Efficient discovery of multiple minimum action pathways using Gaussian process

JaeHwan Shim,† Juyong Lee,‡ and Jaejun Yu†mol

1Department of Physics and Astronomy, Seoul National University
1 Gwanak-ro, Gwanak-gu, Seoul 08826, Republic of Korea
2Department of Chemistry, Division of Chemistry and Biochemistry, Kangwon National University
1, Gangwondaehak-gil, Chuncheon-si, Gangwon-do 24341, Republic of Korea

We present a new efficient transition pathway search method based on the least action principle and the Gaussian process regression method. Most pathway search methods developed so far rely on string representations, which approximate a transition pathway by a series of slowly varying replicas of a system. Since those methods require a large number of replica images, they are computationally expensive in general. Our approach employs the Gaussian process regression method, which takes the Bayesian inference on the shape of a given potential energy surface with a few observed data and Gaussian-shaped kernel functions. Based on the inferred potential, we find multiple low-action pathways by carrying out the action optimization based on the Action-CSA (Conformational space annealing). Here we demonstrate a drastic elevation of computing efficiency about five orders of magnitude for the system with the Müller-Brown potential. Further, for the sake of demonstrating its real-world capabilities, we apply our method to ab initio calculations on alanine dipeptide. The improved efficiency of GPAO makes it possible to identify multiple transition pathways of alanine dipeptide and calculate their transition probabilities with ab initio accuracy. We are confident that our GPAO method is a powerful approach to investigate the mechanisms of complex chemical reactions

I. INTRODUCTION

Finding multiple transition pathways of phase transitions, chemical reactions, and conformational transitions remains challenging in physics and chemistry. In general, it takes a considerably long time to simulate an evolution over an energy barrier between reactants and products. Indeed, for some cases with large energy barriers, it is practically not feasible to carry out a proper simulation even on state-of-the-art computers. To overcome such a limitation, researchers have suggested numerous path sampling approaches [1,6]. Among various attempts, the methods based on the fixed-boundary-value formulation are commonly used. Unlike the initial-value formulation, which awaits a long time to observe rare events, the fixed-boundary-value formulation presumes a virtual trajectory and evaluates the pathways’ quality based on the height of an energy barrier or its action value.

Among many fixed boundary value approaches, several methods based on the least action principles were proposed and considered a promising way to find saddle points on a potential energy surface (PES) [5,7]. Passerone and Parrinello [5] suggested a “so-called” action-derived molecular dynamics (ADMD) method, which looks for low-action pathways via the local optimization of a modified classical action with an energy conservation restraint term. Later, Lee and coworkers [8] developed a kinetically controlled ADMD method by introducing a restraining term to satisfy the equipartition theorem. These classical-action-based methods have an inherent limitation that the classical principle of least action is an extremal, not a minimum principle. In other words, the classical action can be either minimized or maximized, which makes the computational outcome ambiguous.

Another class of pathway sampling approaches uses the Onsager-Machlup (OM) action [9] to obtain the most probable transition pathway of a diffuse system described by the Langevin equation [10–13]. For a diffuse system, the probability $p(x_B | x_A; t)$ for a state $x_A$ to become another state $x_B$ within a given time interval $t$, can be described as follows [9, 11, 14]:

$$\int_{x_A}^{x_B} Dx(s)e^{-\frac{1}{k_B T} S_{OM}(x(s))}$$ (1)

where $S_{OM}$ is the OM action and $D$ is the notation for integrating all the possible pathways connecting two end states. The probability of a transition from the state $x_A$ to $x_B$ can be calculated by summing up the probability of all pathways $\sum_s e^{-\frac{S_{OM}(x(s))}{k_B T}}$. The most dominant pathway may effectively be determined by the probability of that state transition. However, for most cases, a state transition usually incorporates multiple pathways, such as multiple protein folding pathways. Lee and coworkers utilized an efficient global optimization method to identify the most dominant transition pathway corresponding to the global minimum of OM action and sample multiple less prevalent pathways [13]. In their work, the locally optimized transition pathways were sampled by using the modified classical action to avoid the Hessian calculation of a potential function. Lee and Brooks tested the efficiency of the direct minimization of OM action by using the system’s Hessian information. They calculated the gradient of OM action analytically and directly minimized OM action [7]. The di-
rect optimization of the Onsager-Machlup action (DOA) method guarantees to locate the local minima of OM action. However, it did not preserve the total energy of the system.

A major bottleneck of the existing pathway search methods is that they require a significant amount of computational resources. Conventional pathway search methods optimize the transition pathway by calculating the energy and gradient of every configuration image. From this information, they obtain the optimal coordinates, which minimize a given action. However, the calculation of the energies and forces for multiple images increases the computational burden quite seriously. A small enough step size is necessary to generate continuous and smooth trajectories. However, the use of a smaller step size obviously increases computational costs. In other words, there is a trade-off relationship between the accuracy and computational efficiency of trajectories. For instance, to achieve the continuity of a trajectory, Passerone and Parrinello [5] used 300 images to find trajectories on a simple 2D Müller-Brown (MB) potential.

