Title

Geometry optimization made simple with translation and rotation coordinates.

Journal

The Journal of chemical physics, 144(21)

ISSN

0021-9606

Authors

Wang, Lee-Ping

Song, Chenchen

Publication Date

2016-06-01

DOI

10.1063/1.4952956

Geometry Optimization Made Simple with Translation and Rotation Coordinates

Lee-Ping Wang[1] and Chenchen Song[2,3]

[1)] *Department of Chemistry, University of California; 1 Shields Ave; Davis, CA 95616.* [a)]

[2)] *Department of Chemistry and the PULSE Institute, Stanford University; Stanford, CA 94305.*

[3)] *SLAC National Accelerator Laboratory; Menlo Park, CA 94025.*

The effective description of molecular geometry is important for theoretical studies of intermolecular interactions. Here we introduce a new translation-rotation-internal coordinate system (TRIC) which explicitly includes the collective translations and rotations of molecules, or parts of molecules such as monomers or ligands, as degrees of freedom. The translations are described as the centroid position and the orientations are represented with the exponential map parameterization of quaternions. When TRIC is incorporated into geometry optimization calculations, the performance is consistently superior to existing coordinate systems for a diverse set of systems including water clusters, organic semiconductor donor-acceptor complexes, and small proteins, all of which are characterized by nontrivial intermolecular interactions. The method also introduces a new way to scan the molecular orientations while allowing orthogonal degrees of freedom to relax. Our findings indicate that an explicit description of molecular translation and rotation is a natural way to traverse the many-dimensional potential energy surface.

[a)] Electronic mail: leeping@ucdavis.edu

I. INTRODUCTION

Intermolecular interactions are widely recognized for their central role in determining the structure, function, and properties of macromolecules, molecular complexes, and molecular materials.[1–4] In theoretical chemistry, the fundamental tools for studying intermolecular interactions include optimizing the molecular geometry to locate the critical points on the many-dimensional potential energy surface, then calculating and characterizing the interactions at a given geometry. In recent years, significant advances have been made toward the second goal, such as accurate approximations to high-level electronic structure theories[5–8] and energy decomposition analyses.[9–15] However, achieving the first goal of efficiently optimizing these molecular geometries is still challenging; this is because the potential energy surface is relatively flat along the intermolecular directions compared to intramolecular ones, and long sequences of energy and gradient evaluations are often required to reach a local minimum. Geometry optimization methods that efficiently describe intermolecular degrees of freedom may greatly accelerate these nontrivial calculations.

The Cartesian coordinate system is the simplest way to describe the molecular geometry and is needed to carry out the energy and gradient evaluations. Because the potential energy surface is highly nonlinear and coupled in the Cartesian coordinates (denoted as $\mathbf{x}$), only small steps are possible in the downhill direction. Internal coordinates (ICs, denoted as $\mathbf{q}$) are functions of the Cartesian coordinates that better reflect the collective motions of the atoms. ICs can describe displacements along curved pathways and decouple different kinds of molecular displacements, allowing the optimization algorithm to take more efficient steps. Given a displacement in the internal coordinates $\Delta\mathbf{q}$, the corresponding displacement in the Cartesian coordinates $\Delta\mathbf{x}$ is computed as:

$$\Delta\mathbf{x} = \mathbf{B}^T\mathbf{G}^{-1}\Delta\mathbf{q}, \tag{1}$$

where $\mathbf{B}$ is the Wilson B-matrix ($B_{ij} = \partial q_i/\partial x_j$) and $\mathbf{G} = \mathbf{B}\mathbf{B}^T$.[16] When $\mathbf{q}$ is a nonlinear function, Equation 1 is applied iteratively until a desired IC displacement is achieved.

An early IC system is the Z-matrix coordinates[17,18] where the position of an atom is described using a distance, angle and/or dihedral angle with respect to other atoms; we will refer to these individual functions as *primitive* ICs. The Z-matrix coordinates encounters difficulties for cyclic systems where there are more ICs than Cartesian coordinates, which makes $\mathbf{G}$ ill-behaved; Pulay and Fogarasi showed that this can be overcome by using a

generalized inverse of $\mathbf{G}$.[19] Schlegel extended this method by automatically constructing a full set of primitive ICs from the connectivity graph of the molecule, which we refer to as the redundant IC (RIC) system in this paper.[20] The generalized inverse of $\mathbf{G}^{\mathrm{RIC}}$ is expensive due to the numerous primitive ICs. Baker proposed the delocalized internal coordinate system (DLC) to remove the redundancies, where each delocalized coordinate is a linear combination of primitive ICs corresponding to an eigenvector of $\mathbf{G}^{\mathrm{RIC}}$ with a nonzero eigenvalue.[21–23] As a result, $\mathbf{G}^{\mathrm{DLC}}$ has a smaller dimension and is easier to invert.

When the above coordinate systems are applied to multiple molecules, the atoms between molecules are connected using a minimum spanning tree to describe the system as a single "super-molecule".[20] In this procedure, a minimal number of fictitious intermolecular "bonds" are added to connect the molecules at the points of closest contact, which increases the number of distances by the total number of molecules minus one; intermolecular angles and dihedrals are then added in an analogous fashion to the single-molecule case. To avoid introducing potentially ill-behaved primitive ICs between molecules, Baker proposed replacing them with inverse interatomic distances, but an artificial cutoff is necessary to avoid the number of interatomic distances increasing quadratically.[24,25] Alternatively, Billeter and Thiel introduced the hybrid delocalized internal coordinate (HDLC) method where the intermolecular displacements are described by adding all of the Cartesian coordinates into the primitive IC set for the construction of delocalized ICs.[26] In the above methods, displacements in the "intermolecular" primitive ICs will affect the inter- and intra-molecular structures simultaneously. We hypothesize that a coordinate system that separately describes the inter- and intra-molecular degrees of freedom could improve the performance of geometry optimizations.

