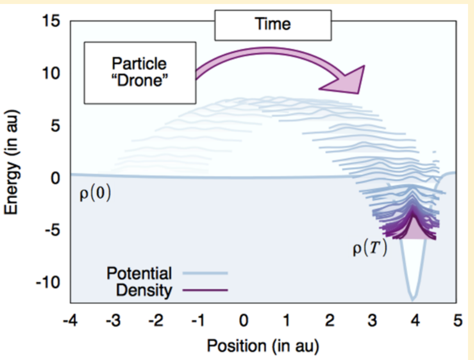
JCTC Journal of Chemical Theory and Computation
Article
Cite This: J. Chem. Theory Comput. 2018, 14, 3351−3362
pubs.acs.org/JCTC

Classical Optimal Control for Energy Minimization Based On Diffeomorphic Modulation under Observable-Response-Preserving Homotopy

Micheline B. Soley,[†] Andreas Markmann, and Victor S. Batista*

Department of Chemistry, Yale University, P.O. Box 208107, New Haven, Connecticut 06520-8107, United States
Energy Sciences Institute, Yale University, P.O. Box 27394, West Haven, Connecticut 06516-7394, United States

ABSTRACT: We introduce the so-called "Classical Optimal Control Optimization" (COCO) method for global energy minimization based on the implementation of the diffeomorphic modulation under observable-response-preserving homotopy (DMORPH) gradient algorithm. A probe particle with time-dependent mass $m(t;\beta)$ and dipole $\mu(r,t;\beta)$ is evolved classically on the potential energy surface $V(r)$ coupled to an electric field $E(t;\beta)$, as described by the time-dependent density of states represented on a grid, or otherwise as a linear combination of Gaussians generated by the k-means clustering algorithm. Control parameters β defining $m(t;\beta)$, $\mu(r,t;\beta)$, and $E(t;\beta)$ are optimized by following the gradients of the energy with respect to β, adapting them to steer the particle toward the global minimum energy configuration. We find that the resulting COCO algorithm is capable of resolving near-degenerate states separated by large energy barriers and successfully locates the global minima of golf potentials on flat and rugged surfaces, previously explored for testing quantum annealing methodologies and the quantum optimal control optimization (QuOCO) method. Preliminary results show successful energy minimization of multidimensional Lennard-Jones clusters. Beyond the analysis of energy minimization in the specific model systems investigated, we anticipate COCO should be valuable for solving minimization problems in general, including optimization of parameters in applications to machine learning and molecular structure determination.



1. INTRODUCTION

Quantum optimal control enables manipulation of dynamics and kinetics with unprecedented specificity through optimization of controls of perturbational electromagnetic fields and laser pulses.[1−6] With this power, quantum control can steer the outcomes of chemical reactions[6−25] and direct dynamics in varied environments.[26−31] These techniques can in turn be applied to advance technologies from nanostructures to quantum computation and communication.[32−41] However, simulations of quantum optimal control are restricted by the capabilities of quantum dynamics propagation methods, particularly in applications to high dimensional problems, since they face the challenge of overcoming the "curse-of-dimensionality" problem.[42] It is, therefore, of great interest to develop *classical analogues* of quantum optimal control methods. A classical analogue to quantum optimal control theory would not only enable applications of the method to more complex chemical and physical systems but would also enable application of the method to multidimensional global optimization problems.

In global optimization, the aim is to locate the greatest maximum or minimum of a function. Global optimization problems are common across fields from protein folding to machine-learning,[43,44] and improvements to global optimization methods have the opportunity to impact many disciplines. A wide variety of optimization algorithms have been developed in this pursuit. Local optimization algorithms,[45−47] including methods which rely on calculation of the position-space gradient,[48] locate maxima or minima on the function but do not guarantee the values constitute global optima. Global optimization methods, such as molecular dynamics,[49−51] simulated annealing,[52−54] potential smoothing,[55−59] and evolutionary algorithms,[60−67] rely on initial parameter choices, and therefore, their success is predicated on the parameters chosen. In addition to this dependence on initial conditions, local "traps" of rugged potential energy surfaces often hinder global optimization. Flat landscapes with localized minima, such as in the "golf problem" (i.e., a shallow parabola with a distant narrow hole), remain challenging benchmark problems for directed optimization algorithms.[57,68−75]

Here, we introduce the classical optimal control optimization (COCO) method to steer the classical dynamics of a probe particle toward the global optimum, as a classical analogue of the quantum optimal control algorithm[76] based on the *diffeomorphic modulation under observable-response-preserving homotopy* (DMORPH).[77−85] In contrast to iterative optimal control methods that rely upon initial guesses near the optimum,[2,86−88] the noniterative DMORPH algorithm strictly follows the control-space gradient allowing for initial guesses far

Received: February 5, 2018
Published: April 20, 2018

from the optimum.[77−85] Classical Liouvillian dynamics is employed, which can be parallelized and applied to high dimensional problems. The k-means clustering algorithm[89−91] is implemented to represent the time-evolved classical density as a linear combination of Gaussians for efficient calculations of the DMORPH gradients.

The COCO method is based on the classical phase-space propagation of the density of states of a probe particle with time-dependent mass $m(t;\beta)$ and dipole $\mu(r,t;\beta)$, evolving on the potential energy surface $V(r)$, coupled to an electric field $E(t;\beta)$. The control parameters β are optimized to minimize the ensemble average potential energy $\langle V(T)\rangle$ of a probe particle at the final propagation time T. The ensemble average value of the particle position $\langle r(T)\rangle$ then yields the location of the global minimum. Given a sufficient number of controls,[77−85] the method locates the global minimum without getting stuck in local minimum traps.[76]

2. METHOD

2.1. Quantum DMORPH.

The quantum DMORPH algorithm[79−85] was used as the basis for formulation of the classical DMORPH algorithm in section 2.2. In the DMORPH algorithm, the derivative of the expectation value of an observable $\hat{O}$ with respect to the controls β_j is evaluated at final time T

$$\frac{\partial\langle\hat{O}(T)\rangle}{\partial\beta_j} = \langle\psi_i|\frac{\partial}{\partial\beta_j}[U^\dagger(T,0;\beta)\hat{O}U(T,0;\beta)]|\psi_i\rangle \tag{1}$$

where $|\psi_i\rangle$ is the initial wave function, $\beta = \{\beta_j\}$ is the complete set of controls, and U is the quantum propagator. Application of the chain rule to the right-hand side of the equation and substitution of the derivative of the propagator U_{β_j} and its complex conjugate $U^\dagger_{\beta_j}$ bestows on the DMORPH algorithm its computational efficiency

$$\frac{\partial\langle\hat{O}(T)\rangle}{\partial\beta_j} = \frac{2}{\hbar}\int_0^T \mathrm{d}t\,\mathrm{Im}[\langle\psi_b|H_{\beta_j}(\beta,t')|\psi_a\rangle] \tag{2}$$

where $|\psi_a\rangle = U(t,0)|\psi_i\rangle$, $|\psi_b\rangle = U(t,0)|\psi_c\rangle$, and $|\psi_c\rangle = U^\dagger(T,0)\hat{O}U(T,0)|\psi_i\rangle$. Whereas calculation of the gradients via finite differencing requires $N + 1$ propagations for N controls, the quantum DMORPH gradient introduced by eq 2 requires only four propagations. Two propagations are needed to form $|\psi_c\rangle$, including the forward propagation of $|\psi_i\rangle$ to the final time T and the backward propagation to time zero after application of the operator $\hat{O}$. The remaining two propagations are the parallel forward propagations of $|\psi_c\rangle$ and $|\psi_i\rangle$ to the final time, with the matrix element of the Hamiltonian gradient evaluated at each intermediate time t. We note that the resulting computational efficiency also reduces memory consumption as only two intermediate wave functions, $|\psi_a\rangle$ and $|\psi_b\rangle$, are held simultaneously in memory.

2.2. Classical DMORPH.

In the spirit of quantum DMORPH, we introduce classical DMORPH as a gradient-based method to determine the value of the controls β that optimize the ensemble average of a classical observable $O(r, p)$. Just as the Schrödinger equation prescribes the time evolution of a wave function ψ in quantum DMORPH, the Liouville equation prescribes the classical time evolution of the density of states ρ in classical DMORPH

$$\frac{\partial\rho}{\partial t} = -\mathrm{i}\mathcal{L}\rho \tag{3}$$

where $i\hat{\mathcal{L}} = \frac{\partial H}{\partial p}\frac{\partial}{\partial x} - \frac{\partial H}{\partial x}\frac{\partial}{\partial p}$ is the Liouvillian operator.

In Liouvillian classical dynamics, the derivative of the ensemble average of the observable $O(r, p)$ with respect to a control β_j is as follows:

$$\frac{\partial\langle O(T)\rangle}{\partial\beta_j} = \frac{\partial}{\partial\beta_j}\int O(r,p)U(T,0)\rho(0;r,p)\mathrm{d}r\,\mathrm{d}p \tag{4}$$

where U is the classical propagator. As shown in Appendix A, in analogy to ref 76, application of the chain rule and substitution of the control-space gradient of the classical propagator U_{β_j} yields the computationally efficient classical DMORPH gradient

$$\frac{\partial\langle O(T)\rangle}{\partial\beta_j} = \int_0^T \mathrm{d}t\int \mathrm{d}r\,\mathrm{d}p O(r,p)U(T,t)(-\mathrm{i}\mathcal{L}_{\beta_j}(t))\rho(t;r,p) \tag{5}$$

where $\mathcal{L}_{\beta_j}$ is the gradient of the Liouvillian operator $\mathcal{L}$ with respect to the control β_j.

