Path optimization by a variational reaction coordinate method. II. Improved computational efficiency through internal coordinates and surface interpolation

Adam B. Birkholz and H. Bernhard Schlegel

Department of Chemistry, Wayne State University, Detroit, Michigan 48202, USA

Reaction path optimization is being used more frequently as an alternative to the standard practice of locating a transition state and following the path downhill. The Variational Reaction Coordinate (VRC) method was proposed as an alternative to chain-of-states methods like nudged elastic band and string method. The VRC method represents the path using a linear expansion of continuous basis functions, allowing the path to be optimized variationally by updating the expansion coefficients to minimize the line integral of the potential energy gradient norm, referred to as the Variational Reaction Energy (VRE) of the path. When constraints are used to control the spacing of basis functions and to couple the minimization of the VRE with the optimization of one or more individual points along the path (representing transition states and intermediates), an approximate path as well as the converged geometries of transition states and intermediates along the path are determined in only a few iterations. This algorithmic efficiency comes at a high per-iteration cost due to numerical integration of the VRE derivatives. In the present work, methods for incorporating redundant internal coordinates and potential energy surface interpolation into the VRC method are described. With these methods, the per-iteration cost, in terms of the number of potential energy surface evaluations, of the VRC method is reduced while the high algorithmic efficiency is maintained. *Published by AIP Publishing.* [http://dx.doi.org/10.1063/1.4948439]

I. INTRODUCTION

In our Paper I,[1] a variational approach to approximating the steepest descent reaction path (SDRP) was discussed as an alternative to the "chain of states" (CoSs) methods.[2–14] The chain of states methods minimize the energies of a series of discrete structures along the path, subject to fictitious forces, constraints or reparameterization schemes that ensure the distribution of structures remains uniform. These path optimization methods are primarily used to avoid the difficult problem of determining the transition state geometry, which is the starting point for more economical and accurate reaction path following methods[15–18] that determine the SDRP by walking downhill on the potential energy surface. The SDRP can be viewed as a simple approximation to the path a reaction follows as it proceeds from reactant to product. The Variational Reaction Coordinate (VRC) method is an alternative to the chain of states class of path optimization methods. In the VRC method, the Variational Reaction Energy (VRE) is functional which is minimized by the SDRP[19,20] and is defined as the line integral of the potential energy gradient norm

$$E_{VRE} = \int_{t_R}^{t_P} \sqrt{\frac{\partial V(\mathbf{x}(t))^{\mathrm{T}}}{\partial x} \frac{\partial V(\mathbf{x}(t))}{\partial x}} \sqrt{\frac{d\mathbf{x}(t)^{\mathrm{T}}}{dt} \frac{d\mathbf{x}(t)}{dt}} dt$$

$$= \int_{t_R}^{t_P} |\mathbf{g}(\mathbf{x}(t))| \, |\boldsymbol{\tau}(t)| \, dt, \tag{1}$$

where V is the potential energy, $\mathbf{x}(t)$ are the coordinates of the reaction path expressed as a linear expansion in a basis

of functions parameterized by t; $t_R = 0$ and $t_P = 1$ are the parameter values corresponding to the reactant and product structures, respectively. For clarity, $\mathbf{g}$ and $\boldsymbol{\tau}$ are used as shorthand for the gradient of the potential, and the tangent to the path. The path is improved by updating the basis set linear expansion coefficients (LECs) to minimize the VRE subject to arc length and coupling constraints. The arc length constraints impose a relationship between the parameter t and the arc length of the path in order to improve the stability of the optimization, whereas the coupling constraints require the path to pass through structures produced by the optimization of individual geometries corresponding to transition states and intermediates along the path. This combined optimization, the Focused VRC (FVRC) method, focuses the effort of the path optimization towards the regions of the potential likely to contain the SDRP method and is able to determine both a good approximation to the SDRP while simultaneously optimizing the geometries of any transition states or intermediate structures along the path with only a small number of iterations required for convergence. This algorithmic efficiency comes at the expense of a high per-iteration cost, due to the necessity of evaluating the VRE and its derivatives by numerical quadrature methods, and our previous tests were limited to demonstrations on analytical test surfaces.[1] This paper focuses on the development of additional methods to improve the applicability of the FVRC method to the study of chemical reactions, as well as to reduce the per-iteration cost to something comparable to existing CoS path optimization approaches.

The FVRC method is summarized in greater depth in Section II. The FVRC was developed using a potential energy surface expressed in Cartesian coordinates. When studying chemical potential energy surfaces (PESs), internal coordinates consisting of a redundant combination of bond stretches, valence angle bends, and dihedral angle torsions typically provide a much more natural description of the relative motion of atoms.[21] This results in less coupling between coordinates, and a potential energy surface that is likely to be more constant in the degrees of freedom not directly involved with a reaction. Incorporating these redundant internal coordinates (RIC) into the VRC method should not only have a positive effect on the efficiency of the algorithm, but should also improve the accuracy of methods used to approximate the PES (i.e., through Hessian updating and interpolation of a finite number of calculations) and help reduce the per-iteration cost.

One challenge that exists in incorporating RIC into the VRC method is that there is not a unique and closed expression for computing the Cartesian embedding of a particular RIC geometry; RIC displacements must be embedded iteratively from a nearby Cartesian geometry. This poses a problem as the Cartesian coordinates are necessary to evaluate the potential energy surface for a given geometry. For this reason, the present work is limited to methods which retain a description of the path in Cartesian coordinates. This avoids the introduction of any possible error/uncertainty in the path due to back-transformation procedures, but may also limit the overall benefit of using RIC as some of the more curvilinear motions from the bends and torsions cannot be fully utilized. Methods that express the path in RIC may be investigated in a future paper. Three approaches to incorporate redundant internal coordinates into the FVRC method are introduced in the following sections:

- Expressing the FVRC coupling constraints in terms of redundant internal coordinate differences in order to avoid having to deal with a separate rotational alignment step (Section III).
- Applying the methods developed for the CVRC method to optimize a least RIC length pathway by minimizing the arc length expressed in RIC. The resulting path should be a better initial guess for the VRC method than a linear Cartesian pathway (Section IV).
- Defining the VRE and its derivatives in terms of an interpolated RIC PES. (Section V).

II. FVRC METHOD SUMMARY

The VRC method represents the path as a linear expansion of n_{basis} basis functions:

$$x_i(t) = \sum_{\mu}^{n_{basis}} C_{i\mu}\phi_\mu(t) \rightarrow \frac{\partial x_i(t)}{\partial C_{i\mu}} = \phi_\mu(t),$$
$$\tau_i(t) = \sum_{\mu}^{n_{basis}} C_{i\mu}\frac{d\phi_\mu(t)}{dt} \rightarrow \frac{\partial \tau_i(t)}{\partial C_{i\mu}} = \frac{d\phi_\mu(t)}{dt}, \quad (2)$$

where the Roman indices are over the $n_{cart} = 3 \times n_{atoms}$ Cartesian coordinates, the Greek indices are over the basis

functions, and the ϕ are quartic B-Spline functions. For a given choice of basis, the linear expansion coefficients (LEC, $C_{i\mu}$) define the path, and derivatives of the VRE with respect to a change in the LEC may be defined as

$$\gamma_{i\mu} = \int_{t_R}^{t_P} \left(\frac{|\boldsymbol{\tau}|}{|\mathbf{g}|} \sum_a H_{ia}g_a\phi_\mu + \frac{|\mathbf{g}|}{|\boldsymbol{\tau}|}\tau_i\frac{d\phi_\mu}{dt} \right) dt, \quad (3)$$

$$\eta_{i\mu j\nu} = \int_{t_R}^{t_P} \left(\frac{|\boldsymbol{\tau}|}{|\mathbf{g}|} \left((Tg+HH)_{ij} - \frac{\sum_{a,b} H_{ia}g_a H_{jb}g_b}{|\mathbf{g}|^2} \right)\phi_\mu\phi_\nu \right.$$
$$+ \frac{\sum_a H_{ia}g_a\tau_i}{|\mathbf{g}||\boldsymbol{\tau}|}\phi_\mu\frac{d\phi_\nu}{dt} + \frac{\tau_i\sum_a H_{ja}g_a}{|\mathbf{g}||\boldsymbol{\tau}|}\frac{d\phi_\mu}{dt}\phi_\nu$$
$$\left. + |\mathbf{g}|\left(\frac{\delta_{ij}}{|\boldsymbol{\tau}|} - \frac{\tau_i\tau_j}{|\boldsymbol{\tau}|^3}\right)\frac{d\phi_\mu}{dt}\frac{d\phi_\nu}{dt} \right) dt, \quad (4)$$

$$Q_{VRE}(\mathbf{C}+\Delta\mathbf{C}) = E(\mathbf{C}) + \Delta\mathbf{C}^T\boldsymbol{\gamma}(\mathbf{C}) + \frac{1}{2}\Delta\mathbf{C}^T\boldsymbol{\eta}(\mathbf{C})\Delta\mathbf{C}, \quad (5)$$

where $\mathbf{H}$ is the potential energy Hessian and $\mathbf{Tg}$ is the product of the third derivative of the potential energy with the potential energy gradient. The $\mathbf{Tg}$ term was found to be unnecessary for good performance when using the focused VRC method,[22] so it is not included in the evaluation of $\boldsymbol{\eta}$ in this work. A positive definite shift matrix $\boldsymbol{\sigma}_{i\mu j\nu}$ is used to shift $\boldsymbol{\eta}$ so that it is positive definite while accounting for some of the curvilinear relationship between the LEC and the gradient

$$\sigma_{i\mu j\nu} = \int_{t_R}^{t_P} \delta_{ij}\frac{d\phi_\mu}{dt}\frac{d\phi_\nu}{dt}dt. \quad (6)$$

A set of n constraints κ_α is used to establish a relationship between the parameter t and the arc length of the path

$$\kappa_\alpha = S(t_{\alpha-1}, t_\alpha) - S(t_\alpha, t_{\alpha+1}), \qquad 1 \le \alpha \le n, \quad (7)$$

where S is the arc length between two points along the path. The t_α are evenly spaced to give $n_{bas} + 1$ segments along the path, where n_{bas} is the number of basis functions used to expand each coordinate

$$S(t_1, t_2) = \int_{t_1}^{t_2} |\boldsymbol{\tau}(t)| \, dt, \quad (8)$$

$$t_\alpha = \frac{\alpha}{n+1}, \qquad 0 \le \alpha \le n+1. \quad (9)$$

In addition to these arc length constraints, coupling constraints are used to focus the path optimization into regions of the potential that are likely to contain transition states and intermediates. These constraints take the following form:

$$\mathcal{F}(\mathbf{C}, \boldsymbol{\theta}, t_{ext}) = \left(\boldsymbol{\theta} + \frac{1}{2}\Delta\mathbf{x}\right)^T\Delta\mathbf{x}, \quad (10)$$

$$\boldsymbol{\Delta}_i x(\mathbf{C}, t_{ext}) = \sum C_{i\mu}\phi_\mu(t_{ext}) - x_{ext,i}, \quad (11)$$

where $\boldsymbol{\theta}$ are Lagrange multipliers for the constraints, t_{ext} is an optimizable parameter corresponding to the location along the path of a local maximum (transition state) or minimum (intermediate) in the energy with respect to t, collectively referred to as the extrema along the path. At the beginning of each VRC iteration, a geometry optimization step is computed at the $\mathbf{x}(t_{ext})$, resulting in $\mathbf{x}_{ext}$. Additionally, step size control for the FVRC method is achieved in a familiar way by limiting

the size of the geometry optimization step rather than imposing a maximum step size on the change to the LEC.