The position of a molecule is easily described using its average Cartesian coordinates (i.e. centroid), but describing the orientation is more challenging. First of all, a suitable definition of rotation for non-rigid molecules is required. Second, an ideal representation of rotation should be differentiable, such that the Wilson-B matrix can be evaluated analytically; non-redundant, such that no additional constraints are needed in the parameter space; and free of singularities. As the rotation group SO(3) is three-dimensional, a non-redundant rotation IC should have three independent variables. These criteria are not satisfied by most existing rotation parameterizations, as will be discussed later.

In this paper, we introduce a well-behaved internal coordinate for molecular orientation,

which accounts for non-rigid molecules with the least-squares superposition method[27–29] and uses the exponential map parameterization of rotations.[30] By including translation and rotation into the primitive IC set and applying the existing delocalization methods, it forms a new coordinate system that we refer to as translation-rotation-internal coordinates (TRIC).

The rest of the paper is organized as follows: We first provide mathematical details of the new translation and rotation coordinates, together with their derivative formulas. We demonstrate the improved efficiency of TRIC in geometry optimization applied to prototypical systems compared to other coordinate systems. Finally, we show how TRIC enables constrained optimization in the molecular orientation, providing a new way to scan the potential energy surface along intermolecular degrees of freedom.

II. THEORY

To build the coordinate system, the atoms are first partitioned into subunits; this could be done automatically by connecting pairs of atoms that are closer than the sum of their covalent radii[31,32], although other divisions are possible.[33]

For each subunit, we define three translation and three rotation internal coordinates. The translation internal coordinate is defined simply as the arithmetic average of the Cartesian coordinates:

$$\overline{x}_i = \frac{1}{N}\sum_{n=1}^{N} x_{in};\ i \in \{1,2,3\}. \tag{2}$$

where the notation x_{in} denotes the i-th Cartesian coordinate of the n-th atom. The derivatives are trivial to compute as $\partial\overline{x}_i/\partial x_{jn} = \delta_{ij}/N$.

Along similar lines, we seek a differentiable function $\mathbf{v}(\mathbf{x},\mathbf{y})$ describing a rotation which brings an ordered list of Cartesian coordinates $\mathbf{x}$ into maximum overlap with a reference $\mathbf{y}$, which is set to the starting geometry. We shall assume without loss of generality that $\mathbf{x}$ and $\mathbf{y}$ have been shifted such that their centroids are located at the origin. The optimal rotation $\mathbf{v}(\mathbf{x},\mathbf{y})$ minimizes the residual squared displacement $D(\mathbf{x},\mathbf{y};\mathbf{v}')$ as:

$$\mathbf{v}(\mathbf{x},\mathbf{y}) = \underset{\mathbf{v}'}{\operatorname{argmin}}\ D(\mathbf{x},\mathbf{y};\mathbf{v}') = \underset{\mathbf{v}'}{\operatorname{argmin}}\ \frac{1}{N}|[\mathbf{U}(\mathbf{v}')\mathbf{x}] - \mathbf{y}|^2, \tag{3}$$

where $\mathbf{U}(\mathbf{v}')$ is a 3×3 rotation matrix parameterized by $\mathbf{v}'$ and $|\cdot|$ is the Euclidean norm.

The trivial representation is to use the elements of $\mathbf{U}$ themselves, but there are nine elements which makes it redundant. The Euler angle representation has the correct number

of parameters, but suffers from singularities when two of the rotation axes are parallel. A third representation is the rotation quaternion: for a rotation through angle θ around the axis defined by a unit vector (u_x,u_y,u_z), the corresponding quaternion is defined as

$$\mathbf{q} = (q_0, q_1, q_2, q_3) = \left(\cos\frac{\theta}{2}, u_x \sin\frac{\theta}{2}, u_y \sin\frac{\theta}{2}, u_z \sin\frac{\theta}{2}\right) \tag{4}$$

and $U(\mathbf{q})$ is given in Eq. 33 of Ref. 28. In this paper, all vectors (such as $\mathbf{u}$) are defined in the global frame, as we do not use molecule-specific local frames.

Following the derivation of Coutsias and coworkers,[28] Equation 3 may be written in terms of quaternions using ordinary linear algebra as

$$\mathbf{q}(\mathbf{x}, \mathbf{y}) = \underset{\mathbf{q}'}{\operatorname{argmin}}\ D(\mathbf{x}, \mathbf{y}; \mathbf{q}') = \underset{\mathbf{q}'}{\operatorname{argmin}}\ \frac{1}{N}\left(|\mathbf{x}|^2 + |\mathbf{y}|^2 - 2\mathbf{q}'^T\mathbf{F}\mathbf{q}'\right) \tag{5}$$

where $\mathbf{F}$ is the symmetric matrix given by

$$\mathbf{F} = \begin{pmatrix} R_{11}+R_{22}+R_{33} & R_{23}-R_{32} & R_{31}-R_{13} & R_{12}-R_{21} \\ R_{23}-R_{32} & R_{11}-R_{22}-R_{33} & R_{12}+R_{21} & R_{13}+R_{31} \\ R_{31}-R_{13} & R_{12}+R_{21} & -R_{11}+R_{22}-R_{33} & R_{23}+R_{32} \\ R_{12}-R_{21} & R_{13}+R_{31} & R_{23}+R_{32} & -R_{11}-R_{22}+R_{33} \end{pmatrix} \tag{6}$$

and $\mathbf{R}$ is calculated from the Cartesian coordinates as

$$R_{ij} = \sum_{n=1}^{N} x_{in} y_{in}, \quad i, j = 1, 2, 3. \tag{7}$$

In Equation 5, D is minimized when the quadratic form $\mathbf{q}'^T\mathbf{F}\mathbf{q}'$ is maximized. Thus the optimal rotation quaternion $\mathbf{q}$ is the eigenvector corresponding to the largest eigenvalue of $\mathbf{F}$.