Calculation of the gradient of the observable via classical DMORPH therefore requires only three propagations, independent of the number N of controls, including backward propagation of the observable $O(r, p)$ and two simultaneous forward propagations of the quantity $O(r, p)U(T,0)$ and the initial density $\rho(0;r, p)$. The classical DMORPH gradient, like the quantum DMORPH gradient of eq 2, offers computational advantages over finite differencing and reduced memory requirements with only two quantities held in memory at intermediate times.

2.3. Classical Optimal Control Optimization.

The classical gradients of the observable $O(r, p) = V(r)$, introduced by eq 34, can be used to locate the global minimum of the potential energy surface $V(r)$. A probe particle with time-dependent mass $m(t;\beta)$ and dipole $\mu(r,t;\beta)$ is placed on the potential energy surface $V(r)$ and acted upon by an electric field $E(t;\beta)$. The dynamics is thus governed by the time-dependent Hamiltonian

$$H(r,t;\beta) = \frac{p^2}{2m(t;\beta)} + V(r) - \mu(r,t;\beta)\cdot E(t;\beta) \tag{6}$$

In multidimensional examples, the dynamics is governed by the generalization of the above Hamiltonian to higher dimensions D

$$H(\mathbf{r},t;\beta) = \sum_i^D \frac{p_i^2}{2m(t;\beta)} + V(\mathbf{r}) - \sum_i^D \mu(r_i,t;\beta)\cdot E(t;\beta) \tag{7}$$

The total propagation time T and integration time step τ are adapted to the problem, while the controls $\beta = (\beta_1, ..., \beta_N)$ are optimized to minimize the ensemble average value of the potential energy at the final time

DOI: 10.1021/acs.jctc.8b00124
J. Chem. Theory Comput. 2018, 14, 3351−3362

$$\langle V(T)\rangle = \int V(r)\rho(T;\, r,\, p;\, \beta)\mathrm{d}r\ \mathrm{d}p = \int V(r)U(T,\, 0;\, \beta)\rho(0;\, r,\, p)\mathrm{d}r\ \mathrm{d}p \tag{8}$$

Such optimization process localizes the density at the global minimum of $V(r)$ at time T. The classical DMORPH gradient eq 34 is employed to minimize $\langle V(T)\rangle$ via the limited memory Broyden–Fletcher–Goldfarb–Shanno with boundaries (L-BFGS-B)[92−94] conjugate-gradient method. The controls that produced the lowest value of the expectation value of the observable were considered as the optimal controls. The ensemble average value $\langle r(T)\rangle$, obtained with the time-evolved density $\rho(T;r,p)$, gives the coordinates of the global minimum of $V(r)$. For global minimization of multidimensional potentials, the location of the global minimum is then refined through a second application of L-BFGS-B in position-space with function tolerance $F_{\mathrm{tol}} = 0$ and projected gradient tolerance $G = 10^{-8}$.

2.3.1. Interaction Hamiltonian. The control parameters include the electric field Fourier coefficients $\beta^{j}_{\mathrm{E,s/c}}$ as defined for the quantum analogue, QuOCO[76]

$$E(t) = [\sum_{j=1}^{n} (\beta^{j}_{\mathrm{E,s}}\sin(\omega_j t) + \beta^{j}_{\mathrm{E,c}}\cos(\omega_j t))] \tag{9}$$

Controls $\beta^{jn}_{\mu,\mathrm{s/c}}$ define the dipole, according to the Fourier series

$$\mu(r,\, t) = \sum_{j,n}\left[\beta^{jn}_{\mu,\mathrm{s}}\sin\left(k_j r + l_n\frac{2\pi}{T}t\right) + \beta^{jn}_{\mu,\mathrm{c}}\cos\left(k_j r + l_n\frac{2\pi}{T}t\right)\right] \tag{10}$$

while the mass controls β^{j}_{m} define the time-dependent mass of the probe particle, according to the linear combination of Dirac delta functions

$$m(t) = \sum_{j=0}^{T/\tau} \beta^{j}_{m}\delta(t - j\tau) \tag{11}$$

for the 30 lowest frequencies on the grid and the lowest combinations of the wavenumber $k_j = (2\pi j)/(r_{\max} - r_{\min})$ and five lowest temporal change rates $l_n = n/T$

COCO optimizations were carried out using 60 electric field control parameters, 12 dipole controls, and 80 mass controls in the interaction Hamiltonian. Poor initial guesses for the controls were employed to demonstrate the success of global minimization even when starting far from the optimal values

$$\beta_{\mathrm{E,s}} = 0.003 \text{ au} \tag{12}$$

$$\beta_{\mathrm{E,c}} = 0.008 \text{ au} \tag{13}$$

$$\beta_{\mu,\mathrm{c/s}} = 1 \text{ au} \tag{14}$$

$$\beta^{i=1,2,...,T/\tau}_{\mathrm{mass}} = 1 + (m_f - 1)\left(\frac{i}{80}\right)^2 \tag{15}$$

where m_f served as an initial guess for the final mass. We note that the mass controls were bounded

$$\beta_{\mathrm{mass}} \geq \frac{4\tau}{(r_{\max} - r_{\min})^2} \tag{16}$$

to ensure the mass could be represented on a grid in benchmark grid-based calculations.

2.3.2. Liouvillian Dynamics and k-Means Density Approximation. We explore the capabilities of COCO for finding the global minima of various potential energy surfaces, after initializing the density of states away from the global minima

$$\rho(0;\, r,\, p) = \frac{1}{2\pi\sigma_r\sigma_p}\exp\left(-\frac{(r - r_k)^2}{2\sigma_r^2} - \frac{(p - p_k)^2}{2\sigma_p^2}\right) \tag{17}$$

with $\sigma_r = \frac{1}{\sqrt{2}}$ au and $\sigma_p = \frac{1}{\sqrt{2}\sigma_r}$ au, initial momentum $p_k = 0.0$ au, and initial position x_k. In multidimensional examples, the density is given by a product of Gaussians. Each component Gaussian corresponds to one dimension in position- and momentum-space as above with $\sigma_r = \frac{1}{1000\sqrt{2}}$ au and $\sigma_p = \frac{1}{1000\sqrt{2}\sigma_r}$ au.

The evolution of the density of states $\rho(T;\, r,\, p)$ was determined by propagating a swarm of trajectories and clustering them with the k-means algorithm,[89−91] to represent the propagated density of states as a sum of Gaussians. Trajectory-based simulations were compared to benchmark calculations based on propagation of the density of states amplitudes on a phase-space grid. The grid-based method also allowed for direct comparison to QuOCO.[76] Where a grid was employed, the time evolved density of states $\rho(T;\, r,\, p)$ was computed on a position grid spanning over the range $r = [-10,10]$ au ($r = [0,10]$ au for multidimensional potentials) with its corresponding momentum grid p with 2^8 equally spaced grid points in both r and p. For efficiency, we set the final simulation time as $T = 8.0$ au and integration time step $\tau = 0.1$ au.

Grid Base. In the grid-based method, the amplitude $\rho(T;\, r_j,\, p_k)$ was obtained for each grid point $(r_j,\, p_k)$ in terms of the initial density of states $\rho(0;\, r_j(0),\, p_k(0))$, after propagating the coordinates and momenta (r_j, p_k) backward in time to the initial coordinates and momenta $(r_j(0),\, p_k(0))$ by using the Velocity–Verlet algorithm.[95]

Gaussian Ansatz. In the simplest (and most approximate) trajectory based method, a single Gaussian ansatz is employed to represent the time-evolved density of states as defined by the first and second moments of the distribution of time-evolved coordinates and momenta (i.e., the density of state $\rho(T;\, r,\, p)$ was approximated as a Gaussian with first and second moments defined by the average position $\langle x(t)\rangle$, average momentum $\langle p(t)\rangle$, and corresponding second moments σ_x and σ_p).

Multi-Gaussian Ansatz. The more accurate trajectory-based method allows for the description of density of states that may have bifurcated or delocalized in phase space. In that implementation, $\rho(T;\, r,\, p)$ is represented as a sum of Gaussians with each Gaussian parametrized by a cluster of time-evolved trajectories within the swarm of trajectories. The clusters are determined by the k-means clustering algorithm, using the Euclidean phase-space distance to the center of each cluster as the classification measure.[89−91] Calculations based on a swarm of trajectories initialized on a 256 × 256 grid were compared to Monte Carlo simulations with 256 initial conditions sampled by the Box–Muller algorithm.[96]

For the implementation of the k-means clustering algorithm, the number of clusters k was determined through the entropy maximization method[97,98] in which the number of clusters k is chosen to maximize the entropy

DOI: 10.1021/acs.jctc.8b00124
J. Chem. Theory Comput. 2018, 14, 3351−3362

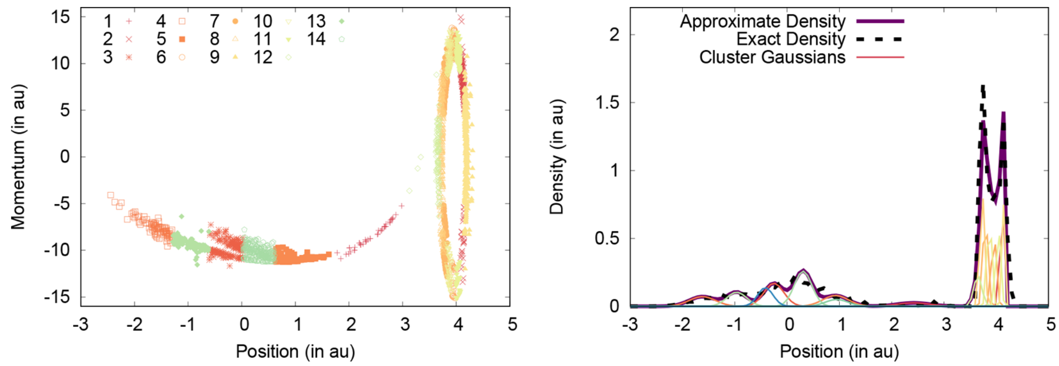


Figure 1. Example of approximate classical density determined via the *k*-means clustering algorithm. The particle swarm is (a) divided into clusters (colored points) in coordinate space, and (b) each cluster is represented by a Gaussian ansatz (thin colored lines); the sum of the Gaussians yields an approximation (thick purple line) to the exact density (dashed black line).

