JCTC Journal of Chemical Theory and Computation

Article

Cite This: J. Chem. Theory Comput. XXXX, XXX, XXX–XXX

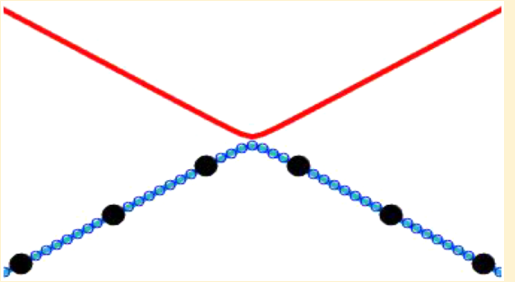
pubs.acs.org/JCTC

A Robust and Unified Solution for Choosing the Phases of Adiabatic States as a Function of Geometry: Extending Parallel Transport Concepts to the Cases of Trivial and Near-Trivial Crossings

Zeyu Zhou, Zuxin Jin, Tian Qiu, Andrew M. Rappe, and Joseph Eli Subotnik*

Department of Chemistry, University of Pennsylvania, Philadelphia, Pennsylvania 19104, United States

ABSTRACT: We investigate a simple and robust scheme for choosing the phases of adiabatic electronic states smoothly (as a function of geometry) so as to maximize the performance of ab initio non-adiabatic dynamics methods. Our approach is based upon consideration of the overlap matrix (**U**) between basis functions at successive points in time and selecting the phases so as to minimize the matrix norm of log(**U**). In so doing, one can extend the concept of parallel transport to cases with sharp curve crossings. We demonstrate that this algorithm performs well under extreme situations where dozens of states cross each other either through trivial crossings (where there is zero effective diabatic coupling), or through non-trivial crossings (when there is a non-zero diabatic coupling), or through a combination of both. In all cases, we compute the time-derivative coupling matrix elements (or equivalently non-adiabatic derivative coupling matrix elements) that are as smooth as possible. Our results should be of interest to all who are interested in either non-adiabatic dynamics, or more generally, parallel transport in large systems.



Downloaded via UNIV OF PENNSYLVANIA on January 29, 2020 at 04:03:23 (UTC).
See https://pubs.acs.org/sharingguidelines for options on how to legitimately share published articles.

1. INTRODUCTION

1.1. Parallel Transport in the Non-adiabatic Regime.

For many problems in chemical physics, one must naturally deal with a quantum subsystem that evolves either in time or according to some external parameter. For instance, the Born–Oppenheimer approximation requires that one consider quantum electronic states as parametrized by nuclear geometry; the eigenvalues of the electronic Schrodinger equation become the potential energy surfaces which are the bedrock of modern chemistry. More generally, however, there is also a long story going back to Longuet-Higgins,[1] Baer,[2] and Mead and Truhlar[3] pointing out that when we make the Born–Opphenheimer approximation, we should not focus exclusively on how quantum *eigenvalues* depend on external parameters, but also on how the *eigenvectors* themselves evolve. In a famous set of papers considering how electronic states depend on the vector potential or magnetic field, Berry showed that the phases of the eigenvectors can have a great deal of rich, topological physics buried within them.[4] More specifically, suppose one propagates a single eigenvector slowly around a closed loop parametrized by $t \in [0, T]$. Berry showed that in the adiabatic limit—which means that two eigenvalues are never close to each other—that eigenvector picks up an extra phase (beyond the dynamical phase) as a function of the classical parameter it depends on.[4−8] Thus, Berry's phase demonstrated the limits of so-called *parallel transport*—which means that $\langle \phi_j(t)|\phi_j(t + \mathrm{d}t)\rangle$ is real and maximized, so that one would expect $\langle \phi_j(t)|\dot{\phi}_j(t + \mathrm{d}t)\rangle \approx 0$; even if one parametrizes adiabatic states according to parallel transport, an adiabatic wave function picks up a phase when the external parameter traces a complete cycle. Therefore, from a different point of view, Berry demonstrated that parallel transport is not consistent with the presence of a globally well-defined adiabatic state (with globally well-defined phases).

Now, despite the failure of parallel transport to account for a global, topological Berry phase, for many practical purposes, parallel transport works well enough and can solve many interesting problems. For instance, in the context of non-adiabatic molecular dynamics, parallel transport is always applied as one propagates nuclear trajectories that explore different nuclear geometries in time; one wants a smooth choice of adiabatic states, and one usually does not ever return to the initial location. In such a case, if one never completes a closed cycle, ignoring Berry's phase (i.e., $\langle \phi_j(t)|\dot{\phi}_j(t + dt)\rangle \approx 0$) is usually a well-behaved, efficient, and accurate approximation. [In fact, noting that, in practice, calculating $T_{jj} = \langle \phi_j|\dot{\phi}_j\rangle$ is very difficult or impossible for almost all ab initio calculations, one can argue that parallel transport is not only sensible but actually required for many modern programs.] Nevertheless, for many problems in chemical physics, implementing parallel transport is not obvious in practice. After all, let **U** be the overlap matrix between electronic states:

$$U_{jk} = \langle \phi_j(t)|\phi_k(t + \mathrm{d}t)\rangle \tag{1}$$

Then, the usual parallel transport approach tells us to make all diagonal elements U_{jj} real and positive, which fixes the sign of each $|\phi_j(t)\rangle$ at each time step. This phase convention is used nowadays routinely for modern ab initio non-adiabatic

Received: September 25, 2019

 © XXXX American Chemical Society

DOI: 10.1021/acs.jctc.9b00952
J. Chem. Theory Comput. XXXX, XXX, XXX–XXX

dynamics calculations (Ehrenfest, FSSH,[9,10] or AIMS[11]) to investigate photoexcited relaxation. And yet, if one moves away from the adiabatic regime, insisting that U_{jj} are real and maximized can be difficult or even unstable: what if $U_{jj} \approx 0$ might be possible in the non-adiabatic regime?

To better understand this failure conceptually, consider how one would apply the concept of parallel transport in the extreme non-adiabatic limit of curve crossings, also called a trivial crossing.[12−20] Consider a two-level model Hamiltonian:

$$\mathbf{H}_{\text{real}} = \begin{bmatrix} 0.1\tanh(R) & \kappa \exp(-R^2) \\ \kappa \exp(-R^2) & -0.1\tanh(R) \end{bmatrix}$$

There is an avoided crossing at $R = 0$, if $\kappa \neq 0$, and let us choose κ to be very small, e.g., 1×10^{-10}. As shown in Figure 1,

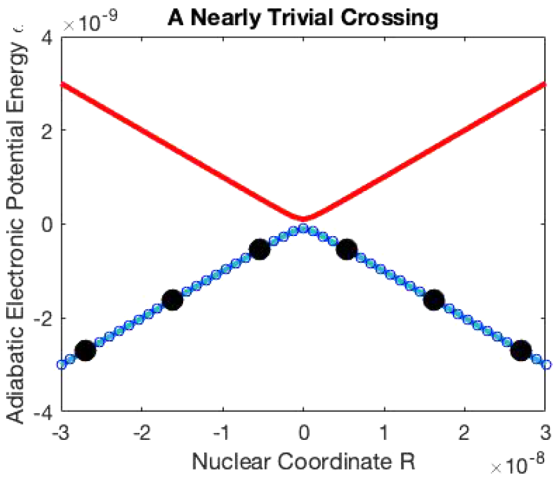


Figure 1. Schematic diagram for a nearly trivial crossing. The red and blue lines represent two adiabatic states with a small gap between them at $R = 0$. The small circles demonstrate that, if we are to rely on parallel transport to determine the phases of the wave functions at each successive point in space, we will need a very dense set of grid points to model the curve crossing at $R = 0$. By contrast, we would like to run dynamics with a larger time step and a more sparse sampling of grid points, e.g., the set of black solid circles. In such a case, we will need a clever algorithm (beyond simple parallel transport) to pick the phases of the adiabatic states because the set of such (black circle) points will miss crucial details of the near-trivial curve crossing.

if we strictly follow parallel transport, we will need many such grid points (as labeled by small circles) to transport our states, and thus a very small time step will be required for the simulation; the computational cost for simulating a curve crossing will be unbearably and unnecessarily large (and without much physical meaning). In practice, we would actually prefer a large time step (as shown by the black circles), and simply recognize that two states switched. But how to choose phases? The overlap matrix $\mathbf{U}$ between the two central black circles in Figure 1 will take the following form:

$$\mathbf{U}_{\text{real}} = \begin{bmatrix} 0 & \mp 1 \\ \pm 1 & 0 \end{bmatrix}$$

The +/− signs should in principle be determined by small time steps and true parallel transport, but the diagonal elements of $\mathbf{U}$ are not useful. Now, admittedly one might wonder: Do we really care about the signs of the off-diagonal matrix elements? Do we really need to get the signs correct? A simple thought experiment, however, will show that these signs can be crucial. Consider the cases of four electronic states with the following two options for the overlap matrix (both with determinant equal to +1):

$$\mathbf{U}_4^{(1)} = \begin{bmatrix} 0 & 0.995 & -0.099 & 0 \\ -0.995 & 0 & 0 & -0.099 \\ 0.099 & 0 & 0 & -0.995 \\ 0 & 0.099 & 0.995 & 0 \end{bmatrix}$$

$$\mathbf{U}_4^{(2)} = \begin{bmatrix} 0 & 0.995 & 0.099 & 0 \\ -0.995 & 0 & 0 & 0.099 \\ 0.099 & 0 & 0 & 0.995 \\ 0 & 0.099 & -0.995 & 0 \end{bmatrix}$$

Clearly one cannot use the diagonal elements of $\mathbf{U}$ to pick the optimal phases of states 1, 2, 3, and 4, and yet, at the same time, these signs cannot be picked arbitrarily because states 1 and 2 do have a sign relative to states 3 and 4, respectively. And so, perhaps not surprisingly, as will be shown below, picking random phases at trivial crossings can lead to very erroneous results in the case of more than two electronic states.