The quaternion representation is free of singularities, but it contains four variables which makes it redundant. Another closely related parameterization that overcomes this problem is the exponential map, which is defined as

$$v_i = 2q_i\left(\frac{\cos^{-1}(q_0)}{\sqrt{1-q_0^2}}\right); \quad i = 1, 2, 3, \tag{8}$$

where $\mathbf{v}$ describes a rotation of $|\mathbf{v}|$ radians about the axis $\hat{\mathbf{v}}$; thus the magnitude and axis of rotation are encoded into a single three-dimensional vector. We note in passing that the exponential map is very closely related to the *axis-angle* representation of rotations.

A well-known fact about the rotation group $SO(3)$ is that it cannot be mapped into $\mathbb{R}^3$ without singularities.[30] In the case of the exponential map, the singularities in $\mathbb{R}^3$ are spherical surfaces with radius $2n\pi$ where n is a positive integer; this is because these vectors correspond to n complete revolutions which leave the coordinates unchanged. Moreover, any rotation with a larger norm than π is equivalent to a smaller rotation in the opposite direction. In practice, we detect when $|\mathbf{v}|$ exceeds 0.9π and reset the reference positions $\mathbf{y}$, and this effectively avoids all such singularities. Therefore, the exponential map is suitable as the rotational internal coordinate.

The partial derivatives $\partial v_i/\partial x_{nj}$ are needed to compute the Wilson B-matrix. First note that the derivative of the exponential map parameters in terms of the quaternion elements are given by:

$$\begin{aligned} \frac{\partial v_i}{\partial q_0} &= q_i \left(2\frac{q_0 \cos^{-1}(q_0)}{(1-q_0^2)^{\frac{3}{2}}} - \frac{2}{(1-q_0^2)} \right); \\ \frac{\partial v_i}{\partial q_i} &= q_0. \end{aligned} \tag{9}$$

In the neighborhood of $q_0 \to 1$ corresponding to vanishingly small rotations, Equations 8 and 9 cannot be used because the numerator and denominator both vanish. This can be addressed by taking the Taylor expansion of Equation 8 around $q_0 = 1$, and the expressions simplify to:

$$v_i = 2 - \frac{2}{3}(q_0 - 1); \tag{10}$$

$$\frac{\partial v_i}{\partial q_0} = -\frac{2}{3}q_i; \quad \frac{\partial v_i}{\partial q_i} = q_0, \tag{11}$$

thus allowing us to compute the exponential map and its derivatives in the entire region of interest including the origin. By the chain rule, the desired derivative formula is given by $\partial v_i/\partial x_{nj} = \sum_{k=1}^{4} \partial v_i/\partial q_k \times \partial q_k/\partial x_{nj}$ with the first term given by Equations 9 and 11 above.

The quaternion derivative $\partial \mathbf{q}/\partial x_{nj}$ is given by the derivative formula for the eigenvector of a matrix:[34]

$$\frac{\partial \mathbf{q}}{\partial x_{jn}} = (\lambda \mathbf{I} - \mathbf{F})^{+} \frac{\partial \mathbf{F}}{\partial x_{jn}} \mathbf{q}, \tag{12}$$

where $(\cdot)^+$ is the generalized matrix inverse, λ is the largest eigenvalue of $\mathbf{F}$ given by Equation 6 above, and $\mathbf{q}$ is the corresponding eigenvector. The derivatives $\partial \mathbf{F}/\partial x_{nj}$ are linear combinations of derivatives of $\mathbf{R}$, for example:

$$\frac{\partial F_{11}}{\partial x_{jn}} = \frac{\partial R_{11}}{\partial x_{jn}} + \frac{\partial R_{22}}{\partial x_{jn}} + \frac{\partial R_{33}}{\partial x_{jn}} \tag{13}$$

where

$$\frac{\partial R_{ij}}{\partial x_{kn}} = \delta_{ik} y_{jn}. \tag{14}$$

III. COMPUTATIONAL METHODS

We developed GeomeTRIC, an open-source optimization program that interfaces to quantum chemistry software by passing Cartesian coordinates as input and retrieving the output energies and Cartesian gradients (see Supporting Information). GeomeTRIC implements Cartesian coordinates, redundant IC (RIC), delocalized IC (DLC), hybrid DLC (HDLC), and translation-rotation-internal coordinates (TRIC). The Q-Chem 4.2 software package[35] was used for the water clusters, and the TeraChem software package[36–38] was used for the others. The details of the optimization algorithm are given in the Supporting Information; we chose a force constant of 0.05 a.u. for the translation and rotation coordinates when constructing the initial guess Hessian matrix.

IV. RESULTS AND DISCUSSION

Copper phthalocyanine (CuPc) and C_{60} are a donor/acceptor pair of materials that are extensively studied as a model system for organic photovoltaics. The molecular packing at the interface in thin films and bulk heterojunctions depends heavily on the intermolecular interactions, and in turn influence the optical and electronic properties.[43] In this example calculation, 23 starting structures were randomly selected from a molecular dynamics simulation of a CuPc/C_{60} mixture which uses the UFF force field[44].