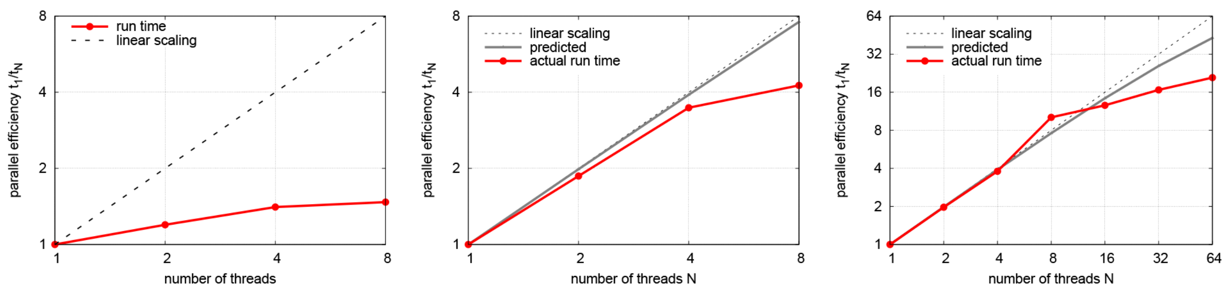


Figure 2. Scaling of computational speed with number of CPUs using the naive parallelization scheme (a), as compared to our approach on an 8 CPU desktop (b) and on a larger cluster with 32 CPUs per node (c).

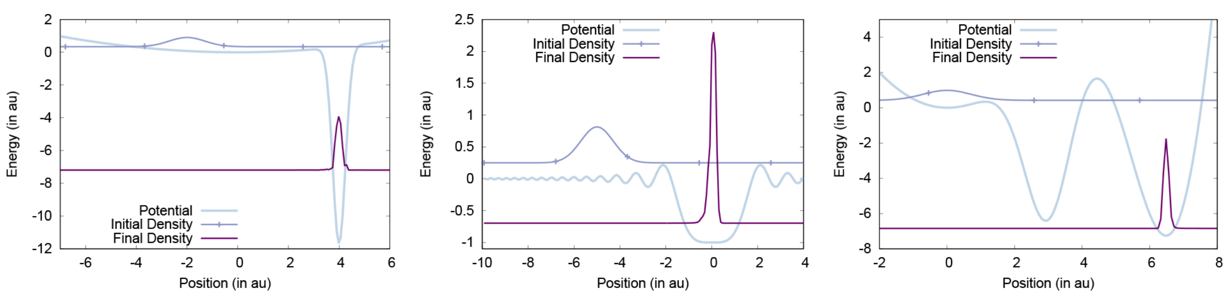


Figure 3. COCO global optimization in (left) single-well golf potential eq 19 (light blue line), (middle) rugged potential eq 20, and (right) triple-well golf potential yielding localization of the initial density (blue) at the final state (purple) after classical propagation for about 200 attoseconds.

$$S(k) = -\sum_{I=1}^{k} \sum_{J=1}^{B} P_I(J) \log P_I(J) \tag{18}$$

where $P_I(J)$ is the probability that a particle in cluster I is in bin J. For the model systems investigated, we chose equally spaced bins covering the area of the particle swarm ($B = 4$ bins in each direction of phase-space for one-dimensional potentials and $B = 4$ or $B = 2$ bins in each direction of coordinate-space area for the first six coordinate-space directions for multidimensional potentials). After the optimum number of clusters k was determined, the swarm of particles was then partitioned into the k clusters according to the k-means clustering algorithm, as shown in Figure 1a for $k = 14$. Each cluster was represented by a Gaussian ansatz with the average position $\langle x(t)\rangle$, momentum $\langle p(t)\rangle$ and standard deviations σ_x and σ_p of the cluster. The sum of Gaussians weighted by the corresponding number of trajectories per cluster defines the multi-Gaussian ansatz for the density of state $\rho(T; r_j, p_k)$. Figure 1b shows the comparison of the resulting k-means ansatz approximation to the exact density for the distribution of trajectories shown in Figure 1a.

All calculations were parallelized according to the parallelization scheme discussed in section 2.4. The Liouvillian operator $\mathcal{L}$ in eq 3, necessary to compute the DMORPH gradients defined by eq 34, was approximated through finite differencing of the density $\rho(t;x,p)$ with respect to x and p.

2.4. Parallelization. The implemented trivial parallelization scheme exploits the simplicity of propagating classical trajectories and achieves almost linear scaling with the number of CPUs by distributing the propagation of individual phase-space points among processing elements (Figure 2).

COCO could simply distribute the propagation of phase space points over CPUs, according to the final value of momentum. However, the final density of states obtained by such a naive approach would have to be sent to the master CPU for evaluation of observables, requiring expensive communication of the entire set of phase space amplitudes. Here, we implement a more efficient parallelization method based on the distribution of partially overlapping phase-space segments among CPUs, each of which includes two additional position rows flanking the individual section of phase space.

DOI: 10.1021/acs.jctc.8b00124
J. Chem. Theory Comput. 2018, 14, 3351−3362

The resulting computational overhead allows for the local finite-difference evaluation of the Liouvillian gradients. The components of the observable defined by eq 8 and DMORPH gradients introduced by eq 34 are thus evaluated locally and summed by an efficient parallel reduction call. As shown in Figure 2, the resulting parallelization scheme yields almost linear scaling when the number of CPUs is much smaller than the grid size. The predicted scaling deviates slightly due to the addition of the two rows of phase space points, with a greater impact for parallelization over more CPUs. Small deviations from linear scaling in the runtime are also seen due to machine specifications, such as system time measurement and network overhead.

3. RESULTS

3.1. Golf and Triple Well Potentials: Benchmark Grid-Base Calculations.

Figure 3 and Table 1 show that the COCO algorithm can successfully locate the global minima of golf potentials with long, flat valleys and near-degenerate minima. Optimal controls are converged within several hundred propagations. The optimized controls successfully directed the motion of the probe particle into the global energy minima within about 200 attoseconds (T = 8.0 au). The ensemble average value of the position at the final time $\langle x(T)\rangle$ provides the position of the global minimum x_{gm} within a standard deviation σ_x.

Table 1. Results (in atomic units) of Global Optimization of Potentials with Ensemble Average Value of the Position at the Final Time $\langle x(T)\rangle$ Localized near Global Minimum Location x_{gm} within Standard Deviation σ_x[a]

potential eq	$\langle x(T)\rangle$	x_{gm}	σ_x	$\langle V(T)\rangle$	props.
19	3.982	3.984	0.156	−10.407	264
20	0.038	0.000	0.151	−0.999	152
21	6.489	6.484	0.091	−7.215	68

[a]The minimum ensemble average value of potential $\langle V(T)\rangle$ is found after given number of propagations (props.).

COCO has successfully determined the global minimum of the golf potential, shown in Figure 3 (left panel), of the form

$$V(x) = \frac{1}{2}kx^2 - D\exp\left(-\frac{(x-x_0)^2}{2\sigma^2}\right) \tag{19}$$

which consists of a shallow harmonic well (k = 0.04 au) at the origin and a deep hole localized at x_0 = 4 au far away from the harmonic well minimum of width σ = 0.25 au and depth D = 12 au. A final mass initial guess m_f = 50 au was employed in eq 15.

Considering the initial position of the probe particle at x_k = −2.0 au, the gradient of the potential followed by steepest descent would lead the particle to the local minimum at x_{lm} = 0.0 au. However, following the gradient of the controls, according to eq 34, the probe particle is "lifted off" the surface like a drone and led to the global minimum at x_{gm} = 4.0 au.

The COCO method has successfully located the global minimum of a rugged potential with a flat surface at the minimum, shown in Figure 3 (middle panel), of the form

$$V(x) = -\frac{\sin(x^2)}{x^2} \tag{20}$$

even when the initial position of the probe particle x_k = −5.0 au was far from the global minimum for initial guess final mass m_f = 50 au.

Furthermore, COCO successfully determined the global minimum of a triple-well potential with near-degenerate minima shown in Figure 3 (right panel)[76,99] of the form

$$V(x) = \frac{k}{2}x^2 - D\exp\left(-\frac{(x-x_0)^2}{2\sigma^2}\right) - D'\exp\left(-\frac{(x-x_0')^2}{2\sigma'^2}\right) \tag{21}$$

with harmonic constant k = 1.0 au; local minimum well depth D = 10.0 au, position x_0 = 3.0 au, and width σ = 0.7 au; and global minimum well depth D' = 30.0 au, position x_0' = 7.0 au, and width σ' = 1.5 au. The global minimum was found for the initial guess final mass m_f = 100 au and initial position x_k = −2.0 au. We note that COCO successfully resolved the global minimum, even though the potential involves near-degenerate minima separated by a large barrier that surpasses the energy of the initial density ρ_i.

The field of COCO succeeds at resolving the global optimization problem of the triple-well without exploiting quantum effects, such as tunneling, by simply lifting the probe particle over the barriers and then localizing the density of states at the global minimum (Figure 4). In fact, even QuOCO turns out to be operating analogously (Figure 4) and thus both methods are expected to exhibit similar capabilities. In fact, the controls evolve similarly when propagated according to COCO and QuOCO, as shown for the model golf potential eq 19 (Figure 4), with k = 1.0 au, D = 8 au, x_0 = 4 au, σ = 1 au, and x_k = −2.0 au in eq 17 and an initial guess final mass m_f = 8 au in eq 15.

