H|\eta_k\Phi_J^{(k)}\rangle$$

This suggests to introduce a reduced, self-adjoint operator for the kth mode

$$\Gamma^{(k)} = \sum_J \langle\Phi_J^{(k)}|H|\Psi_0\rangle\,\langle\Psi_0|H|\Phi_J^{(k)}\rangle$$

(where the scalar products are now over all modes except the kth one) in such a way that it holds

$$\hbar^2||\mathbf{D}^{(k)}||^2 = \langle\eta_k|\Gamma^{(k)}|\eta_k\rangle = \gamma_k$$

Accordingly, the original local-in-time error $\hbar^2\varepsilon^2 = \Delta E_0^2 - ||\dot{\Psi}_0||^2$ transforms, upon spawning, into

$$\hbar^2\varepsilon'^2 = \hbar^2\varepsilon^2 - \sum_k \gamma_k$$

(notice that the added spfs do not modify $|\Psi_0\rangle$, hence neither the average energy nor its variance). One can thus maximize the error reduction by choosing, for each mode, the eigenvectors of maximum value of the operator $\Gamma^{(k)}$ (in the appropriate residual space of the kth mode).