Figure 1 shows the impact of the coordinate system on the optimization results in a representative example. The optimization using TRIC converges in 248 cycles, followed by HDLC (322), DLC (413), RIC (420) and Cartesians (771). Moreover, the primitive coordinates describing intermolecular motion affects the path taken by the optimization and the final geometry. TRIC is the only coordinate system to contain the translation and rotation coordinates, and converges to a stacked CuPc dimer with C_{60} molecules in close contact. The HDLC / Cartesian coordinates both use the atomic Cartesian coordinates, and they converge to two similar structures (RMSD = 0.40 Å) where the CuPc molecules are far apart. The RIC/DLC both use interatomic distances, angles, and dihedral angles,

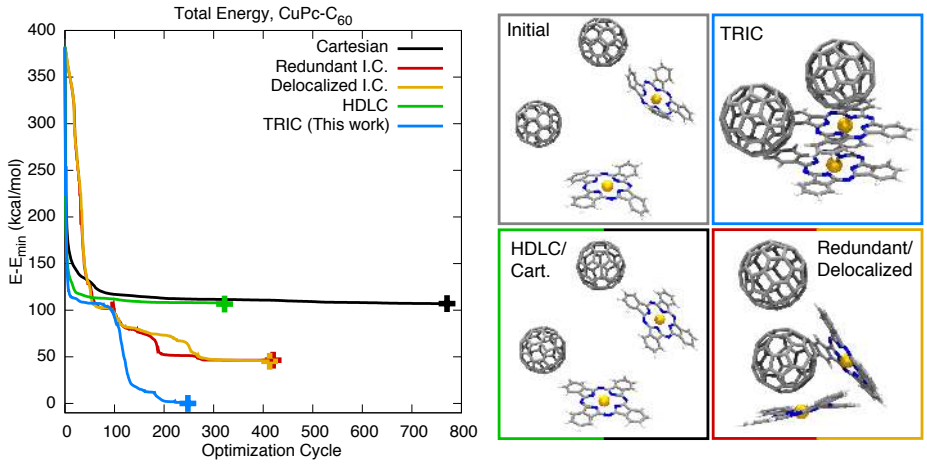

FIG. 1. Optimization results for 2CuPc-2C_{60} (neutral, triplet) using different choices of the coordinate system. Calculations used the PBE0 approximation[39] with the DFT-D3 empirical dispersion correction[40], the LANL2DZ basis set and effective core potential[41,42] for copper, and the 3-21G basis for all other atoms. Energies are referenced to the TRIC-optimized structure which has the lowest energy. The lower panels show initial and optimized structures under different coordinate systems. The HDLC/Cartesian coordinates and RIC/DLC converged to highly similar structures (RMSD $<$ 1.5 Å) so only one structure is shown for each pair. Atoms are colored as C, gray; N, blue; Cu, orange.

and they also converge to two similar structures (RMSD = 1.62 Å) where the two CuPc molecules are perpendicular.

It is well-known that the result of a local optimization largely depends on the choice of initial conditions, and even a small perturbation of the starting structure could cause convergence to a different minimum. Thus, we tested TRIC, HDLC and DLC on the full set of 23 starting structures to assess the statistical significance of the results generated using these coordinate systems. Figure S1 shows that the HDLC calculations have the largest number of cycles (428 ± 231); DLC converges in fewer iterations (250 ± 104) and TRIC requires the least (216 ± 64). All of the TRIC calculations converged in fewer than 400 cycles, and 15 out of 23 calculations reached a lower final energy than either HDLC or DLC. Student's t-test indicates that TRIC converges to the lowest energy structure more often than HDLC or DLC with a 95% confidence level.

Trp-cage Miniprotein. Proteins are diverse polymers with complex nonbonded in-

teractions; the monomers are covalently bonded through the protein backbone and possess flexible side chains, making this an interesting case study for internal coordinate systems. Here we minimized the energy of the Trp-cage miniprotein, starting from the solution NMR structure, PDB ID 1L2Y.[45] The protein was divided into subunits according to the amino acid residue number.

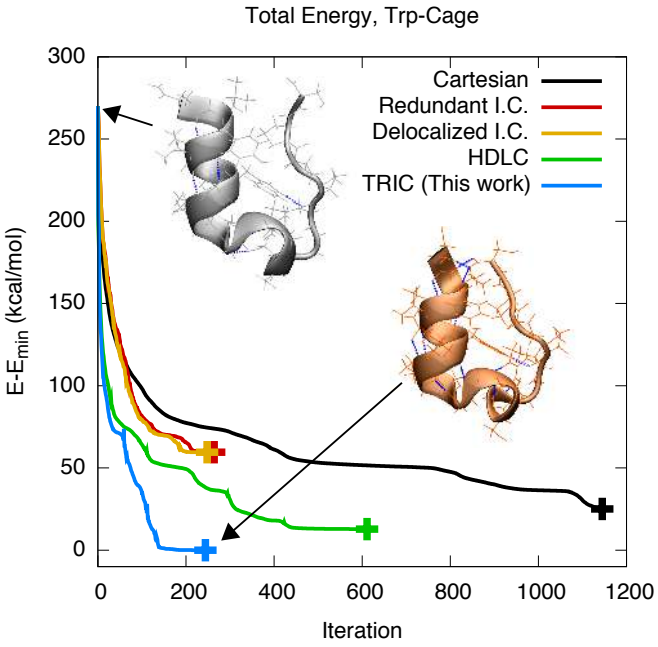

FIG. 2. Optimization results for Trp-cage, using the same electronic structure method as CuPc-C60 but with the COSMO continuum solvent model added.[46,47] The protein backbone is shown as ribbons and atoms are shown as lines; hydrogen bonds are highlighted in blue with a 3.0 Å distance cutoff between heavy atoms and a 30 degree cutoff for the angle between hydrogen, donor atom and acceptor atom.