During every iteration of the Focused VRC method, E, $\boldsymbol{\gamma}$, and $\boldsymbol{\eta}$ are computed via numerical integration, the scaling parameter ξ_σ is determined by the Rational Function Optimization (RFO) method,[23] the extrema along the current path are found and the geometry optimization steps are computed, and the following Lagrangian is minimized:

$$\mathcal{L}_{FVRC} = Q_{VRE}(\mathbf{C} + \Delta\mathbf{C}) - \frac{1}{2}\xi_\sigma \Delta\mathbf{C}^T\boldsymbol{\sigma}\Delta\mathbf{C}$$
$$+ \sum_\alpha \lambda_\alpha \kappa_\alpha(\mathbf{C} + \Delta\mathbf{C}) + \mathcal{F}(\mathbf{C}, \boldsymbol{\theta}, t_{ext}), \quad (12)$$

with respect to the change in the LEC ($\Delta\mathbf{C}$), the Lagrange multipliers for the constraints on the relative arc lengths (λ_α), the location of the extrema along the path (t_{ext}), and the Lagrange multipliers for the coupling constraints ($\boldsymbol{\theta}$). This minimization is performed microiteratively, with the $\boldsymbol{\kappa}$ and $\mathcal{F}$ recomputed using updated values of $\mathbf{C} + \Delta\mathbf{C}$, $\boldsymbol{\theta}$ and t_{ext}. Following the convergence of the microiterations, the resulting path locally minimizes the approximation Q_{VRE}, while ensuring that $\kappa_\alpha = 0$ for all α and that the path passes through the points $\mathbf{x}_{ext}$ at the corresponding values of t_{ext}.

III. RIC COUPLING CONSTRAINTS FOR THE FVRC METHOD

The coupling constants as described in Sec. II (Eq. (10)) have the unfortunate drawback that the goal structures $\mathbf{x}_{ext}$ are defined in terms of the n_{cart} absolute Cartesian coordinates of the structure, while the geometry of non-linear molecular systems is defined by only $n_{cart} - 6$ internal coordinates. The 6 extra degrees of freedom involve the external translation and rotation of the entire molecule. The translation degrees of freedom do not come into play if there is no difference in the overall translation between the reactant and product structures since interpolation of the initial path, minimization of the VRE and the computation of an optimization step will not introduce any overall translation anywhere along the path. Overall rotation can only be defined infinitesimally, however, due to its curvilinear relationship with the geometry. Including any rotational information in the coupling constraints is problematic, so to improve the stability of the optimization, $\mathbf{x}_{ext}$ has to be rotated during each microiteration in order to minimize the overall rotation between $\mathbf{x}_{ext}$ and the current value of $\mathbf{x}(t_{ext})$.

Instead of defining the coupling constraints in terms of Cartesian displacements $\mathbf{x}$, it would be advantageous to use internal coordinates $\mathbf{q}$, but if the internal coordinate definition has any redundancies the $\mathbf{q}$, $\boldsymbol{\theta}$ second derivative of the coupling constraints will be overdetermined, resulting in a singular augmented Hessian. The solution is to define the derivatives of the coupling constraints only in the locally non-redundant space defined as the first $n_{act} = n_{cart} - 6$ left-singular vectors ($\mathbf{U}_{act}$) of the Wilson B-Matrix $B_{ai} = \frac{\partial q_a}{\partial x_i}$, computed by singular value decomposition ($\mathbf{B} = \mathbf{USV}^T$, $\mathbf{U}$ and $\mathbf{V}$ are unitary, and $\mathbf{S}$ is diagonal). The RIC coupling constraints take the form

$$\mathcal{F}_q = \left(\frac{1}{2}\Delta\mathbf{q} + \boldsymbol{\theta}_q\right)^T \Delta\mathbf{q}, \quad (13)$$

$$\Delta_a q_{ext} = \sum_b [\mathbf{U}_{act}]_{ab}(q_b(t_{ext}) - q_{ext,b}), \quad 1 \le a \le n_{act}, \quad (14)$$

where 2π is added or subtracted to the differences corresponding to dihedral coordinates as necessary to account for the discontinuity at $+/-\pi$. To compute the derivatives of Eq. (13), the reduced Lagrange multipliers, B-matrix, B-matrix derivative, and RIC displacements are defined as follows:

$$\Delta\mathbf{q}_r = \mathbf{U}_{act}^T\Delta\mathbf{q}_{ext}, \qquad \boldsymbol{\theta}_r = \mathbf{U}_{act}^T\boldsymbol{\theta}_q, \quad (15)$$

$$\mathbf{B}_r = \mathbf{U}_{act}^T\mathbf{B}, \qquad \frac{\partial\mathbf{B}_r}{\partial\mathbf{x}} = \mathbf{U}_{act}^T\frac{\partial\mathbf{B}}{\partial\mathbf{x}}, \quad (16)$$

where the r subscript corresponds to the reduced quantity. This makes the incorrect assumption that $\mathbf{U}_{act}$ is constant with respect to a change in the Cartesian coordinates, but since the x-derivative of $\mathbf{U}_{act}$ term will vanish in $\mathbf{B}_r$ when $\Delta\mathbf{q}$ is zero, inclusion of this term should not be necessary for convergence during the microiterations. The $\boldsymbol{\theta}_r$, t_{ext}, and LEC derivatives of $\mathcal{F}_q$ are

$$\frac{\partial\mathcal{F}_q}{\partial\boldsymbol{\theta}_a} = \Delta_a q_{ext}, \quad (17)$$

$$\frac{d\mathcal{F}_q}{dt_{ext}} = (\Delta\mathbf{q}_r + \boldsymbol{\theta}_r)^T\mathbf{B}_r\boldsymbol{\tau}_{ext}, \quad (18)$$

$$\frac{\partial\mathcal{F}_q}{\partial C_{i\mu}} = \sum_a (\Delta\mathbf{q}_r + \boldsymbol{\theta}_r)_a[\mathbf{B}_r]_{ai}\phi_\mu(t_{ext}), \quad (19)$$

$$\frac{d\partial\mathcal{F}_q}{dt_{ext}\partial\boldsymbol{\theta}_a} = \sum_i [\mathbf{B}_r]_{ai}[\boldsymbol{\tau}_{ext}]_i, \quad (20)$$

$$[\mathbf{M}_{xx}]_{ij} = \frac{\partial^2\mathcal{F}_q}{\partial x_i\partial x_j} = \sum_b [\mathbf{B}_r]_{bi}[\mathbf{B}_r]_{bj} + (\Delta\mathbf{q}_r + \boldsymbol{\theta}_r)_b\frac{\partial[\mathbf{B}_r]_{bi}}{\partial x_j}, \quad (21)$$

$$\frac{d^2\mathcal{F}_q}{dt_{ext}^2} = \boldsymbol{\tau}_{ext}^T\mathbf{M}_{xx}\boldsymbol{\tau}_{ext} + (\Delta\mathbf{q}_r + \boldsymbol{\theta}_r)^T\mathbf{B}_r\frac{d\boldsymbol{\tau}_{ext}}{dt_{ext}}, \quad (22)$$

$$\frac{d\partial\mathcal{F}_q}{dt_{ext}\partial C_{i\mu}} = [\mathbf{M}_{xx}\boldsymbol{\tau}_{ext}]_i\phi_\mu(t_{ext}) + \left[(\Delta\mathbf{q}_r + \boldsymbol{\theta}_r)^T\mathbf{B}_r\right]_i\frac{d\phi_\mu(t_{ext})}{dt}, \quad (23)$$

$$\frac{\partial^2\mathcal{F}_q}{\partial C_{i\mu}\partial\boldsymbol{\theta}_a} = [\mathbf{B}_{ai}]\phi_\mu(t_{ext}), \quad (24)$$

$$\frac{\partial^2\mathcal{F}}{\partial C_{i\mu}\partial C_{j\nu}} = [\mathbf{M}_{xx}]\phi_\mu(t_{ext})\phi_\nu(t_{ext}). \quad (25)$$

With the derivatives defined in this fashion, $\mathcal{F}_q$ can replace $\mathcal{F}$ in the FVRC Lagrangian.

IV. LEAST LENGTH PATH

A. RIC arc length

The arc length formula (Eq. (8)) generalizes the concept of length to things that are curved in nature, like the length

traveled by a reaction path in Cartesian coordinates. It can also be used to compare lengths of curves in curvilinear spaces, such as the space of redundant internal coordinates. These spaces are different from rectilinear spaces like Cartesian coordinates in that they have a variable metric, which defines the infinitesimal relationship between a change in the coordinates, and the change in the length traveled. A simple conceptual example of the role a metric plays in computing distances is to define two points on a sphere. In this case, a curvilinear coordinate system used to represent those points could be a pair of angles (i.e., a spherical coordinate system with a fixed radius), and the metric would define how to compute the arc of a circle on the cross section of the sphere containing two points.