As a way to get around the computational cost issue, Jónsson and coworkers [16] suggested the one-image evaluation (OIE) method incorporating the Gaussian process (GP) [15] with a nudged elastic band (NEB) [4] to reduce the number of energy and force calculations. The OIE method finds the minimum energy pathway (MEP) by building a Gaussian PES that approximates the actual PES iteratively. The following steps construct the approximate Gaussian PES. First, a posterior distribution is used to evaluate the most uncertain image along the pathway. Second, the most uncertain image is evaluated with actual energy calculation and then added the value to a database to improve the Gaussian PES’s accuracy. Third, on the newly constructed Gaussian PES, the MEP is updated to fit the new Gaussian PES using NEB. After the relaxation of NEB, the whole process is repeated until the level of uncertainty decreases below a certain threshold. With the MB potential, the number of function calls is effectively reduced by two orders of magnitude [17].

Another major limitation of existing pathway search methods, such as NEB, is that they are local optimization methods whose final results heavily depend on the initial guesses on pathways. Generally, conventional pathway search methods start from the linear interpolation between end states and perform local optimization of the initial pathway. These approaches can find the correct minimum energy or action pathways only when a given PES is simple and smooth. However, many complex systems have highly rugged PES, which does not guarantee to locate the most dominant pathway near the linear interpolation between two states. To avoid this problem, an extensive search on a pathway space is essential.

In this work, to overcome these limitations, we present a novel transition-pathway-sampling method by combining the state-of-the-art machine learning techniques and the advantages of both Passerone’s and Lee’s methods. We also carry out direct OM action optimization with a total-energy-conservation restraint without using the original system’s Hessian. One of the core algorithms in our method, Gaussian process action optimization (GPAO), shares a similar principle with the OIE method. To avoid calculations of all the intermediate images, we calculate only the most uncertain images as measured by the variance via GP and then add the calculation result to the database. The potential energies, forces, and Hessian of the intermediate images are estimated using the Gaussian PES. Within the Gaussian PES, the MEP is directly optimized so as to find its minimum action. Also, GPAO samples multiple sub-optimal pathways, local minima of action, because it uses the identical sampling method with Action-CSA [13].

In this work, we also implement a routine that automatically finds suitable target energy throughout the iteration and combines the GP with the conventional ADMD or DOA methods. Target energy is inherently hard to know prior to an actual pathway sampling because proper target energy is closely related to the height of the energy barrier of the MEP. Here we utilize a mechanism that approximates the target energy to the maximum potential energy of the temporal MEP throughout the iteration. As a result, the computation efficiency of our transition-pathway-sampling method shows a dramatic enhancement. The conventional ADMD requires, even for the simple MB example, given with suitable target energy, 1,200,000 calls of force calculations to find the MEP.

On the other hand, GPAO with classical action, without knowledge of the target energy, reduces the number of 1,200,000 actual function calls down to 12. Encouraged by the significant computational efficiency gain by GPAO in five orders of magnitude, we take a step forward to undertake a computationally challenging example: finding multiple transition paths of alanine dipeptide through ab initio calculations. Action-CSA’s capability, which preserves the diversity of distinct pathways while optimizing the OM action globally and efficiently, results in multiple physically accessible trajectories between the two end-states. To the best of our knowledge, our work is the first report, which identifies multiple conformational transition pathways of alanine dipeptide with an ab initio potential.

II. THEORY AND METHOD

A. Gaussian Process

Significant computational improvements of GPAO over the conventional ADMD or DOA methods are attributed to the ability to quantify the variance of specific images in a pathway by using the Bayesian inference. Let us suppose a pathway consisting of $N$ images and $M$ evaluated data points. In our GP model, the probability distribution of a given string of images, $X = \{X_1, X_2, \ldots, X_N\}$,
action with energy conservation restraint, $\Theta_{\text{cls}}$, is,

$$\Theta_{\text{cls}} (\{\mathbf{X}_\alpha\}, E_t) = S_{\text{cls}} + \mu_E \sum_{n=0}^{N-1} \left(E^{(n)} - E_t\right)^2$$

where $\mu_E$ is the coefficient of the total energy conservation restraint term, $E_t$ is the target total energy of the system. The classical action is defined as,

$$S_{\text{cls}} = \sum_{n=0}^{N-1} dt \left[ \frac{1}{2} \sum_{a=0}^{A-1} m_a \left( \mathbf{x}_a^{(n)} - \mathbf{x}_a^{(n+1)}\right)^2 dt^2 - V \left( \mathbf{x}^{(n)} \right) \right]$$

Total energy at the time step $n$ is defined as follows:

$$E^{(n)} = \frac{1}{2} \sum_{a=0}^{A-1} m_a \left( \mathbf{x}_a^{(n)} - \mathbf{x}_a^{(n+1)}\right)^2 + V \left( \mathbf{x}^{(n)} \right)$$

To minimize the action analytically, the derivatives of action are necessary. The derivatives of classical action and the restraint terms with respect to the coordinates of the $n$th image are,

$$\frac{\partial S_{\text{cls}}}{\partial \mathbf{x}^{(n)}} = dt \sum_{a=0}^{A-1} \left( - \frac{\partial}{\partial \mathbf{x}_a^{(n)}} V \left( \mathbf{x}^{(n)} \right) + m_a a_a^{(n)} \right)$$

$$\frac{\partial}{\partial \mathbf{x}^{(n)}} \left( \mu_E \sum_{n=0}^{N-1} \left(E^{(n)} - E_t\right)^2 \right)$$

$$= 2\mu_E dt \sum_{a=0}^{A-1} \left[ \left( \frac{\partial}{\partial \mathbf{x}_a^{(n)}} V \left( \mathbf{x}^{(n)} \right) + \frac{1}{dt} p_a^{(n)} \right) \left( -E_t + V \left( \mathbf{x}^{(n)} \right) + \frac{1}{2} m_a v_a^{(n)T} \cdot v_a^{(n)} \right) \right.$$ 