The time snapshots for the transient density along QuOCO and COCO optimization (Figure 4, top left) and comparison of time-dependent ensemble average value of "kinetic energy" (i.e., $\langle H(r,t;\beta) - V(r)\rangle$ in Figure 4, top right panels) show that there is energy transfer between the perturbational field and the probe particle, as necessary to lift the probe particle like a drone over the potential barrier as seen at time t = 3.0 au, before localizing it at the position of the global minimum at the final propagation time t = 8.0 au. Even the control parameters associated with the time-dependent mass (Figure 4, bottom left), electric field (Figure 4, bottom center) and dipole (Figure 4, bottom right) are quite comparable for QuOCO and COCO optimizations.

3.2. COCO Implementation Based on a Swarm of Particles.

Here, we explore the capabilities of the COCO algorithm as implemented with approximate gradients obtained with a Gaussian ansatz for the propagated density of states. The first and second moments of the single Gaussian ansatz are defined by the coordinates and momenta of a swarm of trajectories propagated according to the classical equations of motion. Initial conditions are sampled by Monte Carlo according to the initial density of states. Such a gridless implementation of COCO is expected to be particularly relevant to optimization of systems with high dimensionality that remain sufficiently localized in phase space. A generalization to an ansatz based on multiple Gaussians (as necessary for density of states that bifurcate or become delocalized in phase space) is discussed later in this section as implemented in conjunction with the k-means algorithm.[89−91]

DOI: 10.1021/acs.jctc.8b00124
J. Chem. Theory Comput. 2018, 14, 3351−3362

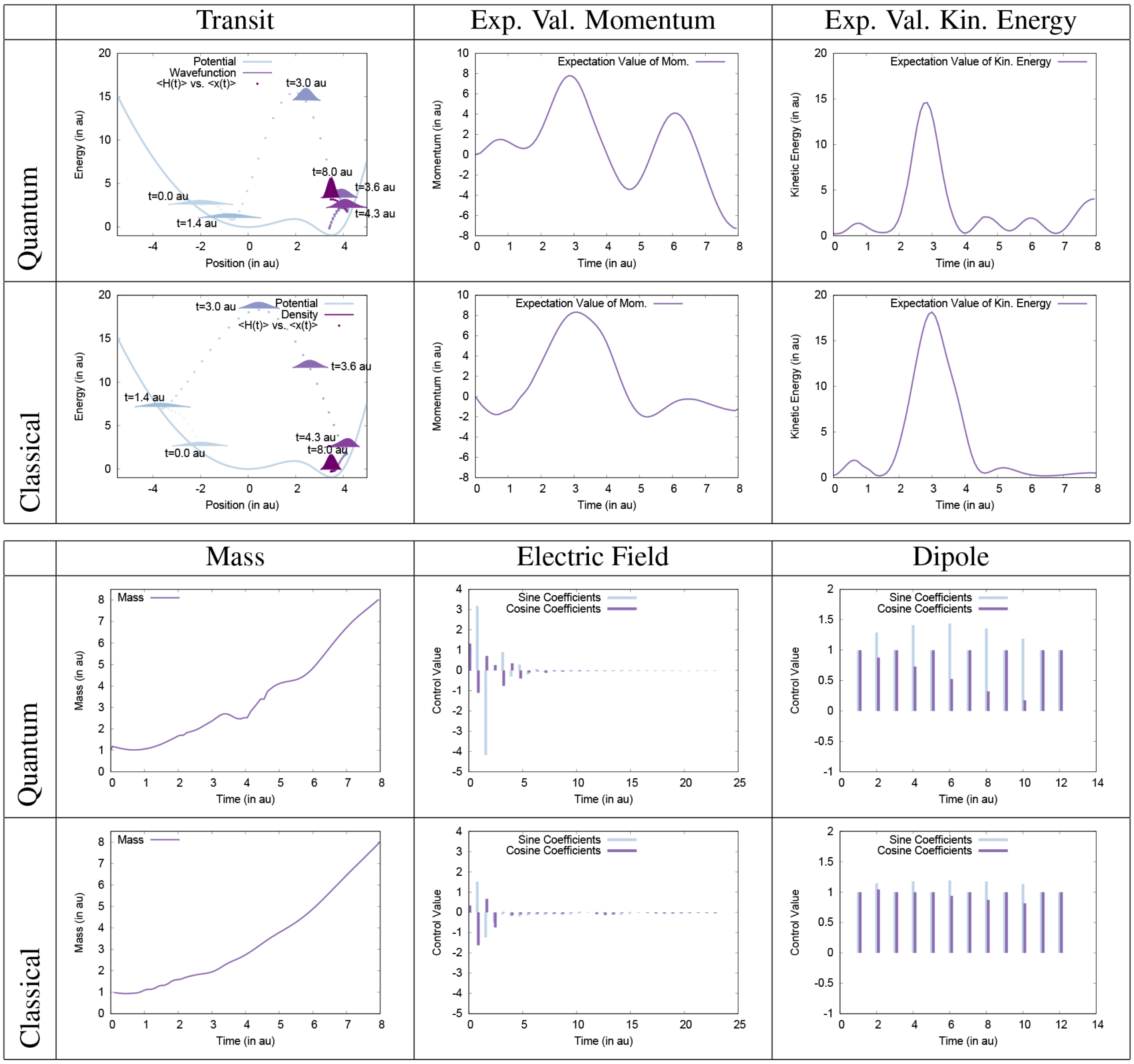


Figure 4. Comparison of results of quantum and classical optimal control optimization.

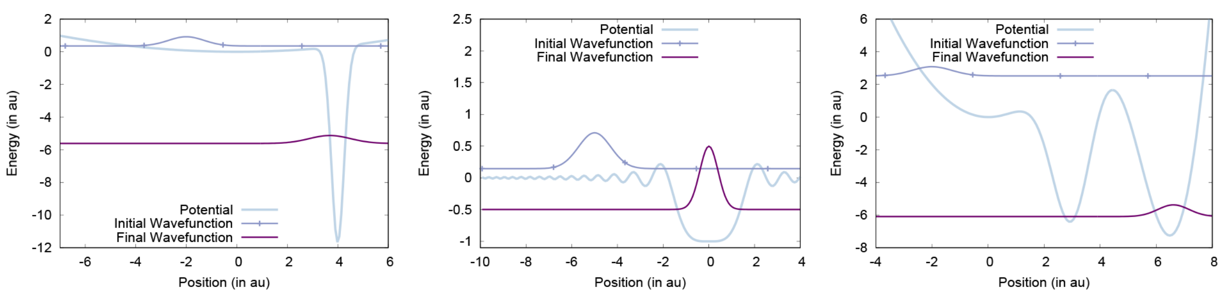


Figure 5. Global optimization based on a swarm of trajectories for the (a) single-well golf potential defined by eq 19 (light blue line); (b) rugged potential defined by eq 20; and (c) triple-well golf potential defined by eq 21. The initial density (blue) is evolved by the control field to the final density in global minimum well (purple).

Figure 5 and Table 2 show that the gridless implementation of COCO, based on a single Gaussian ansatz, successfully locates the global minima of the three benchmark model potentials discussed in Sec. 3.1. The parameters defining the potentials and initial conditions are described in section 3.1, with the exception of m_f in eq 15 for the golf potentials—here, $m_f = 16$ au for the single-well golf potential introduced by eq 19 and $m_f = 150$ au for the triple-well golf potential introduced by eq 21. In all cases, the ensemble average value of the position at the final time $\langle x(T) \rangle$ was located within a standard deviation σ_x

DOI: 10.1021/acs.jctc.8b00124
J. Chem. Theory Comput. 2018, 14, 3351−3362

Table 2. Results (in Atomic Units) of Global Optimization with Grid-Free Propagation Comparing Ensemble Average Value of Position of the Final Density $\langle x(T)\rangle$ and Location of Global Minimum Well x_{gm} within Standard Deviation σ_x and Giving the Ensemble Average Value of Potential at the Final Time $\langle V(T)\rangle$ after Given Number of Propagations (props.) of the Optimizer

potential eq	$\langle x(T)\rangle$	x_{gm}	σ_x	$\langle V(T)\rangle$	props.
19	3.688	3.984	0.821	−7.705	128
20	0.002	0.000	0.401	−0.987	100
21	6.623	6.484	0.533	−6.566	32

of the location of the global minimum x_{gm} within a couple hundred attoseconds and several hundred propagations.

Clustering of trajectories by using the k-means algorithm enabled the generalization of the gridless implementation of COCO to systems whose density of states might spread or bifurcate in phase space as it evolves in time. For the model systems investigated, the implementation of COCO based on clustering of trajectories also led to successful global minimization, as shown in Figure 6 and Table 3. Parameters were identical for the Gaussian ansatz implementation, with the exception of m_f in eq 15 (m_f = 25 au for the single-well golf potential eq 19 and rugged potential eq 20 and m_f = 150 au for the triple-well golf potential eq 21).

Furthermore, clustering yielded successful global minimization of a single-well golf potential eq 19 that confounded the grid-based global optimization method of section 3.1, as shown in Figure 7. Through the iterations of the optimizer from the initial density at position x_k = −5.0 au of m_f = 16 au, the clustering algorithm capitalized on a cluster of particles in the global minimum well at position x_0 = 4.0 au of width σ = 0.25 au and depth D = 12 au despite its distance from the harmonic well (k = 1.0 au). The ability of the clustering algorithm to locate distant minima suggests clustering may be used to locate distant global minima in general.

