Supplemental Materials: Coupled Wavepackets for Non-Adiabatic Molecular Dynamics: A Generalization of Gaussian Wavepacket Dynamics to Multiple Potential Energy Surfaces

Alexander White, Sergei Tretiak, and Dmitry Mozyrsky

Theoretical Division, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA

I. HELLER'S ADIABATIC GAUSSIAN WAVEPACKET DYNAMICS FROM PATH INTEGRAL APPROACH

Here we re-derive the *Thawed Gaussian* approximation, originally derived by Heller, [1] based on a time slicing procedure. This will provide the foundation from which we will extend the derivation to multiple coupled potential energy surfaces. The wavefunction of the system at time t is given by the time dependent Schrödinger Equation and the initial wavefunction:

$$\Psi(\mathbf{x},t) = e^{-i\varepsilon H(\mathbf{x})}e^{-i\varepsilon H(\mathbf{x})}...e^{-i\varepsilon H(\mathbf{x})}e^{-i\varepsilon H(\mathbf{x})}\Psi(\mathbf{x},0)\ . \tag{1}$$

Here we have defined $\varepsilon \equiv \frac{t}{M\hbar}$, where M is a large number. For a single time step, we expand to first order in ε:

$$\begin{aligned}\Psi(\mathbf{x},t) = ...\big\{1 - i\varepsilon K(\mathbf{x}) - i\varepsilon V(\mathbf{x})\big\}\Psi(\mathbf{x},0) = \\ ...\big\{1 - i\varepsilon\big(\sum_i \frac{-\hbar^2}{2m_i}\frac{\partial^2}{\partial {x_i}^2} + V(\mathbf{x})\big)\big\}\Psi(\mathbf{x},0) \\ \approx ...\Psi(\mathbf{x},\hbar\varepsilon)\ .\end{aligned} \tag{2}$$

Using similar notation to Heller the initial wavepacket is given by:[1]

$$\Psi(\mathbf{x},0) = \exp[\,\frac{i}{\hbar}\{\,\gamma_0 + \mathbf{p}_0^T(\mathbf{x}-\mathbf{x}_0) + (\mathbf{x}-\mathbf{x}_0)^T\hat{\alpha}_0(\mathbf{x}-\mathbf{x}_0)\big\}]\ . \tag{3}$$

For Equation 2 we have a second derivative term:

$$\begin{aligned}\sum_i \frac{-\hbar^2}{2m_i}\frac{\partial^2}{\partial {x_i}^2}\Psi(\mathbf{x},0) = \sum_i \frac{-i\hbar}{2m_i}\frac{\partial}{\partial x_i}\big\{[\mathbf{p}_0]_i + [\,\hat{\alpha}_0(\mathbf{x}-\mathbf{x}_0)]_i + [(\mathbf{x}-\mathbf{x}_0)^T\hat{\alpha}_0]_i\big\}\Psi(\mathbf{x},0) \\ = \big\{\,-i\hbar\,\mathrm{Tr}[\hat{\alpha}_0\hat{m}^{-1}] + \frac{1}{2}\mathbf{p}_0^T\hat{m}^{-1}\mathbf{p}_0 + \mathbf{p}_0^T\hat{m}^{-1}\hat{\alpha}_0(\mathbf{x}-\mathbf{x}_0) + \\ (\mathbf{x}-\mathbf{x}_0)^T\hat{\alpha}_0\hat{m}^{-1}\mathbf{p}_0 + 2\,(\mathbf{x}-\mathbf{x}_0)^T\hat{\alpha}_0\hat{m}^{-1}\hat{\alpha}_0(\mathbf{x}-\mathbf{x}_0)\big\}\Psi(\mathbf{x},0)\ ,\end{aligned} \tag{4}$$

and terms from the potential expanded to the quadratic term around $\mathbf{x}_0$ (as in Ref. 1):

$$V(\mathbf{x}) \approx V(\mathbf{x}_0) + \mathbf{V}'(\mathbf{x}_0)^T(\mathbf{x}-\mathbf{x}_0) + \frac{1}{2}(\mathbf{x}-\mathbf{x}_0)^T\hat{V}''(\mathbf{x}_0)(\mathbf{x}-\mathbf{x}_0) \ . \tag{5}$$

All terms in Eqs. 4 and 5 are of order $\varepsilon\hbar$. Thus they can be collected and returned to the exponential form (accurate up to order ε):

$$\begin{aligned}\Psi(\mathbf{x},\varepsilon\hbar) \approx \exp\Big[\frac{i}{\hbar}\Big\{\gamma_0 + \varepsilon\hbar(i\hbar\,\mathrm{Tr}[\hat{\alpha}_0\hat{m}^{-1}] + \frac{1}{2}\mathbf{p}_0\hat{m}^{-1}\mathbf{p}_0 - V(\mathbf{x}_0)) \\ + \big\{\mathbf{p}_0 - \varepsilon\hbar\mathbf{V}'(\mathbf{x}_0)\big\}^T(\mathbf{x}-\mathbf{x}_0-\varepsilon\hbar\hat{m}^{-1}\mathbf{p}_0) \\ +(\mathbf{x}-\mathbf{x}_0-\varepsilon\hbar\hat{m}^{-1}\mathbf{p}_0)^T\big\{\hat{\alpha}_0 - \varepsilon\hbar\big(2\hat{\alpha}_0\hat{m}^{-1}\hat{\alpha}_0 + \frac{\hat{V}''}{2}\big)\big\}(\mathbf{x}-\mathbf{x}_0-\varepsilon\hbar\hat{m}^{-1}\mathbf{p}_0)\Big\} + O(\varepsilon^2)\Big] \ .\end{aligned} \tag{6}$$