$$+ \frac{1}{dt} p_a^{(n-1)} \left( -E_t + V \left( \mathbf{x}^{(n-1)} \right) + \frac{1}{2} m_a v_a^{(n-1)T} \cdot v_a^{(n-1)} \right) \right]$$

where $v_a^{(n)} = (\mathbf{x}_a^{(n+1)} - \mathbf{x}_a^{(n)}) / dt, a_a^{(n)} = (2\mathbf{x}_a^{(n)} - \mathbf{x}_a^{(n-1)} - \mathbf{x}_a^{(n+1)}) / dt^2$ and $p_a^{(n)} = m_a v_a^{(n)}$. Our parameter set for the classical action on the MB potential consists of $dt = 0.02, m_a = 1, \mu_E = 0.1, E_t = -0.45$ and $N = 300$.

If a system propagates according to the Langevin dynamics, the dynamics of the diffusive system can be described
by the OM action \[\Sigma_{\text{OM}}\]. The discretized formula of the OM action and its derivatives are given by,

\[
S_{\text{OM}} = \frac{\Delta V}{2} + \frac{1}{4} \sum_{n=0}^{N} \left[ \frac{dt}{2\gamma} \left( |\nabla V(x^{(n+1)})|^2 + |\nabla V(x^{(n)})|^2 \right) 
- \left( \nabla V(x^{(n+1)}) - \nabla V(x^{(n)}) \right) \cdot v^{(n)} + \frac{\gamma}{dt} v^{(n)} T \cdot v^{(n)} \right]
\]

where \(\partial \nabla V(x^{(n)}) / \partial x^{(n)}\) is the Hessian of the potential, \(\Delta V\) is the energy difference between the initial and the final step, and \(\gamma\) is a damping constant. We used \(dt = 0.02\), \(N = 300\) and \(\gamma = 1\) (unit time)\(^{-1}\) for the MB potential.

The method we propose in this paper imposes the total energy conservation restraint on the OM action, that is,

\[
\Theta_{\text{OM}}(\{x_n\}, E_t) = S_{\text{OM}} + \mu_E \sum_{n=0}^{N-1} \left( E^{(n)} - E_t \right)^2 \quad (12)
\]

We used \(dt = 0.02\), \(N = 300\), \(\gamma = 1\) (unit time)\(^{-1}\), \(\mu_E = 0.1\) and \(E_t = -0.45\) for the MB potential example. The BFGS algorithm in the scipy Python library is used to minimize the action and the minimization is conducted on the sine components of the actual coordinates. That is, we searched \(\tilde{x}^{(k)}\):

\[
\tilde{x}^{(k)} = 2 \sum_{n=0}^{N-2} x^{(n)} \sin \left( \frac{\pi (2k+1)(n+1)}{2N} \right) + (-1)^k x^{(N-1)}
\]

that minimizes a given action.

### C. Action minimization using Gaussian process based on full atom coordinates

Our GPAO method handles two tasks simultaneously. First, it minimizes the uncertainty of a Gaussian PES near the true MEP in an iterative way. Second, the method determines the total energy \(E_t\) of the MEP automatically. In this study, we assumed that the \(E_t\) of a system is close to the maximum potential energy along with the MEP, \(V^{(\max)}\). To gradually approximate the \(E_t\) to \(V^{(\max)}\), we approximated \(E_t\) as \(\left( \mu^{(\max)} - E_t - 0.05 \right) / 2\), where \(\mu^{(\max)}\) is the estimated maximum potential energy of MEP on the Gaussian PES.

How the GPAO method minimizes the MEP’s uncertainty and approximates the \(E_t\) to \(-0.45\) on the MB potential is illustrated in Fig. 1. In Fig. 1(a), plotted are the Gaussian PES and \(\mu\), which are constructed from the evaluated data points marked as \(\times\). In the middle plots of Fig. 1(b), the uncertainty maps, \(\Sigma\), are plotted with the trajectory indicating the most uncertain point at each iteration, \(\Sigma^{(\text{max})}\). Potential energy and force data of the most uncertain point are added to the database, and the Gaussian PES near the trajectory is refined close to the true PES. The bottom graph of Fig. 1(c), shows the potential energy \(V\), total energy \(H\), kinetic energy \(E_k\), and target energy \(E_t\) as a function of distance. The detailed procedure of GPAO is given as follows:

1. Find the most uncertain configuration along the pathway. If the maximum uncertainty is lower than certain tolerance, find the configuration having a maximum potential energy
2. Calculate the energy and force of the configuration obtained at step 1 and add them to the database
3. Tune the hyperparameters using the GP regression algorithm.
4. Adopting a new hyperparameter set and data, construct a new Gaussian PES
5. Set \(E_t \leftarrow \left( \mu^{(\max)} - E_t - 0.05 \right) / 2\)
6. With new target energy and a Gaussian PES, optimize the given action
7. Check the convergence of uncertainty, action convergence, and the maximum energy difference to real data. If any of the three conditions do not meet the requirement, reiterate from step 1. Otherwise, it is considered converged.