While the situation above may appear extreme, in practice trivial crossings arise all the time—especially in large systems. Moreover, the basic premise of setting phases so as to make the diagonal matrix elements real and maximally positive can be also problematic in less obvious cases. Consider the following overlap matrix $\mathbf{U}_{\text{pt}}$ that obeys parallel transport in N-dimensions:

$$\mathbf{U}_{\text{pt}} = \begin{bmatrix} 1-\frac{2}{N} & -\frac{2}{N} & -\frac{2}{N} & \cdots & -\frac{2}{N} \\ -\frac{2}{N} & 1-\frac{2}{N} & -\frac{2}{N} & \cdots & -\frac{2}{N} \\ -\frac{2}{N} & -\frac{2}{N} & 1-\frac{2}{N} & \cdots & -\frac{2}{N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -\frac{2}{N} & -\frac{2}{N} & -\frac{2}{N} & \cdots & 1-\frac{2}{N} \end{bmatrix} \quad (2)$$

This matrix is unitary, and as $N \to \infty$, the off-diagonal matrix elements approach zero and the diagonal matrix elements approach unity. Nevertheless, $\det(\mathbf{U}_{\text{pt}}) = -1$. [A simple proof is as follows: set $\vec{a}$ to be a $1\times N$ vector with all elements +1. Clearly, $\mathbf{U}_{\text{pt}} = \mathbf{I}_N - \frac{2}{N}\vec{a}^{\mathrm{T}}\vec{a}$. By using the Weinstein−Aronszajn identity, we immediately obtain $\det(\mathbf{U}_{\text{pt}}) = \det\left(\mathbf{I}_N - \frac{2}{N}a^{\mathrm{T}}a\right) = \det\left(I_1 - \frac{2}{N}aa^{\mathrm{T}}\right) = 1 - \frac{2}{N}N = -1$.] Therefore, this matrix cannot correspond to a pure rotation. Instead, there must be one eigenvalue equal to −1; such an eigenvalue would correspond to an eigenvector that is reflected by the $\mathbf{U}_{\text{pt}}$ matrix. From this example, we may conclude that even choosing the diagonal matrix elements to be near +1 is not necessarily a good extension of parallel transport (see section 2.1.1).

Thus, we are clearly in need of a *new* protocol to extend parallel transport to the non-adiabatic regime for arbitrary (and

DOI: 10.1021/acs.jctc.9b00952
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

not necessarily diagonally dominant) overlap matrices $\mathbf{U}$. At this point, a mathematically inclined reader might ask: Is it at all reasonable or even possible to make such an extension without taking very small steps? Unfortunately, from the chemists' perspective, there is no known alternative. Many dynamical approaches (especially FSSH and AIMS) are clearly optimal in an adiabatic basis and simple numerical tests can clearly show that choosing phases incorrectly can lead to very bad results. And this failure is not surprising: the choice of phase carries dynamical information quantum mechanically. In general, choosing U_{jj} to be real and maximized is a good idea, since this process forces each eigenvector to change as slowly as possible without distorting the motion of the nuclear degrees of freedom, such that the adiabatic representation becomes a strong framework for smoothly expanding the nuclear Schrödinger equation. In particular, one's choice of $\mathbf{U}$ defines one's choice of the non-adiabatic derivative coupling $\vec{\mathbf{d}}$, which is the key matrix that breaks the Born−Oppenheimer approximation. In turn, $\vec{\mathbf{d}}$ is critically dependent on the choices of phases of $\mathbf{U}$. Thus, choosing $\mathbf{U}$ wisely can enforce smoothness of $\vec{\mathbf{d}}$.

With this sensitivity in mind, there is no choice. For chemical dynamics, we must find a stable and optimal answer to the question: What criteria should be selected for choosing the phases of $\mathbf{U}$? Alternatively, if significant improvement is not possible, we will be forced to take very small time steps.

1.2. Choosing Adiabatic State Phases Specifically in the Context of Non-adiabatic Dynamics. At this point, let us consider the case of non-adiabatic dynamics more explicitly. In this context, $\mathbf{U}$ is of paramount importance because it is related to the time-averaged derivative coupling $\vec{\mathbf{d}}$, which is used to propagate the equations of motion in the electronic degrees of freedom.

$$i\hbar\dot{c}_j = H^{\mathrm{el}}_{jj}c_j - i\hbar\sum_k \vec{d}_{jk}\cdot\vec{v}c_k \tag{3}$$

Here, $\vec{d}_{jk}$ is defined by

$$\vec{d}_{jk} = \langle\phi^{\mathrm{ad}}_j(\vec{R}(t))|\nabla_{\vec{R}}\phi^{\mathrm{ad}}_k(\vec{R}(t))\rangle \tag{4}$$

Of course, chemists have long known that propagating eqs 3 and 4 together is a bad idea. After all, $\vec{d}_{jk}$ is usually obtained by Hellmann−Feynman theorem as

$$\vec{d}_{jk} = \frac{\langle\phi^{\mathrm{ad}}_j|\vec{\nabla}_{\vec{R}}H^{\mathrm{el}}(\vec{R})|\phi^{\mathrm{ad}}_k\rangle}{\varepsilon_k - \varepsilon_j} \tag{5}$$

Thus, $\vec{d}_{jk}$ explodes when there is a crossing between adiabatic states j and k ($\vec{d}_{jk}\to\infty$ as $(\varepsilon_k - \varepsilon_j)\to 0$), and one will need very small time steps near a crossing. With this limitation in mind, one better scheme is to take the logarithm of $\mathbf{U}$, by which the time-averaged time-derivative coupling matrix $\mathbf{T}$ is obtained.[18]

$$\frac{1}{\mathrm{d}t}\int_t^{t+\mathrm{d}t}\mathrm{d}\tau\vec{d}_{jk}(\tau)\cdot\vec{v}(\tau) = T_{jk}(t+\mathrm{d}t/2)$$
$$= \left\langle\phi^{\mathrm{ad}}_j\middle|\frac{\mathrm{d}\phi^{\mathrm{ad}}_k}{\mathrm{d}t}\right\rangle = \frac{[\log(\mathbf{U})]_{jk}}{\mathrm{d}t} \tag{6}$$

As argued by Meek and Levine in a two-state context,[15] this approach is equivalent to calculating the time-averaged derivative coupling and can be used easily in dynamics (and is more stable than the Hammes-Schiffer−Tully method.[12])

Nevertheless, to implement the Meek−-Levine approach[15,18] or other existing trivial crossing approaches,[12,13,16,17,19,20] the inevitable question remains: What are the *phases* of $\mathbf{U}$?[21−23] And, for the skeptical reader who thinks that choosing +/− phases cannot be very important, consider this: What if the Hamiltonian is complex? There will be an entire manifold of possible choices of phases. Specifically at trivial crossing, we expect an overlap matrix $\mathbf{U}_{\mathrm{complex}}$ to be of the form

$$\mathbf{U}_{\mathrm{complex}} = \begin{bmatrix} 0 & e^{i\theta} \\ -e^{-i\theta} & 0 \end{bmatrix}$$

and there will be infinitely many possible choices for θ, each of which implies different physics for coupled nuclear-electron motion, and choosing an incorrect or inconsistent θ can lead to catastrophic consequences even for real $\mathbf{U}$ (where $\theta = 0$ or π).[24]

Obviously, for future purposes, we require a robust solution for picking the phases of the columns of $\mathbf{U}$, one that is applicable in the adiabatic and non-adiabatic limits.

1.3. Outline. In this paper, we will extend the concept of parallel transport to the non-adiabatic regime by choosing the phases of $\mathbf{U}$ such that for each time step, $\mathbf{U}$ is a proper rotation matrix, no matter whether there is one trivial crossing, multiple trivial crossings or no trivial crossings, in both the real and complex regimes. In section 2, we will introduce algorithms both for the real regime (which is easy) and for the complex regime (which is a bit harder). In section 3, we will test our prescription on a modified model problem. In section 4, we conclude and make several observations about the algorithm.

2. METHOD

Consider a path in configuration space, and let $\mathbf{U}$ be the overlap matrix between adiabatic basis sets at different times, as defined in eq 1. Note that $\mathbf{U}$ converges to the identity when the time step dt approaches 0. Of course, for an identity matrix, the logarithm is a null matrix, and so one can argue that the optimal phases of $\mathbf{U}$ should be those phases that minimize the norm of all elements of $\mathbf{T}$. Unfortunately, however, even for the case of real $\mathbf{U}$, it would be very expensive to directly calculate the $\mathbf{T}$ matrices for all possible signs of the columns of $\mathbf{U}$: in principle one would require 2^N logarithm calculations (i.e., matrix diagonalizations) so as to find the matrix $\mathbf{U}$ with the optimal sign choices. Moreover, in the complex regime, it would be impossible to minimize $\sum_{jk}|T_{jk}|^2$ and find the truly optimal $\mathbf{U}$ without having a grid in θ as well (which would require $(N_\theta)^N$ diagonalizations). One would prefer a reasonable set of approximations for minimizing $\sum_{jk}|T_{jk}|^2$ (or equivalently $\mathrm{Tr}(|\log(\mathbf{U})|^2)$).

2.1. Necessary Conditions for Minimizing Tr(|log(U)|^2). *2.1.1. We Must Insist det(U) = 1.* Let us now show that one of the necessary conditions for minimizing $\mathrm{Tr}(|\log(\mathbf{U})|^2)$ is that $\det(\mathbf{U})$ should be +1.

For any unitary $\mathbf{U}$, we can decompose $\mathbf{U}$ as

$$\mathbf{U} = \mathbf{R}\mathbf{\Lambda}\mathbf{R}^\dagger \tag{7}$$

where $\mathbf{\Lambda}$ is diagonal and takes the following form:

DOI: 10.1021/acs.jctc.9b00952
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

$$\mathbf{\Lambda} = \begin{bmatrix} e^{i\delta_1} & & & & \\ & e^{i\delta_2} & & & \\ & & \ddots & & \\ & & & e^{i\delta_{N-1}} & \\ & & & & e^{i\delta_N} \end{bmatrix} \tag{8}$$

Suppose we change $\mathbf{U}$ to $\mathbf{U}'$ by

$$\mathbf{U}' = e^{i\alpha}\mathbf{U} \tag{9}$$

the resulting $\mathbf{U}'$ is still unitary and can be expanded as

$$\mathbf{U}' = \mathbf{R}\mathbf{\Lambda}'\mathbf{R}^{\dagger} \tag{10}$$

where

$$\mathbf{\Lambda}' = e^{i\alpha}\mathbf{\Lambda} = \begin{bmatrix} e^{i(\delta_1+\alpha)} & & & & \\ & e^{i(\delta_2+\alpha)} & & & \\ & & \ddots & & \\ & & & e^{i(\delta_{N-1}+\alpha)} & \\ & & & & e^{i(\delta_N+\alpha)} \end{bmatrix}$$

For $\mathbf{U}'$, the quantity $\mathrm{Tr}(|\log(\mathbf{U}')|^2)$ is

$$\mathrm{Tr}(|\log(\mathbf{U}')|^2) = \sum_N (\delta_N + \alpha)^2 \tag{11}$$

One necessary condition for the quantity to be at minimum is

$$\frac{\partial \mathrm{Tr}(|\log(\mathbf{U}')|^2)}{\partial \alpha} = \sum_N 2(\delta_N + \alpha) = 0 \tag{12}$$

This implies that, if $\mathbf{U}$ is the exact minimum, we must have

$$\sum_N \delta_N = 0 \tag{13}$$

Immediately,

$$\det(\mathbf{U}) = e^{i\sum_N \delta_N} = 1 \tag{14}$$

Hence, one necessary condition for the quantity $\mathrm{Tr}(|\log(\mathbf{U})|^2)$ to be minimized is that $\det(\mathbf{U}) \equiv 1$. However, this is far from a sufficient condition.