By defining updated Gaussian variables we now have equations of motion which are accurate up to first order in the time step $dt = \varepsilon\hbar$:

$$\dot{\mathbf{x}}_0 = \hat{m}^{-1}\mathbf{p}_0 \ , \tag{7}$$

$$\dot{\mathbf{p}}_0 = -\mathbf{V}'(\mathbf{x}_0) \ , \tag{8}$$

$$\dot{\hat{\alpha}}_0 = -2\hat{\alpha}_0\hat{m}^{-1}\hat{\alpha}_0 - \frac{\hat{V}''}{2} \ , \tag{9}$$

$$\dot{\gamma}_0 = i\hbar\,\mathrm{Tr}[\hat{\alpha}_0\hat{m}^{-1}] + \frac{1}{2}\mathbf{p}_0\hat{m}^{-1}\mathbf{p}_0 - V(\mathbf{x}_0) \ . \tag{10}$$

This result is exactly that of Heller's multidimensional thawed Gaussian wavepacket dynamics.[1]

II. GENERALIZATION TO MULTI-STATE SYSTEM

We again begin with the time dependent Schrödinger Equation. This time our wavefuncton is a vector, and the Hamiltonain is a matrix, in electronic state space (defined for a specific geometry $\mathbf{x}$):

$$|\Psi(\mathbf{x},t)\rangle = e^{-i\varepsilon\hat{H}(\mathbf{x})}e^{-i\varepsilon\hat{H}(\mathbf{x})}...e^{-i\varepsilon\hat{H}(\mathbf{x})}e^{-i\varepsilon\hat{H}(\mathbf{x})}|\Psi(\mathbf{x},0)\rangle \ . \tag{11}$$

The initial wavefunction can undergo a *Born-Oppenheimer expansion*, where for each $\mathbf{x}$ the wavefunction is expanded in a basis of eigenstates of $\hat{V}(\mathbf{x})$:

$$|\Psi(\mathbf{x},0)\rangle = \sum_n |n[\mathbf{x}]\rangle\langle n[\mathbf{x}]|\Psi(\mathbf{x},0)\rangle \ . \tag{12}$$

However, here we will take a different approach. For all $\mathbf{x}$ we will expand in the eigenstates of $\hat{V}(\mathbf{x})$ at $\mathbf{x} = \mathbf{x}_0$, where $\mathbf{x}_0$ is the center of the Gaussian wavepacket $\langle n[\mathbf{x}]|\Psi(\mathbf{x},0)\rangle$:

$$|\Psi(\mathbf{x},0)\rangle = \sum_n |n[\mathbf{x}_0]\rangle\langle n[\mathbf{x}_0]|\Psi(\mathbf{x},0)\rangle \ . \tag{13}$$

While this choice if formally legal, $\sum_n |n[\mathbf{x}_0]\rangle\langle n[\mathbf{x}_0]| = 1$, it may seem a strange choice. However, it is fully consistent with the trajectory based branching scheme that will ultimately be used to solve this system of equations, and gives the correct receipt for "hopping" trajectories' boundary conditions. We will consider the wavepacket at time $\hbar\varepsilon$ on state m, which is projected onto the basis of eigenstates of $\hat{V}(\mathbf{x}_1)$, where $\mathbf{x_1}$ is the center of the wavepacket at time $\hbar\varepsilon$. Taking only the first ε step in Eq. 11, and expanding the exponential to first order in ε we have:

$$\langle m[\mathbf{x}_1]|\Psi(\mathbf{x},\hbar\varepsilon)\rangle \approx \sum_n \langle m[\mathbf{x}_1]|\big\{\hat{I} - i\varepsilon\hat{K}(\mathbf{x}) - i\varepsilon\hat{V}(\mathbf{x})\big\}|n[\mathbf{x}_0]\rangle\langle n[\mathbf{x}_0]|\Psi(\mathbf{x},0)\rangle \ . \tag{14}$$

Eq. 14 describes the wavepacket at time $\hbar\varepsilon$ on electronic surface m, which has contributions from wavepackets at time 0 on all surfaces (n). The eigenstate $|m[\mathbf{x}_1]\rangle$ can be projected in the basis of $|\chi[\mathbf{x}_0]\rangle$ eigenstates (up to the first order in ε):

$$\begin{aligned} |m\rangle = \sum_l |l\rangle\langle l|m\rangle &= \sum_l \big\{\delta_{l,m} + \langle l|\nabla_{\mathbf{x}_1} m\rangle \cdot (\mathbf{x}_1 - \mathbf{x}_0)\big\}|l\rangle + ... \\ &\approx |m\rangle + \varepsilon\hbar \sum_l \mathbf{d}_{l,m} \cdot \hat{m}^{-1}\mathbf{p}_0|l\rangle + O(\varepsilon^2) \ . \end{aligned} \tag{15}$$

We have assumed that the difference $\mathbf{x_1} - \mathbf{x_0}$ is proportional to $\hbar\varepsilon$ and that we have constant momentum over that time step. $\mathbf{p}_0$ is the momentum of $\langle n[\mathbf{x}_0]|\Psi(\mathbf{x},0)\rangle$. From this point we will drop the $[\mathbf{x}_0]$ label and assume, unless otherwise labeled, that our electronic basis states are eigenstates of $\hat{V}(\mathbf{x}_0)$.