Figure 2 shows the optimization results; TRIC converges in the fewest iterations (244), followed by DLC (248), RIC (264), HDLC (634) and Cartesian (1146). Here, both TRIC and HDLC include the bonds, angles, and dihedrals between residues. The TRIC optimized structure has a heavy-atom RMSD of 1.94 Å from the starting structure and contains more hydrogen bonds, in line with our expectation that the energy minimum should contain more hydrogen bonds than a structure obtained at room temperature. The DLC and redundant IC behave almost identically and converge to a pair of highly similar structures (RMSD <

0.01 Å). Finally, the HDLC and Cartesian coordinates converge in a much greater number of iterations (> 600) and the two final structures are rather similar (RMSD = 0.60 Å).

The amino acid residues in a protein are connected by the backbone amide bonds, thus removing these bonds from the connectivity graph will disconnect the protein into separate fragments. This in turn prevents the primitive ICs involving atoms on different residues from being generated and leads to a block-diagonal $\mathbf{G}$ matrix, which is suggested in Ref. 26 to speed up the inverse iterations for Cartesian steps. Figure S2 shows that excluding these bonded ICs increases the number of optimization cycles for TRIC by a greater amount than for HDLC, although TRIC still converges in fewer cycles in both cases. This is also observed statistically for all 38 NMR structures using the AMBER99-SB force field[48] in Figure S3. Thus, although TRIC is designed to reduce the number of optimization cycles, decoupling residues from each other is one way to accelerate the inverse iterations when they are a computational bottleneck.

Water clusters contain strong but flexible hydrogen bonds, which leads to a rich diversity of local minimum structures for water clusters as small as the hexamer.[49] We thus examined the influence of the coordinate system on the distribution of optimization results for a large number of water cluster structures drawn from liquid molecular dynamics simulations using the iAMOEBA model[50]. Five cluster sizes were considered (6, 8, 12, 16, and 20 molecules) with 100 structures each. We also present comparisons of our results with other software packages, including DL-FIND[51] (using HDLC) and Q-Chem 4.2[35] (using DLC) in Figure S4.

Figure 3 compares the performance of different coordinate systems for optimizing 12 water molecules. The number of optimization cycles is significantly different between coordinate systems, while the distribution of final energies are quite similar. The calculations with TRIC have the smallest mean and variance in the number of cycles compared to the other coordinate systems. Although the Cartesian coordinates are peaked at the largest number of iterations, DLC has the largest number of outliers requiring more than 500 cycles. These outliers could indicate the risks of using interatomic distances, angles and torsions to describe intermolecular degrees of freedom, as they undergo very different motions from intramolecular ones.

Figure 4 shows how the number of optimization cycles depends on the system size. TRIC is the most efficient for all of the system sizes tested, and also has the lowest slope of 4.1,

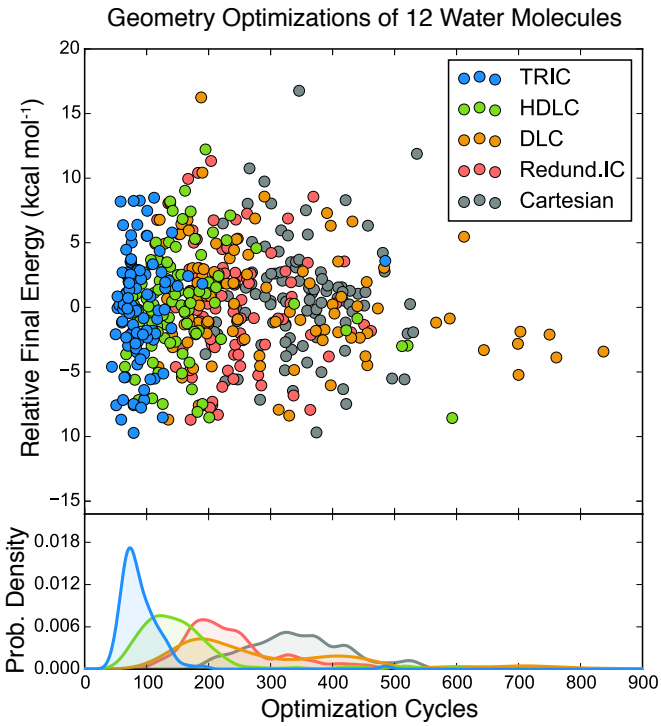

FIG. 3. Optimization results for 100 water clusters (12 molecules) with five different coordinate systems. The PBE0-D3 approximation and 6-31G* basis were used. The scatter plot shows the number of optimization cycles vs. the final minimized energies relative to the average. The lower panel is a kernel density estimate with kernel width 0.2 showing the distribution of optimization cycles for each coordinate system.

indicating the increase in the number of cycles per molecule added. At the largest cluster sizes (20 and 24), the slope decreases to 2.3 cycles per molecule added; the apparent sub-linear scaling of the cycle number is encouraging considering the scaling of the electronic structure method is often quadratic or greater. Taken together with the CuPc-C_{60} and Trp-cage results, the data indicates that TRIC significantly reduces the cost of intermolecular geometry optimization.