For molecular geometries, the RIC is a curvilinear coordinate set that is a more naturally uncoupled representation of the relative motions of atoms, and Wilson's B-Matrix defines the metric. To determine the RIC length of a path that is defined in Cartesian coordinates, the following arc length formula can be used instead:

$$S_{RIC}(t_1, t_2) = \int_{t_1}^{t_2} \sqrt{\tau^T \mathbf{G}_C \tau} \, dt, \qquad (26)$$

where $\mathbf{G}_C = \mathbf{B}^T \mathbf{B}$, $\mathbf{B}$ depends on $\mathbf{x}(t)$, and τ depends on t. $\mathbf{G}_C$ can be thought of as the $n_{crt} \times n_{crt}$ inner product version of Wilson's G-matrix $\mathbf{G}_Q = \mathbf{B}\mathbf{B}^T$, which is commonly used in redundant internal coordinate transformations. Any path between reactant and product that minimizes S_{RIC} will have the smallest possible overall change in the redundant internals. This should provide a more reasonable initial pathway for optimization than the linear path in Cartesians by eliminating nonphysical or unrealistic configurations, such as those with atomic collisions due to very small bonds or angles. With some slight modifications, the methods developed to minimize the VRE can also be used to minimize an arc length formula like S_{RIC}. This will produce a path defined in Cartesian coordinates that is ready to be optimized by VRC or some other path optimization method.

The first step in applying the VRC methods to minimizing S_{RIC} is to compute the derivatives with respect to a change in the LEC. For notational convenience, let $|\tau_G| = \sqrt{\tau^T \mathbf{G}_C \tau}$. The first derivative is given by

$$\frac{\partial S_{RIC}}{\partial C_{i\mu}} = \int_{t_R}^{t_P} \left(\phi_\mu \frac{\partial}{\partial x_i} + \frac{d\phi_\mu}{dt} \frac{\partial}{\partial \tau_i} \right) |\tau_G| \, dt, \qquad (27)$$

$$\frac{\partial}{\partial x_i} |\tau_G| = \frac{\sum_{k,l} \tau_k \frac{\partial [\mathbf{G}_C]_{kl}}{\partial x_i} \tau_l}{2 |\tau_G|}, \qquad (28)$$

$$\frac{\partial}{\partial \tau_i} |\tau_G| = \frac{\sum_k \tau_k ([\mathbf{G}_C]_{ik})}{|\tau_G|}, \qquad (29)$$

with the x-derivatives of $\mathbf{G}_C$ constructed straightforwardly as

$$\frac{\partial G_{kl}}{\partial x_i} = \sum_a B_{ak} \frac{\partial B_{al}}{\partial x_i} + \frac{\partial B_{ak}}{\partial x_i} B_{al}. \qquad (30)$$

The second derivatives are derived in a similar fashion

$$\frac{\partial^2 S_{RIC}}{\partial C_{i\mu} \partial C_{j\nu}} = \int_{t_R}^{t_P} \left(\phi_\mu \phi_\nu \frac{\partial}{\partial x_i} \frac{\partial}{\partial x_j} + \frac{d\phi_\mu}{dt} \phi_\nu \frac{\partial}{\partial \tau_i} \frac{\partial}{\partial x_j} + \phi_\mu \frac{d\phi_\nu}{dt} \frac{\partial}{\partial x_i} \frac{\partial}{\partial \tau_j} + \frac{d\phi_\mu}{dt} \frac{d\phi_\nu}{dt} \frac{\partial}{\partial \tau_i} \frac{\partial}{\partial \tau_j} \right) |\tau_G| \, dt, \qquad (31)$$

$$\frac{\partial}{\partial x_i} \frac{\partial}{\partial x_j} |\tau_G| = \frac{\sum_{k,l} \tau_k \frac{\partial^2 [\mathbf{G}_C]_{kl}}{\partial x_i \partial x_j} \tau_l}{2 |\tau_G|} - \frac{\partial |\tau_G|}{\partial x_i} \frac{\partial |\tau_G|}{\partial x_j} \frac{1}{|\tau_G|}, \qquad (32)$$

$$\frac{\partial}{\partial x_i} \frac{\partial}{\partial \tau_j} |\tau_G| = \frac{\sum_k \tau_k \frac{\partial [\mathbf{G}_C]_{kj}}{\partial x_i}}{2 |\tau_G|} - \frac{\partial |\tau_G|}{\partial x_i} \frac{\partial |\tau_G|}{\partial \tau_j} \frac{1}{|\tau_G|}, \qquad (33)$$

$$\frac{\partial}{\partial \tau_i} \frac{\partial}{\partial \tau_j} |\tau_G| = \frac{[\mathbf{G}_C]_{ij}}{|\tau_G|} - \frac{\partial |\tau_G|}{\partial \tau_i} \frac{\partial |\tau_G|}{\partial \tau_j} \frac{1}{|\tau_G|}, \qquad (34)$$

$$\frac{\partial^2 G_{kl}}{\partial x_i \partial x_j} = \sum_a 2 \frac{\partial B_{ak}}{\partial x_i} \frac{\partial B_{al}}{\partial x_j} + B_{ak} \frac{\partial^2 B_{al}}{\partial x_i \partial x_j} + \frac{\partial^2 B_{ak}}{\partial x_i \partial x_j} B_{al}. \qquad (35)$$

Since $\mathbf{G}_C$ is defined as a $n_{cart} \times n_{cart}$ matrix, no additional consideration of the possible redundancy in the RIC coordinate system is necessary beyond the requirement that the B-Matrix has n_{act} linearly independent rows. However, even when the RIC is well chosen the matrix $\mathbf{G}_C$ will be singular due to the external degrees of freedom, and therefore Eq. (31) will also be singular. In order to ensure the numerical stability of the VRC-like optimization, these external degrees of freedom need to be addressed in some fashion.

B. Handling translation and rotation

The simplest way to account for the external degrees of freedom in a VRC-like optimization to find the least length RIC path is to add to $\mathbf{G}_C$ a pair of orthonormal projectors that span the translation and rotation spaces

$$\mathbf{G}_C^* = \mathbf{G}_C + \mathbf{P}_T + \mathbf{P}_R. \qquad (36)$$

Vectors for the infinitesimal translation ($\mathbf{t}$) and rotation ($\mathbf{r}$) are constructed with the portion of the vectors corresponding to the kth atom given by

$$\mathbf{t}_{l,k} = \mathbf{e}_l, \qquad (37)$$

$$\mathbf{r}_{l,k} = \mathbf{x}_k \times \mathbf{e}_l, \qquad (38)$$

where $\times$ denotes the 3-dimensional cross product, $\mathbf{e}_l$ is the lth row of the 3-dimensional identity matrix, and $\mathbf{x}_k$ are the 3-dimensional Cartesian coordinates for atom k translated so that the origin is located at the center of molecule (i.e., $\sum_k x_{k,l} = 0$ for each l). These vectors can be arranged as the $3 \times n_{cart}$ matrices $\mathbf{T}$ and $\mathbf{R}_x$, which are used to construct the orthonormal projectors

$$\mathbf{P}_T = \frac{1}{n_{atoms}} \sum_l^3 \mathbf{t}_l \mathbf{t}_l^T = \frac{\mathbf{TT}^T}{n_{atoms}}, \tag{39}$$

$$\mathbf{P}_R = \mathbf{R}_x (\mathbf{R}_x^T \mathbf{R}_x)^{-1} \mathbf{R}_x^T = \mathbf{R}_x \mathbf{A}^{-1} \mathbf{R}_x^T, \tag{40}$$

note that the matrix $\mathbf{A}$ is the 3×3 moment of inertia for the geometry defined by x assuming all atomic weights are equal. Since $\frac{\partial \mathbf{P}_T}{\partial \mathbf{x}} = 0$ and since $\mathbf{P}_T$ is contained within the constant null space of $\mathbf{G}_C$, and since $\mathbf{x}(t)$ should never contain any translational motion, it is sufficient to add the following term to the second derivatives of S_{RIC} to ensure that the modes corresponding to overall translation are non-singular:

$$[\mathbf{P_T}]_{i\mu j\nu} = \int_{t_R}^{t_P} [\mathbf{P}_T]_{ij} \frac{d\phi_\mu}{dt} \frac{d\phi_\nu}{dt} dt. \tag{41}$$

The derivatives of Eq. (40) require a bit more work. To begin, the following equivalency is used:

$$\boldsymbol{\tau}^T \mathbf{P}_R \boldsymbol{\tau} \equiv \mathbf{z}^T \mathbf{A}^{-1} \mathbf{z}, \tag{42}$$

where $\mathbf{z}$ is the 3-dimensional sum of cross products between the tangent and position of each atom

$$\mathbf{z} = \sum_k^{n_{atoms}} \mathbf{x}_k \times \boldsymbol{\tau}_k. \tag{43}$$

The LEC derivatives of $\boldsymbol{\tau}^T \mathbf{P}_R \boldsymbol{\tau}$ can be computed using the $\mathbf{x}$ and $\boldsymbol{\tau}$ derivatives of $\boldsymbol{\tau}^T \mathbf{P}_R \boldsymbol{\tau}$

$$\mathbf{z}^T \mathbf{A}^{-1} \frac{\partial \mathbf{z}}{\partial \mathbf{x}} = \mathbf{z}^T \mathbf{A}^{-1} \mathbf{R}_\tau, \tag{44}$$

$$\mathbf{z}^T \mathbf{A}^{-1} \frac{\partial \mathbf{z}}{\partial \boldsymbol{\tau}} = -\mathbf{z}^T \mathbf{A}^{-1} \mathbf{R}_x, \tag{45}$$

$$\mathbf{z}^T \frac{\partial \mathbf{A}^{-1}}{\partial \mathbf{x}} \mathbf{z} = \mathbf{z}^T \mathbf{A}^{-1} \mathbf{R}_{\mathbf{z}^T \mathbf{A}^{-1} \mathbf{R}_x}, \tag{46}$$

where $\mathbf{R}_\tau$ and $\mathbf{R}_{\mathbf{z}^T \mathbf{A}^{-1} \mathbf{R}_x}$ are computed in the same way as Eq. (38), but with $\mathbf{x}$ replaced by $\boldsymbol{\tau}$ and $\mathbf{z}^T \mathbf{A}^{-1} \mathbf{R}_x$, respectively. Numerical tests indicated that the contributions from Eqs. (44) and (45) were typically a few orders of magnitude larger than those from Eq. (46) and that the second derivatives of $\mathbf{A}^{-1}$ were even smaller. Since these terms appear to be negligible, $\mathbf{A}^{-1}$ is assumed to be constant to simplify the construction of the second derivatives of $\boldsymbol{\tau}^T \mathbf{P}_R \boldsymbol{\tau}$. The second derivatives are computed as follows:

$$\mathbf{z}^T \mathbf{A}^{-1} \frac{\partial^2 \mathbf{z}}{\partial \boldsymbol{\tau} \partial \mathbf{x}} = -\mathbf{z}^T \mathbf{A}^{-1} \mathbf{L}, \tag{47}$$

$$\frac{\partial \mathbf{z}}{\partial \boldsymbol{\tau}}^T \mathbf{A}^{-1} \frac{\partial \mathbf{z}}{\partial \mathbf{x}} = -\mathbf{R}_x^T \mathbf{A}^{-1} \mathbf{R}_\tau, \tag{48}$$

$$\frac{\partial \mathbf{z}}{\partial \mathbf{x}}^T \mathbf{A}^{-1} \frac{\partial \mathbf{z}}{\partial \mathbf{x}} = \mathbf{R}_\tau^T \mathbf{A}^{-1} \mathbf{R}_\tau, \tag{49}$$

$$\frac{\partial \mathbf{z}}{\partial \boldsymbol{\tau}}^T \mathbf{A}^{-1} \frac{\partial \mathbf{z}}{\partial \boldsymbol{\tau}} = \mathbf{R}_x^T \mathbf{A}^{-1} \mathbf{R}_x, \tag{50}$$

where $\mathbf{L}$ is a $3 \times n_{cart} \times n_{cart}$ tensor that satisfies

$$z_i = \mathbf{x}^T \mathbf{L}_i \boldsymbol{\tau}, \tag{51}$$

with each of the n_{atoms} 3×3 blocks on the diagonal of $\mathbf{L}_i$ defined by the appropriate elements of the 3-dimensional Levi-Civita symbol.

In addition to the LEC derivatives for the S_{RIC} using $\mathbf{G}_C^*$, the following shift matrix is used in the place of the overlap of the basis set derivatives used in the standard VRC method:

$$[\boldsymbol{\sigma}_S]_{i\mu j\nu} = \int_{t_R}^{t_P} \left(\frac{d\phi_\mu}{dt} \frac{d\phi_\nu}{dt} \frac{[\mathbf{G}_C^*]_{ij}}{|\boldsymbol{\tau}_G|} \right) dt, \tag{52}$$

where $|\boldsymbol{\tau}_G| = \sqrt{\boldsymbol{\tau}^T \mathbf{G}_C \boldsymbol{\tau}}$.

C. Improved λ_σ initialization

In the Constrained VRC (CVRC) and FVRC methods described in the Paper I,[1] ξ_σ was determined once per macro-iteration prior to beginning the microiterations by applying the RFO method to the unconstrained VRE derivatives. While this approach did result in a working algorithm, ξ_σ did not display the correct convergence behavior (i.e., it did not decrease to zero at convergence) and the resulting over-correction was suspected to play a role in the slowdown observed in the method in the final few iterations. Here, an iterative method to initialize the values of both ξ_σ and the $\boldsymbol{\lambda}$ is described that results in the correct convergence behavior. This is accomplished by alternately computing ξ_σ followed by updating the $\boldsymbol{\lambda}$ until no further change in the $\boldsymbol{\lambda}$ is observed.

To begin, the $\boldsymbol{\lambda}$ are set to zero. Each iteration begins by computing the first and second derivatives of the S_{RIC} Lagrangian with ξ_σ set to 0

$$\frac{\partial \mathcal{L}_S}{\partial \mathbf{C}} = \frac{\partial S_{RIC}}{\partial \mathbf{C}} + \sum_\alpha \lambda_\alpha^\kappa \frac{\partial \kappa_\alpha}{\partial \mathbf{C}}, \tag{53}$$

$$\frac{\partial^2 \mathcal{L}_S}{\partial \mathbf{C}^2} = \frac{\partial^2 S_{RIC+}}{\partial \mathbf{C}^2} + \sum_\alpha \lambda_\alpha^\kappa \frac{\partial^2 \kappa_\alpha}{\partial \mathbf{C}^2}, \tag{54}$$

and using these derivatives along with $\boldsymbol{\sigma}_S$ to compute ξ_σ. Then, the following system of equations can be constructed:

$$\begin{bmatrix} \mathbf{X} & \mathbf{Y}^T \\ \mathbf{Y} & \mathbf{0} \end{bmatrix} \begin{pmatrix} \Delta \mathbf{C} \\ \Delta \boldsymbol{\lambda} \end{pmatrix} = \begin{pmatrix} \mathbf{w} \\ \mathbf{v} \end{pmatrix}, \tag{55}$$

$$\mathbf{X} = \frac{\partial^2 \mathcal{L}_S}{\partial \mathbf{C}^2} - \xi_\sigma \boldsymbol{\sigma}_S, \qquad \mathbf{w} = \frac{\partial \mathcal{L}_S}{\partial \mathbf{C}}, \tag{56}$$

$$\mathbf{Y} = \frac{\partial \boldsymbol{\kappa}}{\partial \mathbf{C}}, \qquad \mathbf{v} = \boldsymbol{\kappa} \tag{57}$$

and the Schur complement of $\mathbf{X}$ can be used to solve for $\Delta \boldsymbol{\lambda}$

$$\Delta \boldsymbol{\lambda} = -(\mathbf{Y}^T \mathbf{X}^{-1} \mathbf{Y})^{-1} (\mathbf{Y}^T \mathbf{X}^{-1} \mathbf{w} - \mathbf{v}), \tag{58}$$

which is used to update $\boldsymbol{\lambda}$. This process repeats until the RMS of $\Delta \boldsymbol{\lambda}$ is less than 10^{-6}. This method also shows convergent behavior when additional constraints are added to $\mathbf{Y}$, $\Delta \boldsymbol{\lambda}$, and $\mathbf{v}$, so long as $\mathbf{v}$ is small. These conditions are always met for the arc length and rotation constraints as long as the initial path satisfies those constraints but may not be met by the coupling constraints in the FVRC method since the initial path will have a non-zero displacement in the coupling constants.

D. RIC arc length minimization algorithm

For the present algorithm, the redundant internal coordinate set was constructed by merging the bonding

skeleton from the reactant and product structures, and then including the standard set of bond stretches between bonded atoms, angle bends between pairs of stretch coordinates, and torsion coordinates between pairs of bends. The methods described in this section are general for any complete set of redundant internal coordinates, and it may be the case that different choices of coordinate set may provide some benefit for interpolation. The exploration of different coordinate definitions is a worthwhile area for future study:

1. Input initial (linear Cartesian) path.
2. Compute S_{RIC}, S_{RIC} derivatives, and $\boldsymbol{\sigma}_S$.
3. Initialize ξ_σ and the $\boldsymbol{\lambda}$ according to Sec. IV C.
4. Begin microiterations
   (a) Compute the constraints $\kappa\,(\mathbf{C} + \boldsymbol{\Delta C})$ (Eq. (7)) and their derivatives with respect to a change in the LEC.
   (b) Compute the (augmented) $\boldsymbol{\Delta C}$ and $\boldsymbol{\lambda}$ derivatives of the $\mathcal{L}_S$ as outlined in Sections II and IV.
   (c) Update $\boldsymbol{\Delta C}$ and $\boldsymbol{\lambda}$ using Newton's method to produce an augmented displacement.
   (d) Check augmented gradient and augmented displacement for convergence of microiterations, end microiterations if converged.
   (e) goto 4a.
5. Update LEC for path, and recompute S_{RIC}, S_{RIC} derivatives and $\boldsymbol{\sigma}_S$.
6. Check the predicted change in the S_{RIC} for convergence, end macroiterations if converged
7. goto 3.

The predicted change in the S_{RIC} was found to be a more reliable convergence criteria for this algorithm than a more typical optimization criteria such as the RMS or magnitude of the S_{RIC} gradient. When a change in the arc length of less than 10^{-2} Bohr or radians is achieved, the path has been cleaned up enough to avoid the atom collisions that can be present in the linear Cartesian path, while also satisfying the constraint conditions $\kappa_\alpha = 0$ and $\rho_{\alpha,i} = 0$ for all α and all i.

This approach can also be combined with the coupling constants used in the FVRC method in order to produce an initial path that minimizes S_{RIC} and also travels through one or more particular geometries, such as known intermediates or guess transition state structures. In the case of bi-molecular reactions, it was previously demonstrated[22] that interpolation of the bond order for bonds being broken or formed is an effective approach for approximating the structure at the transition state, and the methods outlined in Section III can be modified to define the goal geometry only in terms of those bonds by setting $n_{act} = n_{cts}$, where n_{cts} is the number of bonds being broken or formed, and defining the goal geometry only in terms of those bonds.