2.1.2. We Must Minimize a Polynomial Function of U. We can find another approximate, necessary condition besides the constraint $\det(\mathbf{U}) = 1$ by expanding $\log(\mathbf{U})$ into a Taylor series and truncating at second order around the identity. Note that this approximation represents a slightly dangerous approach, because we are interested in the phases of adiabatic states around trivial or near trivial crossings, where $\mathbf{U}$ is far from the identity, and there is no reason to presume that a Taylor series around the identity should converge; and a second-order approximation need not be accurate at all. Nevertheless, bearing in mind this caveat, we will proceed and test this approximation numerically. To second order,

$$\begin{aligned} &\mathrm{Tr}(|\log(\mathbf{U}')|^2) \approx \\ &\sum_{jk} ((\mathbf{U} - \mathbf{I}) - (\mathbf{U} - \mathbf{I})^2/2)_{jk}((\mathbf{U}^* - \mathbf{I}) - (\mathbf{U}^* - \mathbf{I})^2/2)_{jk} \\ &= \sum_{jk} \left(-\frac{\mathbf{U}^2}{2} + 2\mathbf{U} - \frac{3}{2}\mathbf{I}\right)_{jk} \left(-\frac{(\mathbf{U}^*)^2}{2} + 2\mathbf{U}^* - \frac{3}{2}\mathbf{I}\right)_{jk} \end{aligned} \tag{15}$$

Since $\mathbf{U}$ is a unitary matrix, i.e.,

$$\mathbf{U}\mathbf{U}^{\dagger} = \mathbf{U}^{\dagger}\mathbf{U} = \mathbf{I} \tag{16}$$

thus,

$$\mathrm{Tr}(\mathbf{I}) = \mathrm{Tr}(\mathbf{U}^{\dagger}\mathbf{U}) = \sum_{jk} U_{jk}U_{kj}^{\dagger} = \sum_{jk} U_{jk}U_{jk}^* \tag{17}$$

Here, $\mathrm{Tr}(\mathbf{I})$ equals the number of electronic states. Similarly,

$$\sum_{jk} [\mathbf{U}^2]_{jk}[\mathbf{U}^*]_{jk}^2 = \mathrm{Tr}(\mathbf{I}) \tag{18}$$

Hence, eq 15 becomes

$$\begin{aligned} &\mathrm{Tr}(|\log(\mathbf{U}')|^2) \approx \\ &\frac{3}{4}\mathrm{Tr}((\mathbf{U}^*)^2) + \frac{3}{4}\mathrm{Tr}(\mathbf{U}^2) - 4(\mathrm{Tr}(\mathbf{U}) + \mathrm{Tr}(\mathbf{U}^*)) + \frac{13}{2}\mathrm{Tr}(\mathbf{I}) \\ &= \mathrm{Re}\left(\frac{3}{2}\mathrm{Tr}(\mathbf{U}^2) - 8\mathrm{Tr}(\mathbf{U})\right) + \frac{13}{2}\mathrm{Tr}(\mathbf{I}) \end{aligned} \tag{19}$$

Since we want to minimize this function by changing the phase of each column of $\mathbf{U}$, we can drop the constant term.

Thus, in the end, up to second order, the overlap matrix $\mathbf{U}$ should satisfy the following two conditions:

- $\det(\mathbf{U}) = 1$
- $\mathrm{Re}(\mathrm{Tr}(3\mathbf{U}^2 - 16\mathbf{U}))$ is minimized

2.2. The Two Algorithms for Real and Complex Regimes. For completeness, we will now present a step-by-step algorithm for picking the phases of $\mathbf{U}$ in both real and complex regimes.

2.2.1. Real Regime. For an N-state problem, we will find the optimal combination of signs between columns of $\mathbf{U}$ by minimizing $\mathrm{Tr}(3\mathbf{U}^2 - 16\mathbf{U})$ through Jacobi sweeps[25] while maintaining $\det(\mathbf{U}) = +1$. Specifically, a flowchart is as follows.

Step 1: For a real overlap matrix $\mathbf{U}$, the determinant must obey $\det(\mathbf{U}) = \pm 1$. If $\det(\mathbf{U}) = -1$, we change the sign of the first eigenvector $|\phi_1^{\mathrm{ad}}(\vec{R}(t + \mathrm{d}t_c))\rangle$.

Step 2: To maintain $\det(\mathbf{U}) = +1$, we check whether we should simultaneously flip the signs of a pair of columns of $\mathbf{U}$. There are $N(N-1)/2$ pairs of indices for an N-state system. For each pair j,k, we minimize $\mathrm{Tr}(3\mathbf{U}^2 - 16\mathbf{U})$ by calculating the difference $\Delta_{\mathrm{real}}^{jk}$ (see Appendix A for the derivation of $\Delta_{\mathrm{real}}^{jk}$) as follows:

```
% flagc is a binary variable to keep track of
 whether Jacobi sweeps have converged
set flagc = 1
loop j = 1 : N
  loop k = (j + 1) : N
    Δ^jk_real ≡ 3(U_jj^2 + U_kk^2) + 6(U_jk U_kj)
                +8(U_jj + U_kk) − Σ_l 3(U_jl U_lj + U_kl U_lk)
    if Δ^jk_real < 0
                |φ_j^ad(R(t+dt_c))⟩ = −|φ_j^ad(R(t+dt_c))⟩
                |φ_k^ad(R(t+dt_c))⟩ = −|φ_k^ad(R(t+dt_c))⟩
                set flagc = 0
  end
end
```
(20)

Step 3: If flagc = = 0, return to Step 2.

DOI: 10.1021/acs.jctc.9b00952
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

2.2.2. Complex Regime. To extend the ansatz above into the complex regime, there are two major differences. First, strictly speaking a complex logarithm function is multivalued, and one might worry about whether our approach is even well defined. Nevertheless, for our purposes (i.e., minimization), we need only construct $\mathrm{Tr}(|\log(\mathbf{U})|^2) = \sum_{j=1}^{N}|\delta_j|^2$ in eqs 7 and 8, and the principal value of a complex logarithm is always well defined. In other words, we need only insist that $\forall j$, $\delta_j \in (-\pi, \pi]$, which should solve this first problem.

A second, more important difference is that, whereas the phases of $\mathbf{U}$ are arbitrary up to a +/− sign in the real regime, in the complex regime each adiabatic state can carry a complex phase $\exp(i\theta_j)$ for each adiabatic state $|\phi_j^{\mathrm{ad}}(\vec{R}(t + \mathrm{d}t_c))\rangle$. If we want to explore changing the relative phases of two states j and k, while maintaining $\det(\mathbf{U}) = 1$, we will need to sweep over the following phase possibilities:

$$|\phi_j^{\mathrm{ad}}(\vec{R}(t + \mathrm{d}t_c))\rangle \rightarrow \exp(i\theta_{jk})|\phi_j^{\mathrm{ad}}(\vec{R}(t + \mathrm{d}t_c))\rangle$$
$$|\phi_k^{\mathrm{ad}}(\vec{R}(t + \mathrm{d}t_c))\rangle \rightarrow \exp(-i\theta_{jk})|\phi_k^{\mathrm{ad}}(\vec{R}(t + \mathrm{d}t_c))\rangle \quad (21)$$

Compare with eq 20 above. The final flowchart is then as follows.

Step 1: We must start with a reasonable initial guess for the phases of $\mathbf{U}$. For each column l, we search over all rows indexed by m, and we find the matrix element U_{ml} with the greatest absolute value $[\mathrm{abs}(U_{ml})]$ (most often, as from parallel transport, we will find $m = l$); we then insist that U_{ml} should be real and positive by multiplying the whole column l (i.e., the eigenvector $|\phi_l^{\mathrm{ad}}(\vec{R}(t + \mathrm{d}t_c))\rangle$) by the complex conjugate of its phase $\frac{\mathrm{conj}(U_{ml})}{\mathrm{abs}(U_{ml})}$. This is the ansatz of standard parallel transport.

Step 2: As a complex overlap matrix, $\det(\mathbf{U})$ can be complex. If $\det(\mathbf{U}) = \exp(i\alpha)$, we change the phase of the first eigenvector,

$$|\phi_1^{\mathrm{ad}}(\vec{R}(t + \mathrm{d}t_c))\rangle = \exp(-i\alpha)|\phi_1^{\mathrm{ad}}(\vec{R}(t + \mathrm{d}t_c))\rangle \quad (22)$$

By changing the phase of the first column of $\mathbf{U}$, we now have $\det(\mathbf{U}) = +1$.

Step 3: To maintain $\det(\mathbf{U}) = +1$, we change the phases of a pair of columns of $\mathbf{U}$ simultaneously. There are $N(N-1)/2$ pairs of indices for an N-state system. For each pair j, k, we multiply eigenvector $|\phi_j^{\mathrm{ad}}(R(t + \mathrm{d}t_c))\rangle$ by $\exp(i\theta_{jk})$ and multiply eigenvector $|\phi_k^{\mathrm{ad}}(R(t + \mathrm{d}t_c))\rangle$ by $\exp(-i\theta_{jk})$; we choose θ_{jk} such that we minimize $\mathrm{Re}(\mathrm{Tr}(3\mathbf{U}^2 - 16\mathbf{U}))$ (see eq 19). The Jacobi sweeps are performed as follows:

```
% flagc is a binary variable to keep track of
whether Jacobi sweeps have converged
Set flagc = 1
loop j = 1 : N
  loop k = (j + 1) : N
    %% Calculate the following four intermediate quantities
    Γ_1^jk = [Σ_l 6Re(U_ij U_jl + U_ik U_kl)] − 12Re(U_jk U_kj) − 6Re(U_jj^2 + U_kk^2) − 16Re(U_ij + U_kk)
    Γ_2^jk = 3Re(U_jj^2 + U_kk^2)
    Ξ_1^jk = [Σ_l 6Im(U_ik U_kl − U_ij U_jl)] − 6Im(U_kk^2 − U_jj^2) − 16Im(U_kk − U_jj)
    Ξ_2^jk = 3Im(U_kk^2 − U_jj^2)
    Δ_complex^jk = Γ_1^jk cos(θ_jk) + Γ_2^jk cos(2θ_jk) + Ξ_1^jk sin(θ_jk) + Ξ_2^jk sin(2θ_jk)
    %% Calculate θ_jk by the four intermediate quantities
    θ_jk = θ_jk(Γ_1^jk, Γ_2^jk, Ξ_1^jk, Ξ_2^jk)
    %% Change the phases of state j and k
        |φ_j^ad(R(t + dt_c))⟩ = exp(iθ_jk)|φ_j^ad(R(t + dt_c))⟩
        |φ_k^ad(R(t + dt_c))⟩ = exp(−iθ_jk)|φ_k^ad(R(t + dt_c))⟩
    if θ_jk ≠ 0
      set flagc = 0
  end
end
```
(23)

Step 4: If flagc = = 0, return to Step 3.