If we insert Eq. 15 into Eq. 14:

$$\begin{aligned} \langle m[\mathbf{x}_1]|\Psi(\mathbf{x},\hbar\varepsilon)\rangle \approx \sum_n \Big[\langle m| + \hbar\varepsilon \sum_l \mathbf{d}_{l,m} \cdot \hat{m}^{-1}\mathbf{p}_0\langle l|\Big] \times \\ \big\{1 - i\varepsilon\big(\hat{K}(\mathbf{x}) + \hat{V}(\mathbf{x})\big)\big\}|n\rangle\langle n|\Psi(\mathbf{x},0)\rangle \ , \end{aligned} \tag{16}$$

and separate out the term $m = n$ in the sum, keep only terms up to $O(\varepsilon)$, we have:

$$\begin{aligned} &\langle m[\mathbf{x}_1]|\Psi(\mathbf{x},\hbar\varepsilon)\rangle \approx \\ &\langle m|\big\{\hat{I} - i\varepsilon\hat{K}(\mathbf{x}) - i\varepsilon\hat{V}(\mathbf{x})\big\}|m\rangle\langle m|\Psi(\mathbf{x},0)\rangle \\ &- i\varepsilon \sum_{n\neq m} \Big[i\hbar\mathbf{d}_{n,m} \cdot \hat{m}^{-1}\mathbf{p}_0 + \langle m|\big\{\hat{K}(\mathbf{x}) + \hat{V}(\mathbf{x})\big\}|n\rangle\Big]\langle n|\Psi(\mathbf{x},0)\rangle \ . \end{aligned} \tag{17}$$

Here we have used the fact that $\mathbf{d}_{m,m} = 0$ and $|m/n\rangle$ is an eigenstate in order to reduce terms. We now expand the potential energy matrix operator around $\mathbf{x}_0$:

$$\begin{aligned} \hat{V}(\mathbf{x}) &\approx \hat{V}(\mathbf{x}_0) + \hat{\mathbf{V}}'(\mathbf{x}_0)^T(\mathbf{x} - \mathbf{x}_0) + \frac{1}{2}(\mathbf{x} - \mathbf{x}_0)^T\underline{\hat{\mathbf{V}}}''(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) \\ &= \sum_{\alpha,\beta} \Big\{V_{\alpha,\beta}(\mathbf{x}_0) + \mathbf{V}'_{\alpha,\beta}(\mathbf{x}_0)^T(\mathbf{x} - \mathbf{x}_0) + \frac{1}{2}(\mathbf{x} - \mathbf{x}_0)^T\underline{\mathbf{V}}''_{\alpha,\beta}(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)\Big\}|\alpha\rangle\langle\beta| \ , \end{aligned} \tag{18}$$

where the α/β electronic basis set is $\mathbf{x_0}$ independent, *e.g* the atomic orbital basis set in realistic calculations, or the diabatic basis set in most model problems. Inserting Eq. 18 into 17 leads to:

$$\langle m[\mathbf{x}_1]|\Psi(\mathbf{x},\hbar\varepsilon)\rangle \approx \tag{19}$$
$$\big\{1 - i\varepsilon\big[\sum_i \frac{-\hbar^2}{2M_i}\frac{\partial^2}{\partial {x_i}^2} + E_m(\mathbf{x}_0) - \mathbf{F}_m(\mathbf{x}_0)^T(\mathbf{x}-\mathbf{x}_0) + (\mathbf{x}-\mathbf{x}_0)^T\underline{\tilde{\mathbf{H}}}_m(\mathbf{x}_0)(\mathbf{x}-\mathbf{x}_0)\big]\big\}\langle m|\Psi(\mathbf{x},0)\rangle$$
$$- i\varepsilon\sum_{n\neq m}\Big[i\hbar\mathbf{d}_{n,m}\cdot\hat{m}^{-1}\mathbf{p}_0 + \mathbf{V}'_{m,n}(\mathbf{x}_0)^T(\mathbf{x}-\mathbf{x}_0) + \frac{1}{2}(\mathbf{x}-\mathbf{x}_0)^T\underline{\mathbf{V}}''_{m,n}(\mathbf{x}_0)(\mathbf{x}-\mathbf{x}_0)\big\}\Big]\langle n|\Psi(\mathbf{x},0)\rangle \ .$$

Now we consider the form

$$\langle n|\Psi(\mathbf{x},0)\rangle \equiv \exp[\frac{i}{\hbar}\{\gamma_0 + \mathbf{p}_0^T(\mathbf{x}-\mathbf{x}_0) + (\mathbf{x}-\mathbf{x}_0)^T\hat{\alpha}_0(\mathbf{x}-\mathbf{x}_0)\}] \times N_n \ . \tag{20}$$

The first line ($m = n$ case) of Eq. 19 is nearly identical to Eq. 2 with some minor differences. First, the real weight of the wavepacket (N_n) arises since the initial state need not be pure. Second while the first expansion term is the force vector $\langle m|\hat{\mathbf{V}}'(\mathbf{x}_0)|m\rangle = \frac{\partial}{\partial\mathbf{x}}E_m(\mathbf{x})|_{\mathbf{x}=\mathbf{x}_0} = -\mathbf{F}_m(\mathbf{x}_0)$ due to the Hellman Feynman theorem, the second expansion term is not the true Hessian matrix of the potential energy surface ($\underline{\tilde{\mathbf{H}}}_m(\mathbf{x}_0) \neq \frac{\partial^2}{\partial\mathbf{x}^2}E_m(\mathbf{x})|_{\mathbf{x}=\mathbf{x}_0}$), because we first expanded in the basis which is $\mathbf{x_0}$ invariant, then rotated into the eigenbasis of $\hat{V}(\mathbf{x}_0)$. Thus we have:

$$\sum_{\alpha,\beta}\langle m|\alpha\rangle\underline{\mathbf{V}}''_{\alpha,\beta}(\mathbf{x}_0)\langle\beta|m\rangle \equiv \underline{\tilde{\mathbf{H}}}_m(\mathbf{x}_0) \ . \tag{21}$$