3.3. Lennard-Jones Clusters: Multidimensional COCO.

Grid-free COCO was implemented for the optimization of Lennard-Jones clusters. COCO was found to determine the minimum energy configuration of Lennard-Jones clusters of the form[99,100]

Table 3. Results (in Atomic Units) of Global Optimization with Clustered Particle Swarm Propagation Comparing the Ensemble Average Value of Position of the Final Density $\langle x(T)\rangle$ and Location of Global Minimum Well x_{gm} within Standard Deviation σ_x and Giving the Ensemble Average Value of Potential at the Final Time $\langle V(T)\rangle$ after Given Number of Propagations (props.) of the Optimizer

potential eq	$\langle x(T)\rangle$	x_{gm}	σ_x	$\langle V(T)\rangle$	props.
19	3.989	3.984	0.382	−8.712	232
20	0.058	0.000	0.282	−0.995	216
21	6.438	6.484	0.147	−7.118	48

$$V(\mathbf{r}) = \sum_{i>j=1}^{N} V(r_{ij}) \tag{22}$$

$$= \sum_{i>j=1}^{N} 4\epsilon\left[\left(\frac{\sigma}{r_{ij}}\right)^{12} - \left(\frac{\sigma}{r_{ij}}\right)^{6}\right] \tag{23}$$

Results were determined for N particles of interatomic distances $\mathbf{r} = \{r_{ij}\}$, atomic diameter σ = 2. au, and well depth ϵ = 1. au (ϵ = 10. au for 4 to 13 particles) at minimum $r_{ij} = \sqrt[6]{2}\sigma$. To ensure the particles were optimized as a single cluster, a harmonic trapping potential of strength constant k = 0.1 au was included for clusters of less than 4 particles with the form[99,101,102]

$$V(\mathbf{r}) = \sum_{i>j=1} \frac{k}{2} r_{ij}^2 \tag{24}$$

A cutoff of $V(r_{ij}) = V(0.5\sigma^2)$ for small interatomic distances $r_{ij}^2 < 0.5\sigma^2$ prevented numeric overflow. The cluster position was also fixed by the choice of particle i coordinates (x_i, y_i, z_i) of $x_1 = y_1 = z_1 = y_2 = z_2 = z_3 = 0$.[103]

The global minimum was located for various Lennard-Jones clusters, as shown in Figure 8 and Table 4. The global minima were found for an initial guess final mass of m_f = 1024 au in all cases. The optima were found with a Gaussian initial density eq 17 centered at

$$\mathbf{r}_k = [3, 3, 3, -3, 3, 3] \text{ au} \tag{25}$$

for 2 to 4 particles or

$$\mathbf{r}_k = [2, 1, 2, -1, 2, 1, 1, -2, 1, 1, 1, -2 \tag{26}$$

$$-1, -1, 2, 1, -2, -1, -1, 2, -1, 1, 1, 2 \tag{27}$$

$$-2, 0, 0, -1, -1, -2, -1, -2, 0] \text{ au} \tag{28}$$

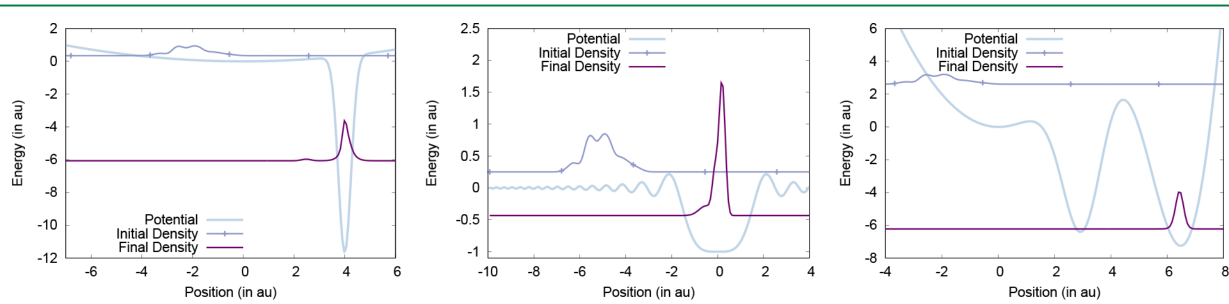


Figure 6. Global optimization with clustered particle swarm propagation of (a) single-well golf potential eq 19 (light blue line), (b) rugged potential eq 20, and (c) triple-well golf potential eq 21 localizes initial density (blue) as final density in global minimum well (purple).

DOI: 10.1021/acs.jctc.8b00124
J. Chem. Theory Comput. 2018, 14, 3351−3362

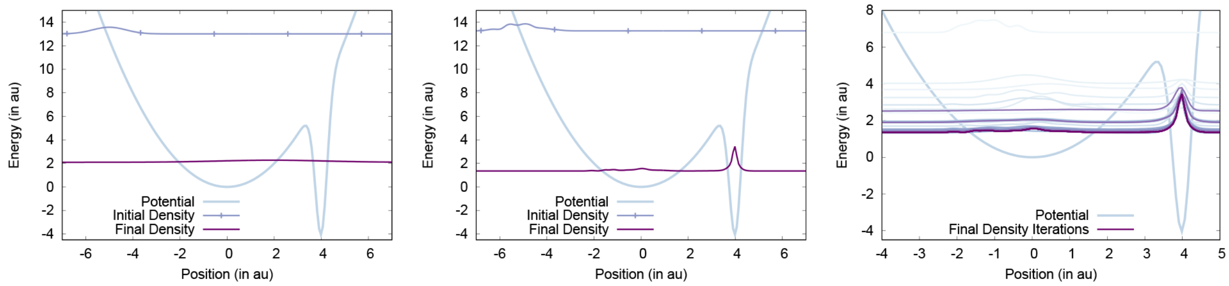


Figure 7. Global optimization of benchmark single-well golf potential eq 19 (light blue line) with particle swarm propagation (a) without and (b) with clustering algorithm localizes initial density (blue) as final density in global minimum well (purple). (c) The final density successively locates the global minimum through iteration from the initial guess result (lightest blue) to the final optimized result (dark purple).

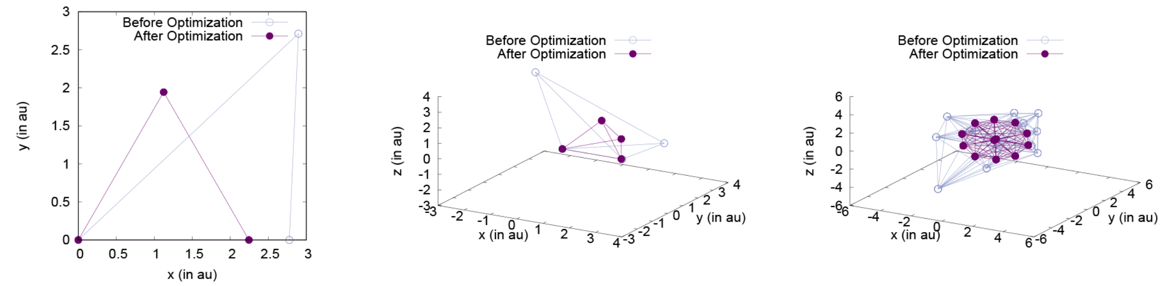


Figure 8. COCO optimization of Lennard-Jones clusters eq 23 determines global minimum configuration (dark purple) from initial guess (light purple).

Table 4. Results (as Multiple of Depth Parameter ϵ) of Global Optimization of Lennard-Jones Clusters Eq 23 with Differing Number of Particles (No. Particles) Comparing Ensemble Average Value of Potential at the Final Time $\langle V(T)\rangle$ to the Previously Reported Global Minimum (GM)[104] after Given Number of Propagations (Props.) of the Optimizer

no. particles	$\langle V(T)\rangle$	GM	props.
3	-3.000ϵ	-3.000ϵ	240
4	-6.000ϵ	-6.000ϵ	112
13	-44.327ϵ	-44.327ϵ	84

for 13 particles. As shown in Figure 8 and Table 4, the global minima found via COCO agreed with the lowest reported global minima.[104]

4. TOWARD APPLICATIONS

The current method is expected to be applicable for optimization of any function in position-space. However, the method as stated has limitations in control space. Successful optimization requires a sufficient number of controls and a sufficient final time T for propagation. Moreover, the current method is primarily applicable to smooth functions of control space due to the inclusion of the L-BFGS-B optimizer in application of the D-MORPH gradient eq 34. For example, in the case of a 13-particle Lennard-Jones cluster, the L-BFGS-B algorithm located the global minimum but the solution did not meet the conditions for completion of the L-BFGS-B line search. The early termination of the algorithm suggested the surface may have been sufficiently nonsmooth or ill-conditioned to cause difficulties with the L-BFGS-B algorithm. In the future, these limitations can be overcome with a sufficient number of controls, conversion of the final time T from a parameter to a control, and use of a nonsmooth alternative to the L-BFGS-B optimizer.