V. REDUNDANT INTERNAL COORDINATE VRC METHOD

A. RIC VRE definition

As mentioned in the introduction, the VRE is the line integral of the gradient norm (Eq. (1)). Since the gradient norm has units of energy/displacement, and the tangent norm

has units of displacement/dt, the integral of the gradient norm times the tangent norm over dt must have units of energy. The term in the RIC arc length formula (Eq. (26)) has units of RIC displacement/dt, so it makes sense that an RIC VRE can be constructed by incorporating a redundant internal coordinate gradient norm into S_{RIC}

$$E_{VRE}^q = \int_{t_R}^{t_P} \sqrt{\mathbf{g}_x^T \mathbf{G}_C^- \mathbf{g}_x}\sqrt{\boldsymbol{\tau}^T \mathbf{G}_C^* \boldsymbol{\tau}}\,dt, \tag{59}$$

where $\mathbf{G}_C^-$ indicates the pseudoinverse of the inner-product G-matrix defined in Section IV, and the modified version of $\mathbf{G}_C$ from Section II is only used in the tangent norm expression. The derivatives of $\mathbf{G}_C^-$ are defined straightforwardly as

$$\frac{\partial[\mathbf{G}_C^-]_{ab}}{\partial x_i} = \sum_j \frac{\partial[\mathbf{B}^-]_{ja}}{\partial x_i}[\mathbf{B}^-]_{jb} + [\mathbf{B}^-]_{ja}\frac{\partial[\mathbf{B}^-]_{jb}}{\partial x_i}, \tag{60}$$

where the derivative of the pseudoinverse is given by[24]

$$\frac{\partial[\mathbf{B}^-]_{ja}}{\partial x_i} = \sum_{k,b} -[\mathbf{B}^-]_{jb}\frac{\partial B_{bk}}{\partial x_i}[\mathbf{B}^-]_{ka} + [\mathbf{G}_C^-]_{jk}\frac{\partial B_{bk}}{\partial x_i}[\mathbf{P}_I^\perp]_{ab}$$
$$+ [\mathbf{P}_C^\perp]_{jk}\frac{\partial B_{bk}}{\partial x_i}[\mathbf{G}_Q^-]_{ab}. \tag{61}$$

In order to avoid having to compute the pseudoinverse of $\mathbf{B}$ and its derivative at every evaluation of the RIC VRE, an alternative formulation of Eq. (59) can be used instead

$$E_{VRE}^q = \int_{t_R}^{t_P} \sqrt{\mathbf{g}_q^T \mathbf{g}_q}\sqrt{\boldsymbol{\tau}^T \mathbf{G}_C \boldsymbol{\tau}}\,dt. \tag{62}$$

This expression is equivalent to (59), but assuming that some form of surface fitting or interpolation is used to compute $\mathbf{g}_q$ directly as a function of t, the cost of computing $\mathbf{G}_C^-$ during the integration of Eq. (62) and its derivatives may be avoided by using the chain rule, $\frac{\partial}{\partial x_i} = \sum_a B_{ai}\frac{\partial}{\partial q_a}$. The derivatives of the RIC VRE may be computed in a similar fashion to the derivatives of S_{RIC}, using Eqs. (28), (29), and (32)-(34) along with the first and second x-derivatives of the RIC gradient norm

$$\frac{\partial|\mathbf{g}_q|}{\partial x_i} = \frac{1}{|\mathbf{g}_q|}\sum_a B_{ai}[\mathbf{H}_q \mathbf{g}_q]_a, \tag{63}$$

$$\frac{\partial^2|\mathbf{g}_q|}{\partial x_i \partial x_j} = \frac{1}{|\mathbf{g}_q|}\sum_a \left(\frac{\partial B_{ai}}{\partial x_j}[\mathbf{H}_q \mathbf{g}_q]_a + B_{ai}B_{bj}\left[\frac{\partial \mathbf{H}_q}{\partial q_b}\mathbf{g}_q\right]_a\right.$$
$$\left. - \sum_b B_{ai}B_{bj}\left([\mathbf{H}_q \mathbf{H}_q]_{ab} + \frac{[\mathbf{H}_q \mathbf{g}_q]_a[\mathbf{H}_q \mathbf{g}_q]_b}{|\mathbf{g}_q|^2}\right)\right). \tag{64}$$

B. Approximating the RIC PES

In order for the RIC VRE definition in Eq. (62) to be useful, a method for approximating the RIC PES as a function of t is necessary. This can be accomplished efficiently by curve fitting, treating each unique element of the gradient and Hessian as a 1-dimensional function of t, and then approximating that function with a method for interpolating a curve based upon a limited number of known points. One such interpolation method is involves using polyharmonic splines.[25]

Polyharmonic splines are often used to approximate functions of more than one variable. They have some features that make them attractive to use here, including efficient evaluation of the fit spline, strict interpolation of data points (i.e., the fit curve is guaranteed to evaluate to the exact data at the known points), and a smooth and well behaved interpolation between points. A polyharmonic spline function $f(t)$ is evaluated as

$$f(t) = u_1 + u_2 t + \sum_i w_i \phi_r (|t - c_i|), \qquad (65)$$

where ϕ_r is a radial basis function, the c are the values of t where the data are known (also called the centers of the interpolation), and the w and u are expansion weights determined by solving the following system of equations:

$$\begin{bmatrix} \mathbf{A} & \mathbf{V}^T \\ \mathbf{V} & 0 \end{bmatrix} \begin{bmatrix} \mathbf{w} \\ \mathbf{u} \end{bmatrix} = \begin{bmatrix} \mathbf{f} \\ \mathbf{0} \end{bmatrix}, \qquad (66)$$

$$A_{ij} = \phi_r (|c_i - c_j|), \qquad (67)$$

$$V_{1i} = 1, \qquad (68)$$

$$V_{2i} = c_i, \qquad (69)$$

$$f_i = f(c_i), \qquad (70)$$

where the f are the different functions being interpolated (e.g., the energy, gradient, and Hessian). In the present work, the following radial function is used:

$$\phi_r(c) = c^3 \ln c^c. \qquad (71)$$

Generally speaking, more complex target functions require more centers to achieve the same degree of accuracy. For this reason, only the energy, gradient, and Hessian terms will be fit, rather than fitting some of the intermediate terms like $\mathbf{H}_q \mathbf{g}_q$ or $\mathbf{H}_q \mathbf{H}_q$. Additionally, when transforming the internal coordinate Hessian, additional terms that are normally neglected in geometry optimization must also be included. Normally, only the first term in Eq. (61) is used, which is sufficient to accurately transform the Hessian in the $n_{act} \times n_{act}$ space. The final two terms are necessary to accurately transform the part of the Hessian contained in the $n_{act} \times n_{red}$ space, where $n_{red} = n_{RIC} - n_{act}$ is the number of redundant coordinates. The transformed internal coordinate Hessian expression becomes

$$[\mathbf{H}_q]_{ab} = \sum_{ij} \left([\mathbf{B}^-]_{ia} [\mathbf{B}^-]_{jb} \left(H_{ij} + \sum_c [\mathbf{g}_q]_c \frac{\partial B_{ci}}{\partial x_j} \right) \right.$$
$$\left. + [\mathbf{g}_x]_i \left(\frac{\partial [\mathbf{B}^-]_{ia}}{\partial x_j} [\mathbf{B}^-]_{jb} + \frac{\partial [\mathbf{B}^-]_{ib}}{\partial x_j} [\mathbf{B}^-]_{ja} \right) \right), \quad (72)$$

where subscripts x and q are used to differentiate between the Cartesian and RIC gradients, and the plus sign in the first term is necessary to cancel out the double-counting of the $n_{act} \times n_{act}$ part in the derivative of the pseudoinverse terms.

In the algorithm outlined in Section V D, the energy, the n_{RIC} elements of the RIC gradient, and the $\frac{1}{2} n_{RIC} \times (n_{RIC} + 1)$ unique elements of the RIC Hessian are fit as functions of t using polyharmonic splines at the start of each macroiteration. The PES data used in the fit are evaluated at the reactant, product, and an additional number of geometries that are equally spaced along the current path.

The spline approximations are used to compute the VRE and VRE derivatives. Since the energy and gradient are fit independently of one another, the t_{ext} are optimized to satisfy $\widetilde{\mathbf{g}}_q^T \mathbf{B} \boldsymbol{\tau} = 0$, where $\widetilde{\mathbf{g}}$ is the approximated gradient, as appropriate. Then, prior to verifying the extrema and computing the displacements for the coupling constants, the PES is evaluated at the t_{ext} to ensure that the gradient and Hessian are both accurate enough to reliably converge to the intermediates and transition states.

C. Coupled partitioned RFO

In the initial Focused VRC method, optimization of the geometry at the extrema was carried out by a standard Newton step ($\Delta \mathbf{x} = -\mathbf{H}^{-1} \mathbf{g}$) if the Hessian at the extrema had the correct number of negative eigenvalues (one for transition state, zero for minima), and the length of the computed optimization step was less than a maximum allowed step size. Otherwise, a projected, downhill step was computed as follows:

$$\mathbf{P}_\tau^{\parallel} = \frac{\boldsymbol{\tau} \boldsymbol{\tau}^T}{\boldsymbol{\tau}^T \boldsymbol{\tau}}, \qquad \mathbf{P}_\tau^{\perp} = \mathbf{I} - \mathbf{P}_\tau^{\parallel}, \qquad (73)$$

$$\Delta \mathbf{x} = -a_{scl} \left(\mathbf{P}_\tau^{\perp} \mathbf{H} \mathbf{P}_\tau^{\perp} - \xi_{rfo} \mathbf{I} + \mathbf{P}_\tau^{\parallel} \right)^{-1} \mathbf{P}_\tau^{\perp} \mathbf{g}, \qquad (74)$$

where ξ_{rfo} is the RFO correction to find a minimum, and $a_{scl} \leq 1$ is a scale factor that reduces $\Delta \mathbf{x}$ to the maximum allowed step size. Once $\Delta \mathbf{x}$ is computed by either methods, an updated geometry $\mathbf{x}_{ext} = \mathbf{x}(t_{ext}) + \Delta \mathbf{x}$ is computed, the energy and gradient are computed at $\mathbf{x}_{ext}$ and a line search is performed to find the approximate minimum between the two points. This process was successful because $\mathbf{g}^T \boldsymbol{\tau}$ was very small in magnitude at the t_{ext} as they were determined using the data computed during the numerical integration of the VRE and its derivatives. If the t_{ext} are determined using the approximated gradient $\widetilde{\mathbf{g}}(t)$ as in Section V B, it is very likely that $\widetilde{\mathbf{g}}_q^T \mathbf{B} \boldsymbol{\tau} \neq \mathbf{g}_q^T \mathbf{B} \boldsymbol{\tau}$, and $\mathbf{g}_q^T \mathbf{B} \boldsymbol{\tau}$ will be non-negligible and Eq. (74) will need to be modified to account for this change when computing a step towards a transition state.

A common method used to compute the optimization step when seeking a transition state is the partitioned RFO (pRFO) method. In this method, the eigenspace of the Hessian is divided into the transition vector (TV), and the minimization space, and separate RFO corrections are computed for each

$$\Delta \mathbf{x} = \Delta \mathbf{x}_{TV} + \Delta \mathbf{x}_{min} = - \frac{g_{TV} \mathbf{v}_{TV}}{(\lambda_{TV} - \xi_{RFOmax})}$$
$$- \sum_i^{n-1} \frac{g_i \mathbf{v}_i}{(\lambda_i - \xi_{RFOmin})}, \qquad (75)$$

where ξ_{RFOmax} is the RFO correction that maximizes the energy along the transition vector, while ξ_{RFOmin} is the RFO correction that minimizes the energy in the remaining $n - 1$ eigenvectors orthogonal to the transition vector. Since this approach operates in the eigenspace of the Hessian, the TV and minimization spaces are orthogonal, and there is no interaction between the spaces that needs to be accounted for. In the case of the focused VRC method, it would be ideal to use the tangent to the path as an approximation to the TV, and compute a step that maximizes the energy along the current path while minimizing energy in the space orthogonal to the

current path. Since there is no guarantee that the tangent to the path will be an eigenvector of the Hessian, there will be an interaction that must be accounted for. This is done by using a coupled partitioned RFO method (cpRFO).