Fig. 1(a) and (b) show how Gaussian PES is gradually approximated to the true PES (Fig. 1(a)) as data accumulates throughout the iteration. At iteration 1, the pathway is optimized with the three data points: the initial, the final, and a random intermediate states between two. Due to the lack of data points, potentials estimated along the trajectory are still uncertain. Fig. 1(c) illustrates that the uncertain range (light blue intervals) of the estimation at iteration 1 is much wider than the one
FIG. 1. Demonstrative images of GPAO process. The expectation value of the Gaussian PES map, uncertainty map, and energy graph along the pathway are plotted from top to bottom. From left to right, map, and graph at iteration 1, 2, 3, and 9. The red line at the Gaussian PES indicates trajectory, the black cross position where data acquired and the white arrows are force data at that coordinates. On the uncertainty map, we marked additional red crosses corresponding to the latest position where data being acquired in addition to the white crosses marked at the identical locations on all Gaussian PESs. In the energy graph, the blue line shows potential energy, sky blue means 95% confidence range of potential energy, the red dot indicates kinetic energy, the orange dashed line is the total energy and the green diamond line expresses target energy.

at iteration 9. Furthermore, Fig. 1(b) shows that maximum uncertainty at iteration 1 ($\Sigma^{(\text{max})} = 5.15 \times 10^{-1}$) is smaller than that at iteration 9 ($\Sigma^{(\text{max})} = 9.47 \times 10^{-2}$). At iteration 2, including the most uncertain data point of the previous iteration (marked with a red cross in Fig. 1(b)), the pathway is optimized with four data points. Most uncertain points in the pathway are marked with red crosses in Fig. 1(b) at iteration 2, which are added to the next iteration database. At iteration 9, the pathway is optimized with 11 data points and $\Sigma^{(\text{max})} < 0.05$.

Fig. 1(c) shows how GPAO approximates the $E_t$ to $V^{(\text{max})}$ iteratively. At iteration 1, $E_t$ was set to $-1.47$, which is the minimum energy between the initial and the final image. GPAO had $\mu^{(\text{max})} = -0.101$ from the trajectory optimized with three data points. Using this information, at iteration 2, $E_t$ was raised to $-0.81$, which is a half of $E_t - \mu^{(\text{max})}$ of the previous iteration. GPAO resulted in $\mu^{(\text{max})}$ of $-0.20$ with four data points. Similarly, at iteration 3, $E_t$ was set to half of $E_t - \mu^{(\text{max})}$ of the previous iteration, $-0.368$.

Additional information of the trajectory, SOM, $E_t$, $\mu^{(\text{max})}$, $\Sigma^{(\text{max})} = 1.94 \times \sqrt{\text{max} diag(\Sigma_t)}$ and the hyperparameters of the kernels $l^2, \sigma_n^E, \sigma_n^f$ of each iteration are displayed on Fig. 1(a) and Fig. 1(b). $\sigma_n^E$ and $\sigma_n^f$ are noise parameters corresponding to the energy and forces of data points, respectively. The parameter set for the action was assigned to the same as the conventional ADMD with $\Theta_{\text{OM}}$ calculation.

We used the squared exponential kernel as the Gaussian kernel for the MB model:

$$k \left( \mathbf{x}^{(i)}, \mathbf{x}^{(j)} \right) = \sigma_f \exp \left( -\frac{1}{2 l^2} \left( \mathbf{x}^{(i)} - \mathbf{x}^{(j)} \right)^T \left( \mathbf{x}^{(i)} - \mathbf{x}^{(j)} \right) \right)$$

(14)

where $l$ and $\sigma_f$ are isotropic real hyperparameters. We used the multivariate Gaussian distribution with zero mean for our GP regression model, $\mathbf{m} = \mathbf{m}_s = 0$. 
D. Action optimization using the Gaussian process and reduced coordinates

![Alanine Dipeptide PES (eV)](image)

**FIG. 2.** (a) The PES of alanine dipeptide as a function of reduced coordinates, φ and ψ angles, V(ψ, φ). (b) The C7eq and C7ax conformations of alanine dipeptide are drawn. The brown balls correspond to carbon, the light blue ones are nitrogen, the red ones are oxygen, and the ivory ones are hydrogen.

Alanine dipeptide is a small molecule consisting of 22 atoms, and it has been widely used as a test system for computational methods in biophysics. As displayed in Fig. 2, alanine dipeptide has two rotatable bonds associated with two dihedral angles, the φ and ψ angles. The φ angle is a dihedral angle between C−1 NCα C, and the ψ angle is defined as a dihedral angle between NCα CN+1. These two dihedral angles almost exclusively determine alanine dipeptide’s conformation because the bond angles and bond lengths are much more rigid than dihedral angles. Alanine dipeptide has two stable conformations, C7eq and C7ax.

The PES’s shape shows that multiple conformational transition pathways from one conformation to the other are possible (Fig. 2). Lee and coworkers discovered all possible conformational transition pathways through the action optimization using the conformation space annealing (Action-CSA) method.

The potential energies and forces of alanine dipeptide are calculated using the ab initio based density-functional-theory package VASP. The constrained MD method of VASP is used to optimize the conformations of C7ax and C7eq with fixed φ and ψ angles. With fixed both φ and ψ dihedral angles, MD simulations are performed with the Anderson thermostat. With random initial velocities, the heat bath temperature decreased from 80 K to 0 K for 500 steps. The probability that a system interacts with the heat bath was given to 5%. We use the generalized gradient approximation (GGA) exchanged functional and a van der Waals scheme considering three-body interactions. The energy cutoff is set to 520 eV, and a convergence rate for the SCF scheme is 10−5 eV. The system is slowly annealed to a stable structure with the fixed φ and ψ angles.