COCO presents an efficient method for global optimization of multidimensional functions. In contrast to the calculation of the function at every point, the method requires only calculation of the function "on-the-fly" along the dynamical path of a particle swarm. The path of the particles may constitute a small subset of the full points on the surface, which can lead to computational savings in the number of function evaluations. As a global optimization method, COCO also exhibits advantages over local optimization methods. In global minimization applications of local optimization methods, the local method may be used repeatedly to determine a set of local minima of which the lowest is understood to be the global minimum. However, unless the function is calculated everywhere, the true global minimum may not be among this set. In contrast, COCO uses the control-space gradient to find the global minimum through the directed motion of particles. In this way, the number of necessary function evaluations is reduced and, given sufficient number of controls, the optimizer is led directly to the global minimum.[79]

5. CONCLUSIONS

We have introduced the COCO algorithm for energy minimization, based on classical dynamics steered by a controllable external adaptive field. COCO is a classical analogue of the recently introduced QuOCO method that exploits the diffeomorphic modulation under observable-response-preserving homotopy (DMORPH) gradient and the Broyden Fletcher Goldfarb Shanno (BFGS) iterative scheme for nonlinear optimization. We have shown that the classical analogue DMORPH gradients of the ensemble average values with respect to the controls can be obtained in terms of gradients of the classical Liouvillian, requiring only 3 propagations independent of the number N of control parameters. The classical DMORPH gradients thus introduce significant computational advantages relative to finite difference

DOI: 10.1021/acs.jctc.8b00124
J. Chem. Theory Comput. 2018, 14, 3351−3362

methods that require $N + 1$ propagations. We have compared benchmark grid-based implementations (i.e., by propagating the amplitudes of the density of states on a phase-space grid in the Eulerian frame) to implementations based on classical trajectories in the Lagrangian frame with time-evolved density of states approximated by a single-Gaussian or a multiple-Gaussian ansatz generated by the *k*-means clustering algorithm. We have shown the capabilities of the COCO algorithm as applied to resolving the global minima of golf potentials, rugged surfaces, and multiwells with near degenerate minima separated by high energy barriers. COCO has also been shown to successfully locate the global minima of multidimensional Lennard-Jones clusters. The reported results show promise for practical implementations.

APPENDIX A. DERIVATION OF CLASSICAL DMORPH

The classical DMORPH algorithm can be derived in analogy to the quantum DMORPH algorithm through the method presented in ref 76. The classical propagator $U(t, 0) = e^{-i\int_0^t dt' \mathcal{L}(t')}$ for Liouvillian operator $\mathcal{L}$ fulfills the Liouville equation, as follows:

$$i\frac{\partial}{\partial t}U(t, 0) = \mathcal{L}(t)U(t, 0) \tag{29}$$

Application of the product rule to eq 29 and the adjoint Liouville equation derived in Appendix B yields

$$i\frac{\partial}{\partial t}[U^{-1}(t, 0)U_{\beta_j}(t, 0)] = i\left[\frac{\partial U^{-1}(t, 0)}{\partial t}U_{\beta_j}(t, 0) + U^{-1}(t, 0)\frac{\partial U_{\beta_j}(t, 0)}{\partial t}\right] \tag{30}$$

$$= -U^{-1}(t, 0)\mathcal{L}(t)U_{\beta_j}(t, 0) + U^{-1}(t, 0)[\mathcal{L}_{\beta_j}(\beta, t)U(t, 0) + \mathcal{L}(\beta, t)U_{\beta_j}(t, 0)] \tag{31}$$

$$= U^{-1}(t, 0)\mathcal{L}_{\beta_j}(\beta, t)U(t, 0) \tag{32}$$

Integrating eq 32, we obtain the classical DMORPH expression for the gradient propagator

$$U_{\beta_j}(T, 0) = -iU(T, 0)\int_0^T U^{-1}(t, 0)\mathcal{L}_{\beta_j}(\beta, t)U(t, 0)dt \tag{33}$$

in complete analogy to the corresponding quantum expression.[76]

Substitution of the control-space gradient of the classical propagator eq 33 into the equation for the control-space gradient of the ensemble average of the observable eq 4 yields the classical DMORPH gradient, as follows:

$$\begin{aligned}\frac{\partial\langle O(T)\rangle}{\partial\beta_j} &= \int drdp O(r, p)U_{\beta_j}(T, 0)\rho(0; r, p) \\ &= \int drdp O(r, p)U(T, 0)\int_0^T dt U^{-1}(t, 0)(-i\mathcal{L}_{\beta_j}(t))U(t, 0) \rho(0; r, p) \\ &= \int_0^T dt\left[\int drdp O(r, p)U(T, 0)U^{-1}(t, 0)(-i\mathcal{L}_{\beta_j}(t))U(t, 0) \rho(0; r, p)\right] \\ &= \int_0^T dt \int drdp O(r, p)U(T, t)(-i\mathcal{L}_{\beta_j}(t))\rho(t; r, p)\end{aligned} \tag{34}$$

in analogy to the quantum DMORPH gradient eq 2.

APPENDIX B. ADJOINT LIOUVILLE EQUATION

We derive the adjoint Liouville equations as a classical analogue of the adjoint Schrödinger equation. The backward propagator from time t to time 0 is

$$U^\dagger(t, 0) = (U(t, 0))^{-1} = U^{-1}(t, 0) = U(0, t) \tag{35}$$

where

$$U^\dagger(t, 0)U(t, 0) = U^{-1}(t, 0)U(t, 0) = \mathrm{Id} = \mathrm{const} \tag{36}$$

$$\begin{aligned} i\hbar\frac{\partial}{\partial t}(U^\dagger(t, 0)U(t, 0)) &= i\hbar\frac{\partial}{\partial t}U^\dagger(t, 0)U(t, 0) + U^\dagger(t, 0)i\hbar\frac{\partial}{\partial t}U(t, 0) \\ &= 0\end{aligned} \tag{37}$$

where Id is the identity operator. For the quantum propagator

$$i\hbar\frac{\partial}{\partial t}U(t, 0) = H(\beta, t)U(t, 0) \tag{38}$$

Therefore, the propagator can be shown to obey the equation

$$i\hbar\frac{\partial}{\partial t}U^\dagger(t, 0)U(t, 0) + U^\dagger(t, 0)H(t)U(t, 0) = 0 \tag{39}$$

$$-i\hbar\frac{\partial}{\partial t}U^\dagger(t, 0)U(t, 0) = U^\dagger(t, 0)H(t)U(t, 0) \tag{40}$$

$$-i\hbar\frac{\partial}{\partial t}U^\dagger(t, 0) = U^\dagger(t, 0)H(t) \tag{41}$$

where the last line is obtained by applying $U^\dagger(t,0)$ from the right, giving the adjoint Schrödinger equation which was derived using the inverse property of the backward propagator, never invoking the adjoint operation or the self-adjoint property of the Hamilton operator.

In analogy to the adjoint Schrödinger eq 41, a temporal inverse of the classical propagator $U(t, 0) = e^{-i\int_0^t dt' \mathcal{L}(t')}$ introduced by eq 29 can be obtained, as follows:

$$0 = \frac{\partial}{\partial t}(\mathrm{Id}) \tag{42}$$

$$= i\frac{\partial}{\partial t}(U^{-1}(t, 0)U(t, 0)) \tag{43}$$

$$= i\frac{\partial}{\partial t}U^{-1}(t, 0)U(t, 0) + U^{-1}(t, 0)\mathcal{L}(t)U(t, 0) \tag{44}$$

DOI: 10.1021/acs.jctc.8b00124
J. Chem. Theory Comput. 2018, 14, 3351−3362

$$-i\frac{\partial}{\partial t}U^{-1}(t, 0)U(t, 0) = U^{-1}(t, 0)\mathcal{L}(t)U(t, 0) \tag{45}$$

$$-i\frac{\partial}{\partial t}U^{-1}(t, 0) = U^{-1}(t, 0)\mathcal{L}(t) \tag{46}$$

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: victor.batista@yale.edu.

ORCID

Victor S. Batista: 0000-0002-3262-1237

Present Address

†Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA 02138

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

V.S.B. acknowledges support from the National Science Foundation grant CHE 1465108 and supercomputer time from NERSC and the Yale University Faculty of Arts and Sciences High Performance Computing Center partially funded by the National Science Foundation grant CNS 08-21132. M.S. acknowledges financial support from the Arnold and Mabel Beckman Foundation, the National Science Foundation Graduate Research Fellowship, the Harvard Merit/Graduate Society Term-time Research Fellowship, and the Fulbright U.S. Student Program to Germany. This work used Extreme Science and Engineering Discovery Environment (XSEDE), which was supported by National Science Foundation grant number DGE-1144152. We would like to thank Dr. Erik T. J. Nibbering for his support.

■ REFERENCES

(1) Shi, S.; Woody, A.; Rabitz, H. Optimal control of selective vibrational excitation in harmonic linear chain molecules. *J. Chem. Phys.* **1988**, *88*, 6870−6883.

(2) Peirce, A. P.; Dahleh, M. A.; Rabitz, H. Optimal control of quantum-mechanical systems: Existence, numerical approximation, and applications. *Phys. Rev. A: At., Mol., Opt. Phys.* **1988**, *37*, 4950−4964.

(3) Shi, S.; Rabitz, H. Selective excitation in harmonic molecular systems by optimally designed fields. *Chem. Phys.* **1989**, *139*, 185−199.

(4) Kosloff, R.; et al. Wavepacket dancing: Achieving chemical selectivity by shaping light pulses. *Chem. Phys.* **1989**, *139*, 201−220.

(5) Jakubetz, W.; et al. Theory of optimal laser pulses for selective transitions between molecular eigenstates. *Chem. Phys. Lett.* **1990**, *165*, 100−106.

(6) Brif, C.; Chakrabarti, R.; Rabitz, H. Control of quantum phenomena: past, present and future. *New J. Phys.* **2010**, *12*, 075008.

(7) Chen, C.; Yin, Y.-Y.; Elliott, D. S. Interference between optical transitions. *Phys. Rev. Lett.* **1990**, *64*, 507−510.

(8) Chen, C.; Elliott, D. S. Measurements of optical phase variations using interfering multiphoton ionization processes. *Phys. Rev. Lett.* **1990**, *65*, 1737−1740.

(9) Park, S. M.; Lu, S.-P.; Gordon, R. J. Coherent laser control of the resonance-enhanced multiphoton ionization of HCl. *J. Chem. Phys.* **1991**, *94*, 8622−8624.

(10) Zhu, L.; Kleiman, V. D.; Li, X.; Lu, S.; Trentelman, K.; Wada, R.; Gordon, R. J. Coherent Laser Control of the Product Distribution Obtained in the Photoexcitation of HI. *Science* **1995**, *270*, 77−80.