Similarly to the pRFO approach, a maximizing step and an RFO correction are computed after projection onto the tangent

$$\Delta \mathbf{x}_\tau = -\frac{\mathbf{g}^T \tau \tau}{(\tau^T \mathbf{H} \tau - \xi_{RFO\tau})}, \tag{76}$$

where ξ_{RFO} is the positive eigenvalue of the augmented Hessian $\mathbf{H}_{aug,\tau}$

$$\mathbf{H}_{aug,\tau} = \begin{bmatrix} \tau^T \mathbf{H} \tau & \mathbf{g}^T \tau \\ \mathbf{g}^T \tau & 0 \end{bmatrix}. \tag{77}$$

The minimization portion of the step is then computed using a shifted version of the Schur complement to the tangent

$$\mathbf{g}_{cmp} = \mathbf{P}_\tau^\perp (\mathbf{g} + \mathbf{H} \Delta \mathbf{x}_\tau), \tag{78}$$

$$\mathbf{H}_{cmp} = \mathbf{P}_\tau^\perp \left(\mathbf{H} - \frac{\mathbf{H} \tau \tau^T \mathbf{H}}{(\tau^T \mathbf{H} \tau - \xi_{RFO\tau})} \right) \mathbf{P}_\tau^\perp + \mathbf{P}_\tau^\parallel, \tag{79}$$

$$\Delta \mathbf{x}_{min} = -(\mathbf{H}_{cmp} - \xi_{RFOmin}\mathbf{I})^{-1} \mathbf{g}_{cmp}, \tag{80}$$

where ξ_{RFOmin} is the most negative eigenvalue of the augmented Hessian

$$\mathbf{H}_{aug,min} = \begin{bmatrix} \mathbf{H}_{cmp} & \mathbf{g}_{cmp}^T \\ \mathbf{g}_{cmp} & 0 \end{bmatrix} \tag{81}$$

and the final step is the combination of the maximization and minimization steps, scaled such that a maximum step size is not exceeded

$$\Delta \mathbf{x} = a_{scl} (\Delta \mathbf{x}_\tau + \Delta \mathbf{x}_{min}). \tag{82}$$

The coupling between the two spaces is accounted for by the additional terms in both the Schur complement gradient and Schur complement Hessian. The additional term in the Schur complement gradient adds the expected change in the gradient resulting from a step to the maximum along the current path, while the additional projector in the Schur complement Hessian is positive (as the denominator will always be negative), and functions as a penalty towards motion in the coupling between the TV and the minimization space. As this coupling goes to zero, the second term drops out, and the cpRFO method reproduces the standard pRFO result.

The methods described in this section are used to compute displacements in RIC, using the RIC tangent $\tau_q = \mathbf{B}\tau_x$, and all of the standard methods for handling single geometry optimizations using redundant internal coordinates including projection into the locally non-redundant space and iterative embedding of the computed displacement back into Cartesians. Additionally, a standard RFO step was found to be better when searching for a minimum when the incorrect number of eigenvalues were computed in the Hessian, and a standard Newton step scaled back to a maximum step size was used rather than the RFO method whenever the Hessian had the correct number of negative eigenvalues.

D. RIC-FVRC algorithm

Like with the standard FVRC method outlined in Paper I,[1] the RIC-FVRC method is considered converged when the gradient at the intermediates and transition states is sufficiently small.

1. Input initial path (converged RIC-CTS path).
2. Evaluate the PES to fit the polyharmonic splines.
3. Locate, optimize, and verify the t_{ext} corresponding to transition states and intermediates, and compute the $\mathbf{q}_{ext}$ goal geometries.
4. Compute VRE, VRE derivatives, VRE error ϵ_E, and σ_S.
5. Set ξ_σ and $\boldsymbol{\lambda}$ according to Sec. IV C, and set $\Delta\mathbf{C}$, $\boldsymbol{\theta}$, and μ_ϵ to $\mathbf{0}$.
6. Begin microiterations

   (a) Compute the constraints $\kappa (\mathbf{C} + \Delta\mathbf{C})$ and their derivatives with respect to a change in the LEC.
   (b) Update the $\mathbf{U}_r$ and $\Delta\mathbf{q}_r$ for all of the extrema.
   (c) Compute $\epsilon_E (\mathbf{C} + \Delta\mathbf{C})$ and turn on optimization of μ_ϵ if $\epsilon < 0$ and $|\Delta\boldsymbol{q}_r| \approx 0$.
   (d) Construct derivatives of the FVRC Lagrangian and Update $\Delta\mathbf{C}$, $\boldsymbol{\lambda}$, $\boldsymbol{\theta}$, t_e, and μ_ϵ.
   (e) Check augmented gradient and augmented displacement for convergence of microiterations, end microiterations if converged.
   (f) goto 6a.

7. Check the gradient at the $\mathbf{q}_{ext}$ for convergence, and end macroiterations if converged.
8. Update LEC for path, refit splines and compute the VRE, VRE derivatives and ϵ_E.
9. Locate, optimize, and verify the t_{ext}, and compute the $\mathbf{q}_{ext}$ goal geometries
10. goto 4.

VI. RESULTS AND DISCUSSION

The methods in this study were implemented in Mathematica,[26] using energies, gradients, and Hessians computed with the Gaussian 09 electronic structure program.[27] While a more rigorous benchmarking and comparison to existing chain of states methods is planned for a future publication, a proof of concept is provided using eight example reactions (see Figure 1 for a scheme of the reactions; geometries for the reactants and products are included in the supplementary material[32]). For each reaction, the least length RIC and RIC-CTS paths were computed, aside from the bispidine reaction which does not involve the breaking and forming of bonds and only the RIC path was computed. Figure 2 shows the geometry for the maximum energy structure along the Cartesian interpolation, the least length RIC path, the RIC-CTS path, and the final converged TS for the ene and cope reactions. The Cartesian path was usually significantly worse than either of the other interpolations, mostly due to an unrealistic shortening of bonds not involved with the reaction, and an unrealistic lengthening of the bonds
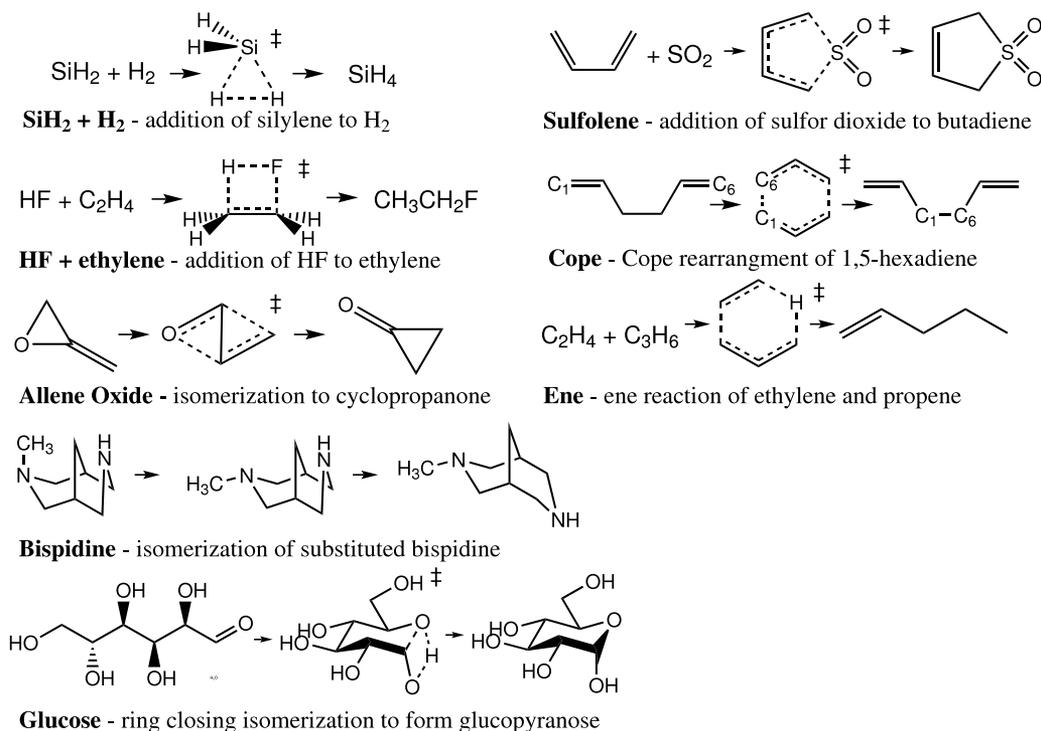
FIG. 1. Reactant, TS, and product structures for test reactions. For the bispidine reaction, the intermediate structure is shown instead of a transition state.
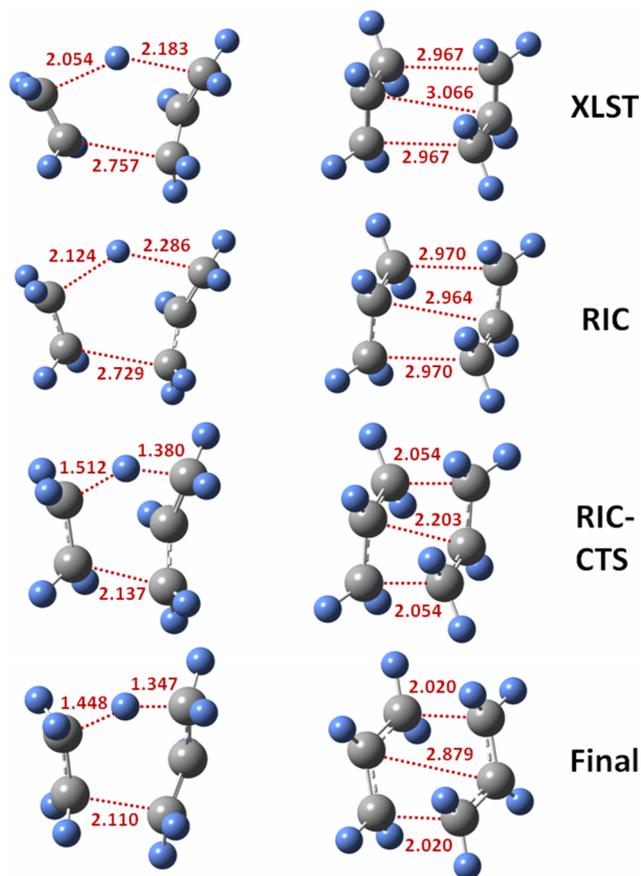

FIG. 2. Comparison of the maximum energy structures along the path for the ene (left) and cope (right) reactions. From top to bottom, Cartesian interpolation (XLST), least length RIC path, RIC-CTS path, final converged path. Distances are in Angstroms.

that are breaking or forming. The least length RIC path cleans up the former but not the latter, while the RIC-CTS path typically gives reasonable lengths for the breaking and forming bonds as well. The coordinates that differ the most between the RIC-CTS maximum and the converged transition state includes the bends and torsions involving the breaking/forming bonds, as well as any intramolecular distances which tend to be underestimated in the RIC-CTS structure.