Since the structures of alanine dipeptide are almost exclusively defined by φ and ψ angles, its potential can be expressed as the function of them as follows: $V(\{\mathbf{x}^{(n)}\}) = V(\phi^{(n)}, \psi^{(n)})$. This reduction on coordinates $\mathbf{x}^{(n)} = [\phi^{(n)}, \psi^{(n)}]$.$\mathbf{v}^{(n)} = [\phi^{(n+1)} - \phi^{(n)}, \psi^{(n+1)} - \psi^{(n)}]$ is extended on the formulation of the modified OM action for the alanine dipeptide system. Our $\Theta_{OM}$ for the alanine dipeptide system is,
\[
\Theta_{OM} = S_{OM} + \mu_E \sum_{n=0}^{N-1} \left( E^{(n)} - E_i \right)^2 \\
= \frac{\Delta V}{2} + \frac{1}{4} \sum_{n=0}^{N} \left[ \frac{dt}{2\gamma} \left( \frac{1}{I^{(n)}_{\phi}(\phi, \psi)} + \frac{1}{I^{(n)}_{\psi}(\phi, \psi)} \right) \left| \nabla V \left( \left\{ x^{(n+1)} \right\} \right) \right|^2 + \left| \nabla V \left( \left\{ x^{(n)} \right\} \right) \right|^2 \right] \\
- \left( \nabla V \left( \left\{ x^{(n+1)} \right\} \right) - \nabla V \left( \left\{ x^{(n)} \right\} \right) \right) \cdot \psi^{(n)} \\
+ \frac{\gamma}{dt} \left( I^{(n)}_{\phi}(\phi, \psi) \left( \frac{\phi^{(n+1)} - \phi^{(n)}}{dt} \right)^2 + I^{(n)}_{\psi}(\phi, \psi) \left( \frac{\psi^{(n+1)} - \psi^{(n)}}{dt} \right)^2 \right) \\
+ \mu_E \sum_{n=0}^{N-1} \left( \frac{1}{2} I^{(n)}_{\phi}(\phi, \psi) \left( \frac{\phi^{(n+1)} - \phi^{(n)}}{dt} \right)^2 + \frac{1}{2} I^{(n)}_{\psi}(\phi, \psi) \left( \frac{\psi^{(n+1)} - \psi^{(n)}}{dt} \right)^2 \right) \\
+ V \left( \left\{ x^{(n)} \right\} \right) - E_i \right)^2 
\]

where \( \nabla V \left( \left\{ x^{(n)} \right\} \right) = \left[ \partial_\phi V \left( \left\{ x^{(n)} \right\} \right), \partial_\psi V \left( \left\{ x^{(n)} \right\} \right) \right] \)

indicates the potential difference with respect to \( \phi \) and \( \psi \), \( I^{(n)}_{\phi}(\phi, \psi) \) and \( I^{(n)}_{\psi}(\phi, \psi) \) indicate the moments of inertia associated with \( \phi \) and \( \psi \) angles of \( n \)th image, respectively.

In case we obtain \( \nabla V \left( \left\{ x^{(n)} \right\} \right) \) with full atomic coordinates, we map them onto the reduced coordinates by projecting each force vector to a corresponding rotation axis. The average values of \( I_\phi \) and \( I_\psi \) are 112.96 uÅ² and 75.93 uÅ², respectively. The derivatives of the moments of inertia, \( \partial_\phi I_\phi \), \( \partial_\psi I_\phi \), \( \partial_\phi I_\psi \), and \( \partial_\psi I_\psi \) are calculated with VASP using the data we used for generating the PES map. Except for high energy regions, where \( \phi \approx 0^\circ \) and \( \psi \approx 180^\circ \), the changes in the moment of inertia are negligible. About 99% of the data show that \( |\partial_\phi I_\psi| \) is less than 2.211, which is less than 2%, and 95% of the data show \( |\partial_\psi I_\phi| < 0.727 \), less than 0.6%. Similarly, 99% of \( |\partial_\phi I_\phi| \), \( |\partial_\psi I_\phi| \), and \( |\partial_\psi I_\psi| \) are less than 2.702 uÅ², 0.572 uÅ², and 1.258 uÅ², respectively. About 95% of the data are less than 0.948 uÅ², 0.388 uÅ², and 0.540 uÅ². We assume that the changes in the moments of inertia, \( \partial_\phi I_\phi \), \( \partial_\psi I_\phi \), \( \partial_\phi I_\psi \), and \( \partial_\psi I_\psi \) are negligible because a trajectory rarely visits high energy regions where the moment of inertia may change rapidly. We use the GP to estimate the moment of inertia with the same database to construct the Gaussian PES. The action for the relaxation was \( \Theta_{OM} \) with parameters \( \gamma = 0.001 \left( \sqrt{\text{Å}^2/\text{eV}} \right)^{-1}, \) \( dt = 0.1, 0.2, 0.3 \left( \sqrt{\text{Å}/\text{eV}} \right) \) and \( \mu_E = 0.1 \).

### E. Efficient global action optimization using Action-CSA

We combined GPAO with the Action-CSA method [13] to search multiple transition pathways because stochastic reactions occur through multiple pathways in general. In this work, the Action-CSA is set to start with randomly generated 40 initial pathways, followed by local optimizations using GPAO. Afterward, we proceed with the following steps:

1. Pick two pathways from the candidates randomly.
2. Generate a new pathway from two pathways using a crossover operator.
3. Apply random mutations/perturbations to the new pathway with a certain probability.
4. Locally minimize the new pathway.
5. Pick the most similar pathway to the new pathway from the candidates. If the distance between the new pathway and a similar pathway is lower than a cutoff distance, compare the OM action between them. Otherwise, compare the highest OM pathway in the candidates.
6. If a new one has lower action than its counterpart, replace it with the new one.
7. Reduce the cutoff distance. If the cutoff distance is lower than a certain threshold, stop the iteration. Otherwise, go back to step 1.