A. Orientation constrained optimization

The TRIC coordinates enable a new kind of constrained optimization where the relative orientations of molecules are constrained while allowing the orthogonal degrees of freedom to relax, which is useful for characterizing the intermolecular potential energy surface. To

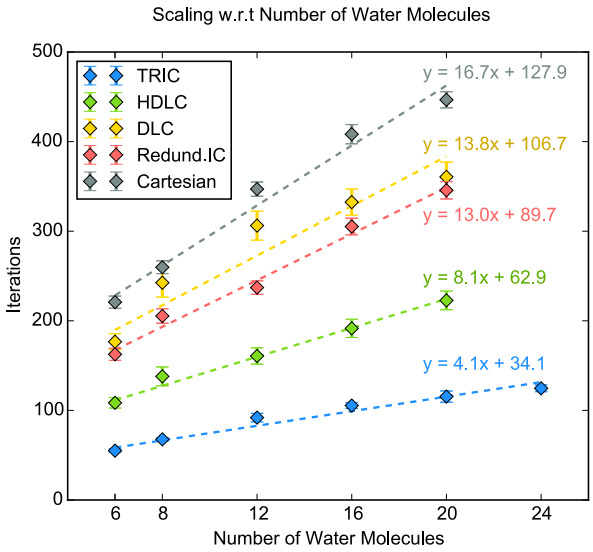

FIG. 4. Scaling behavior of the number of optimization cycles with respect to the cluster size. The error bars are one standard error. The dotted lines are linear fits.

demonstrate this capability we carried out orientation-constrained optimizations on a pentacene dimer, where one molecule is held to its original orientation while rotating the other around the chosen axis. The constraints are implemented using Lagrange multipliers as described in Ref. 52.

Although the idea of constraining the molecular orientation is intuitive, care must be taken to select the correct internal coordinates to constrain. Because the energy of the overall system is invariant to overall rotations, we fix the orientation of one pentacene molecule relative to its original structure and scan the orientation of the other molecule. The rotation is specified using a rotation axis and a rotation angle; the optimization then searches for a geometry where the molecule is rotated by the specified angle with respect to the starting structure. By scaling the unit vector of the specified rotation axis by the specified angle, one obtains the constraint values for the exponential map parameters corresponding to v_i in Equation 8.

Figure 5 shows a sequence of optimized structures of the pentacene dimer as the second molecule is is rotated through 180 degrees in 5 degree increments. We scanned the orientation first with dispersion-corrected PBE0[39,40] and then with the GAFF force field.[53], and the results are significantly different. The final structure in DFT has a much smaller horizontal offset compared to the force field, which indicates stronger intermolecular interactions in

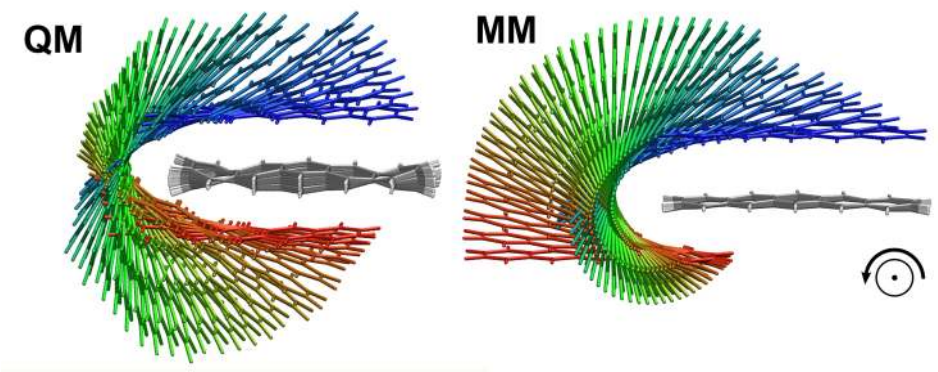

FIG. 5. Optimized structures of pentacene dimer using PBE0-D3/6-311++G(d,p) (top) and the GAFF force field (bottom). The rotation axis is perpendicular to the plane of the image. The structures are aligned using the monomer in gray/white as a reference, and the other monomer is colored according to sequence number (blue, initial; red, final). The details of this calculation are provided in the Supporting Information.

DFT. DFT also predicts much greater distortions in the monomers, which could arise from the relatively strong intermolecular interactions, or could indicate that molecules are less rigid. The stronger interactions are also evident in Figure S5, which shows that DFT predicts much higher energy barriers to rotation.

Relaxed potential energy scans are a commonly used approach in theoretical chemistry, principally as a means for testing hypothetical reaction coordinates.[54–56] With the introduction of the rotational coordinate in TRIC, it is now possible to perform a relaxed potential energy scan over the molecular orientation. We conjecture that this could open up new directions in the study of intermolecular forces, particularly concerning the interplay between molecular distortions and intermolecular forces that have been proposed for inclusion into polarizable force fields.[57]

V. CONCLUSION

As electronic structure theory is increasingly applied to study intermolecular interactions, corresponding advances in geometry optimization methods are essential. The TRIC coordinate system significantly reduces the number of optimization cycles needed for these complicated systems and also enables exploring the potential energy surface while con-

straining the intermolecular orientations. Such calculations can provide valuable data for parameterizing empirical potentials which sample the intermolecular degrees of freedom in finite-temperature simulations.[58] We are optimistic that further applications of TRIC will include basin-hopping[33,59,60] and reaction path-finding methods[61–67] as they all involve geometry optimizations on molecular potential energy surfaces.

VI. ACKNOWLEDGEMENTS

LPW acknowledges support from the UC Davis Department of Chemistry and a gift from Walt Disney Imagineering. CCS is grateful for an E. K. Potter Stanford Graduate Fellowship and support through NSF ACI-1450179. We would also like to acknowledge Chi-Yuen Wang and Yudong Qiu for helpful discussions.