The derivation of $\Delta_{\mathrm{complex}}^{jk}$ is in Appendix A. To solve for θ_{jk}, see Appendix B and Appendix C.

In the end, we have calculated the overlap matrix $\mathbf{U}$ with the new eigenvectors having well-defined phases. [Therefore, if needed, we can numerically compute the logarithm of $\mathbf{U}$ by Schur decomposition, see eq 6.[26]]

2.3. An Approximate Solution Based on Parallel Transport. For most dynamical calculations, choosing the signs of adiabatic states needs to be very fast: it should be *much* faster, for instance, than diagonalization itself. Now, the protocol above suggests looping over all pairs of adiabatic states and performing Jacobi sweeps until convergence is attained (and the signs are fixed). In practice, for a very large matrix, this scheme could take time—although so far, in our experience, the protocol above is always much faster than diagonalizing the electronic Hamiltonians.

Nevertheless, in practice, one might want to fix some adiabatic state phases using parallel transport (i.e., set $U_{jj} \approx 1$ for some j) and then pick other adiabatic state phases with the more advanced scheme above. For this purpose, one would need a cutoff. In our experience, a natural cutoff should be $U_{jj} \leq 1 - \frac{2}{N}$. A great deal of numerical investigation suggests that if $|U_{jj}| > 1 - \frac{2}{N}$, $\forall j$, then parallel transport (i.e., setting U_{jj} positive) already minimizes $\mathrm{Tr}(|\log(\mathbf{U})|^2)$. If some U_{jj} satisfy $|U_{jj}| > 1 - \frac{2}{N}$ (Class I) and other U_{jj} satisfy $|U_{jj}| < 1 - \frac{2}{N}$ (Class II), there is no guarantee that parallel transport is good enough for any state. Nevertheless, if the computation demands are heavy enough, we would recommend setting U_{jj} to be real and positive for states in Class I and deciding U_{jj} with Jacobi sweeps for Class II states. This approximation can be implemented in both the real and the complex regimes.

3. RESULTS

3.1. Numerical Test on the Algorithm in the Real Regime. Before applying the algorithm to a real dynamical

DOI: 10.1021/acs.jctc.9b00952
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

Table 1. Results of a Simple Test of Our Algorithm with 1000 Random Unitary Matrices[a]

matrix dimension N	no. of matrices locating the global minimum of $\mathrm{Tr}(3\mathbf{U}^2 - 16\mathbf{U})$	$\langle \mathrm{Tr}(3\mathbf{U}^2 - 16\mathbf{U})\rangle/13N - \langle \mathrm{Tr}(3\mathbf{U}^2_{\mathrm{global}} - 16\mathbf{U}_{\mathrm{global}})\rangle/13N$	no. of matrices locating the global minimum of $\mathrm{Tr}(\|\log(\mathbf{U})\|^2)$	$\langle \mathrm{Tr}(\|\log(\mathbf{U})\|^2)\rangle/N - \langle \mathrm{Tr}(\|\log(\mathbf{U}_{\mathrm{global}})\|^2)\rangle/N$
2	1000	0	1000	0
3	1000	0	1000	0
4	997	0.106	980	0.075
5	995	0.0726	975	0.063
6	988	0.0367	950	0.04
8	985	0.0251	865	0.0673
10	982	0.0114	720	0.0761

[a]Note that for all cases and all dynamics, our algorithm finds a real proper rotation matrix $\mathbf{U}$ with a real matrix logarithm (i.e., $\mathbf{U}$ has no eigenvalues equal to -1). However, our algorithm does not necessarily locate the globally optimal $\mathbf{U}$ for either the direct target function $\mathrm{Tr}(3\mathbf{U}^2 - 16\mathbf{U})$ or the indirect target function $\mathrm{Tr}(\|\log(\mathbf{U})\|^2)$, especially when N grows larger. Nevertheless, Jacobi sweeps fail for only 18/1000 test cases at finding the global minimum of $\mathrm{Tr}(3\mathbf{U}^2 - 16\mathbf{U})$, and the total deviation $(\langle \mathrm{Tr}(3\mathbf{U}^2 - 16\mathbf{U})\rangle - \langle \mathrm{Tr}(3\mathbf{U}^2_{\mathrm{global}} - 16\mathbf{U}_{\mathrm{global}})\rangle)/13N$ above the optimal $\mathbf{U}_{\mathrm{global}}$ is very small in all cases, suggesting that the Jacobi sweeps algorithm is fairly robust. Admittedly, the failure rate for finding the true global minimum $\mathbf{U}$ is higher for the $\mathrm{Tr}(\|\log(\mathbf{U})\|^2)$ criterion, which is a clear indication of the shortcomings of approximating a matrix logarithm with a polynomial. Nevertheless, as shown in the last column, the matrices $\mathbf{U}$ as obtained from our algorithm do not have very large deviations, and we believe that they should be good enough for dynamics. See the results below and Appendix D.

model problem, we will generate a set of real random orthogonal matrices $\{\mathbf{U}\}$,[27] and test if our algorithm can find the optimal rotation matrix by optimizing the signs of all of the columns of $\mathbf{U}$. In principle, we might not be able to find the proper rotation matrix with the smallest $\mathrm{Tr}(\|\log(\mathbf{U})\|^2)$ by truncating the Taylor series at second order in $\mathbf{U}$ (see section 2.1.2). Rather, we should really calculate the direct target function $\mathrm{Tr}(\|\log(\mathbf{U})\|^2)$ explicitly (which is an infinite order Taylor series in $\mathbf{U}$). However, to show statistically that our algorithm works well enough, for a set of 1000 random unitary matrices of dimension N in Table 1, we will assess how well our algorithm indeed recovers the optimal matrix with the smallest $\mathrm{Tr}(\|\log(\mathbf{U})\|^2)$. Note further that, in principle, the algorithm in eqs 20 and 23 could fail also by finding a local minimum (as opposed to a global minimum) of $\mathrm{Re}(\mathrm{Tr}(3\mathbf{U}^2 - 16\mathbf{U}))$, and so to assess our approach, we will also benchmark how well Jacobi sweeps find the global minimum of our target function $\mathrm{Re}(\mathrm{Tr}(3\mathbf{U}^2 - 16\mathbf{U}))$. As a side note, in all cases, our method is able to find a proper rotation matrix $\mathbf{U}$ with a real matrix logarithm, i.e., $\mathbf{U}$ has no eigenvalues equal to -1.

In Table 1, we benchmark our algorithm as a function of vector space dimension N. As the dimension N grows larger, we find that it does become more difficult to locate the global minimum of $\mathbf{U}$ for either the direct or indirect target functions. For $N = 10$, our algorithm fails to find the global minimum of $\mathrm{Re}(\mathrm{Tr}(3\mathbf{U}^2 - 16\mathbf{U}))$ with probability 1.8%, and we fail to find the global minimum of $\mathrm{Tr}(\|\log(\mathbf{U})\|^2)$ with the approximate polynomial with probability 28%. Nevertheless, in all cases, Table 1 also demonstrates that the $\mathbf{U}$ found by our algorithm is probably good enough. For instance, the deviation $\mathrm{Tr}(\|\log(\mathbf{U})\|^2) - \mathrm{Tr}(\|\log(\mathbf{U}_{\mathrm{global}})\|^2)/N$ is never very large after we perform a Jacobi sweeps minimization: the deviation is only 0.0761 for the case of a 10×10 matrix. Overall, our belief is that the algorithm above should perform very well in practice. Although we have not rigorously tested how large a deviation we can tolerate for accurate dynamics (but see section 3.2), all data so far indicates that, if there are nearly equivalent sign conventions with small $\mathrm{Tr}(\|\log(\mathbf{U})\|^2)$ or $\mathrm{Re}(\mathrm{Tr}(3\mathbf{U}^2 - 16\mathbf{U}))$, the exact choice of sign will not have large consequences; however, large dynamical errors will arise if we select a sign convention that is not one of the nearly equivalent minima (and with a signficantly larger value of $\mathrm{Tr}(\|\log(\mathbf{U})\|^2)$. For the algorithm presented above, we find that small deviations from the target function global minima will arise only when $\mathbf{U}$ is large and very dense (with few zeros); thus, these particular matrices represent the case in which all N states are interacting strongly within the same time step, instead of the more standard case where only a handful of states interact and cross each other. These are not expected to be common situations with reasonable simulation time steps. Note that we would not expect any problems if, e.g., we run a trajectory with 1000 states and there are 250 quartets of states crossings near trivially with each other at one time step; small problems might arise, however, if all 1000 states crossed with each other in one time step. For details, see Appendix D.

3.2. Testing the Algorithm with a Simple Model Problem in the Complex Regime. In a companion paper, we tested Floquet Fewest Switch Surface Hopping (F-FSSH) on a few model problems with real Hamiltonians.[28] In that paper, we used the algorithm above to compute $\mathbf{U}$ in the context of real Hamiltonians. In this paper, we will focus on a similar F-FSSH model problem but now with complex diabatic couplings in the Floquet picture, so that we can test the algorithm for complex Hamiltonians.