Finally there is an additional zeroth order in $(\mathbf{x}-\mathbf{x}_0)$ term, coming from the $K_{m,m}(\mathbf{x}_0)$.

$$K_{m,m}(\mathbf{x}_0)\langle m|\Psi(\mathbf{x},0)\rangle = -\frac{\hbar^2}{2m_i}\frac{\partial^2}{\partial x_i^2}\langle m|\Psi(\mathbf{x},0)\rangle \ . \tag{22}$$

Thus this first line tells us that the diagonal term is just Heller's *Thawed Gaussian* wavepacket dynamics, but with the second derivative matrix given by Eq. 21.

Now we turn our attention to the $m \neq n$ terms $\mathbf{V}'_{m,n}(\mathbf{x}_0)$ and $\underline{\hat{\mathbf{V}}}''_{m,n}(\mathbf{x}_0)$. The first derivative term is related to the non-adiabatic coupling vectors through the Hellman-Feynman theorem:

$$\mathbf{V}'_{m,n}(\mathbf{x}_0) = \langle m[\mathbf{x}]|\hat{\mathbf{V}}'(\mathbf{x})|n[\mathbf{x}]\rangle\Big|_{\mathbf{x}=\mathbf{x}_0} = \mathbf{d}^T_{n,m}[E_m(\mathbf{x}_0) - E_n(\mathbf{x}_0)] \ . \tag{23}$$

The second derivative is similar to the diagonal case:

$$\underline{\mathbf{V}}''_{m,n}(\mathbf{x}_0) = \sum_{\alpha,\beta}\langle m|\alpha\rangle\underline{\mathbf{V}}''_{\alpha,\beta}(\mathbf{x}_0)\langle\beta|n\rangle \ . \tag{24}$$

Collecting all the terms, up to $O(\varepsilon)$, and defining $D_{n,m} \equiv \mathbf{d}^T_{n,m}\hat{m}^{-1}\mathbf{p}_0$ we have:

$$\langle m[\mathbf{x}_1]|\Psi(\mathbf{x},\hbar\varepsilon)\rangle \approx = ... + \hbar\varepsilon\sum_{n\neq m}D_{n,m}\Big[1 + \frac{i}{\hbar}D^{-1}_{n,m}[E_n(\mathbf{x}_0) - E_m(\mathbf{x}_0)]\mathbf{d}^T_{n,m}(\mathbf{x}-\mathbf{x}_0) \tag{25}$$
$$- \frac{i}{\hbar}\frac{1}{2}(\mathbf{x}-\mathbf{x}_0)^T D^{-1}_{n,m}\hat{V}''_{m,n}(\mathbf{x}_0)(\mathbf{x}-\mathbf{x}_0)\Big]\langle n|\Psi(\mathbf{x},0)\rangle \ .$$

We can reduce the expression to :

$$\langle m[\mathbf{x}_1]|\Psi(\mathbf{x},\hbar\varepsilon)\rangle \approx \equiv ... + \hbar\varepsilon \sum_{n\neq m} D_{n,m}\Big[1 + \frac{i}{\hbar}\mathbf{\Delta P}^T_{m,n}(\mathbf{x}-\mathbf{x}_0) + \frac{i}{\hbar}(\mathbf{x}-\mathbf{x}_0)^T\Delta\hat{\alpha}_{m,n}(\mathbf{x}-\mathbf{x}_0)\Big] \ . \tag{26}$$

The term inside the bracket can be exponentiated (assuming all terms are small):

$$\begin{aligned}
&\exp\Big\{\frac{i}{\hbar}\mathbf{\Delta P}^T_{m,n}(\mathbf{x}-\mathbf{x}_0) + \frac{i}{\hbar}(\mathbf{x}-\mathbf{x}_0)^T\Delta\hat{\alpha}_{m,n}(\mathbf{x}-\mathbf{x}_0)\Big\} \qquad (27)\\
&\quad\approx \Big[1 + \frac{i}{\hbar}\mathbf{\Delta P}^T_{m,n}(\mathbf{x}-\mathbf{x}_0) + \frac{i}{\hbar}(\mathbf{x}-\mathbf{x}_0)^T\Delta\hat{\alpha}_{m,n}(\mathbf{x}-\mathbf{x}_0)\Big] \ ,\\
&\quad\text{where } \Delta\hat{\alpha}_{m,n} = -\frac{1}{2}\frac{\hat{V}''_{m,n}(\mathbf{x}_0)}{\mathbf{d}^T_{n,m}\hat{m}^{-1}\mathbf{p}_0} \ ,\\
&\quad\text{and}\quad \mathbf{\Delta P}_{m,n} = \frac{[E_n(\mathbf{x}_0) - E_m(\mathbf{x}_0)]}{\mathbf{d}^T_{n,m}\hat{m}^{-1}\mathbf{p}_0}\mathbf{d}^T_{n,m} \ .
\end{aligned}$$