VII. SUPPORTING INFORMATION

Supporting Information Available. The open-source code that implements TRIC is available online at https://github.com/leeping/geomeTRIC.

REFERENCES

[1]T. P. Knowles, A. W. Fitzpatrick, S. Meehan, H. R. Mott, M. Vendruscolo, C. M. Dobson, and M. E. Welland, Science **318**, 1900 (2007).

[2]A. Natarajan, L. S. Kaanumalle, S. Jockusch, C. L. D. Gibb, B. C. Gibb, N. J. Turro, and V. Ramamurthy, J. Am. Chem. Soc. **129**, 4132 (2007).

[3]S. Difley, L.-P. Wang, S. Yeganeh, S. R. Yost, and T. Van Voorhis, Acc. Chem. Res. **43**, 995 (2010).

[4]A. Stone, *The Theory of Intermolecular Forces* (Oxford University Press, 2013).

[5]G. Knizia, T. B. Adler, and H.-J. Werner, J. Chem. Phys. **130** (2009), 10.1063/1.3054300.

[6]M. S. Gordon, D. G. Fedorov, S. R. Pruitt, and L. V. Slipchenko, Chem. Rev. **112**, 632 (2012).

[7]D. G. Liakos and F. Neese, J. Chem. Theory Comput. **11**, 4054 (2015).

[8]C. Riplinger, P. Pinski, U. Becker, E. F. Valeev, and F. Neese, J. Chem. Phys. **144**, 24109 (2016).

[9]K. Kitaura and K. Morokuma, Int. J. Quantum Chem. **10**, 325 (1976).

[10]K. Morokuma, Acc. Chem. Res. **10**, 294 (1977).

[11]S. Rybak, B. Jeziorski, and K. Szalewicz, J. Chem. Phys. **95**, 6576 (1991).

[12]W. Chen and M. S. Gordon, J. Phys. Chem. **100**, 14316 (1996).

[13]E. G. Hohenstein and C. D. Sherrill, J. Chem. Phys. **132** (2010), 10.1063/1.3426316.

[14]R. J. Azar and M. Head-Gordon, J. Chem. Phys. **136** (2012), 10.1063/1.3674992.

[15]C. D. Sherrill, Acc. Chem. Res. **46**, 1020 (2013).

[16]E. B. Wilson, J. C. Decius, and P. C. Cross, *Molecular Vibrations: The Theory of Infrared and Raman Vibrational Spectra* (Dover, 1980).

[17]R. M. Badger, J. Chem. Phys. **2** (1934), 10.1063/1.1749433.

[18]R. M. Badger, J. Chem. Phys. **3**, 710 (1935).

[19]G. Fogarasi, X. F. Zhou, P. W. Taylor, and P. Pulay, J. Am. Chem. Soc. **114**, 8191 (1992).

[20]C. Y. Peng, P. Y. Ayala, H. B. Schlegel, and M. J. Frisch, J. Comput. Chem. **17**, 49 (1995).

[21]J. Baker, A. Kessi, and B. Delley, J. Chem. Phys. **105**, 192 (1996).

[22]J. Baker, J. Comput. Chem. **18**, 1079 (1997).

[23]J. Baker, D. Kinghorn, and P. Pulay, J. Chem. Phys. **110**, 4986 (1999).

[24]J. Baker and P. Pulay, J. Chem. Phys. **105**, 11100 (1996).

[25]J. Baker and P. Pulay, J. Comput. Chem. **21**, 69 (2000).

[26]S. R. Billeter, A. J. Turner, and W. Thiel, Phys. Chem. Chem. Phys. **2**, 2177 (2000).

[27]W. Kabsch, Acta Crystallographica Section A **32**, 922 (1976).

[28]E. A. Coutsias, C. Seok, and K. A. Dill, J. Comput. Chem. **25**, 1849 (2004).

[29]D. L. Theobald, Acta Crystallographica Section A **61**, 478 (2005).

[30]F. S. Grassia, Journal of Graphics Tools **3**, 29 (1998).

[31]L. Pauling, J. Am. Chem. Soc. **69**, 542 (1947).

[32]B. Cordero, V. Gomez, A. E. Platero-Prats, M. Reves, J. Echeverria, E. Cremades, F. Barragan, and S. Alvarez, Dalton Trans. , 2832 (2008).

[33]S. Heiles and R. L. Johnston, Int. J. Quantum Chem. **113**, 2091 (2013).

[34]K. B. Petersen and M. S. Pedersen, Technical University of Denmark **7**, 15 (2008).

[35]Y. Shao, Z. Gan, E. Epifanovsky, A. T. B. Gilbert, M. Wormit, J. Kussmann, A. W. Lange, A. Behn, J. Deng, X. Feng, D. Ghosh, M. Goldey, P. R. Horn, L. D. Jacobson, I. Kaliman, R. Z. Khaliullin, T. Kus, A. Landau, J. Liu, E. I. Proynov, Y. M. Rhee, R. M.