Figure 3 compares the energy profiles for the different approximate pathways for a selection of the reactions and the final converged results from the RIC-FVRC method. In the HF + Ethylene reaction, even the Cartesian approximation is reasonably good and neither the RIC nor the RIC-CTS path offers an improvement. For the remaining reactions, however, the Cartesian interpolation results in an energy profile that is significantly worse than the other options. For the cope and ene reactions, the RIC-CTS path is also an improvement over the RIC. The maximum energy is slightly higher for the cope RIC-CTS path due to the CTS approach underestimating the lengths of the bonds breaking/forming at the TS, but both the structure and the energy profile more closely resemble those of the final path than the RIC path. The RIC path for the bispidine reaction does not involve an intermediate structure as the final path does, but it is nonetheless a clear improvement over the Cartesian path.

To demonstrate the effectiveness of the RIC-FVRC method, three optimizations were carried out for each of these reactions using 9 basis functions per coordinate, with the PES sampled at 4, 9, or 19 equally spaced points along the path. For the first three reactions, the PES was defined at
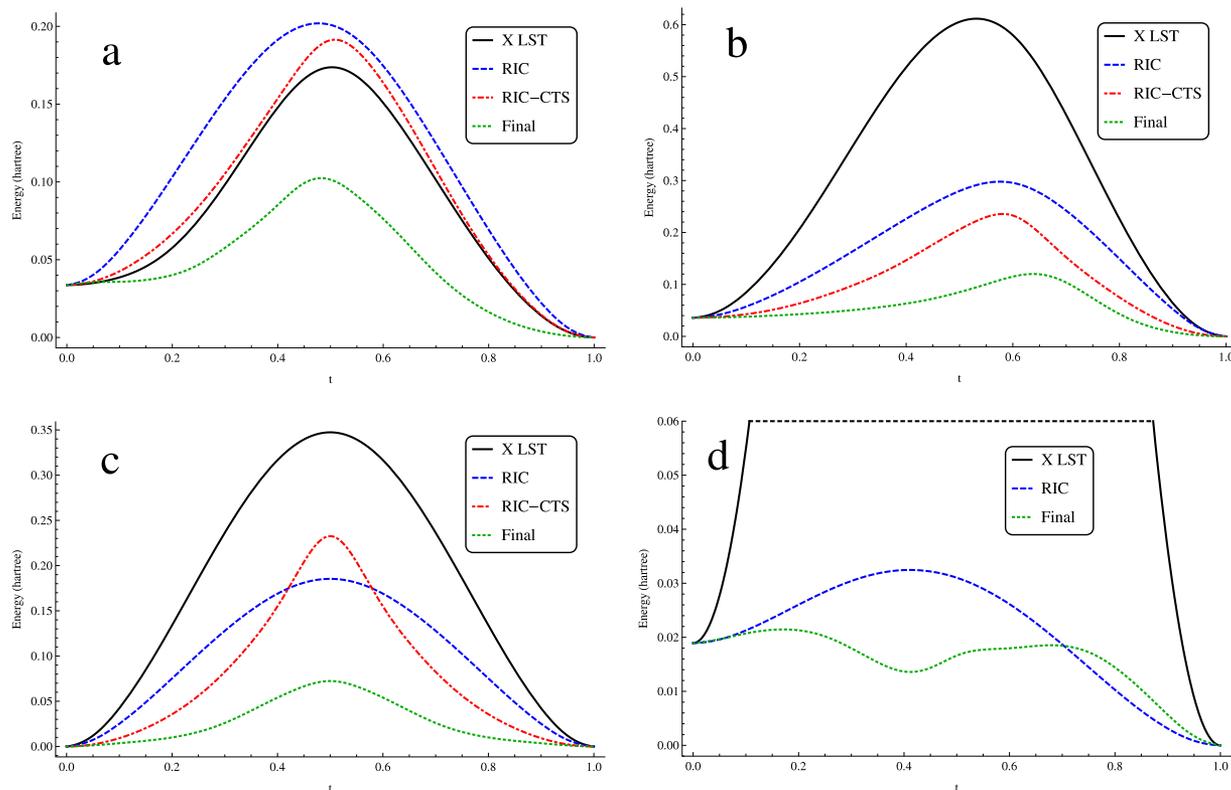
FIG. 3. Comparison of energy profiles for the Cartesian (XLST), least length RIC (RIC), RIC-CTS, and final converged path. (a) HF + Ethylene, (b) ene, (c) cope, (d) bispidine. The plot for bispidine is zoomed in to show the fine structure of the RIC and final paths, the maximum along the XLST path is approximately 0.5 hartree.

the B3LYP[28–31]/6-31G(d,p) level of theory, while the 5 larger reactions used the HF/3-21G level of theory. A convergence criterion of 10^{-4} hartree/bohr was used for the RMS gradient at the q_{ext}, with an additional requirement that no false maxima were observed along the current path. The gradient at q_{ext} rather than geometry prior to the computation of the optimization step ($q(t_{ext})$) was used as the latter is not guaranteed to converge due to the error in the interpolation of the gradient. All of the optimizations began with the RIC-CTS path aside from the bispidine optimization, which began with the RIC path. In each case, a maximum step size at intermediates and transition states of 0.5 bohr/rad was allowed. The total number of iterations required for convergence are recorded

in Table I. Also shown in the table are the total number of PES evaluations (energy, gradient, and Hessian) required for convergence, which is equal to

$$n_{tot} = 2 + n_{int} \times n_{vrc} + n_{found} + n_{used}, \qquad (83)$$

where the first 2 are the evaluation of the Hessian at the reactant and product, which may be reused throughout the optimization, n_{int} is the number of sampling points used during the interpolation, n_{vrc} is the number of VRC iterations required to converge, n_{found} are the number of TS/Intermediates located along the path, while n_{used} is the number of TS/Intermediates that were included in the VRC microiterations. Most of the optimizations were able to successfully converge even

TABLE I. Number of iterations and PES evaluations required to converge.

	VRC iterations			n_{ext}			n_{keep}			Total PES evals[a]		
# of interpolation points:	4	9	19	4	9	19	4	9	19	4	9	19
$SiH_2 + H_2$	2	2	2	2	2	2	2	2	2	14	24	44
HF + Ethylene	8	6	5	14	8	5	8	6	5	66	70	117
Allene Oxide	…	4	4	…	4	4	…	4	4	…	46	86
Sulfolene	6	6	6	6	6	6	6	6	6	38	68	128
Ethylene + Propene (Ene)	4	6	4	4	6	4	4	6	4	26	68	86
1,5-hexene (Cope)	4	4	4	4	4	4	4	4	4	26	46	86
Bispidine isomerization	…	7	7	…	17	17	…	17	17	…	99	169
Glucose isomerization	5	6	6	5	8	8	5	6	6	32	70	130

[a]Each PES evaluation consists of an energy, gradient, and Hessian calculation. The total number of PES evaluations is 2 (for reactant and product) + (# of interpolation points) × (VRC iterations) + number of extrema found (n_{ext}) + number of extrema included in the microiterations (n_{keep}).

when as few as 4 interpolation points were used, and even when the 4 or 9 interpolation point optimizations required more iterations to converge, the total number of Hessian evaluations was generally much smaller than the 19 interpolation point optimization. It is likely that the density of evaluations is less important than the accuracy of the interpolated gradient curves, as the problem encountered with the failed optimizations had to do with the incorrect placement of the transition state structures, resulting in optimization steps that included too much motion along the current path. These steps failed to systematically improve the transition state, or caused significant instability in the microiterations. In the case of the ene reaction, the 9 sampling point optimization took more iterations to converge than the 4 sampling point optimization for a similar reason. The polyharmonic splines used to fit the sample points are also well suited to fit data with non-uniform sampling, and it is expected that the development of a non-uniform sampling process that chooses sample points based upon the curvature in the path or the predicted location of the transition state could improve the quality of the fit and stability of the VRC optimization while keeping the number of PES evaluations per iteration low.

Figure 4 compares the energy profile, represented by a polyharmonic spline fit to 19 sampling points along the path, for each the final converged paths against the number of sampling points used in the interpolation during the optimization. In each case, it is clear that the same transition state was found and that for the most part, the pathways from the 9 and 19 interpolation point optimizations were very similar when the same number of VRC optimizations

were required for convergence. While the 4 interpolation point optimizations that converged did successfully locate the correct transition state, the final pathway typically contained more error, and in the case of the HF + Ethylene reaction actually smoothed over some false maxima near the reactant and product. Depending on the needs of the user carrying out the VRC optimization, this error may be considered negligible, but visual inspection of an animation of the resulting path and/or a VRC optimization step carried out using a larger number of interpolation points should be considered if the pathway is to be used for anything more than a verification that the TS connects the reactant and product.