With this goal in mind, consider Tully's simple avoided crossing model problem modified to be time-dependent as follows:

$$
\begin{aligned}
H^{\mathrm{el}}_{00}(R) &= A[1 - \exp(-B \times R)] \quad R > 0,\\
H^{\mathrm{el}}_{00}(R) &= -A[1 - \exp(B \times R)], \quad R < 0\\
H^{\mathrm{el}}_{11}(R) &= -H^{\mathrm{el}}_{00}(R)\\
H^{\mathrm{el}}_{10}(R, t) &= H^{el}_{01}(R, t) = C \exp(-D \times R^2) \cos(\omega t + \zeta)
\end{aligned}
\tag{24}
$$

Unless stated otherwise, all parameters will be chosen the same as in ref 28, $A = 0.01$, $B = 1.6$, $C = 0.005$, $D = 1.0$, $\omega = 0.012$, and we set $dt = 1$. In ref 28, we set $\zeta = 0$ such that the only periodic function was the cosine (which has real Fourier components). When an arbitrary phase ζ is introduced, however,

$$
\begin{aligned}
H^{\mathrm{el}}_{10}(R, t) &= H^{\mathrm{el}}_{01}(R, t) = C \exp(-D \times R^2) \cos(\omega t + \zeta)\\
&= V(R) \cos(\omega t + \zeta)
\end{aligned}
\tag{25}
$$

DOI: 10.1021/acs.jctc.9b00952
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

At this point, to run F-FSSH, we apply the relevant Fourier transform (see eq 18 in ref 28) to generate the Floquet Hamiltonian:

$$[\hat{\mathcal{H}}^{F}(R)]_{(n\nu)(m\mu)} = \frac{1}{T_0}\int_0^{T_0} dt\langle\nu|\hat{H}^{el}(R,t)|\mu\rangle \times \exp[-i(n-m)\omega t] + \delta_{\mu\nu}\delta_{mn}n\hbar\omega \quad (26)$$

Thus, the Floquet Hamiltonian after the Fourier-type transformation becomes complex although the electronic Hamiltonian is still real.

$$\cos(\omega t+\zeta) = \frac{\cos(\zeta)+i\sin(\zeta)}{2}\exp(i\omega t) + \frac{\cos(\zeta)-i\sin(\zeta)}{2}\exp(-i\omega t) \quad (27)$$

F-FSSH trajectories are then run as follows: (i) The trajectories are initialized at $R_0 = -9.0$ on Floquet state $\mu = 0$ and $m = 0$. (ii) The infinitely big complex Floquet Hamiltonian is truncated at $m = 4$ and diagonalized to obtain the adiabatic potential quasi-energy surfaces and the eigenvectors. (iii) The nuclear degree of freedom is propagated along the relevant adiabatic Floquet quasi-energy surfaces and the electronic degrees of freedom are propagated according to eq 3. (iv) The time-averaged time-derivative matrix **T** is obtained according to eq 5. (v) Phases of all electronic states along each trajectory is recorded by calculating the product of the relevant overlap matrices **U** (see details in section 3.8 in ref 28). (vi) All trajectories are propagated for certain amounts of time, i.e. enough for them to leave the interaction zone completely. The transmission/reflection probabilities are recorded and calculated (see eq 46 in ref 28 for details).

Now, for the calculations above, another approach would be to simply shift the time coordinate with $t' = t + \zeta/\omega$ and change t to t' accordingly in eq 46 in ref 28. In so doing, one could in fact transform the complex Floquet Hamitlonian in eq 26 into an entirely real form. Nevertheless, for our purposes, benchmarking F-FSSH with a trivially complex Floquet Hamiltonian will be a straightforward test of our algorithm for complex Hamiltonians. Note that if we were to invoke a non-trivially complex Hamiltonian, we would need to discuss Berry's force and an approximate hopping direction,[29,30] which would only complicate the present paper (and will be addressed in a future publication). Note that, when running F-FSSH, one would find that, unless we implement a robust ansatz in the complex regime, multiple trivial crossings at the origin can result in transitions to the wrong dressed states.

Exact and F-FSSH results for $\zeta = \pi/3$ are shown in Figure 2, in comparison to the results with $\zeta = 0$ (red dotted line, which is the same as the black line in Figure 2f in ref 28). Apparently, changing ζ leads to a shift of the oscillation on the exact black line. As in ref 28, F-FSSH performs well if we adopt our phase convention, but the algorithm fails completely if we use simple parallel transport (where we force U_{jj} to be real and positive no matter how small U_{jj} is) and a reasonable time step. These results confirm that in practice, choosing the phases of **U** following the algorithm in section 2.2.2 is robust, efficient, and essential for accuracy. Note also that we recover exactly the same dynamics if we use the approximate scheme for **U**, whereby only some eigenvector signs are optimized beyond parallel transport (see section 2.3).

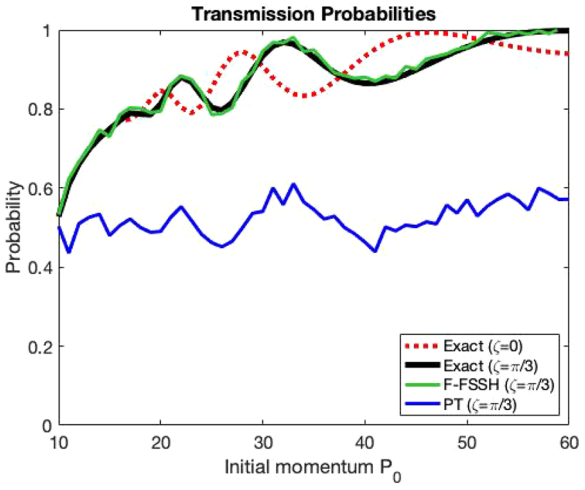


Figure 2. Transmission probabilities on diabat $|0\rangle$ for modified complex simple avoided crossing problem. The black line shows the exact results with phase $\zeta = \frac{\pi}{3}$. The red dotted line represents the exact results with $\zeta = 0$. When $\zeta \neq 0$, there is a clear shift of the phase of oscillation. The green line plots "F-FSSH ($\zeta = \frac{\pi}{3}$)" data. Note that F-FSSH predicts accurately the transmission probabilities using the sign convention outlined above. However, as shown by the blue line labeled "PT ($\zeta = \frac{\pi}{3}$)", straightforward parallel transport results demonstrate that one can find results that are very incorrect if one uses a different convention (where one forces U_{jj} to be real and positive even if U_{jj} is small). As is well known, choosing the phases of adiabatic states is critically important here.[24]

4. CONCLUSIONS

In conclusion, we have presented an ansatz that can efficiently and universally pick phases for parametrized sets of eigenvectors so as to evaluate the time-derivative coupling matrix **T** as smoothly as possible for both avoided crossings and trivial crossings and we have tested our ansatz for both real and complex model problems that contain multiple pairwise trivial crossings at the same time step. To minimize $\mathrm{Tr}(|\log(\mathbf{U})|^2)$, we make the ansatz that we must enforce (i) the exact constraint that $\det(\mathbf{U}) = +1$ and (ii) the approximate constraint that $\mathrm{Re}(\mathrm{Tr}(3\mathbf{U}^2 - 16\mathbf{U}))$ should be minimized. Despite any limitations from the uncontrolled quadratic Taylor series approximation of $\mathrm{Tr}(|\log(\mathbf{U})|^2)$, all results confirm that the ansatz is robust and efficient. In general, our constraints yield a **U** that has very small value of $\mathrm{Tr}(|\log(\mathbf{U})|^2)$, and very often we reach the global minimum. Overall, we are quite confident that our ansatz should be very powerful as far as calculating time-derivative couplings or non-adiabatic couplings for surface hopping calculations or other non-adiabatic dynamics formalisms.

As far as performance, the Jacobi sweeps in section 2.2 (in both the real and complex regimes) converge fast so that there is no significant additional computational cost to implementing our ansatz in surface hopping; one can even apply Jacobi sweeps to a subgroup of the set of adiabatic states and achieve effectively the same results with basically zero cost (see section 2.3). Alternatively, in the future, one can also imagine running Monte Carlo to minimize $\mathrm{Re}(\mathrm{Tr}(3\mathbf{U}^2 - 16\mathbf{U}))$, rather than performing Jacobi sweeps (which is basically steepest descent and is usually not optimal for large systems). [Monte Carlo should also limit any propensity to find local minima for $\mathrm{Re}(\mathrm{Tr}(3\mathbf{U}^2 - 16\mathbf{U}))$ rather than the global minimum.

DOI: 10.1021/acs.jctc.9b00952
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

Admittedly, however, Monte Carlo cannot solve the problem of an inadequate Taylor series, or help us minimize $\mathrm{Tr}(|\log(\mathbf{U})|^2)$ globally if $\mathrm{Tr}(|\log(\mathbf{U})|^2)$ is not minimized by $\mathrm{Re}(\mathrm{Tr}(3\mathbf{U}^2 - 16\mathbf{U}))$.] In practice, running nuclear dynamics and evaluating electronic structure will remain the only bottleneck in dynamics calculations. In the end, the present algorithm, combined with the methods of ref 18, should allow us to run FSSH in a blackbox manner with reasonably large time steps and achieve real gains in cost, while never worrying about trivial or non-trivial crossings.

APPENDIX A: DERIVATION OF $\Delta^{JK}_{\mathrm{REAL}}$, $\Delta^{JK}_{\mathrm{COMPLEX}}$, Γ^{JK}_1, Γ^{JK}_2, Ξ^{JK}_1, AND Ξ^{JK}_2 IN SECTION 2.2

Let us calculate the difference in $\mathrm{Re}(\mathrm{Tr}(3\mathbf{U}^2 - 16\mathbf{U}))$ between overlap matrices $\mathbf{U}$ before and after we multiply state j by $\exp(i\theta_{jk})$ and state k by $\exp(-i\theta_{jk})$. We will start with $\Delta^{jk}_{\mathrm{complex}}$, and then $\Delta^{jk}_{\mathrm{real}}$ can be obtained naturally by assigning $\theta_{jk} = \pi$. We denote the original overlap matrix as $\mathbf{U}$ and the overlap matrix with the phases of two column changed as $\mathbf{U}'$, i.e.,

$$U'_{mn} = \begin{cases} U_{mn} & \text{if } n \neq j,\ k \\ U_{mn}\exp(i\theta_{jk}) & \text{if } n = j \\ U_{mn}\exp(-i\theta_{jk}) & \text{if } n = k \end{cases} \tag{28}$$

Then,

$$\mathrm{Tr}(\mathbf{U}' - \mathbf{U}) = U_{jj}\exp(i\theta_{jk}) + U_{kk}\exp(-i\theta_{jk}) - U_{jj} - U_{kk} \tag{29}$$

$$\begin{aligned}\mathrm{Tr}(\mathbf{U}'^2 - \mathbf{U}^2) = \sum_l \Big\{ & 2U_{jl}U_{lj}(\exp(i\theta_{jk}) - 1) \\ & + 2U_{kl}U_{lk}(\exp(-i\theta_{jk}) - 1)\Big\} \\ & - 2U_{jk}U_{kj}(\exp(i\theta_{jk}) + \exp(-i\theta_{jk}) - 2) \\ & + U_{jj}^2(\exp(i\theta_{jk}) - 1)^2 + U_{kk}^2(\exp(-i\theta_{jk}) - 1)^2\end{aligned} \tag{30}$$