This condition for the shift in momentum on hop is the same as previously derived. It conserves energy approximately (exactly in infinitely high momentum limit). To ensure that all trajectories conserve energy exactly for all momenta, we make the approximate transformation $\mathbf{d}^T_{n,m}\hat{m}^{-1}\mathbf{p}_0 = \mathbf{d}^T_{n,m}\hat{m}^{-1}\{\mathbf{p}_0+\mathbf{p}_1\}/2 + \mathbf{d}^T_{n,m}\hat{m}^{-1}\{\mathbf{p}_0-\mathbf{p}_1\}/2 \approx \mathbf{d}^T_{n,m}\hat{m}^{-1}\{\mathbf{p}_0+\mathbf{p}_1\}/2 \times \exp\Big[\frac{\mathbf{d}^T_{n,m}\{\mathbf{p}_0-\mathbf{p}_1\}}{\mathbf{d}^T_{n,m}\{\mathbf{p}_0+\mathbf{p}_1\}}\Big]$. With this consideration, and assuming $\mathbf{d}^T_{n,m}\{\mathbf{p}_0-\mathbf{p}_1\} << \mathbf{d}^T_{n,m}\{\mathbf{p}_0+\mathbf{p}_1\}$ our final result for the $(m \neq n)$ case:

$$\begin{aligned}
\langle m[\mathbf{x}_1]|\Psi(\mathbf{x},\hbar\varepsilon)\rangle \approx ... + \sum_{n\neq m}\frac{1}{2}\mathbf{d}^T_{n,m}\hat{m}^{-1}\{\mathbf{p}_0+\mathbf{p}_1\} \times \exp\Big[\frac{\mathbf{d}^T_{n,m}\{\mathbf{p}_0-\mathbf{p}_1\}}{\mathbf{d}^T_{n,m}\{\mathbf{p}_0+\mathbf{p}_1\}}\Big] \times N_n \qquad (28)\\
\times \exp[\ \frac{i}{\hbar}\{\ \gamma_0 + \mathbf{p}_1^T(\mathbf{x}-\mathbf{x}_0) + (\mathbf{x}-\mathbf{x}_0)^T\hat{\alpha}_1(\mathbf{x}-\mathbf{x}_0)\}] \ ,\\
\text{where } \mathbf{p}_1 = \mathbf{p}_0 + \frac{[E_n(\mathbf{x}_0) - E_m(\mathbf{x}_0)]}{\mathbf{d}^T_{n,m}\hat{m}^{-1}\{\frac{\mathbf{p}_0+\mathbf{p}_1}{2}\}}\mathbf{d}^T_{n,m} \ ,\\
\text{and}\quad \hat{\alpha}_1 = \hat{\alpha}_0 + \frac{\hat{V}''_{m,n}}{\mathbf{d}^T_{n,m}\hat{m}^{-1}\{\mathbf{p}_0+\mathbf{p}_1\}} \ .
\end{aligned}$$

III. WAVEPACKET RECONSTRUCTION

For Coupled Propagation and GWP Consolidation (See Section IV) we need to calculate overlap of two normalized Gaussians. We seek to define a single Gaussian which closely approximates two separate but similar Gaussians:

$$N_G e^{i\gamma_G}|G(\mathbf{x}_G,\mathbf{p}_G,\hat{\alpha}_G)\rangle \approx N_1 e^{i\gamma_1}|g_1(\mathbf{x}_1,\mathbf{p}_1,\hat{\alpha}_1)\rangle + N_2 e^{i\gamma_2}|g_2(\mathbf{x}_2,\mathbf{p}_2,\hat{\alpha}_2)\rangle \equiv |\psi\rangle \tag{29}$$

where G, g_1 and g_2 are normalized unphased complex Gaussians. In the limit that the superposition $|\psi\rangle$ is indeed a Gaussian then the mapping is exact:

$$N_G = |\langle\psi|\psi\rangle|^{\frac{1}{2}} \,, \tag{30}$$

$$\mathbf{x}_G = \frac{\langle\psi|\mathbf{x}|\psi\rangle}{\langle\psi|\psi\rangle} \,, \tag{31}$$

$$\mathbf{p}_G = -i\frac{\langle\psi|\partial_{\mathbf{x}}|\psi\rangle}{\langle\psi|\psi\rangle} \,, \tag{32}$$

$$\mathrm{Im}[\hat{\alpha}_G] = \frac{1}{4}\Big[\frac{\langle\psi|(\mathbf{x}-\mathbf{x}_G)^2|\psi\rangle}{\langle\psi|\psi\rangle}\Big]^{-1} \,, \tag{33}$$

$$\frac{\langle\psi|(-i\partial_{\mathbf{x}}-\mathbf{p}_G)^2|\psi\rangle}{\langle\psi|\psi\rangle} - \mathrm{Im}[\hat{\alpha}_G] = \mathrm{Re}[\hat{\alpha}_G]\,\mathrm{Im}[\hat{\alpha}_G]^{-1}\mathrm{Re}[\hat{\alpha}_G] \,. \tag{34}$$

Equation 34 can be solved using the Geometric mean for positive definite matrices $\mathrm{Im}[\hat{\alpha}_G]$ and $\frac{\langle\psi|(-i\partial_{\mathbf{x}}-\mathbf{p}_G)^2|\psi\rangle}{\langle\psi|\psi\rangle} - \mathrm{Im}[\hat{\alpha}_G]$:

$$\mathrm{Re}[\hat{\alpha}_G] = \mathrm{Im}[\hat{\alpha}_G]^{\frac{1}{2}}\Big[\mathrm{Im}[\hat{\alpha}_G]^{-\frac{1}{2}}\{\frac{\langle\psi|(-i\partial_{\mathbf{x}}-\mathbf{p}_G)^2|\psi\rangle}{\langle\psi|\psi\rangle} - \mathrm{Im}[\hat{\alpha}_G]\}\mathrm{Im}[\hat{\alpha}_G]^{-\frac{1}{2}}\Big]^{\frac{1}{2}}\mathrm{Im}[\hat{\alpha}_G]^{\frac{1}{2}} \,. \tag{35}$$