(11) Zhu, L.; Suto, K.; Fiss, J. A.; Wada, R.; Seideman, T.; Gordon, R. J. Effects of resonances on the coherent control of the photoionization and photodissociation of HI and DI. *Phys. Rev. Lett.* **1997**, *79*, 4108−4111.

(12) Kleiman, V. D.; Zhu, L.; Li, X.; Gordon, R. J. Coherent phase control of the photoionization of H_2S. *J. Chem. Phys.* **1995**, *102*, 5863−5866.

(13) Kleiman, V. D.; Zhu, L.; Allen, J.; Gordon, R. J. Coherent control over the photodissociation of CH_3I. *J. Chem. Phys.* **1995**, *103*, 10800−10803.

(14) Xing, G.; Wang, X.; Huang, X.; Bersohn, R.; Katz, B. Modulation of resonant multiphoton ionization of CH_3I by laser phase variation. *J. Chem. Phys.* **1996**, *104*, 826−831.

(15) Shnitman, A.; Sofer, I.; Golub, I.; Yogev, A.; Shapiro, M.; et al. Experimental observation of laser control: Electronic branching in the photodissociation of Na_2. *Phys. Rev. Lett.* **1996**, *76*, 2886−2889.

(16) Baumert, T.; Helbing, J.; Gerber, G. Coherent control with femtosecond laser pulses. *Adv. Chem. Phys.* **2007**, *101*, 47−82.

(17) Herek, J. L.; Materny, A.; Zewail, A. H. Femtosecond controla of an elementary unimolecular reaction fromt the transition-state region. *Chem. Phys. Lett.* **1994**, *228*, 15−25.

(18) Potter, E. D.; Herek, J. L.; Pedersen, S.; Liu, Q.; Zewail, A. H. Femtosecond laser control of a chemical reaction. *Nature* **1992**, *355*, 66−68.

(19) Baumert, T.; Gerber, G. Fundamental interaction of molecules (Na_2, Na_3) with intense femtosecond laser pulses. *Isr. J. Chem.* **1994**, *34*, 103−114.

(20) Manz, J.; Sundermann, K.; de Vivie-Riedle, R. Quantum optimal control strategies for photoisomerization via electronically excited states. *Chem. Phys. Lett.* **1998**, *290*, 415−422.

(21) Artamonov, M.; Ho, T.-S.; Rabitz, H. Quantum optimal control of ozone isomerization. *Chem. Phys.* **2004**, *305*, 213−222.

(22) Artamonov, M.; Ho, T.-S.; Rabitz, H. Quantum optimal control of HCN isomerization. *Chem. Phys.* **2006**, *328*, 147−155.

(23) Artamonov, M.; Ho, T.-S.; Rabitz, H. Quantum optimal control of molecular isomerization in the presence of a competing dissociation channel. *J. Chem. Phys.* **2006**, *124*, 064306.

(24) Kurosaki, Y.; Artamonov, M.; Ho, T.-S.; Rabitz, H. Quantum optimal control of isomerization dynamics of a one-dimensional reaciton-path model dominated by a competing dissociation channel. *J. Chem. Phys.* **2009**, *131*, 044306.

(25) Kurosaki, Y.; Ho, T.-S.; Rabitz, H. The role of dissociation channels of excited electronic states in quantum optimal control of ozone isomerization: A three-state dynamical model. *Chem. Phys.* **2016**, *469-470*, 115−122.

(26) Solas, F.; Ashton, J. M.; Markmann, A.; Rabitz, H. A. Toward adaptive control of coherent electron transport in semiconductors. *J. Chem. Phys.* **2009**, *130*, 214702.

(27) Beltrani, V.; Rabitz, H. Exploiting time-independent Hamiltonian structure as controls for manipulating quantum dynamics. *J. Chem. Phys.* **2012**, *137*, 094109.

(28) De Chiara, G.; Calarco, T.; Anderlini, M.; Montangero, S.; Lee, P. J.; Brown, B. L.; Phillips, W. D.; Porto, J. V. Optimal control of atom transport for quantum gates in optical lattices. *Phys. Rev. A: At., Mol., Opt. Phys.* **2008**, *77*, 052333.

(29) Hohenester, U.; Rekdal, P. K.; Borzì, A.; Schmiedmayer, J. Optimal quantum control of Bose−Einstein condensates in magnetic microtraps. *Phys. Rev. A: At., Mol., Opt. Phys.* **2007**, *75*, 023602.

(30) Doria, P.; Calarco, T.; Montangero, S. Optimal control technique for Many Body Quantum Systems dynamics. 2010, arxiv:1003.3750. arXiv.org e-Print archive. https://arxiv.org/abs/1003.3750.

(31) Li, G.; Welack, S.; Schreiber, M.; Kleinekathöfer, U. Tailoring current flow patterns through molecular wires using shaped optical pulses. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2008**, *77*, 075321.

(32) Grigorenko, I.; Rabitz, H. Optimal control of the local electromagnetic response of nanostructured materials: Optimal detectors and quantum disguises. *Appl. Phys. Lett.* **2009**, *94*, 253107.

(33) Sanders, G. D.; Kim, K. W.; Holton, W. C. Quantum computing with complex instruction sets. *Phys. Rev. A: At., Mol., Opt. Phys.* **1999**, *59*, 1098−1101.