Next, we take the real part of each difference:

$$\begin{aligned}\mathrm{Re}(\mathrm{Tr}(\mathbf{U}' - \mathbf{U})) = & \mathrm{Re}(U_{jj} + U_{kk})\cos\theta_{jk} \\ & + \mathrm{Im}(U_{\mathrm{kk}} - U_{\mathrm{jj}})\sin\theta_{jk} - U_{jj} - U_{kk}\end{aligned} \tag{31}$$

$$\begin{aligned}\mathrm{Re}(\mathrm{Tr}(\mathbf{U}'^2 - \mathbf{U}^2)) = & \sum_l \{2\mathrm{Re}(U_{jl}U_{lj} + U_{kl}U_{lk}) \\ & (\cos\theta_{jk} - 1) + 2\mathrm{Im}(U_{kl}U_{lk} - U_{jl}U_{lj})\sin\theta_{jk}\} \\ & - 4\mathrm{Re}(U_{jk}U_{kj})(\cos\theta_{jk} - 1) + \mathrm{Re}(U_{jj}^2 + U_{kk}^2) \\ & (\cos(2\theta_{jk}) - 2\cos\theta_{jk} + 1) + \mathrm{Im}(U_{kk}^2 - U_{jj}^2) \\ & (\sin(2\theta_{jk}) - 2\sin\theta_{jk})\end{aligned} \tag{32}$$

Since we want to minimize $\Delta^{jk}_{\mathrm{complex}} \approx \mathrm{Re}(\mathrm{Tr}(3\mathbf{U}'^2 - 16\mathbf{U}'))$, we want the difference $\mathrm{Re}(\mathrm{Tr}(3(\mathbf{U}'^2 - \mathbf{U}^2) - 16(\mathbf{U}' - \mathbf{U})))$ to be minimized, and so all constants for a given $\mathbf{U}$ can be dropped. By combining the coefficients of the cosine or sine functions, we obtain

$$\begin{aligned}\Delta^{jk}_{\mathrm{complex}} \equiv & \\ & 3\mathrm{Re}(U_{jj}^2 + U_{kk}^2)\cos(2\theta_{jk}) + 3\mathrm{Im}(U_{kk}^2 - U_{jj}^2)\sin(2\theta_{jk}) \\ & + \Big\{\Big(\sum_l 6\mathrm{Re}(U_{jl}U_{lj} + U_{kl}U_{lk})\Big) - 12\mathrm{Re}(U_{jk}U_{kj}) \\ & - 6\mathrm{Re}(U_{jj}^2 + U_{kk}^2) - 16\mathrm{Re}(U_{jj} + U_{kk})\Big\}\cos\theta_{jk} \\ & + \Big\{\Big(\sum_l 6\mathrm{Im}(U_{kl}U_{lk} - U_{jl}U_{lj})\Big) - 6\mathrm{Im}(U_{kk}^2 - U_{jj}^2) \\ & - 16\mathrm{Im}(U_{kk} - U_{jj})\Big\}\sin\theta_{jk} \\ \equiv & \Gamma_2^{jk}\cos(2\theta_{jk}) + \Xi_2^{jk}\sin(2\theta_{jk}) \\ & + \Gamma_1^{jk}\cos\theta_{jk} + \Xi_1^{jk}\sin\theta_{jk}\end{aligned} \tag{33}$$

Here, the coefficients are the same as those in section 2.2.2.

For the real case, rather than minimizing a function, we would like simply to check whether a difference $\Delta^{jk}_{\mathrm{real}}$ is positive or negative. To that end, we will keep the constants that we dropped above between eqs 31,32 and eq 33. By setting $\theta_{jk} = \pi$, we can obtain the relevant difference quickly by combining eqs 31 and 32 in the real regime:

$$\begin{aligned}\Delta^{jk}_{\mathrm{real}} \equiv & 3(U_{jj}^2 + U_{kk}^2) + 6(U_{jk}U_{kj}) + 8(U_{jj} + U_{kk}) \\ & - \sum_l 3(U_{jl}U_{lj} + U_{kl}U_{lk})\end{aligned} \tag{34}$$

APPENDIX B: MINIMIZATION OF $\Delta^{JK}_{\mathrm{COMPLEX}}$

In order to find θ_{jk} by minimizing $\Delta^{jk}_{\mathrm{complex}}$ (eq 33),

$$\begin{aligned}\Delta^{jk}_{\mathrm{complex}} \equiv & \Gamma_2^{jk}\cos(2\theta_{jk}) + \Xi_2^{jk}\sin(2\theta_{jk}) + \\ & \Gamma_1^{jk}\cos\theta_{jk} + \Xi_1^{jk}\sin\theta_{jk}\end{aligned} \tag{35}$$

we take the derivative with respect to θ_{jk}:

$$\begin{aligned}\frac{\mathrm{d}\Delta^{jk}_{\mathrm{complex}}}{\mathrm{d}\theta_{jk}} \equiv & 2\Xi_2^{jk}\cos(2\theta_{jk}) - 2\Gamma_2^{jk}\sin(2\theta_{jk}) \\ & - \Gamma_1^{jk}\sin\theta_{jk} + \Xi_1^{jk}\cos\theta_{jk}\end{aligned} \tag{36}$$

At a minimum, this derivative must equal 0.

To solve this equation, we set $x = \cos\theta_{jk}$. Eq 36 becomes

$$\begin{aligned}& 2\Xi_2^{jk}(2x^2 - 1) - 4\Gamma_2^{jk}x\sqrt{1 - x^2} - \Gamma_1^{jk}\sqrt{1 - x^2} + \Xi_1^{jk}x \\ & = 0\end{aligned} \tag{37}$$

By rearranging the equation, we obtain a quartic equation:

$$\begin{aligned}& 16[(\Xi_2^{jk})^2 + (\Gamma_2^{jk})^2]x^4 + 8(\Xi_1^{jk}\Xi_2^{jk} + \Gamma_1^{jk}\Gamma_2^{jk})x^3 \\ & + [(\Xi_1^{jk})^2 + (\Gamma_1^{jk})^2 - 16(\Xi_2^{jk})^2 - 16(\Gamma_2^{jk})^2]x^2 \\ & + 4(\Xi_1^{jk}\Xi_2^{jk} + 2\Gamma_1^{jk}\Gamma_2^{jk})x + [4(\Xi_2^{jk})^2 - (\Gamma_1^{jk})^2] = 0\end{aligned} \tag{38}$$

There are three scenarios. First, when $(\Xi_2^{jk})^2 + (\Gamma_2^{jk})^2 = 0$, eq 36 reduces to

DOI: 10.1021/acs.jctc.9b00952
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

$$\Gamma_1^{jk} \sin \theta_{jk} = \Xi_1^{jk} \cos \theta_{jk} \tag{39}$$

Note that θ_{jk} has a period 2π while arctan (x) has a period π, and thus, there are two roots,

$$\theta_{jk}^1 = \arctan\left(\frac{\Xi_1^{jk}}{\Gamma_1^{jk}}\right) \tag{40}$$

$$\theta_{jk}^2 = \arctan\left(\frac{\Xi_1^{jk}}{\Gamma_1^{jk}}\right) + \pi \tag{41}$$

We pick the θ_{jk} that makes eq 35 smallest. The physical meaning is that we have encountered a pairwise trivial crossing between states j and k.

The second scenario is that we have to solve this quartic equation without any possible reduction. We can either apply the general form of the solution to any quartic equation, or by constructing the companion matrix of eq 38 (see Appendix C). For a quartic equation, we may have four real roots at most. Again, θ_{jk} has a period 2π. Thus, for each real root, there are two possible θ_{jk}:

$$\theta_{jk}^1 = \arccos(x) \tag{42}$$

$$\theta_{jk}^2 = -\arccos(x) \tag{43}$$

In total, we have eight roots at most. Since we are trying to find the minimum of $\Delta_{\text{complex}}^{jk}$, we may simply calculate $\Delta_{\text{complex}}^{jk}$ for all real roots and choose the root with smallest $\Delta_{\text{complex}}^{jk}$.

In principle, there is a third scenario that is the most tricky: all four coefficients could effectively be zeros. In this case, the physical meaning is that within the time step, we encounter one or more multistate trivial crossing between state j, k and at least one other state, and there are no other adiabatic states $\{m\}$ that talk to either j or k: $\sum_m U_{jm}U_{mj} = \sum_k U_{jk}U_{kj} = 0$. For this situation, the ansatz above, based on truncation at second order, is no longer valid. In theory, one could derive a similar but higher-order algorithm to solve for the phases of $\mathbf{U}$ to minimize $\text{Tr}(|\log(\mathbf{U})|^2)$ more accurately. Nevertheless, we believe this scenario should not be very physically relevant, since the phases of different adiabatic states (j and k) can matter only when two systems interact directly or indirectly and in such a situation, it seems very unlikely there will not be one single state that interacts with either j or k at the same time (so that $(\mathbf{U}^2)_{jj} = (\mathbf{U}^2)_{kk} = 0$). As a practical matter, we believe truncating at second order should be sufficient.

APPENDIX C: FINDING ROOTS OF A POLYNOMIAL

For a monic polynomial equation with real coefficients,

$$p(x) = x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0 = 0 \tag{44}$$

there will be n roots (and some of them may be complex). The roots can be obtained by constructing the Frobenius companion matrix $\mathbf{C}(p)$:

$$\mathbf{C}(p) = \begin{bmatrix} 0 & 0 & \ldots & 0 & -a_0 \\ 1 & 0 & \ldots & 0 & -a_1 \\ 0 & 1 & \ldots & 0 & -a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & -a_{n-1} \end{bmatrix}$$

The eigenvalues which can be obtained by performing a Schur decomposition are the roots of the polynomial equation $p(x)$.

APPENDIX D: LIMITATIONS OF OUR ALGORITHM

In this appendix, we want to explicitly explore the different cases where our algorithm breaks down. We will consider explicitly three different 4×4 unitary matrices **A**, **B**, and **C**.