With two randomly selected pathways, we choose a random image along a pathway. With the selected position, we slice both trajectories at the selected image and merge them. After a new pathway is generated, a random mutation/perturbation is performed with a probability of 30%. For the pathway’s mutation, the pathway is modified by adding another random number at the sine components of a pathway. That is, the \( \tilde{x}^{(k)} \) in the Eq. (13) is altered. After the local minimization of the pathway using the GPAO method, we search the nearest neighbor by measuring the Fréchet distance between the new pathway and the current candidates. The nearest neighbor is considered similar, readily accessible to each other if the Fréchet distance is shorter than the current cutoff distance \( d_{cut} \).

The initial cutoff distance is set to half of the average distance between the initial pathways. After an iteration,
we reduce $d_{\text{cut}}$ by 2% and repeat the whole procedure from step 1 until $d_{\text{cut}}$ becomes below 0.05.

III. RESULT AND DISCUSSION

A. Action minimization on the Müller-Brown potential

![Muller Brown](image)

**FIG. 3.** A comparison of the minimum action pathways obtained with classical action with energy conservation restraint ($S_{\text{cls}}$, blue dashed line), OM action ($S_{\text{OM}}$, orange dash-dotted line), and OM action with energy conservation constraint ($S_{\text{OM}}^\text{OM}$, green solid line). (a) the minimum action pathways between two stable potential energy minima on the Müller-Brown potential and a zoomed image near the saddle point. (b) energy profile.

To compare the characteristics of pathways obtained with different actions, we carry out action optimization calculations on the MB potential, defined as follows:

$$V(x, y) = \sum_{\mu=1}^{4} A_{\mu} e^{\alpha_{\mu}(x-x_{\mu}^{0})^2+b_{\mu}(x-x_{\mu}^{0})(y-y_{\mu}^{0})+c_{\mu}(y-y_{\mu}^{0})^2},$$

where $x_{\mu}^{0} = (1, 0, -0.5, -1)$, $y_{\mu}^{0} = (0, 0.5, 1.5, 1)$, $\alpha_{\mu} = (-1, -1, -6.5, 0.7)$, $b_{\mu} = (0, 0, 11, 0.6)$, $c_{\mu} = (-10, -10, -6.5, 0.7)$, and $A_{\mu} = (-2, -1, -1.7, 0.15)$.

The result of direct action optimization on the MB potentials demonstrates that $S_{\text{OM}}^\text{OM}$, the pathway with the minimum modified OM action $\Theta_{\text{OM}}$, successfully conserves the system’s total energy throughout the pathway (Fig. 3). The magnitude of total energy fluctuations of the pathway was reduced by 98.7% compared to that of $S_{\text{OM}}^\text{OM}$, the pathway optimized using $S_{\text{OM}}$. The OM value of $S_{\text{OM}}^\text{OM}$ increased only about 19%, and the Fréchet distance to $S_{\text{OM}}^\text{OM}$ is less than 0.067 (Table I). This result demonstrates that the total energy conservation term does not change the shape of the pathway significantly while keeping the total energy of the system constant.

Fig. 3(b) shows the total energy profile of three trajectories and their energy gaps, the differences between the maximum and the minimum total energies of pathways. The energy gaps of $C_{\Theta}^\text{OM}$, $C_{S}^\text{OM}$, and $C_{\Theta}^\text{OM}$ are 0.09, 3.85, and 0.05, respectively. The large spikes of the total energy occurred near the region where the trajectory climbs up the energy barriers of the PES between the timesteps 0.8 1.1.

One reason for this huge spike of the total energy is the second term of the $S_{\text{OM}}$ equation (Eq. (10)). The first term in Eq. (10) prefers to minimize atom’s forces, and the third term keeps the velocities between images slow. The second term in Eq. (10) favors following a concave potential energy surface along its ongoing direction. To be more specific, if a pathway has to climb up a PES during a transition, it prefers to climb up the surface fast. This tendency is clearly manifested in Fig. 3(a). Right before $S_{\text{OM}}$ reaches the saddle point with the maximum potential energy of the PES, the images are largely separated because of large velocities of images.

In contrast, the images of $C_{\Theta}^\text{OM}$ near the saddle point are almost continuous and uniformly distributed. This uniform distance between images is because the total energy spike is mainly due to the second term in Eq. (10). Our results demonstrate that unphysical increases of atomic velocities to lower the second term can be suppressed effectively by the energy conservation restraint without changing the trajectory direction.

Our method avoids the occurrence of unphysical energy spikes observed with $S_{\text{OM}}$, but still maintains $S_{\text{OM}}$ low. In Table I $S_{\text{OM}}$ values of $C_{\Theta}^\text{OM}$, $C_{\Theta}^\text{OM}$ and $C_{\Theta}^\text{OM}$ are 14.49, 1.387, and 1.652, respectively. The potential energies at the saddle point, $V^{(\text{max})}$, of $C_{\Theta}^\text{OM}$, $C_{\Theta}^\text{OM}$, and $C_{\Theta}^\text{OM}$ are $-4.00$, $-0.406$, and $-0.406$, respectively. The results indicate that the enforcement of the energy conservation restraint on $S_{\text{OM}}$ gives only little effect on finding low $V^{(\text{max})}$ and $S_{\text{OM}}$. Furthermore, Fig. 3(a) shows how two pathways of $C_{\Theta}^\text{OM}$ and $C_{\Theta}^\text{OM}$ overlap each other. The Fréchet distance between two trajectories is only 0.067. Also, 95% of the intermediate images of $C_{\Theta}^\text{OM}$ are located within a distance of 0.013 from $C_{\Theta}^\text{OM}$.