For both the reactant and product of the cope and the reactant of the glucose isomerization reactions, the minima used for the VRC optimization had backbones with more of a compact shape which results in the reacting groups being near to one another. In each case, there was also a lower energy extended isomer that would have to undergo a significant torsional change to line up the reacting groups. VRC optimizations were also carried out using the extended minima instead. The initial cartesian interpolation resulted in very poor quality TS estimates with high energies, and the RIC optimization began with gradients that were an order of magnitude or two larger than for the other reactions, and took significantly more macroiterations to converge than for the other reactions. The resulting RIC pathways had very long arc lengths due to the considerable amount of rearrangement required to represent the backbone torsion in the Cartesian representation, and the VRC optimizations did not fully converge even when a large number of interpolation
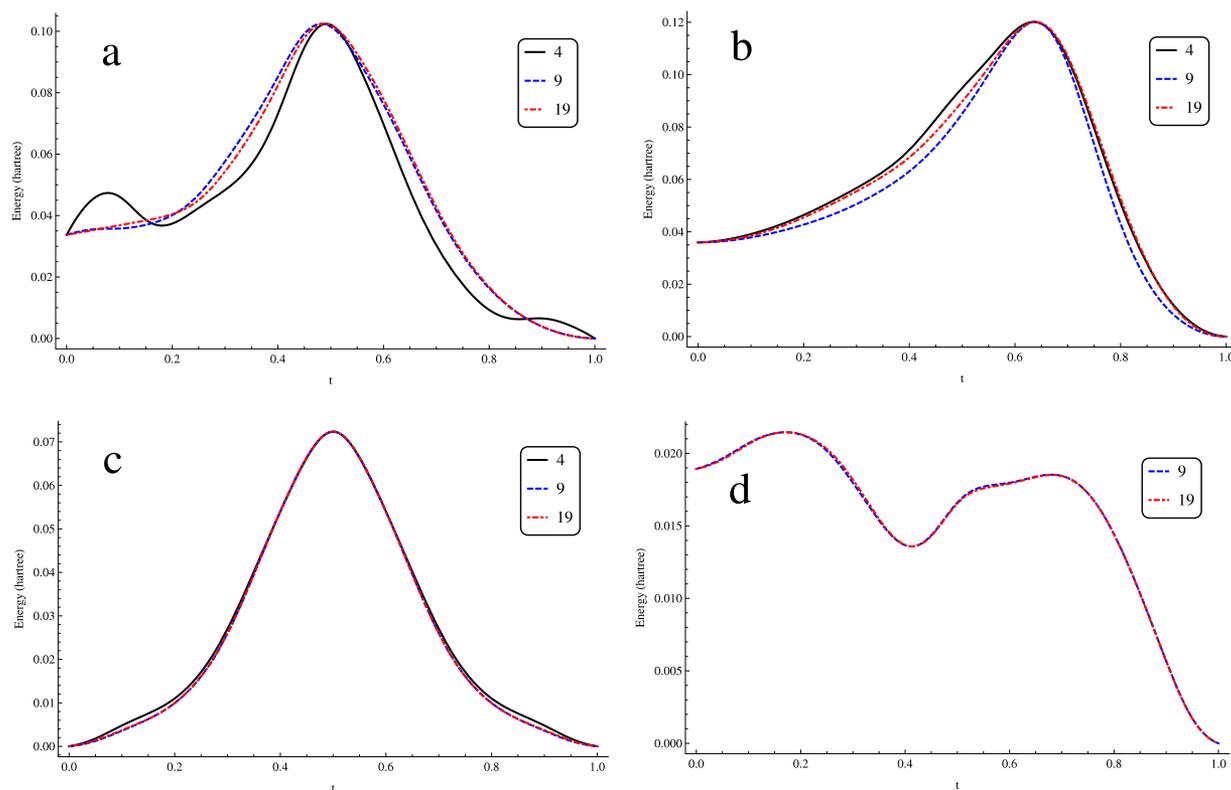

FIG. 4. Comparison of the converged pathways found using different numbers of interpolation points along the path. (a) HF+ethane, (b) ene, (c) cope, (d) bispidine.

points was used. Even though tight convergence was not possible, the VRC optimizations did quickly locate the compact intermediate. This suggests that the results of the RIC optimization may be helpful in determining whether or not the reactant and product are likely separated by a single transition state, and if they are not, a coarse VRC optimization may be used to locate an intermediate that is more suitable than the current reactant or product for optimization.

VII. SUMMARY

The RIC-FVRC method is able to produce an approximate SDRP along with the fully converged geometries of any intermediates and transition states along the path in a small number of iterations. By approximating the PES derivatives using curve fitting techniques, this algorithmic efficiency is maintained while the per-iteration cost is reduced to something more comparable with existing chain of states methods. Inclusion of Hessian updating to avoid the computation of analytical Hessians should be investigated in the future, but it is expected to be non-trivial due to the role Hessian information plays in computing the VRE gradient and identifying false minima/maxima. Additionally, the development of adaptive and/or systematic approaches to fitting the PES curves could improve the accuracy of those curves while using as few evaluations of the chemical PES as possible per iteration. Once improvements such as these are included in the present methodology, the efficiency and reliability of the VRC method should compare very favorably to existing published reaction path optimization methods.

ACKNOWLEDGMENTS

[1]A. B. Birkholz and H. B. Schlegel, "Path optimization by a variational reaction coordinate method. I. Development of formalism and algorithms," J. Chem. Phys. **143**, 244101 (2015).

[2]R. Elber and M. Karplus, "A method for determining reaction paths in large molecules: Application to myoglobin," Chem. Phys. Lett. **139**, 375–380 (1987).

[3]G. Mills and H. Jónsson, "Quantum and thermal effects in H_2 dissociative adsorption: Evaluation of free energy barriers in multidimensional quantum systems," Phys. Rev. Lett. **72**, 1124 (1994).

[4]P. Y. Ayala and H. B. Schlegel, "A combined method for determining reaction paths, minima, and transition state geometries," J. Chem. Phys. **107**, 375–384 (1997).

[5]G. Henkelman and H. Jónsson, "Improved tangent estimate in the nudged elastic band method for finding minimum energy paths and saddle points," J. Chem. Phys. **113**, 9978–9985 (2000).

[6]G. Henkelman, B. P. Uberuaga, and H. Jónsson, "A climbing image nudged elastic band method for finding saddle points and minimum energy paths," J. Chem. Phys. **113**, 9901–9904 (2000).

[7]W. E, W. Ren, and E. Vanden-Eijnden, "String method for the study of rare events," Phys. Rev. B **66**, 052301 (2002).

[8]B. Peters, A. Heyden, A. T. Bell, and A. Chakraborty, "A growing string method for determining transition states: Comparison to the nudged elastic band and string methods," J. Chem. Phys. **120**, 7877–7886 (2004).

[9]S. K. Burger and W. Yang, "Quadratic string method for determining the minimum-energy path based on multiobjective optimization," J. Chem. Phys. **124**, 054109 (2006).

[10]I. V. Khavrutskii, K. Arora, and C. L. Brooks III, "Harmonic Fourier beads method for studying rare events on rugged energy surfaces," J. Chem. Phys. **125**, 174108 (2006).

[11]A. Behn, P. M. Zimmerman, A. T. Bell, and M. Head-Gordon, "Incorporating linear synchronous transit interpolation into the growing string method: Algorithm and applications," J. Chem. Theory Comput. **7**, 4019–4025 (2011).

[12]M. U. Bohner, J. Meisner, and J. Kästner, "A quadratically-converging nudged elastic band optimizer," J. Chem. Theory Comput. **9**, 3498–3504 (2013).

[13]P. Plessow, "Reaction path optimization without NEB springs or interpolation algorithms," J. Chem. Theory Comput. **9**, 1305–1310 (2013).

[14]P. M. Zimmerman, "Single-ended transition state finding with the growing string method," J. Comput. Chem. **36**, 601–611 (2015).

[15]K. Fukui, "The path of chemical reactions—The IRC approach," Acc. Chem. Res. **14**, 363–368 (1981).

[16]W. Quapp and D. Heidrich, "Analysis of the concept of minimum energy path on the potential energy surface of chemically reacting systems," Theor. Chim. Acta **66**, 245–260 (1984).

[17]C. Gonzalez and H. B. Schlegel, "An improved algorithm for reaction path following," J. Chem. Phys. **90**, 2154–2161 (1989).

[18]H. P. Hratchian and H. B. Schlegel, "Accurate reaction paths using a hessian based predictor–corrector integrator," J. Chem. Phys. **120**, 9918–9924 (2004).

[19]W. Quapp, "Chemical reaction paths and calculus of variations," Theor. Chem. Acc. **121**, 227–237 (2008).

[20]R. Crehuet and J. M. Bofill, "The reaction path intrinsic reaction coordinate method and the Hamilton–Jacobi theory," J. Chem. Phys. **122**, 234105 (2005).

[21]P. Pulay and G. Fogarasi, "Geometry optimization in redundant internal coordinates," J. Chem. Phys. **96**, 2856–2860 (1992).

[22]A. B. Birkholz and H. B. Schlegel, "Using bonding to guide transition state optimization," J. Comput. Chem. **36**, 1157–1166 (2015).

[23]A. Banerjee, N. Adams, J. Simons, and R. Shepard, "Search for stationary points on surfaces," J. Phys. Chem. **89**, 52–57 (1985).

[24]G. H. Golub and V. Pereyra, "The differentiation of pseudo-inverses and nonlinear least squares problems whose variables separate," SIAM J. Numer. Anal. **10**, 413–432 (1973).

[25]W. R. Madych and S. Nelson, "Polyharmonic cardinal splines," J. Approximation Theory **60**, 141–156 (1990).

[26]Mathematica version 9.0, Wolfram Research, Inc. Champaign, IL, 2013.

[27]M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox, GAUSSIAN 09, Revision D.01, Gaussian, Inc., Wallingford, CT, 2009.

[28]A. D. Becke, "Density-functional thermochemistry. III. The role of exact exchange," J. Chem. Phys. **98**, 5648–5652 (1993).

[29]C. Lee, W. Yang, and R. G. Parr, "Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density," Phys. Rev. B **37**, 785 (1988).

[30]P. Stephens, F. Devlin, C. Chabalowski, and M. J. Frisch, "Ab initio calculation of vibrational absorption and circular dichroism spectra using density functional force fields," J. Phys. Chem. **98**, 11623–11627 (1994).

[31]S. Vosko, L. Wilk, and M. Nusair, "Accurate spin-dependent electron liquid correlation energies for local spin density calculations: A critical analysis," Can. J. Phys. **58**, 1200–1211 (1980).

[32]See supplementary material at http://dx.doi.org/10.1063/1.4948439 for Cartesian structures for the reactant and product structures in Angstroms.