These expectation values and overlap of a superposition of multivariate Gaussians can be calculated analytically. Finally the phase can be found by maximizing the overlap of $\langle\psi|N_G e^{i\gamma_G}|G(\mathbf{x}_G,\mathbf{p}_G,\hat{\alpha}_G)\rangle$, under the constraint of Equation 30:

$$e^{i\gamma_G} = \frac{\langle G(\mathbf{x}_G,\mathbf{p}_G,\hat{\alpha}_G)|\psi\rangle}{|\langle\psi|G(\mathbf{x}_G,\mathbf{p}_G,\hat{\alpha}_G)\rangle|} \,. \tag{36}$$

The expectation values of the superposition are given as a sum over combinations of the Gaussians:

$$\langle\psi|O|\psi\rangle = \sum_{i,j\in 1,2}\langle g_j|O|g_i\rangle N_i N_j \exp[i\{\gamma_i-\gamma_j\}] \,. \tag{37}$$

Through Equations 31-32 the dynamics of the Coupled GWP depends strongly on the Thawed Gaussian Approximation, and the particular value of $\hat{\alpha}$. As the Coupled GWPs separate and the approximation of Equations 30-36 become less valid, the dependence of Equations 31-32 on $\hat{\alpha}$ can lead to unstable dynamics. This is particularly true when the value of $\mathrm{Im}[\hat{\alpha}]$ becomes small (a very wide wavepacket). One valuable feature of the Thawed Gaussian Approximation is that the dynamics are fully classical, and independent of the phase and width. In the same spirit, here we seek to add further approximations to the dynamics which will add stability, with minimal sacrifice of accuracy. We note that typically we will break coupling when $|\langle g_1|g_2\rangle|$ is much less than unity. The quantity $\langle g_1|g_2\rangle$ appears in the evaluation of all Equations 30-36.

$$\langle g_1|g_2\rangle = (\mathrm{Det}[\frac{2}{\pi}\hat{\alpha}_2])^{\frac{1}{4}}(\mathrm{Det}[\frac{2}{\pi}\hat{\alpha}_1])^{\frac{1}{4}}\int \mathbf{dx}\exp\big[i\{(\mathbf{x}-\mathbf{x}_2)\hat{\alpha}_2(\mathbf{x}-\mathbf{x}_2)-(\mathbf{x}-\mathbf{x}_1)\hat{\alpha}_1^*(\mathbf{x}-\mathbf{x}_1)\}\big] \tag{38}$$
$$\times\exp\big[i\{\mathbf{p}_1\mathbf{x}_1-\mathbf{p}_2\mathbf{x}_2\}\big]\times\exp\big[i\{\mathbf{p}_2-\mathbf{p}_1\}\mathbf{x}\big].$$

To stabilize the dynamics we make approximations when finding $\mathbf{x}_G$ and $\mathbf{p}_G$. For high momentum and close narrow wavepackets:

$$(\mathrm{Det}[\frac{2}{\pi}\hat{\alpha}_2])^{\frac{1}{4}}(\mathrm{Det}[\frac{2}{\pi}\hat{\alpha}_1])^{\frac{1}{4}}\exp\big[i\{(\mathbf{x}-\mathbf{x}_2)\hat{\alpha}_2(\mathbf{x}-\mathbf{x}_2)-(\mathbf{x}-\mathbf{x}_1)\hat{\alpha}_1^*(\mathbf{x}-\mathbf{x}_1)\}\big] \approx \delta(\mathbf{x}-[\frac{\mathbf{x}_1+\mathbf{x}_2}{2}]) \ ,$$

which leads to:

$$\langle g_1|g_2\rangle \approx \exp\big[i\{\frac{\mathbf{p}_2+\mathbf{p}_1}{2}\}\{\mathbf{x}_1-\mathbf{x}_2\}\big] \ . \tag{39}$$

Use of Equation 39 leads to:

$$\langle g_1|\mathbf{x}|g_2\rangle = -i\partial_{\mathbf{p_2}}\langle g_1|g_2\rangle + \langle g_1|\mathbf{x}_1|g_2\rangle \approx \frac{\mathbf{x}_1+\mathbf{x}_2}{2}\exp\big[i\{\frac{\mathbf{p}_2+\mathbf{p}_1}{2}\}\{\mathbf{x}_1-\mathbf{x}_2\}\big] \ , \tag{40}$$

$$\begin{aligned} -i\langle g_1|\partial_{\mathbf{x}}|g_2\rangle = -\frac{i}{2}\big[\langle g_1|\partial_{\mathbf{x}}g_2\rangle - \langle\partial_{\mathbf{x}}g_1|g_2\rangle\big] &= \frac{\mathbf{p}_1+\mathbf{p}_2}{2}\langle g_1|g_2\rangle + \hat{\alpha}_2\langle g_1|\mathbf{x}-\mathbf{x}_2|g_2\rangle + \hat{\alpha}_1^*\langle g_1|\mathbf{x}-\mathbf{x}_1|g_2\rangle \\ &\approx \frac{\mathbf{p}_1+\mathbf{p}_2}{2}\exp\big[i\{\frac{\mathbf{p}_2+\mathbf{p}_1}{2}\}\{\mathbf{x}_1-\mathbf{x}_2\}\big] \ . \end{aligned} \tag{41}$$

Equations 40 and 41 are used to in the calculations shown in the main text. In the calculation of Equation 41 we discard the terms which are proportional to $\hat{\alpha}$. This is consistent with the approximation leading to Equation 39, that momentum is high and wave packets are narrow $\big(\frac{\mathbf{p}_1+\mathbf{p}_2}{2} \gg \frac{1}{2}\{\mathbf{x}_1-\mathbf{x}_2\}\{\hat{\alpha}_2-\hat{\alpha}_1^*\}\big)$. For the calculation of $\hat{\alpha}_G$ the full set of Equations 30-34 are used, but with the GWPs following the dynamics guided by Equations 40 and 41. Using Equation 39 in the evaluation of N_G and γ_G (Equations 30 and 36) provides similar results to using the full $\langle g_1|g_2\rangle$ calculation for the models considered in the main text.