B. Efficiency of GPAO

The most significant advantage of the GPAO methods is a dramatic reduction of the number of force calls by four-to-five orders of magnitude than the conventional action optimization method while preserving their accuracy (Table I). The numbers of force calls to find $C_{\Theta}^\text{OM}$,
The significant reduction of the computational cost by the GPAO method allows us to find transition pathways of a molecule with a quantum-mechanical (QM) potential. As discussed in the next section, a single force calculation of alanine dipeptide with VASP requires around 1741.4 CPU seconds. Using our method, the optimal transition pathway of alanine dipeptide within a single day. However, it will take five years to complete the same calculation by the conventional action optimization method and the QM potential.

The significant computational gain is attainable because the computation cost of GPAO is independent of the number of images and atoms. The number of potential energy evaluations of GPAO depends only on the convergence of a Gaussian PES. In contrast to GPAO, the calculation cost of the conventional action optimization method, such as ADM and DOA, is strictly proportional to the number of images in a pathway. To ensure the continuity of a pathway, we followed the same number of steps sizes, 300 images for a pathway, with the previous study on the MB potential. Thus, for a single action evaluation, the energy and forces of 300 images must be calculated.

On the contrary, the number of force calls required for the GPAO method depends on the number of data points, which should be large enough to accurately approximate the PES. Overall, the GPAO required similar numbers of force calls regardless of the kinds of action (Table I). In addition, the results of the GP-assisted NEB methods, GP-C<sup>N</sup>NEB, showed similar numbers of force calls to reproduce the result of the conventional NEB method. Indeed, it turns out that the number of data points required to find an accurate MEP on the MB potential is also in a range between 10 and 20. These results imply that the type of local relaxation method does not significantly affect the efficiency of the GPAO methods.

Our GPAO methods lead to the transition pathways remarkably close to those obtained by the conventional action optimization method. Overall, the discrepancy between the pathways obtained with the two classes of the methods is less than 0.08% of the pathways’ total distances (Table II). The distance between C<sup>G</sup> and C<sup>OM</sup> was 0.006. Similarly, the distances between C<sup>G</sup> and C<sup>OM</sup> and the corresponding pathways obtained by the GPAO method were 0.082 and 0.042, respectively. In general, the results of all five GP-assisted methods show a deviation of less than a Fréchet distance of 0.1 compared to the conventional action optimization result. Even further, the GP-C<sup>N</sup>NEB results deviate from the conventional result less than a Fréchet distance of 0.01. These results support that approximation of PES by GP is accurate enough to locate the transition pathways of low action on a given PES.

GPAO gives us a way to improve the Hessian’s accuracy by accumulating data points, which allows us to utilize calculation tools that are hard to acquire.
sian. The conventional way of obtaining the Hessian of a DFT potential is using a finite difference method. That is, to acquire a Hessian at an arbitrary point \( \phi^{(n)}, \psi^{(n)} \), we need to compute \( \nabla V(\phi^{(n)}, \psi^{(n)}), \nabla V(\phi^{(n)} + \varepsilon, \psi^{(n)}) \) and \( \nabla V(\phi^{(n)}, \psi^{(n)} + \varepsilon) \). These two additional force calls triples the computation cost to evaluate \( \Theta_{OM} \), which prevents the use of finite difference methods from being applied to relatively large systems. Thus, such a fast evaluation of a potential enables us to use Hessian information for practical problems.

C. Multiple transition pathways of alanine dipeptide

The PES of alanine dipeptide is periodic because two dihedral angles, \( \phi \) and \( \psi \), are chosen as the reaction coordinates. We use the following periodic kernel to describe \( \phi, \psi \) configurations. This result is consistent with previous studies conducted with molecular mechanics potentials [13, 22].

\[
k \left( x^{(m)}, x^{(n)} \right) = \sigma_f \exp \left( -\frac{2}{l^2} \sum_{d=1}^{D} \sin^2 \left( \frac{x^{(m)}_d - x^{(n)}_d}{2} \right) \right)
\]

where \( x = (\phi, \psi) \) is a two-dimensional reaction coordinate, \( l, \sigma_f \) are isotropic hyperparameters.

We sampled multiple transition pathways of alanine dipeptide using GP-CSA calculations with three different total transition times, \( T_{total} = N dt = 3, 6, \) and 9 ps. Among the pathways obtained with GP-CSA, the distinctive pathways having low action values are plotted in Fig. 4. Path A in Fig. 4(b) has the lowest \( S_{OM} \) value for all transition times, suggesting that Path A is the most dominant pathway between \( C7_{eq} \) and \( C7_{ax} \) conformations. This result is consistent with previous studies conducted with molecular mechanics potentials [13, 22]. The maximum potential energy of a pathway estimated with GP, \( \mu^{(max)} \) is near -128.48 eV (Table II). For all three different \( T_{total} \), our method consistently located the correct saddle points.