DOI: 10.1021/acs.jctc.8b00124
J. Chem. Theory Comput. 2018, 14, 3351−3362

(34) Tesch, C. M.; Kurtz, L.; de Vivie-Riedle, R. Applying optimal control theory for elements of quantum computation in molecular systems. *Chem. Phys. Lett.* **2001**, *343*, 633−641.
(35) D'Alessandro, D.; Dahleh, M. Optimal control of two-level quantum systems. *IEEE Trans. Autom. Control* **2001**, *45*, 866−876.
(36) Palao, J. P.; Kosloff, R. Quantum Computing by an Optimal Control Algorithm for Unitary Transformations. *Phys. Rev. Lett.* **2002**, *89*, 188301.
(37) Grace, M. D.; Dominy, J.; Kosut, R. L.; Brif, C.; Rabitz, H. Environment-invariant measure of distance between evolutions of an open quantum system. *New J. Phys.* **2010**, *12*, 015001.
(38) Chou, Y.; Huang, S.-Y.; Goan, H.-S. Optimal control of fast and high-fidelity quantum gates with electron and nuclear spins of a nitrogen-vacancy center in diamond. *Phys. Rev. A: At., Mol., Opt. Phys.* **2015**, *91*, 052315.
(39) Xi, Z. Entanglement capacity of ising type systems with control. In *Proceedings of the Chinese Control Conference*; IEEE, 2006; 2109−2114.
(40) Murphy, M.; Montangero, S.; Giovannetti, V.; Calarco, T. Communication at the quantum speed limit along a spin chain. *Phys. Rev. A: At., Mol., Opt. Phys.* **2010**, *82*, 022318.
(41) Zhang, X.-P.; Shao, B.; Hu, S.; Zou, J.; Wu, L.-A. Optimal control of fast and high-fidelity quantum state transfer in spin-1/2 chains. *Ann. Phys.* **2016**, *375*, 435−443.
(42) Greene, S. M.; Batista, V. S. The TT-SOFT Method: Multidimensional Nonadiabatic Quantum Dynamics. *J. Chem. Theory Comput.* **2017**, *13*, 4034−4042.
(43) Levitt, M.; Warshel, A. Computer-simulation of Protein Folding. *Nature* **1975**, *253*, 694−698.
(44) Samuel, A. Some studies in machine learning using the game of checkers. *IBM J. Res. Dev.* **1959**, *3*, 210−229.
(45) Hooke, R.; Jeeves, T. A. "Direct Search" Solution of Numerical and Statistical Problems. *J. Assoc. Comput. Mach.* **1961**, *8*, 212−229.
(46) Spendley, W.; Hext, G. R.; Himsworth, F. R. Sequential Application of Simplex Designs in Optimisation and Evolutionary Operation. *Technometrics* **1962**, *4*, 441−461.
(47) Nelder, J. A.; Mead, R. A Simplex Method for Function Minimization. *Comput. J.* **1965**, *7*, 308−313.
(48) Lee, E. T. Optimization by a Gradient Technique. *Ind. Eng. Chem. Fundam.* **1964**, *3*, 373−380.
(49) Stillinger, F. H.; Rahman, A. Molecular Dynamics Study of Temperature Effects on Water Structure and Kinetics. *J. Chem. Phys.* **1972**, *57*, 1281−1292.
(50) Witt, A.; Sebastianelli, F.; Tuckerman, M. E.; Bacic, Z. Path Integral Molecular Dynamics Study of Small H(2) Clusters in the Large Cage of Structure II Clathrate Hydrate: Temperature Dependence of Quantum Spatial Distributions. *J. Phys. Chem. C* **2010**, *114*, 20775−20782.
(51) Yu, T. Q.; Tuckerman, M. E. Temperature-Accelerated Method for Exploring Polymorphism in Molecular Crystals Based on Free Energy. *Phys. Rev. Lett.* **2011**, *107*, 015701.
(52) Levitt, M.; Lifson, S. Refinement of protein conformations using a macromolecular energy minimization procedure. *J. Mol. Biol.* **1969**, *46*, 269−279.
(53) Minary, P.; Tuckerman, M. E.; Martyna, G. J. Dynamical Spacial Warping: A Novel Method for the Conformational Sampling of Biophysical Structure. *SIAM* **2008**, *30*, 2055−2083.
(54) Summa, C. M.; Levitt, M. Near-native structure refinement using in vacuo energy minimization. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 3177−3182.
(55) Amara, P.; Hsu, D.; Straub, J. E. Global energy minimum searches using an approximate solution of the imaginary time Schroedinger Equation. *J. Phys. Chem.* **1993**, *97*, 6715−6721.
(56) Finnila, A. B.; et al. Quantum annealing: A new method for minimizing multidimensional functions. *Chem. Phys. Lett.* **1994**, *219*, 343−348.
(57) Andricioaei, I.; Straub, J. E. Finding the needle in the haystack: Algorithms for conformational optimization. *Comput. Phys.* **1996**, *10*, 449−454.
(58) Farhi, E.; Goldstone, J.; Gutmann, S.; Sipser, M. Quantum Computation by Adiabatic Evolution. 2000, arXiv:quant-ph/0001106. arXiv.org e-Print archive. https://arxiv.org/abs/quant-ph/0001106.
(59) Liu, P.; Berne, B. J. Quantum path minimization: An efficient method for global optimization. *J. Chem. Phys.* **2003**, *118*, 2999−3005.
(60) Fogel, L. Autonomous Automata. *Industrial Research Magazine* **1962**, *4*, 14−19.
(61) Cavicchio, D. J. Adaptive search using simulated evolution. Ph.D. thesis, Dept. Computer and Communication Sciences, University of Michigan, 1970.
(62) Holland, J. H. *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*; University of Michigan Press, 1975.
(63) Koza, J. R. Hierarchical Genetic Algorithms Operating on Populations of Computer Programs. In *Proceedings of the 11th International Joint Conference on Artificial Intelligence*; Kaufman, M., Ed.; American Association for Artificial Intelligence: San Mateo, 1989.
(64) Koza, J. R. *Genetic Programming: A Paradigm for Genetically Breeding Populations of Computer Programs to Solve Problems*; Stanford University: Stanford, CA, 1990.
(65) Kennedy, J.; Eberhart, R. Particle Swarm Optimization. *ICNN Proceedings* **1995**, *1−6*, 1942−1948.
(66) Glover, F.; McMillan, C.; Novick, B. Interactive design software and computer graphics for architectural and space planning. *Ann. Oper. Res.* **1985**, *5*, 557−573.
(67) Glover, F. Tabu Search Methods in Artificial Intelligence and Operations Research. *ORSA Artif. Intell.* **1987**, *1*, 6.
(68) Sejnowski, T. J.; Hinton, G. E. In *Vision, Brain, and Computer*; Arbib, M. A., Hanson, A. R., Eds.; MIT Press: Cambridge, MA, 1986.
(69) Baum, E. B. Intractable Computations without Local Minima. *Phys. Rev. Lett.* **1986**, *57*, 2764−2767.
(70) Bernasconi, J. Low autocorrelation binary sequences: statistical mechanics and configuration space analysis. *J. Phys. (Paris)* **1987**, *48*, 559−567.
(71) Dill, K. A. Polymer principles and protein folding. *Protein Sci.* **1999**, *8*, 1166−1180.
(72) Lee, Y. H.; Berne, B. J. Global Optimization: Quantum Thermal Annealing with Path Integral Monte Carlo. *J. Phys. Chem. A* **2000**, *104*, 86−95.
(73) Tovchigrechko, A.; Vakser, I. A. How common is the funnel-like energy landscape in protein-protein interactions. *Protein Sci.* **2001**, *10*, 1572−1583.
(74) Zhou, H.-X. Disparate Ionic-Strength Dependencies of On and Off Rates in Protein-Protein Association. *Biopolymers* **2001**, *59*, 427−433.
(75) Locatelli, M. In *Handbook of Global Optimization*; Pardalos, P. M., Romeijn, H. E., Eds.; Kluwer Academic Publishers: Dordrecht, The Netherlands, 2002; Vol. 2, pp 179−229.
(76) Soley, M.; Markmann, A.; Batista, V. S. Steered Quantum Dynamics for Energy Minimization. *J. Phys. Chem. B* **2015**, *119*, 715−727.
(77) Boltyanskii, V. G.; Gamkrelidze, R. V.; Pontryagin, L. S. On the theory of optimal processes. *Dokl. Akad. Nauk SSSR* **1956**, *110*, 7−10.
(78) Pontryagin, L. S.; Boltyanskii, V. G.; Gamkrelidze, R. V.; Mishechenko, E. F. In *The Mathematical Theory of Optimal Processes*; Bittner, L., Ed.; John Wiley and Sons, 1962.
(79) Rabitz, H. A.; Hsieh, M. M.; Rosenthal, C. M. Quantum Optimally Controlled Transition Landscapes. *Science* **2004**, *303*, 1998−2001.
(80) Rothman, A.; Ho, T.-S.; Rabitz, H. Observable-preserving control of quantum dynamics over a family of related system. *Phys. Rev. A: At., Mol., Opt. Phys.* **2005**, *72*, 023416.
(81) Rothman, A.; Ho, T.-S.; Rabitz, H. A. Quantum observable homotopy tracking control. *J. Chem. Phys.* **2005**, *123*, 134104.
(82) Ho, T.; Rabitz, H. Why do effective quantum controls appear easy to find? *J. Photochem. Photobiol., A* **2006**, *180*, 226−240.
(83) Wu, R.; Rabitz, H.; Hsieh, M. Characterization of the critical submanifolds in quantum ensemble control landscapes. *J. Phys. A: Math. Theor.* **2008**, *41*, 015006.

DOI: 10.1021/acs.jctc.8b00124
J. Chem. Theory Comput. 2018, 14, 3351−3362

(84) Hsieh, M.; Wu, R.; Rabitz, H. Topology of the quantum control landscape for observables. *J. Chem. Phys.* **2009**, *130*, 104109.
(85) Beltrani, V.; Dominy, J.; Ho, T.-S.; Rabitz, H. Exploring the top and bottom of the quantum control landscape. *J. Chem. Phys.* **2011**, *134*, 194106.
(86) Zhao, M.; Rice, S. A. Comment concerning the optimal control of transformations in an unbounded quantum system. *J. Chem. Phys.* **1991**, *95*, 2465−2472.
(87) Demiralp, M.; Rabitz, H. Optimally controlled quantum molecular dynamics: A perturbation formulation and the existence of multiple solutions. *Phys. Rev. A: At., Mol., Opt. Phys.* **1993**, *47*, 809−816.
(88) Botina, J.; Rabitz, H.; Rahman, N. A new approach to molecular classical optimal control: Application to the reaction HCN → HC + N. *J. Chem. Phys.* **1995**, *102*, 226−36.
(89) MacQueen, J. Some methods for classification and analysis of multivariate observations. In *Proceedings of the Berkeley Symposium on Mathematical Statistics and Probability*; University of California, Berkeley, 1967; pp 281−297.
(90) Lloyd, S. P. Least squares quantization in PCM. *IEEE Trans. Inf. Theory* **1982**, *28*, 129−137.
(91) Mirkin, B. Chapter K-Means Clustering and Related Approaches. *Clustering: A Data Recovery Approach: Computer Science and Data Analysis*; CRC Press, 2013; pp 87−132.
(92) Byrd, R. H.; Lu, P.; Nocedal, J.; Zhu, C. A Limited Memory Algortihm for Bound Constrained Optimization. *SIAM Journal on Sci. Stat. Comp.* **1995**, *16*, 1190−1208.
(93) Zhu, C.; Byrd, R. H.; Lu, P.; Nocedal, J. L-BFGS-B: Algorithm 778: L-BFGS-B, FORTRAN routines for large scale bound constrained optimization. *ACM Trans. Math. Softw.* **1997**, *23*, 550−560.
(94) Morales, J. L.; Nocedal, J. L-BFGS-B: Remark on Algorithm 778: L-BFGS-B, FORTRAN routines for large scale bound constrained optimization. *ACM Trans. Math. Softw.* **2011**, *38*, 7.
(95) Swope, W. C.; Andersen, H. C.; Berens, P. H.; Wilson, K. R. A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters. *J. Chem. Phys.* **1982**, *76*, 637−649.
(96) Box, G. E. P.; Muller, M. E. A Note on the Generation of Random Normal Deviates. *Ann. Math. Stat.* **1958**, *29*, 610−611.
(97) Shannon, C. E. A Mathematical Theory of Communication. *Bell Syst. Tech. J.* **1948**, *27*, 379−423.
(98) Aldana-Bobadilla, E.; Kuri-Morales, A. A Clustering Method Based on the Maximum Entropy Principle. *Entropy* **2015**, *17*, 151−180.
(99) Andricioaei, I.; Straub, J. E. Global Optimization Using Bad Derivatives: Derivative-Free Method for Molecular Energy Minimization. *J. Comput. Chem.* **1998**, *19*, 1445−1455.
(100) Hoare, M. R. Structure and dynamics of simple microclusters. *Adv. Chem. Phys.* **2007**, *40*, 49−135.
(101) Ma, J.; Straub, J. E. Simulated annealing using the classical density distribution. *J. Chem. Phys.* **1994**, *101*, 533−541.
(102) Ma, J.; Straub, J. E. Erratum: Simulated annealing using the classical density distribution [J. Chem. Phys. 101, 533 (1994)]. *J. Chem. Phys.* **1995**, *103*, 9113.
(103) Northby, J. A. Structure and binding of Lennard-Jones clusters: $13 \leq N \leq 147$. *J. Chem. Phys.* **1987**, *87*, 6166−6177.
(104) Wales, D. J.; Doye, J. P. K.; Dullweber, A.; Hodges, M. P.; Clavo, F. Y. N. F.; Hernández-Rojas, J.; Middleton, T. F. The Cambridge Cluster Database. http://www-wales.ch.cam.ac.uk/CCD.html.

DOI: 10.1021/acs.jctc.8b00124
J. Chem. Theory Comput. 2018, 14, 3351−3362