Richard, M. A. Rohrdanz, R. P. Steele, E. J. Sundstrom, I. Woodcock, H. Lee, P. M. Zimmerman, D. Zuev, B. Albrecht, E. Alguire, B. Austin, G. J. O. Beran, Y. A. Bernard, E. Berquist, K. Brandhorst, K. B. Bravaya, S. T. Brown, D. Casanova, C.-M. Chang, Y. Chen, S. H. Chien, K. D. Closser, D. L. Crittenden, M. Diedenhofen, J. DiStasio, Robert A., H. Do, A. D. Dutoi, R. G. Edgar, S. Fatehi, L. Fusti-Molnar, A. Ghysels, A. Golubeva-Zadorozhnaya, J. Gomes, M. W. D. Hanson-Heine, P. H. P. Harbach, A. W. Hauser, E. G. Hohenstein, Z. C. Holden, T.-C. Jagau, H. Ji, B. Kaduk, K. Khistyaev, J. Kim, J. Kim, R. A. King, P. Klunzinger, D. Kosenkov, T. Kowalczyk, C. M. Krauter, K. U. Lao, A. D. Laurent, K. V. Lawler, S. V. Levchenko, C. Y. Lin, F. Liu, E. Livshits, R. C. Lochan, A. Luenser, P. Manohar, S. F. Manzer, S.-P. Mao, N. Mardirossian, A. V. Marenich, S. A. Maurer, N. J. Mayhall, E. Neuscamman, C. M. Oana, R. Olivares-Amaya, D. P. O'Neill, J. A. Parkhill, T. M. Perrine, R. Peverati, A. Prociuk, D. R. Rehn, E. Rosta, N. J. Russ, S. M. Sharada, S. Sharma, D. W. Small, A. Sodt, *et al.*, Mol. Phys. **113**, 184 (2015).

[36]I. S. Ufimtsev and T. J. Martinez, J. Chem. Theory Comput. **4**, 222 (2008).

[37]I. S. Ufimtsev and T. J. Martinez, J. Chem. Theory Comput. **5**, 2619 (2009).

[38]I. S. Ufimtsev and T. J. Martinez, J. Chem. Theory Comput. **5**, 1004 (2009).

[39]C. Adamo and V. Barone, J. Chem. Phys. **110**, 6158 (1999).

[40]S. Grimme, S. Ehrlich, and L. Goerigk, J. Comput. Chem. **32**, 1456 (2011).

[41]P. J. Hay and W. R. Wadt, J. Chem. Phys. **82**, 299 (1985).

[42]C. Song, L.-P. Wang, T. Sachse, J. Preiss, M. Presselt, and T. J. Martinez, J. Chem. Phys. **143** (2015), 10.1063/1.4922844.

[43]B. P. Rand, D. Cheyns, K. Vasseur, N. C. Giebink, S. Mothy, Y. P. Yi, V. Coropceanu, D. Beljonne, J. Cornil, J. L. Bredas, and J. Genoe, Adv. Funct. Mater. **22**, 2987 (2012).

[44]A. K. Rappe, C. J. Casewit, K. S. Colwell, W. A. Goddard, and W. M. Skiff, J. Am. Chem. Soc. **114**, 10024 (1992).

[45]J. W. Neidigh, R. M. Fesinmeyer, and N. H. Andersen, Nat. Struct. Mol. Biol. **9**, 425 (2002).

[46]J. Tomasi, B. Mennucci, and R. Cammi, Chem. Rev. **105**, 2999 (2005).

[47]F. Liu, N. Luehr, H. J. Kulik, and T. J. Martinez, J. Chem. Theory Comput. **11**, 3131 (2015).

[48]V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg, and C. Simmerling, Proteins-Structure Function and Bioinformatics **65**, 712 (2006).

[49]D. M. Bates and G. S. Tschumper, J. Phys. Chem. A **113**, 3555 (2009).

[50]L.-P. Wang, T. Head-Gordon, J. W. Ponder, P. Ren, J. D. Chodera, P. K. Eastman, T. J. Martinez, and V. S. Pande, The Journal of Physical Chemistry B **117**, 9956 (2013).

[51]J. Kstner, J. M. Carr, T. W. Keal, W. Thiel, A. Wander, and P. Sherwood, J. Phys. Chem. A **113**, 11856 (2009).

[52]H. B. Schlegel, Wiley Interdisciplinary Reviews-computational Molecular Science **1**, 790 (2011).

[53]J. M. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, and D. A. Case, J. Comput. Chem. **25**, 1157 (2004).

[54]M. K. Beyer, Journal of Chemical Physics **112**, 7307 (2000).

[55]L. Salassa, C. Garino, G. Salassa, R. Gobetto, and C. Nervi, Journal of the American Chemical Society **130**, 9590 (2008).

[56]L.-P. Wang and T. V. Voorhis, The Journal of Physical Chemistry Letters **2**, 2200 (2011).

[57]K. Palmo, B. Mannfors, N. G. Mirkin, and S. Krimm, Chemical Physics Letters **429**, 628 (2006).

[58]L. P. Wang, T. J. Martinez, and V. S. Pande, J. Phys. Chem. Lett. **5**, 1885 (2014).

[59]D. J. Wales and J. P. K. Doye, J. Phys. Chem. A **101**, 5111 (1997).

[60]H. Kusumaatmaja, C. S. Whittleston, and D. J. Wales, J. Chem. Theory Comput. **8**, 5159 (2012).

[61]G. Henkelman, B. P. Uberuaga, and H. Jonsson, J. Chem. Phys. **113**, 9901 (2000).

[62]E. Weinan, W. Q. Ren, and E. Vanden-Eijnden, Phys. Rev. B **66** (2002), 10.1103/PhysRevB.66.052301.

[63]B. Peters, A. Heyden, A. T. Bell, and A. Chakraborty, J. Chem. Phys. **120**, 7877 (2004).

[64]W. Quapp, J. Chem. Phys. **122** (2005), 10.1063/1.1885467.

[65]D. Sheppard, R. Terrell, and G. Henkelman, J. Chem. Phys. **128** (2008), 10.1063/1.2841941.

[66]P. Zimmerman, J. Chem. Theory Comput. **9**, 3043 (2013).

[67]T. Mori and T. J. Martinez, J. Chem. Theory Comput. **9**, 1155 (2013).