IV. SUMMARY OF THE ALGORITHM

Initial Propagation:

Step 1: Initialize Gaussian wavepacket (GWP) with desired parameters, on state m.

Step 2: Propogate the GWP forward in time using Eq. 5 from the main text.

Step 3: If outside region of non-adiabatic coupling (NAC) repeat step 2. If GWP reaches a region of significant NAC, $\mathbf{d}_{n,m}^T\frac{\hat{m}^{-1}\{\mathbf{p}_0+\mathbf{p}_1\}/2}{|\hat{m}^{-1}\{\mathbf{p}_0+\mathbf{p}_1\}/2|} > D_{min}$ (a user set threshold), then generate new wavepacket on state n using Eq. 6 from main text and energy conserving change in momentum. If energy cannot be conserved, generation is not allowed. Establish a connection between these GWPs on m and n, and begin Coupled Propogation.

Coupled Propogation:

Step 4: Propogate all GWP forward in time using Eq. 5 from main text.

Step 5: Calculate overlap of normalized GWP with it's connections normalized "hopped" GWP (with width and momentum shift described above). If the amplitude of the overlap is greater than an accuracy controlling threshold, O_{min}, reconstruct new wavepacket as described in Section III . If the amplitude of the overlap is lower than O_{min}, eliminate connection between these GWPs. Continue with Step 2 for each wavepacket independently.

Step 6: If a GWP leaves the region of significant NAC, $\mathbf{d}_{n,m}^{T} \frac{\hat{m}^{-1}\{\mathbf{p}_0+\mathbf{p}_1\}/2}{|\hat{m}^{-1}\{\mathbf{p}_0+\mathbf{p}_1\}/2|} < D_{min}$, eliminate connection between these GWPs. Continue with Step 2 for each wavepacket independently.

End of Propogation:

Step 7: Once final (or output) time is reached, the wavefunction is given as a sum over the weighted complex GWPs. Expectation values can be calculated directly from the wavefunction.

Consolidation and Filtering (Optional Consideration):

Periodically, on some predefined number of time step, one could attempt to condense the trajectories by using the same reconstruction method as used in coupled propagation. Additionally after such a consolidation step one could discard trajectories which are insignificant (by real weight).

Step A: Remove connection between all coupled trajectories.

Step B: Loop through PES,

Step C: Find the largest weight GWP on the PES, this GWP is the initial value of the reconstructed GWP (GWP-New).

Step D: Loop through all available GWPs on the PES. If the normalized GWP has overlap with the normalized GWP-New is higher than O_{min}, then add the GWP to GWP-New (by reconstruction). Continue to update GWP-New until fully looped through the GWP's. Repeat Loop until no new GWPs are added to GWP-New.

Step E: Repeat Step C, until all GWPs have been added to new GWPs.

Step F: Return to Step C for next PES, continuing until all PES are checked.

Step G: Once all new, consolidated GWPs are generated, calculate average weight of GWPs. For all GWPs, if the weight is less than a threshold percentage of the average weight (we have used 3%) then discard the GWP from further propagation.

Step H: Procedure (starting from Step B) can be repeated, until no further change in the GWPs occurs, to attempt further consolidation.

Step I: For all GWPs on different PES: If the overlap of the normalized GWPs is greater than a threshold, they are re-connected and undergo Coupled Propagation (Step 4), else if no "partner" is found then the GWP undergoes independent propagation (Step 2).

Monte-Carlo Sampling of many branch GWPs (Optional Consideration):

If a GWP is the result of many branches one expects its contribution to be small, with many similar GWPs contributing. Thus for branches of higher order than a defined number B_{max}, one may choose to sample the branching by Monte-Carlo rather than explicitly propagating both branches.

Step 5/6-a: If connection between GWPs is ended, and the GWPs have branched more than B_{max} times, choose one trajectory to propagate by random number generation (Monte Carlo), based on the weights of the GWP. The propagated GWP will have new weight which is equal to the sum of the weights of the two GWPs.

Step 5/6-b: New GWPs generated by the Consolidation and Filtering procedure will be considered as having branched zero times.

Miscellaneous considerations to limit branches (Optional Considerations):

6-M: (Modified Step 6) If GWP leave the region of significant non-adiabatic coupling (NAC), $\mathbf{d}^T_{n,m}\frac{\hat{m}^{-1}\{\mathbf{p}_0+\mathbf{p}_1\}/2}{|\hat{m}^{-1}\{\mathbf{p}_0+\mathbf{p}_1\}/2|} < D_{min}$, eliminate connection in single direction, allowing GWP which is still inside NAC region to add to the GWP which has left, but not the reverse. This can help prevent the generation of many small GWPs as the coupled GWPs leave the region of NAC at slightly different times.