The first matrix **A** has the following sign possibilities:

$$\mathbf{A}_1 = \begin{bmatrix} 0.6575 & -0.3565 & -0.6354 & -0.1920 \\ 0.1351 & 0.6081 & -0.4038 & 0.6700 \\ 0.0916 & 0.6991 & -0.0847 & -0.7041 \\ 0.7355 & 0.1199 & 0.6527 & 0.1363 \end{bmatrix}$$

$$\mathbf{A}_2 = \begin{bmatrix} 0.6575 & -0.3565 & 0.6354 & 0.1920 \\ 0.1351 & 0.6081 & 0.4038 & -0.6700 \\ 0.0916 & 0.6991 & 0.0847 & 0.7041 \\ 0.7355 & 0.1199 & -0.6527 & -0.1363 \end{bmatrix}$$

$$\mathbf{A}_3 = \begin{bmatrix} 0.6575 & 0.3565 & 0.6354 & -0.1920 \\ 0.1351 & -0.6081 & 0.4038 & 0.6700 \\ 0.0916 & -0.6991 & 0.0847 & -0.7041 \\ 0.7355 & -0.1199 & -0.6527 & 0.1363 \end{bmatrix}$$

$$\mathbf{A}_4 = \begin{bmatrix} -0.6575 & -0.3565 & 0.6354 & -0.1920 \\ -0.1351 & 0.6081 & 0.4038 & 0.6700 \\ -0.0916 & 0.6991 & 0.0847 & -0.7041 \\ -0.7355 & 0.1199 & -0.6527 & 0.1363 \end{bmatrix}$$

$$\mathbf{A}_5 = \begin{bmatrix} -0.6575 & -0.3565 & -0.6354 & 0.1920 \\ -0.1351 & 0.6081 & -0.4038 & -0.6700 \\ -0.0916 & 0.6991 & -0.0847 & 0.7041 \\ -0.7355 & 0.1199 & 0.6527 & -0.1363 \end{bmatrix}$$

$$\mathbf{A}_6 = \begin{bmatrix} 0.6575 & 0.3565 & -0.6354 & 0.1920 \\ 0.1351 & -0.6081 & -0.4038 & -0.6700 \\ 0.0916 & -0.6991 & -0.0847 & 0.7041 \\ 0.7355 & -0.1199 & 0.6527 & -0.1363 \end{bmatrix}$$

$$\mathbf{A}_7 = \begin{bmatrix} -0.6575 & 0.3565 & 0.6354 & 0.1920 \\ -0.1351 & -0.6081 & 0.4038 & -0.6700 \\ -0.0916 & -0.6991 & 0.0847 & 0.7041 \\ -0.7355 & -0.1199 & -0.6527 & -0.1363 \end{bmatrix}$$

$$\mathbf{A}_8 = \begin{bmatrix} -0.6575 & 0.3565 & -0.6354 & -0.1920 \\ -0.1351 & -0.6081 & -0.4038 & 0.6700 \\ -0.0916 & -0.6991 & -0.0847 & -0.7041 \\ -0.7355 & -0.1199 & 0.6527 & 0.1363 \end{bmatrix}$$

The next matrix **B** has the following sign possibilities:

DOI: 10.1021/acs.jctc.9b00952
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

$$\mathbf{B}_1 = \begin{bmatrix} 0.5987 & 0.1138 & 0.5219 & 0.5969 \\ -0.5288 & 0.5520 & 0.6321 & -0.1274 \\ -0.5694 & -0.1139 & -0.2188 & 0.7842 \\ 0.1942 & 0.8182 & -0.5294 & 0.1121 \end{bmatrix}$$

$$\mathbf{B}_2 = \begin{bmatrix} 0.5987 & 0.1138 & -0.5219 & -0.5969 \\ -0.5288 & 0.5520 & -0.6321 & 0.1274 \\ -0.5694 & -0.1139 & 0.2188 & -0.7842 \\ 0.1942 & 0.8182 & 0.5294 & -0.1121 \end{bmatrix}$$

$$\mathbf{B}_3 = \begin{bmatrix} 0.5987 & -0.1138 & -0.5219 & 0.5969 \\ -0.5288 & -0.5520 & -0.6321 & -0.1274 \\ -0.5694 & 0.1139 & 0.2188 & 0.7842 \\ 0.1942 & -0.8182 & 0.5294 & 0.1121 \end{bmatrix}$$

$$\mathbf{B}_4 = \begin{bmatrix} -0.5987 & 0.1138 & -0.5219 & 0.5969 \\ 0.5288 & 0.5520 & -0.6321 & -0.1274 \\ 0.5694 & -0.1139 & 0.2188 & 0.7842 \\ -0.1942 & 0.8182 & 0.5294 & 0.1121 \end{bmatrix}$$

$$\mathbf{B}_5 = \begin{bmatrix} -0.5987 & 0.1138 & 0.5219 & -0.5969 \\ 0.5288 & 0.5520 & 0.6321 & 0.1274 \\ 0.5694 & -0.1139 & -0.2188 & -0.7842 \\ -0.1942 & 0.8182 & -0.5294 & -0.1121 \end{bmatrix}$$

$$\mathbf{B}_6 = \begin{bmatrix} 0.5987 & -0.1138 & 0.5219 & -0.5969 \\ -0.5288 & -0.5520 & 0.6321 & 0.1274 \\ -0.5694 & 0.1139 & -0.2188 & -0.7842 \\ 0.1942 & -0.8182 & -0.5294 & -0.1121 \end{bmatrix}$$

$$\mathbf{B}_7 = \begin{bmatrix} -0.5987 & -0.1138 & -0.5219 & -0.5969 \\ 0.5288 & -0.5520 & -0.6321 & 0.1274 \\ 0.5694 & 0.1139 & 0.2188 & -0.7842 \\ -0.1942 & -0.8182 & 0.5294 & -0.1121 \end{bmatrix}$$

$$\mathbf{B}_8 = \begin{bmatrix} -0.5987 & -0.1138 & 0.5219 & 0.5969 \\ 0.5288 & -0.5520 & 0.6321 & -0.1274 \\ 0.5694 & 0.1139 & -0.2188 & 0.7842 \\ -0.1942 & -0.8182 & -0.5294 & 0.1121 \end{bmatrix}$$

The third matrix **C** has the following sign possibilities:

$$\mathbf{C}_1 = \begin{bmatrix} 0.1451 & 0.6731 & -0.7116 & 0.1397 \\ -0.9396 & -0.0885 & -0.3019 & -0.1351 \\ 0.1431 & 0.1265 & -0.0437 & -0.9806 \\ -0.2751 & 0.7233 & 0.6329 & 0.0249 \end{bmatrix}$$

$$\mathbf{C}_2 = \begin{bmatrix} -0.1451 & -0.6731 & 0.7116 & -0.1397 \\ 0.9396 & 0.0885 & 0.3019 & 0.1351 \\ -0.1431 & -0.1265 & 0.0437 & 0.9806 \\ 0.2751 & -0.7233 & -0.6329 & -0.0249 \end{bmatrix}$$

$$\mathbf{C}_3 = \begin{bmatrix} 0.1451 & 0.6731 & 0.7116 & -0.1397 \\ -0.9396 & -0.0885 & 0.3019 & 0.1351 \\ 0.1431 & 0.1265 & 0.0437 & 0.9806 \\ -0.2751 & 0.7233 & -0.6329 & -0.0249 \end{bmatrix}$$

$$\mathbf{C}_4 = \begin{bmatrix} -0.1451 & -0.6731 & -0.7116 & 0.1397 \\ 0.9396 & 0.0885 & -0.3019 & -0.1351 \\ -0.1431 & -0.1265 & -0.0437 & -0.9806 \\ 0.2751 & -0.7233 & 0.6329 & 0.0249 \end{bmatrix}$$

$$\mathbf{C}_5 = \begin{bmatrix} 0.1451 & -0.6731 & 0.7116 & 0.1397 \\ -0.9396 & 0.0885 & 0.3019 & -0.1351 \\ 0.1431 & -0.1265 & 0.0437 & -0.9806 \\ -0.2751 & -0.7233 & -0.6329 & 0.0249 \end{bmatrix}$$

$$\mathbf{C}_6 = \begin{bmatrix} 0.1451 & -0.6731 & -0.7116 & -0.1397 \\ -0.9396 & 0.0885 & -0.3019 & 0.1351 \\ 0.1431 & -0.1265 & -0.0437 & 0.9806 \\ -0.2751 & -0.7233 & 0.6329 & -0.0249 \end{bmatrix}$$

$$\mathbf{C}_7 = \begin{bmatrix} -0.1451 & 0.6731 & 0.7116 & 0.1397 \\ 0.9396 & -0.0885 & 0.3019 & -0.1351 \\ -0.1431 & 0.1265 & 0.0437 & -0.9806 \\ 0.2751 & 0.7233 & -0.6329 & 0.0249 \end{bmatrix}$$

$$\mathbf{C}_8 = \begin{bmatrix} -0.1451 & 0.6731 & -0.7116 & -0.1397 \\ 0.9396 & -0.0885 & -0.3019 & 0.1351 \\ -0.1431 & 0.1265 & -0.0437 & 0.9806 \\ 0.2751 & 0.7233 & 0.6329 & -0.0249 \end{bmatrix}$$

In Tables 2, 3 ,and 4, we analyze the matrices **A**, **B**, and **C** together with different sign conventions. The fact that **A**, **B**, and **C** are dense 4×4 matrices implies that four states are crossing with each other strongly (which is uncommon in reality). We report the two relevant quantities $\mathrm{Tr}(3\mathbf{U}^2 - 16\mathbf{U})$ and $\mathrm{Tr}(|\log(\mathbf{U})|^2)$ for all possible sign conventions.

Table 2. An Example for Which Our Jacobi Sweeps Method Successfully Locates the Global Minimum of $\mathrm{Tr}(|\log(\mathbf{A})|^2)$[a]

sign convention	$\mathrm{Tr}(3\mathbf{A}^2 - 16\mathbf{A})$	$\mathrm{Tr}(\lvert\log(\mathbf{A})\rvert^2)$
$\mathbf{A}_1$	−24.0463	6.8250
$\mathbf{A}_2$	−17.5786	7.7976
$\mathbf{A}_3$	−1.4704	11.8361
$\mathbf{A}_4$	5.4582	14.0330
$\mathbf{A}_5$	7.1824	14.6045
$\mathbf{A}_6$	10.9494	16.8259
$\mathbf{A}_7$	18.1041	17.1368
$\mathbf{A}_8$	21.2694	22.2017

[a]The matrix $\mathbf{A}_1$ is the global minimum of both quantities $\mathrm{Tr}(|\log(\mathbf{A})|^2)$ and $\mathrm{Tr}(3\mathbf{A}^2 - 16\mathbf{A})$.

In Table 2, for the **A** matrix, we show that the two quantities $\mathrm{Tr}(3\mathbf{A}^2 - 16\mathbf{A})$ and $\mathrm{Tr}(|\log(\mathbf{A})|^2)$ share the same trends for different sign conventions: minimizing $\mathrm{Tr}(3\mathbf{A}^2 - 16\mathbf{A})$ is consistent with minimizing $\mathrm{Tr}(|\log(\mathbf{A})|^2)$. In Table 3, however,

DOI: 10.1021/acs.jctc.9b00952
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

Table 3. An Example for Which Our Jacobi Sweeps Method Fails To Locate the Global Minimum of $\mathrm{Tr}(|\log(\mathbf{B})|^2)$[a]

sign convention	$\mathrm{Tr}(3\mathbf{B}^2 - 16\mathbf{B})$	$\mathrm{Tr}(\|\log(\mathbf{B})\|^2)$
$\mathbf{B}_2$	−18.6547	7.5890
$\mathbf{B}_1$	−19.5302	7.7673
$\mathbf{B}_4$	−2.1964	11.6490
$\mathbf{B}_3$	1.6531	12.6090
$\mathbf{B}_6$	6.8980	14.4798
$\mathbf{B}_7$	13.8778	16.0589
$\mathbf{B}_5$	13.7363	19.3474
$\mathbf{B}_8$	21.5821	21.2692

[a]The quantity $\mathrm{Tr}(|\log(\mathbf{B})|^2)$ is minimized by sign convention $\mathbf{B}_2$, but the quantity $\mathrm{Tr}(3\mathbf{B}^2 - 16\mathbf{B})$ is minimized by sign convention $\mathbf{B}_1$. The difference (or deviation) arises from a failure of the polynomial approximation $\mathrm{Tr}(3\mathbf{B}^2 - 16\mathbf{B}) \approx \mathrm{Tr}(|\log(\mathbf{B})|^2)$.