The \( S_{OM} \) value of Path D with a \( T_{total} \) of 3 ps is 11.282 eV, which is high enough to make the path improbable. However, the action value decreases drastically to 5.551 eV and 4.887 eV when \( T_{total} \) becomes 6 ps and 9 ps. As \( T_{total} \) increases, Path D becomes longer, which allows the pathway to detour the high potential region and to seek a lower valley of the PES (see Path D in Fig. 4(b), (c)). This tendency is also observed in the \( \mu^{(max)} \) values in Table II. \( \mu^{(max)} \) of Path D at 3 ps is 0.2 eV higher than the other transition times.

It is noticeable that the relatively longer pathways, i.e., Path E, F, and G, were not observed in simulations with \( T_{total} = 3 \) ps. As \( T_{total} \) becomes longer, GPAO can access additional pathways, which require longer \( T_{total} \) to occur. \( S_{OM} \) values of Path E obtained with \( T_{total} = 6 \) ps and 9 ps were 8.547 eV and 6.745 eV, and \( \mu^{(max)} \) values were \(-129.237 \) eV and \(-129.316 \) eV, respectively. The maximum potential energy at the saddle point of Path
E of $T_{\text{total}} = 6$ ps was 0.1 eV higher than Path E with $T_{\text{total}} = 9$ ps. This difference arises because Path E with $T_{\text{total}} = 6$ ps did not pass the saddle point due to a short transition time. However, Path E with $T_{\text{total}} = 9$ ps correctly passed the saddle point (see E in the Fig. 3(c)).

With $T_{\text{total}} = 6$ ps, we find two additional pathways, Path F and Path G, similar to the most dominant pathway, Path A. These two pathways are longer pathways with higher $S_{\text{OM}}$ values, 4.56 eV and 5.293 eV, that of Path A. These results clearly indicate that our GPAO method can accurately find multiple transition pathways of the conformational change of alanine dipeptide with a QM potential, which have not been reported previously.

IV. CONCLUSION

In this study, we demonstrate that the use of the GP algorithm dramatically enhances the efficiency of searching minimum action pathways between potential energy minima by the accurate approximation of PES with much-reduced numbers of energy evaluations. Our results imply that pathway search using the modified OM action with a total energy conservation restraint is a prominent approach. Compared to the sole optimization of OM action, large fluctuations of kinetic energies are effectively removed, without affecting the OM action of a pathway.

OM action gradient requires the Hessian of a potential function, which is computationally expensive to obtain in general. Combining GP with the action optimization method gives not only superior computational efficiency but also the gradients of OM action without actual Hessian calculations. This advantage is particularly important for potential energies whose Hessians are challenging to obtain. For example, the Hessian of DFT-based potentials is often incorrect. However, with GPAO, the accumulation of gradient information at multiple data points leads to accurate estimation of the Hessian of a PES from a Gaussian PES.

Since the Hessian of a Gaussian PES is analytically given, we can achieve a way to obtain the OM action of DFT-based potentials and enhance computational efficiency by five orders of magnitude. Additionally, we also showed that our GPAO methods find suitable external parameters, such as target energy, automatically.

This study also presents a new multiple-pathway-sampling method by combining the Action-CSA method [12] with the GPAO methods. Due to the GPAO method’s efficiency, multiple conformational transition pathways between the two stable conformations of alanine dipeptide can be successfully obtained using a DFT-based potential. This is the first work that sampled multiple low action pathways of alanine dipeptide with a DFT potential to the best of our knowledge.

Undertaking higher-dimensional problems in an efficient way is still a big challenge in machine learning fields and others. This paper proves that our GPAO method properly works for small-dimensional problems. However, the method does not guarantee its efficiency on problems with a large degrees of freedom yet. Fortunately, many methods have been suggested to solve problems on a high dimensional space using various types of descriptors [23, 24]. Finding reaction pathways of the systems with many degrees of freedom can be applied to various problems such as finding drug binding pathways, battery, and protein folding. Thus, we believe that the GPAO methods will open up new possibilities for such problems.

ACKNOWLEDGMENTS

JL acknowledges the support of the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (No. 2018R1C1B600543513). JY acknowledges the support by the National Research Foundation of Korea (NRF)(No. 2020R1F1A1066548). Additional financial support from Samsung Electronics is also acknowledged. This work was also supported by the National Supercomputing Center with supercomputing resources and technical supports (KSC-2020-CRE-0026).

V. DATA AVAILABILITY

The source code of GPAO can be found in the GitHub repository (https://github.com/schinavro/taps). GPAO consists of multiple tools in the open-source package TAPS. TAPS incorporates tools needed for data-driven pathway search methods such as a database or atomic calculator. For atomic calculations, we use the atomic simulation environment (ASE) [27] package, which can easily link with atomic calculators such as VASP [20, 21], QUANTUMESPRESSO [28], OPENMX [29], and SIESTA [30].

[1] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, Optimization by simulated annealing, Science 220, 671 (1983)
[2] B. H. Good, Y.-A. de Montjoye, and A. Clauset, Performance of modularity maximization in practical contexts, Physical Review E 81, 046106 (2010)
[3] R. Guimerá and L. A. Nunes Amaral, Functional cartography of complex metabolic networks, Nature 433, 895 (2005)
[4] H. Jónsson, G. Mills, and K. W. Jacobsen, Nudged elastic band method for finding minimum energy paths of transitions, in Classical and Quantum Dynamics in Condensed Phase Simulations (World Scientific, Singapore, 1998) pp. 385–404.