Step CP-M: (Modification to Coupled Propagation) Only check whether to break connection after GWPs are connected for a finite time ($t_{min} = 10$ a.u.). This helps prevent generation

of many small wave packets in difficult, highly-chaoitc, regions or near edges of NAC region.

Step 3-M: (Modification to Step 3) One can define a region of significant NAC where the Massey parameter[2] for each GWP:

$$\zeta_0 = |\frac{\mathbf{d}_{n,m}^T \hat{m}^{-1} \mathbf{p}_0}{E_n - E_m}| \; / \; \zeta_1 = |\frac{\mathbf{d}_{m,n}^T \hat{m}^{-1} \mathbf{p}_1}{E_m - E_n}| > \zeta_{min} \; . \tag{42}$$

Wwe use $\zeta_{min} = 1E-3$. One may also limit the NAC region to areas where $\frac{\mathbf{d}_{n,m}^T\{\mathbf{p}_0 - \mathbf{p}_1\}}{\mathbf{d}_{n,m}^T\{\mathbf{p}_0 + \mathbf{p}_1\}} < \Delta_{max}$ (we use $\Delta_{max} = 1.5$). When this ratio is large rapidly oscillating phase differences between generated GWPs is expected to cancel out.

Note: With all the optional considerations, there is a threshold parameter ($O_{min}, D_{min}, B_{max}$, $t_{min}, \zeta_{min}, \Delta_{max}$) which can be used to tune accuracy vs efficiency. Convergence with the thresholds can be checked to determine if the information loss is acceptable.

V. NON-ORTHOGONALITY OF BASIS

During Coupled Propagation, we must take a superposition of two Gaussians which are in non-orthogonal basis states. For two coupled GWPs, which generate two "hopped" GWPs we have:

$$\begin{aligned} |\Psi\rangle = g_1|1[\mathbf{x}_1]\rangle + g_{1,2}|2[\mathbf{x}_1]\rangle + g_{2,1}|1[\mathbf{x}_2]\rangle + g_2|2[\mathbf{x}_2]\rangle \\ \approx G_1|1[\mathbf{x}_1]\rangle + G_2|2[\mathbf{x}_2]\rangle \, . \end{aligned} \tag{43}$$

Here, $g_{1(2)}$ is the GWP initially on PES 1 (2) which stays on PES 1(2), and $g_{1,2(2,1)}$ is the GWP which is initially on PES 1 (2) and hops to PES 2(1). The GWPs $g_{1,2}$ and $g_{2,1}$ must be rotated from the electronic states $|2[\mathbf{x}_1]\rangle$ and $|1[\mathbf{x}_2]\rangle$ to $|2[\mathbf{x}_2]\rangle$ and $|1[\mathbf{x}_1]\rangle$ respectively. Projection of $|2[\mathbf{x}_1]\rangle$ onto the orthogonal basis $|2[\mathbf{x}_2]\rangle + |1[\mathbf{x}_2]\rangle$ will result in the wavepacket being "duplicated":

$$g_{1,2}|2[\mathbf{x}_1]\rangle = g_{1,2}\cos\theta \, |2[\mathbf{x}_2]\rangle + g_{1,2}\sin\theta \, |1[\mathbf{x}_2]\rangle \; , \tag{44}$$

and similarly, we have:

$$g_{2,1}|1[\mathbf{x}_2]\rangle = g_{2,1}\cos\theta \, |1[\mathbf{x}_1]\rangle + g_{2,1}\sin\theta \, |2[\mathbf{x}_1]\rangle \; . \tag{45}$$

where $\cos\theta \equiv \langle 1[\mathbf{x}_2]|1[\mathbf{x}_1]\rangle \equiv \langle 2[\mathbf{x}_2]|2[\mathbf{x}_1]\rangle$ and $\sin\theta \equiv \langle 1[\mathbf{x}_2]|2[\mathbf{x}_1]\rangle \equiv \langle 2[\mathbf{x}_1]|1[\mathbf{x}_2]\rangle$. This step can be repeated, and since by definition $|\sin\theta| < 1$ the infinite series of rotations will converge. Thus

Equation 43 can be re-expressed as:

$$|\Psi\rangle = g_1|1[\mathbf{x}_1]\rangle + g_{1,2}|2[\mathbf{x}_1]\rangle + g_{2,1}|1[\mathbf{x}_2]\rangle + g_2|2[\mathbf{x}_2]\rangle \tag{46}$$
$$= \Big[g_1 + \frac{1}{\cos\theta}g_{2,1} + \frac{\sin\theta}{\cos\theta}g_{1,2}\Big]|1[\mathbf{x}_1]\rangle + \Big[g_2 + \frac{1}{\cos\theta}g_{1,2} + \frac{\sin\theta}{\cos\theta}g_{2,1}\Big]|2[\mathbf{x}_2]\rangle$$
$$\approx G_1|1[\mathbf{x}_1]\rangle + G_2|2[\mathbf{x}_2]\rangle\,.$$

Thus, not only does the "hopped" GWP contribute to the reconstructed GWP on the final surface, but it also has a contribution to the re-constructed GWP on the same surface. However this contributions is small for shifts in position $\mathbf{x}_1 - \mathbf{x}_2$ in which the Coupled GWP dynamics works ($\cos\theta \approx 1$) . That is, GWPs should branch before this non-orthogonal basis effect becomes relevant. In the calculations presented in the paper, inclusion of this effect does not change results significantly. We present it here for formal completeness, and for possible future importance.

---

[1] E. J. Heller, The Journal of Chemical Physics **62** (1975).

[2] H. S. W. Massey, Reports on Progress in Physics **12**, 248 (1949).