Table 4. An Example for Which Our Jacobi Sweeps Method Fails To Locate the Global Minimum of Either $\mathrm{Tr}(3\mathbf{C}^2 - 16\mathbf{C})$ or $\mathrm{Tr}(|\log(\mathbf{C})|^2)$[a]

sign convention	$\mathrm{Tr}(3\mathbf{C}^2 - 16\mathbf{C})$	$\mathrm{Tr}(\|\log(\mathbf{C})\|^2)$
$\mathbf{C}_1$	−9.686	10.2505
$\mathbf{C}_2$	−8.4764	10.5041
$\mathbf{C}_3$	−6.9735	10.7389
$\mathbf{C}_4$	−4.5607	11.2894
$\mathbf{C}_5$	3.515	13.2578
$\mathbf{C}_6$	4.2351	13.5732
$\mathbf{C}_7$	9.5151	15.8454
$\mathbf{C}_8$	13.1854	18.9589

[a]Both quantities $\mathrm{Tr}(3\mathbf{C}^2 - 16\mathbf{C})$ and $\mathrm{Tr}(|\log(\mathbf{C})|^2)$ are minimized by $\mathbf{C}_1$; however, it requires a simultaneous change of signs of all columns to transform from $\mathbf{C}_2$ to $\mathbf{C}_1$, which is not possible with a pair-wise Jacobi sweeps method. From Table 1, this situation occurs only with 3/1000 probabilities. In the future, such a situation could be easily addressed with a Monte Carlo simulation.

we show that our method cannot find the global minimum of $\mathrm{Tr}(|\log(\mathbf{B})|^2)$, because $\mathrm{Tr}(3\mathbf{B}^2 - 16\mathbf{B})$ is not completely consistent with $\mathrm{Tr}(|\log(\mathbf{B})|^2)$. Nevertheless, the difference is small and our method does locate the second best sign convention as far as $\mathrm{Tr}(|\log(\mathbf{B})|^2)$ is concerned. Lastly, in Table 4, our method fails to find the global minimum of either $\mathrm{Tr}(3\mathbf{C}^2 - 16\mathbf{C})$ or $\mathrm{Tr}(|\log(\mathbf{C})|^2)$. Although $\mathrm{Tr}(3\mathbf{C}^2 - 16\mathbf{C})$ is completely consistent with $\mathrm{Tr}(|\log(\mathbf{C})|^2)$, a pairwise Jacobi sweeps cannot transform $\mathbf{C}_2$ to $\mathbf{C}_1$, since it requires simultaneous changes of the signs of all four columns. Nevertheless, the difference is small and this situation is highly unlikely with an occurrence probability of roughly 3/1000.

Overall, we expect that the overlap matrix should usually (i) be very close to the identity when there is no trivial crossing or (ii) have maximal element in each column to be close to ±1 (in magnitude) when there is at least one trivial crossing. According to the results in Tables 2, 3, and 4, we conclude that our sign convention is mostly reliable and consistent, unless there are too many states crossing with each other at the same time step with strong diabatic couplings. Even in such rare cases, however, we do still recover a proper rotation matrix with a real matrix logarithm. We do not believe there should ever (in practice) be a need to reduce time step or derive higher order terms in the Taylor series on account of trivial crossings.

AUTHOR INFORMATION

Corresponding Author

*E-mail: subotnik@sas.upenn.edu.

ORCID

Zeyu Zhou: 0000-0001-8681-9460

Tian Qiu: 0000-0001-8510-894X

Andrew M. Rappe: 0000-0003-4620-6496

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by the U.S. Air Force Office of Scientific Research (USAFOSR) AFOSR Grants No. FA9550-18-1-0497 and FA9550-18-1-0420. T.Q. acknowledges the support from the Vagelos Institute for Energy Science and Technology (VIEST). A.M.R. acknowledges the support of the U.S. Department of Energy, Office of Basic Energy Sciences, under grant DE-SC0019281.

REFERENCES

(1) Longuet-Higgins, H. C.; Öpik, U.; Pryce, M. H. L.; Sack, R. Studies of the Jahn-Teller effect. II. The dynamical problem. *Proc. R. Soc. London, Ser. A* **1958**, *244*, 1−16.

(2) Baer, M. Adiabatic and diabatic representations for atom-molecule collisions: Treatment of the collinear arrangement. *Chem. Phys. Lett.* **1975**, *35*, 112−118.

(3) Mead, C. A.; Truhlar, D. G. On the determination of Born−Oppenheimer nuclear motion wave functions including complications due to conical intersections and identical nuclei. *J. Chem. Phys.* **1979**, *70*, 2284−2296.

(4) Berry, M. V. Quantal phase factors accompanying adiabatic changes. *Proc. R. Soc. London, Ser. A* **1984**, *392*, 45−57.

(5) Simon, B. Holonomy, the quantum adiabatic theorem, and Berry's phase. *Phys. Rev. Lett.* **1983**, *51*, 2167.

(6) Yarkony, D. R. Diabolical conical intersections. *Rev. Mod. Phys.* **1996**, *68*, 985.

(7) Yarkony, D. R. Current issues in non-adiabatic chemistry. *J. Phys. Chem.* **1996**, *100*, 18612−18628.

(8) Baer, M. *Beyond Born-Oppenheimer: electronic non-adiabatic coupling terms and conical intersections*; Wiley: New York, 2006.

(9) Tully, J. C. Molecular dynamics with electronic transitions. *J. Chem. Phys.* **1990**, *93*, 1061−1071.

(10) Zaari, R. R.; Varganov, S. A. Nonadiabatic transition state theory and trajectory surface hopping dynamics: intersystem crossing between 3B1 and 1A1 states of SiH2. *J. Phys. Chem. A* **2015**, *119*, 1332−1338.

(11) Ben-Nun, M.; Quenneville, J.; Martínez, T. J. Ab initio multiple spawning: Photochemistry from first principles quantum molecular dynamics. *J. Phys. Chem. A* **2000**, *104*, 5161−5175.

(12) Hammes-Schiffer, S.; Tully, J. C. Proton transfer in solution: Molecular dynamics with quantum transitions. *J. Chem. Phys.* **1994**, *101*, 4657−4667.

(13) Fabiano, E.; Keal, T.; Thiel, W. Implementation of surface hopping molecular dynamics using semiempirical methods. *Chem. Phys.* **2008**, *349*, 334−347.

(14) Barbatti, M.; Pittner, J.; Pederzoli, M.; Werner, U.; Mitrić, R.; Bonačić-Koutecký, V.; Lischka, H. Non-adiabatic dynamics of pyrrole: Dependence of deactivation mechanisms on the excitation energy. *Chem. Phys.* **2010**, *375*, 26−34.

(15) Meek, G. A.; Levine, B. G. Evaluation of the time-derivative coupling for accurate electronic state transition probabilities from numerical simulations. *J. Phys. Chem. Lett.* **2014**, *5*, 2351−2356.

(16) Wang, L.; Prezhdo, O. V. A simple solution to the trivial crossing problem in surface hopping. *J. Phys. Chem. Lett.* **2014**, *5*, 713−719.

DOI: 10.1021/acs.jctc.9b00952
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

(17) Wang, L.; Akimov, A.; Prezhdo, O. V. Recent progress in surface hopping: 2011−2015. *J. Phys. Chem. Lett.* **2016**, *7*, 2100−2112.
(18) Jain, A.; Alguire, E.; Subotnik, J. E. An efficient, augmented surface hopping algorithm that includes decoherence for use in large-scale simulations. *J. Chem. Theory Comput.* **2016**, *12*, 5256−5268.
(19) Dell'Angelo, D.; Hanna, G. Importance of eigenvector sign consistency in computations of expectation values via mixed quantum-classical surface-hopping dynamics. *Theor. Chem. Acc.* **2017**, *136*, 75.
(20) Lee, E. M.; Willard, A. P. Solving the trivial crossing problem while preserving the nodal symmetry of the wavefunction. *J. Chem. Theory Comput.* **2019**, *15*, 4332.
(21) Knowledge of the signs of the columns of **U** is also essential for using the local diabatization protocol for propagation,[22,23] which is an alternative to eq 6.
(22) Granucci, G.; Persico, M.; Spighi, G. Surface hopping trajectory simulations with spin-orbit and dynamical couplings. *J. Chem. Phys.* **2012**, *137*, 22A501.
(23) Plasser, F.; Granucci, G.; Pittner, J.; Barbatti, M.; Persico, M.; Lischka, H. Surface hopping dynamics using a locally diabatic formalism: Charge transfer in the ethylene dimer cation and excited state dynamics in the 2-pyridone dimer. *J. Chem. Phys.* **2012**, *137*, 22A514.
(24) Akimov, A. V. A Simple Phase Correction Makes a Big Difference in Nonadiabatic Molecular Dynamics. *J. Phys. Chem. Lett.* **2018**, *9*, 6096−6102.
(25) Edmiston, C.; Ruedenberg, K. Localized atomic and molecular orbitals. *Rev. Mod. Phys.* **1963**, *35*, 457.
(26) Loring, T. A. Computing a logarithm of a unitary matrix with general spectrum. *Numer. Linear Algebr. Appl.* **2014**, *21*, 744−760.
(27) Diaconis, P.; Shahshahani, M. The subgroup algorithm for generating uniform random variables. *Probab. Eng. Inf. Sci.* **1987**, *1*, 15−32.
(28) Zhou, Z.; Chen, H.-T.; Nitzan, A.; Subotnik, J. E. Non-adiabatic Dynamics in a Laser Field: Using Floquet Fewest Switches Surface Hopping to Calculate Electronic Populations for Slow Nuclear Velocities. To be submitted, 2019.
(29) Miao, G.; Bellonzi, N.; Subotnik, J. An extension of the fewest switches surface hopping algorithm to complex Hamiltonians and photophysics in magnetic fields: Berry curvature and "magnetic" forces. *J. Chem. Phys.* **2019**, *150*, 124101.
(30) Subotnik, J.; Miao, G.; Bellonzi, N.; Teh, H.-H.; Dou, W. A demonstration of consistency between the quantum classical Liouville equation and Berry's phase and curvature for the case of complex Hamiltonians. *J. Chem. Phys.* **2019**, *151*, 074113.

DOI: 10.1021/acs.jctc.9b00952
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX