METADYNAMICS: EXPLORING POTENTIAL ENERGY LANDSCAPES THROUGH LANGEVIN DYNAMICS WITH THE MILSTEIN SCHEME

Bachelor Final Project

Nakshatra Rishi 1578758

Supervisor Björn Baumeier

Final Report

Eindhoven, July 2023



Department of Applied Mathematics and Computer Science $$\mathsf{CASA}$$

Abstract

Metadynamics, a computational method, makes it possible to investigate potential energy surfaces and examine particle systems in complex environments. In this research, the behaviour of a particle controlled by a potential energy surface with two unique minima is explored. The goal is to identify the equilibrium arrangements and transitions between these minima.

Metadynamics is combined with Free Energy Surface (FES) reconstruction methods, Langevin dynamics, and the Milstein scheme approach of numerical integration. The ability to accurately reconstruct the free energy surface is essential for gaining insights into the system's thermodynamics, stability, and equilibrium properties. For a realistic simulation of stochastic processes, it is essential to investigate the Milstein scheme and its appropriate numerical integration in metadynamics. The Milstein scheme offers a higher order of accuracy in handling stochastic differential equations, providing more reliable results compared to simpler integration schemes.

As for the exploration process, the energy landscape is navigated by biasing it systematically and capturing the dynamics of transitions. Moreover, the importance of the parameters of the bias potential is analysed. This is crucial for optimizing the metadynamics simulation and the exploration of the potential energy surface. Understanding them makes it possible to guarantee an accurate depiction of energy barriers and transition states, efficient sampling, and reliable results.

To achieve the objectives, Python simulations are run. The findings from these simulations shed light on the equilibrium states and rare events in 2-dimensional systems by revealing how deterministic and stochastic forces interact. The results demonstrate the effectiveness of metadynamics, FES reconstruction, and the Milstein scheme in understanding and manipulating complex systems. Furthermore, the equilibrium arrangements implying the existence of two minima are correctly identified.

This research has implications across many fields of science and engineering, including prospects for the development of new materials, chemical optimization, research into protein folding, and the examination of intricate physical systems. The multidimensional systems of the real world are nevertheless simplified in this project. A potential energy surface can be investigated using a variety of other techniques, which could lead to a higher degree of accuracy during the reconstruction. Additionally, due to the particle's limited degrees of freedom, interactions and dynamics that take place in additional dimensions cannot be investigated or represented. However, by simulating metadynamics in Python, this research offers guidelines for sampling a particle in a 2D potential energy surface. It also reveals information about the particle's behaviour as it transitions through the minima.

Contents

1	Introduction	3
2	Theoretical Background 2.1 Particle Systems and Potential Energy Landscapes	5 5 5 6 6
3	System Description and Potential Energy Landscape Analysis 3.1 Potential Energy Function 3.1.1 Conservative Force Equation 3.1.2 Minima of the PES 3.2 Bias Potential Energy Function 3.2.1 Bias Force Equation	7 9 9 9 10
4	Understanding Particle Motion and Dynamics4.1Brownian Motion4.2Langevin Dynamics	11 11 11
5	Numerical Integration5.1 Explicit Methods with Additive Noises5.2 Milstein Scheme	13 13 14
6	Simulation Methodology6.1 Defining Constants and Parameters6.2 Defining Initial Conditions and Functions6.3 Simulating Motion of the Particle	16 16 16 17
7	Results 7.1 Parameter Analysis 7.1.1 Weight Analysis 7.1.2 Bias Deposition Interval Analysis 7.1.3 Friction Coefficient Analysis 7.2 Base Case Simulation (100,000 Steps) 7.3 Reconstruction of the PES	 18 18 21 23 24 27
8	Discussion and Conclusion	29

1 Introduction

In a world of complex energy landscapes and elusive transitions, understanding the intricate behaviour of particle systems becomes paramount. This is where Metadynamics comes in. It is a powerful computational method used to investigate complex potential energy landscapes and analyse particle system behaviour. As for its applications, researchers may now explore the intricate features of high-dimensional systems and identify their hidden dynamics by utilising metadynamics. This revolutionizes the understanding and creation of new opportunities for innovation and discovery.

Moreover, metadynamics has become a well-known tool in recent years in materials science, chemistry, biophysics, and computational physics, allowing scientists to study complicated systems at the atomic or molecular level [2]. Many of these systems, in the actual world, display high-dimensional landscapes with numerous stable configurations. The environments of the systems are frequently divided by significant energy barriers and basins. Using conventional simulation techniques, referred to as "standard dynamics" or molecular dynamics, to investigate the rich features of these environments can become excessively time-consuming.

In more complex scenarios, it is common to encounter systems where the exact locations of minima and the pathways connecting them are unknown. The existence and locations of these hidden minima can be determined with the use of metadynamics. Additionally, it enables the investigation of transition pathways between various minima in the potential energy landscape. By providing insights into the kinetics and energetics of processes, this understanding of transition paths facilitates the design of control or optimisation techniques for those processes. Depending on the particular system of interest, metadynamics can use a variety of reaction coordinates for this. For instance, it can be applied to i) the distance between two points or atoms on a landscape or ii) the angle between specific bonds that determine how molecules are oriented relative to one another.

In addition, the energy barriers that exist between the minima are a key factor in influencing the reaction rates. The energy landscape can be manipulated to favour particular minima or transition states by changing variables like temperature, pressure, or reactant concentrations. This knowledge allows for the design of reaction conditions that enhance reaction selectivity, yield, and efficiency. Researchers can identify the most thermodynamically stable products under particular reaction conditions by assessing the stability of various configurations and taking into account the energy barriers between them.

When using molecular dynamics simulations, the system may spend a colossal amount of time examining the area around one local minimum, moving barely towards the global minimum or other important states. This limitation presents a significant challenge in various scientific and engineering domains. In these fields, understanding equilibrium configurations and rare events is critical for designing materials, optimizing chemical reactions, predicting protein folding pathways, and studying the behaviour of complex physical systems. Here, metadynamics emerges as a revolutionary technique that provides a solution by effectively biasing the exploration process, permitting the systematic sampling of the energy landscape, and making it easier to find states and transitions that were previously elusive. However, there are substantial issues with the computing cost and scalability of metadynamics simulations for large-scale systems. The processing demands rise exponentially as the number of particles rises.

Moreover, examining actual systems of interacting particles with molecular dynamics can be very challenging. Therefore, a simplified 2D model potential energy surface (PES) system is used for this

investigation; there are two minima in this landscape. A single particle that moves throughout this surface will be studied. The distance between two points on the landscape serves as the single reaction coordinate, representing the progress of the particle between the two states. This model allows for the exploration of the behaviour and dynamics of the system while avoiding the complexities involved in studying real-world interacting particle systems.

To model the motion of the particle within the simplified PES, Langevin dynamics, a variation that includes stochastic forces, is used in place of molecular dynamics. It blends stochastic motion caused by random thermal fluctuations with deterministic motion caused by forces derived from the potential energy surface. The usage of Langevin dynamics for simulating the particle's motion will be explained further in section 4.2.

In this report, there are several objectives under the concept of metadynamics that will be explored. These can be described as the following questions:

- Given that the two minima of a potential energy surface are known, can the reconstruction of the free energy surface through metadynamics correctly identify the existence of these equilibrium configurations?
- What is the role of parameters during the bias potential deposition and how do they influence the exploration of the potential energy surface?
- What is the Milstein scheme and how can it be correctly implemented for numerical integration during metadynamics?

To achieve these objectives, first, the background and context are explained in section 2. Following that, section 3 defines and provides a description of the model system and potential energy landscape that the particle will explore. The dynamics that affect the particle's motion as it moves through the landscape are then introduced in section 4. The Milstein scheme is derived in section 5 to numerically integrate the particle's position and velocity during this journey. After that, the methodology of the Python simulation is briefly mentioned in section 6. The simulation parameters are then examined in the results section 7 together with the reconstruction of the free energy surface. Finally, the discussion and conclusion section 8 addresses whether the goals were accomplished. Additionally, it summarizes the extent to which the project's objectives were achieved.

2 Theoretical Background

2.1 Particle Systems and Potential Energy Landscapes

Particle systems are collections of atoms or particles that communicate with one another through forces. Molecules, gases, liquids, and solids can all be a part of these systems. The behaviour of these particles is determined by the potential energy landscape, which represents the potential energy of the system as a function of the particles' positions.

The particle system's energy distribution and configuration space are described by the potential energy landscape. Any given arrangement of the particles is represented by a point in the configuration space; this arrangement is given by the potential energy function based on interactions between the particles.

This research project will examine the potential energy landscape of a single particle with two degrees of freedom. It means that the particle can move in the x and y direction. The coordinates of the particle in the form (x, y) are a function of these two degrees of freedom; the x-coordinate corresponds to the x-position of the particle and the y-coordinate is for the y-position. These particle positions can also be denoted by microscopic coordinates of the form R = (x, y). These coordinates are typically used to describe the particle's position in terms of its atomic or molecular constituents and represent the particle's location at a certain time t. Moreover, let S be a set of 2 functions of the microscopic coordinates R of the system, where $S(R) = (S_1(R), S_2(R))$. $S_1(R)$ then corresponds to the x-coordinates and $S_2(R)$ corresponds to the y-coordinates [2]. These two notations for the coordinates of the particle will be used interchangeably.

Furthermore, a potential energy landscape can have minima that are associated with equilibrium configurations. These minima can be divided into a global minimum, which represents the system's most stable configuration, and local minima, which represent stable but less favourable configurations. Energy barriers, corresponding to higher potential energy regions, separate different minima, and particles need to overcome these barriers to transition between equilibrium configurations.

2.2 Mathematical Formulation

The coordinates of the particles that correspond to the configuration with the lowest potential energy are needed to solve the mathematical issue of obtaining the minimal energy state in a two-dimensional system. For the system of one particle in this project and a potential energy function defined as U(x, y), the minimum can be found by setting the first-order partial derivatives of U(x, y) to zero.

2.2.1 Role of Forces and Connection to Potential Energy Landscapes

To establish equilibrium configurations in particle systems, forces are essential. The forces acting on each particle are derived from the potential energy function through the gradient of potential energy with respect to the particle positions. Such forces are also called deterministic forces. This conservative force, due to the potential energy, exerted on the particle is the negative gradient of energy; it is provided by:

$$F_U(x,y) = -\nabla U(x,y). \tag{1}$$

Moreover, the forces acting on the particles are balanced in equilibrium configurations, resulting in a net force of zero. This indicates that there is no acceleration or movement and that the system is at rest. Equilibrium configurations refer to locations where the forces are minimised, or alternatively,

where the local minimum of potential energy is reached.

By analyzing the forces acting on the particles and their relationship with the potential energy landscape, one can gain insights into the stability and behaviour of the system. To bring the system to equilibrium configurations, forces direct the particles towards areas of lower potential energy.

2.3 Metadynamics

2.3.1 The Bias Potential

Since the forces due to the potential energy push the particles toward the minima, there needs to be a force which works against it for the particle to explore other regions of the potential energy landscape. This force can be derived from an additional term called the bias potential. This is denoted by $V_G(x, y)$.

To help with the exploration of the system's configuration space, the bias potential is added to the potential energy function of the system. It is used to overcome energy barriers, especially in systems with complicated energy landscapes. The bias potential's purpose is to encourage the system to go across configuration space areas that are energetically unfavourable or infrequently sampled under standard dynamics.

2.3.2 Role of the Bias Potential in Metadynamics

In metadynamics, the system is periodically influenced by a history-dependent bias potential to promote the exploration of various regions. The energy minima or basins that are encountered are gradually filled by the bias potential during this simulation. It develops during the simulation as a sum of Gaussian functions, also referred to as Gaussian hills or Gaussian potentials. Each Gaussian function contributes to the bias potential with a certain height and width and is centred at a specific configuration that the system visited during the simulation [2].

The system's energy barriers are lowered as a result of the Gaussian hills' effect of increasing the energy of the visited regions. As a result, the system is biased towards discovering new areas and avoiding local energy minima, resulting in improved sampling and a thorough investigation of the energy landscape. To balance the exploration of the energy landscape without excessively biasing the system, the shape, height, and width of the Gaussian hills are essential. The influence of these parameters will be discussed in section 7.1.

Similar to the force due to potential energy, the force due to the bias potential can be calculated as follows:

$$F_{V_C}(x,y) = -\nabla V_G(x,y). \tag{2}$$

This bias force will act on the particle and influence its movement as it traverses the potential energy landscape.

3 System Description and Potential Energy Landscape Analysis

In this section, a description of the system under investigation is given. The potential energy function is defined along with the bias potential for metadynamics. As mentioned in section 1, the set S of 2 functions of the microscopic coordinates R of the system, where $S(R) = (S_1(R), S_2(R))$, will be used here for the bias potential function. $S_1(R)$ and $S_2(R)$ denote the x-coordinates and y-coordinates respectively for the particle at time t.

3.1 Potential Energy Function

As previously mentioned, metadynamics strives to break down energy barriers and make it easier to explore other system states. Since the particle has to transition between the two minima, the potential energy function (PES) chosen should exhibit sufficient sensitivity to perturbations. The intended states might not be effectively sampled by the metadynamics simulation if the potential energy surface is too flat or lacks barriers. Several attempts were made for setting up a working potential energy function for metadynamics was researched for and found in a JCTC article [5]. It has been modified and utilised in this report.

Additionally, the potential energy function's computational effectiveness needs to be taken into account. The importance of this decision can be attributed to the fact that less complex or computationally expensive potentials can enable effective exploration of the energy landscape, allowing for simulation times within the parameters of this study and system sizes within realistic computational limits. Considering the above, for this project, the equation of potential energy function is as follows:

$$U(x,y) = 1 \times 10^{-6} (1.34549x^4 + 1.90211x^3y + 3.92705x^2y^2 - 6.44246x^2 - 1.90211xy^3 + 5.58721xy + 1.33481x + 1.34549y^4 - 5.55754y^2 + 0.904586y + 18.55980).$$
(3)

It can be seen in the equation above that a scaling factor of 1×10^{-6} has been applied. The original potential energy landscape's influence in relation to the bias potential is lessened by this scaling factor. The bias potential becomes more noticeable when the potential energy function is scaled down, enabling it to effectively explore and sample the energy landscape without dominating the system.

The PES as a function of x can be seen below in figure 1. It also shows the two minima of the system.



Figure 1: Potential Energy Landscape as a function of x

The figure 2 below illustrates the original potential energy landscape defined in equation 3. The contour lines on the plot represent regions of equal potential energy. Every contour line represents a different energy level. The lines that are more closely spaced show a gradient or potential energy shift that is steeper or faster. A softer gradient or a slower shift in potential energy, on the other hand, is shown by lines that are farther apart.



Figure 2: Contour Plot of the Potential Energy Landscape

3.1.1 Conservative Force Equation

The force due to the potential energy function defined in 3.1, with respect to each component, is computed below:

$$f_{U_x} = -\frac{\partial U}{\partial x}$$

$$= -1 \times 10^{-6} \cdot (5.38196x^3 + 5.70633x^2y + 7.85410xy^2 - 12.88492x - 1.90211y^3 + 5.58721y + 1.33481)$$

$$f_{U_y} = -\frac{\partial U}{\partial y}$$

$$= -1 \times 10^{-6} \cdot (1.90211x^3 + 7.85410x^2y - 5.70633xy^2 + 5.58721x + 5.38196y^3 - 11.11508y + 0.90459).$$
(4)

3.1.2 Minima of the PES

The two minima of the potential energy function can be found by equating the first-order derivatives of the U(x, y) function to zero:

$$\frac{\partial U}{\partial x} = 1 \times 10^{-6} \cdot (5.38196x^3 + 5.70633x^2y + 7.85410xy^2 - 12.88492x - 1.90211y^3 + 5.58721y + 1.33481) = 0$$

$$\frac{\partial U}{\partial y} = 1 \times 10^{-6} \cdot (1.90211x^3 + 7.85410x^2y - 5.70633xy^2 + 5.58721x + 5.38196y^3 - 11.11508y + 0.90459) = 0.$$
(5)

Solving the above first-order derivative equations, the following minima are found:

Global Minimum:
$$(x_1, y_1) = (-1.87991, 0.78404)$$

Local Minimum: $(x_2, y_2) = (1.78624, -0.83124)$ (6)

3.2 Bias Potential Energy Function

After the definition of the PES, a bias potential energy function has to be derived to perform metadynamics. In this section, functions to run the metadynamics simulation will be defined.

Let N denote the total number of time steps and dt denote the step size for the simulation. Suppose that the bias potential is deposited at every t' time step. Then, at the deposition step, the x and y coordinates correspond to $S_1(R(t'))$ and $S_2(R(t'))$ respectively.

Since the bias potential has the form of Gaussian summations, there is a weight and width that contributes to the function. Let the weight be denoted by ω . This determines the strength of the Gaussian hill. The width, on the other hand, controls the extent or spread of the bias potential along the coordinates. The same value is taken for x and y for the spread. This is denoted by σ_x and σ_y where $\sigma = \sigma_x = \sigma_y$.

Then, the equation for the bias potential is given by:

$$V_G(S,t) = \int_0^N dt' \cdot \omega \cdot \exp\left[-\left(\left(\frac{(S_1(R) - S_1(R(t')))^2}{2 \cdot (\sigma_x)^2}\right) + \left(\frac{(S_2(R) - S_2(R(t')))^2}{2 \cdot (\sigma_y)^2}\right)\right)\right]$$
(7)

3.2.1 Bias Force Equation

Similar to the conservative force equation, the bias force is calculated by the negative gradient of the bias potential energy function. Component-wise, the bias force is as follows:

$$\begin{split} f_{V_{G_x}} &= \omega \cdot \exp\left[-\left(\left(\frac{\left(S_1(R) - S_1(R(t'))\right)^2}{2 \cdot (\sigma_x)^2}\right) + \left(\frac{\left(S_2(R) - S_2(R(t'))\right)^2}{2 \cdot (\sigma_y)^2}\right)\right)\right] \\ &\quad \cdot \left(\frac{S_1(R) - S_1(R(t'))}{(\sigma_x)^2}\right) \\ f_{V_{G_y}} &= \omega \cdot \exp\left[-\left(\left(\frac{\left(S_1(R) - S_1(R(t'))\right)^2}{2 \cdot (\sigma_x)^2}\right) + \left(\frac{\left(S_2(R) - S_2(R(t'))\right)^2}{2 \cdot (\sigma_y)^2}\right)\right)\right] \\ &\quad \cdot \left(\frac{S_2(R) - S_2(R(t'))}{(\sigma_y)^2}\right) \end{split}$$
(8)

In the above equations, it should be noted that the negative sign for force has been cancelled during the derivative computations.

4 Understanding Particle Motion and Dynamics

In this section, the concepts of Brownian Motion and Langevin dynamics will be introduced to define the behaviour of the particle in the configuration space.

4.1 Brownian Motion

For this project, a requirement is that the particle's behaviour imitates Brownian motion. This allows for the particle to explore the configuration space more extensively, sampling rare events and overcoming energy barriers more efficiently. To achieve modelling random movements or fluctuations of the particle, a 2-dimensional Wiener process will be used.

A Wiener process is a continuous-time stochastic process. The defining characteristic of a Wiener process is that it has independent and identically distributed (i.i.d.) increments. In other words, the change in the process over a given period is unrelated to the changes that took place earlier. Furthermore, Wiener process increments have a normal distribution with a mean of zero and a variance proportionate to the time interval [3]. For a 2D Wiener process, the increments at each time step can be represented by a two-dimensional vector.

Let the Wiener process be denoted by:

$$W(t) = \begin{pmatrix} W_1(t) \\ W_2(t) \end{pmatrix}.$$
(9)

where $W_1(t)$ and $W_2(t)$ are the components of the process at time t. The increments of the Wiener process over a small time interval Δt can be represented by:

$$dW(t) = \begin{pmatrix} dW_x \\ dW_y \end{pmatrix} = \Delta W(t) = \begin{pmatrix} \Delta W_1 \\ \Delta W_2 \end{pmatrix} = \begin{pmatrix} \xi_1 \sqrt{\Delta t} \\ \xi_2 \sqrt{\Delta t} \end{pmatrix}$$
(10)

where ξ_1 and ξ_2 are independent random variables drawn from a standard normal distribution with mean = 0 and variance = 1. So, $\xi_1, \xi_2 \sim \mathcal{N}(0, 1)$. These steps are the random variations in the Wiener process during the time t. By averaging these increments across smaller periods, one can approximate the Wiener process's progress over time.

4.2 Langevin Dynamics

In a system where a particle experiences random fluctuations and deterministic forces, to model its dynamics, the simulation method of Langevin dynamics can be employed. To describe the time evolution of a system undergoing these random fluctuations or noise, a stochastic differential equation (SDE) is used. This project considers the case of overdamped Langevin dynamics influencing the system; it is based on the paper by Giovanni Bussi and Davide Branduard [4]. The following is the stochastic differential equation (SDE):

$$dx = -\beta \cdot D \cdot \nabla U dt + \sqrt{2D} \cdot dW(t) \tag{11}$$

where $\beta = \frac{1}{k_B T}$ is the inverse temperature and D is the diffusion coefficient. Here k_B is the Boltzmann constant $(1.38 \times 10^{-23} (J \times K^{-1}))$ and T is the temperature (298K). β is a parameter often used in statistical mechanics and thermodynamics to relate the temperature of a system to its energy. Moreover, ∇U denotes the first-order derivatives of the potential energy function U as computed in

section 3.1.2.

To understand the diffusion coefficient, it is important to establish a friction coefficient. The friction coefficient measures the resistance or dampening that particles encounter when they move through a material. It calculates the dissipative forces that are exerted on the particles, such as those brought on by interactions with the solvent or surrounding medium. For a friction coefficient γ , the diffusion coefficient can be defined as $D = \frac{k_b \cdot T}{\gamma}$. According to this equation, the diffusion coefficient is inversely proportional to the friction coefficient and directly proportional to the product of temperature and Boltzmann's constant. In the setting of Brownian motion, this equation is derived from the Stokes-Einstein relationship [6]. The equation emphasises how frictional forces and thermal energy interact to affect the rate of diffusion in a particular system.

As for metadynamics, the bias potential is considered a dynamic variable and is included in equation 11. The equation then upgrades to:

$$dx = -\beta \cdot D \cdot \nabla (U + V_G) dt + \sqrt{2D} \cdot dW(t)$$
(12)

$$dx = -\beta \cdot D \cdot (\nabla U + \nabla V_G) dt + \sqrt{2D} \cdot dW(t)$$
⁽¹³⁾

In the equations above, the new term ∇V_G represents the gradient of the bias potential function. For simpler representation of equations, the following function is defined:

$$E(S(R)) = \left(\left(\frac{(S_1(R) - S_1(R(t')))^2}{2 \cdot (\sigma_x)^2} \right) + \left(\frac{(S_2(R) - S_2(R(t')))^2}{2 \cdot (\sigma_y)^2} \right) \right).$$
(14)

Furthermore, the gradients of both the potential energy functions are calculated below:

$$\nabla U(x,y) = \begin{pmatrix} 1 \times 10^{-6} \cdot (5.38196x^3 + 5.70633x^2y + 7.85410xy^2 - 12.88492x \\ -1.90211y^3 + 5.58721y + 1.33481) \\ 1 \times 10^{-6} \cdot (1.90211x^3 + 7.85410x^2y - 5.70633xy^2 + 5.58721x \\ +5.38196y^3 - 11.11508y + 0.90459) \end{pmatrix}$$
(15)
$$\nabla V_G(S(R)) = \begin{pmatrix} -\omega \cdot \exp\left(-E(S(R))\right) \cdot \left(\frac{S_1(R) - S_1(R(t'))}{(\sigma_x)^2}\right) \\ -\omega \cdot \exp\left(-E(S(R))\right) \cdot \left(\frac{S_2(R) - S_2(R(t'))}{(\sigma_y)^2}\right) \end{pmatrix}.$$
(16)

The functions computed above are also used in the numerical integration scheme discussed in the next section.

5 Numerical Integration

In this section, an overview of explicit methods with additive noises is given. Then, the Milstein scheme is derived for updating the position and velocity of the particle as it travels the potential energy landscape.

5.1 Explicit Methods with Additive Noises

Explicit methods of integration refer to algorithms that directly calculate the future state of a system based on its current state. They do not require iterative or implicit calculations. When compared to implicit approaches, these methods often have reduced computing costs, making them computationally efficient for many applications.

Considering additive noise in explicit integration methods entails including stochastic or random forces in the dynamics of the system. The random fluctuations or uncertainties that have an impact on the system's behaviour are represented by additive noise. In this instance, the computations of the system's future state at each time step explicitly take the random forces from the Brownian motion into account. It also takes deterministic forces into account. These methods allow for computations on the position and velocities of the particle at the next step.

An approach to explicit methods with additive noises is based on Taylor-type expansion. This is specific to the numerical integration of SDEs and is discussed in section 3.1 of the book Numerical Integration of SDEs [7]. Consider the following system of stochastic differential equations with additive noises from this book:

$$dX = a(t, X)dt + \sum_{r=1}^{q} \sigma_r(t)dw_r(t)$$
(17)

where a(t, X) is the deterministic drift term of the stochastic differential equation, q is the number of independent Wiener processes, $\sigma_r(t)$ is the diffusion term associated with the *r*-th Wiener process and $w_r(t)$ is the *r*-th Wiener process.

The above SDEs have to be adapted to two dimensions for taking into account the two degrees of freedom the particle has in the system under consideration. Nevertheless, the system above can be compared to SDE derived for the Langevin dynamics in equation 13. The deterministic drift term in this case is comparable to $-\beta \cdot D \cdot (\nabla U + \nabla V_G)$, q = 2, and the diffusion term associated with the *r*-th Wiener process is the constant $\sqrt{2D}$. The two dimension adaption looks as follows:

$$dX_{i}(t) = a_{i}(t, X(t))dt + \sum_{r=1}^{2} \sigma_{r}(t)dW_{r}(t),$$

$$X_{i}(0) = X_{i}^{(0)}.$$
(18)

As a result of this comparison, it is possible to use a simple numerical technique to approximate the solution of stochastic differential equations. This stochastic Euler scheme, which is also known as the Euler-Maruyama scheme, only makes use of the first two terms of the Taylor expansion [1]. The Euler-Maruyama approximation scheme for position updates of equation 18 is given by:

$$\begin{pmatrix} x_{k+1} \\ y_{k+1} \end{pmatrix} = \begin{pmatrix} x_k \\ y_k \end{pmatrix} - \beta \cdot D \cdot \nabla (U + V_G) dt + \sqrt{2D} \cdot dW(t),$$
(19)

5.2 Milstein Scheme

The Milstein scheme is an extension of the Euler-Maruyama method. It is used because it can incorporate new terms that result from the expansion of drift and diffusion terms to higher orders. In comparison to the Euler-Maruyama approach, the expansion enables higher-order accuracy in the numerical integration of the SDE and produces more precise results. It extends the accuracy by including the first-order correction term in the expansion of the diffusion term.

In both the x and y position updates in equation 19, the Milstein scheme would add quadratic terms to improve the accuracy of the approximation. However, these terms would depend on the partial derivatives of the diffusion term. The diffusion term is the constant $\sqrt{2D}$ in this case, and so, for the position update of the particle, the equations are the same as the Euler-Maruyama method:

$$x_{k+1} = x_k - \beta \cdot D \cdot \left[\frac{\partial U}{\partial x} + \left(-\omega \cdot \exp\left(-E(S(R))\right) \cdot \left(\frac{S_1(R) - S_1(R(t'))}{(\sigma_x)^2}\right) \right) \right] dt + \sqrt{2D} \cdot dW_x$$
$$y_{k+1} = y_k - \beta \cdot D \cdot \left[\frac{\partial U}{\partial y} + \left(-\omega \cdot \exp\left(-E(S(R))\right) \cdot \left(\frac{S_2(R) - S_2(R(t'))}{(\sigma_y)^2}\right) \right) \right] dt + \sqrt{2D} \cdot dW_y.$$
(20)

The importance of the Milstein scheme can be seen during the velocity update of the particle. The first step is to derive the SDE for velocity from the position equation 13. This is simply done by taking the derivative of equation 13 with respect to time. Then, the velocity SDE is as follows:

$$dv = -\beta \cdot D \cdot (\nabla U + \nabla V_G).$$
⁽²¹⁾

However, the above equation does not take into account the drag force experienced by the particle. The drag force f_{drag} is a force that opposes the motion of an object moving through a fluid or gas. Mathematically, it is described as:

$$f_{drag} = \begin{pmatrix} f_{drag_x} \\ f_{drag_y} \end{pmatrix} = \begin{pmatrix} -\gamma \cdot v_x \\ -\gamma \cdot v_y \end{pmatrix}$$
(22)

where v_x and v_y are velocities of the particle in x and y directions at the current state.

Furthermore, the term $\sqrt{2D} \cdot dW(t)$ representing the stochastic force or the random fluctuations due to the Brownian motion, affects the velocity indirectly through the position update equation. Thus, the new Langevin dynamics equation for velocity becomes:

$$dv = -\beta \cdot D \cdot (\nabla U + \nabla V_G) + f_{drag} + \sqrt{2D} \cdot dW(t).$$
(23)

To numerically integrate the above equation, the second-order partial derivatives of the potential energy and bias potential function need to be calculated because they influence the velocity at the next time step. These are computed as follows:

$$U_{xx} = \frac{\partial^2 U}{\partial x^2} = 1 \times 10^{-6} \left(5.70633x^2 + 11.41266xy + 7.85410y^2 - 12.88492 \right)$$

$$U_{yy} = \frac{\partial^2 U}{\partial y^2} = 1 \times 10^{-6} \left(7.85410x^2 - 11.41266xy + 16.14588y^2 - 11.11508 \right)$$

$$b_{xx} = -\omega \cdot \exp\left(-E(S(R))\right) \left[\frac{1}{(\sigma_x)^2} - \left(\frac{S_1(R) - S_1(R(t'))}{(\sigma_x)^2} \right)^2 \right]$$

$$b_{yy} = -\omega \cdot \exp\left(-E(S(R))\right) \left[\frac{1}{(\sigma_y)^2} - \left(\frac{S_2(R) - S_2(R(t'))}{(\sigma_y)^2} \right)^2 \right]$$
(24)

where b_{xx} and b_{yy} are the second-order partial derivatives of $V_G(S,t)$ in terms of S_1 and S_2 respectively.

Now, the Milstein scheme for updating the velocities has been derived from Yousef Alnafisah's research article [1]. This gives the following equations:

$$v_{x_{k+1}} = v_{x_k} + f_{drag_x} \cdot dt + \beta \cdot D \cdot \left(f_{U_x} + f_{VG_x} \right) \cdot dt + \sqrt{2D} \cdot dW(t) + \frac{1}{2} \beta \cdot D \cdot (U_{xx} + b_{xx}) \cdot (dW_x)^2 \cdot dt - \frac{1}{2} D \cdot (U_{xx} + b_{xx}) v_{y_{k+1}} = v_{y_k} + f_{drag_y} \cdot dt + \beta \cdot D \cdot \left(f_{U_y} + f_{VG_y} \right) \cdot dt + \sqrt{2D} \cdot dW(t) + \frac{1}{2} \beta \cdot D \cdot (U_{yy} + b_{yy}) \cdot (dW_y)^2 \cdot dt - \frac{1}{2} D \cdot (U_{yy} + b_{yy})$$
(25)

6 Simulation Methodology

 $sig_y = 0.25$

In this section, the steps of the metadynamics simulation are briefly discussed. All simulations are done in Python.

6.1 Defining Constants and Parameters

The code begins with importing the required packages numpy, math and matplotlib. Then, the following constants and bias potential parameters such as the weight and width of the Gaussian hill are assigned to variables:

```
# Define constants
kB = 1.38e-23 # Boltzmann constant
T = 298 # Temperature
dt = 1e-03 # Time step
gamma = 6 * np.pi * 0.001 * 1e-5 # Friction coefficient
D = (kB * T) / gamma # Diffusion coefficient
beta = 1 / (kB * T) #Inverse temperature
# Define time array
t = np.arange(0, 100, dt) #Total 100,000 steps
#Define bias potential parameters
weight = 6 * 1e-7 #Weight of the Gaussian hill
sig_x = 0.25 #Width of the Gaussian hill
```

The length of the time array t denotes the total number of steps in the simulation.

In addition to the above parameters, there are also variables assigned for the efficiency of the code. These are for the constants that appear often in the numerical integration steps. An example is the variable half_BD_dt that equals to $\frac{1}{2}\beta \cdot D \cdot dt$ which appears in equation 25. By defining this constant, the re-computation is avoided at every time step.

6.2 Defining Initial Conditions and Functions

To update the particle's position, the coordinates must be initialised such that the particle begins at the global minimum (-1.87991, 0.78404). The particle must have zero starting velocities in both directions in order to be at rest. This guarantees that the particle is initially in a state of stability and would resist moving towards other states without the bias potential.

Moreover, the positions and velocities are tracked during the simulation. This is because the Milstein scheme and the bias potential function both need previous values to calculate the current position and velocity. Arrays of the shape t are used to store these values. Additionally, to record the particle's position at the time step when the bias potential is deposited, two lists called x_loc and y_loc are created.

Next, the potential energy function is defined. The function takes the coordinates (x, y) as input and returns the value of potential energy at that point. As for the definition of the bias potential function, in addition to the current position, it also takes the point at which the Gaussian hill is centred as an input.

Further, the forces derived from both of these functions are defined to return an array of size 2 when the input of the current position is given. This outputs a tuple of the form (force in x-direction, force in y-direction). Similarly, functions that calculate the second-order partial derivatives of the potential energies are also defined.

6.3 Simulating Motion of the Particle

Every particle position and velocity update, along with the bias potential deposition, happens during the for loop that iterates over the length of the time array t. All the code described in this section happens within this for loop. For the base case parameters that will be described in section 7.1, the deposition of the bias potential occurs every 10 steps. This is also called the bias deposition rate. Consequently, the first step of the for loop checks if the iteration step is divisible by the bias deposition rate. If that is the case, it appends the current position of the particle to the lists x_loc and y_loc.

Then, the forces due to the potential energy and the bias force at the current position are calculated by calling the functions defined before. The second-order partial derivatives are also calculated and assigned to new variables. Following that, the Wiener process increments are determined by values generated from a standard normal distribution and scaling them with the square root of the time step:

```
# Calculate random force due to Brownian motion (Wiener process increments)
    dW_x = np.random.normal() * np.sqrt(dt)
    dW_y = np.random.normal() * np.sqrt(dt)
```

Next, the drag force in the x and y directions at the current position is measured according to equation 22. Furthermore, a few efficiency variables are established that sum the potentials and second-order derivatives under the requirements of the Milstein scheme as shown in equation 25. Finally, variables for the position and velocity updates of the Milstein scheme are defined. These add to the created position and velocity arrays mentioned in 6.2.

These simulation steps enable the storage of the bias depositions, particle locations, and velocities, which can then be utilised to create plots that will be shown in the results section 7.

7 Results

In this section, the results of the simulations done in Python will be discussed. The run-time for these simulations is N = 20,000 steps.

7.1 Parameter Analysis

For the analysis of parameters, the following variables will be monitored:

- ω : Weight of Gaussian
- σ : Width of Gaussian
- γ : Friction Coefficient
- τ : Bias Deposition Rate

By changing combinations of the above, the bias deposition graphs are plotted and evaluated. A base case has been set to compare the plots and modify the parameters. The values for the base case are:

- ω : 6×10^{-7}
- σ : 0.25
- $\gamma : 1 \times 10^{-5}$
- *τ* : 10

7.1.1 Weight Analysis

The intensity or amplitude of the Gaussian hill, and thus, the strength of the bias potential introduced during simulation, are both dependent on the weight of the Gaussian. A stronger bias potential results from a higher weight, which causes a more pronounced disruption in the system. There are three values of weight considered 4×10^{-7} , 6×10^{-7} and 8×10^{-7} . The resulting bias deposition graphs are:







Figure 3: Bias Depositions for weight $\omega = 4 \times 10^{-7}$

<u>Case 2:</u> Weight of 6×10^{-7}





Figure 4: Bias Depositions for weight $\omega=6\times 10^{-7}$

<u>Case 3:</u> Weight of 8×10^{-7}





Figure 5: Bias Depositions for weight $\omega = 8 \times 10^{-7}$

Across all plots of varying weight, it can be noticed that for the smallest width of $\sigma = 0.25$, the particle exploration is the narrowest. The shape of the plot, however, shows the widest investigation with a width of $\sigma = 0.8$. This is consistent with the idea that the Gaussian breadth determines how large the bias region is. While a smaller Gaussian concentrates the sampling on particular areas, a wider Gaussian allows for a more thorough investigation of the configuration space. The energy level values on the graphs' right are also higher as weight increases. This heavier weight can aid in overcoming energy obstacles and make it easier to explore various PES locations. However, excessively high hills can cause oversampling or distortion of the underlying dynamics, so it is important to choose an appropriate weight for effective sampling without introducing significant bias. This distortion can be seen clearly in plot (d) of Figure 5.

7.1.2 Bias Deposition Interval Analysis

In metadynamics, the bias potential is gradually accumulated throughout the simulation to improve sampling in the PES's under-explored areas. The frequency with which Gaussian potentials are added to the system depends on the bias deposition rate. For the bias deposition rate analysis, the base case parameters are kept the same except for τ . The bias is modified to be deposited every 5, 10, 100 or 1000 steps. The effects of this change can be seen below:





Figure 6: Various Bias Deposition Intervals for the Base Case

The addition of Gaussian potentials occurs more frequently with a larger bias deposition rate, speeding convergence but perhaps increasing noise in the simulations as seen in (a) of Figure 6. In contrast, a lower deposition rate slows convergence while delivering more well-rounded bias potential landscapes. This can be seen in (d) of Figure 6. To balance the exploration of the configuration space, convergence rate, and computing effectiveness, care must be taken when selecting these values. To attain the necessary sampling and convergence features in metadynamics simulations, these parameters might need to be tuned for various systems and research goals. Although the graphs (c) and (d) in Figure 6 show well-rounded bias potential landscapes, simulating with those parameters will take days and weeks to obtain the optimum results. A deposition rate that helps the particle reasonably explore the PES in the given simulation time needs to be chosen.

7.1.3 Friction Coefficient Analysis

As introduced in section 4.2, the friction coefficient γ measures the resistance or damping experienced by particles as they move through a medium. It also represents the interaction between the particle and its surrounding environment. During trial and error for values of γ , a drastic change in graphs was seen for the particle trajectory and the bias deposition. As a result, the following four values of γ were chosen to be investigated: 1×10^{-3} , 1×10^{-4} , 1×10^{-5} and 1×10^{-6} . These produced the bias deposition graphs below:



Figure 7: Various Friction Coefficient values for the Base Case

The mobility of the particles is negatively correlated with the friction coefficient. A larger value of γ implies greater resistance to particle motion, resulting in slower diffusion. This is depicted in Figure 8's plot (a). Additionally, greater friction coefficients result in a sharper dampening of the particle's velocity, which has an impact on how the particle diffuses and explores the PES.

 γ also determines the strength of the dissipative forces acting on the particle's motion. It is clear that the system went through too many uncontrolled energy fluctuations in plots (d) of both Figures 7 and 8. Small friction coefficients indicate weak dissipative forces influencing the motion of the particle. As a result, the particle's velocity may rise quickly, resulting in a disorderly buildup of energy and unpredictable behaviour. Without adequately sampling the relevant parts of the potential energy surface (PES), the particle may move too easily and quickly. This is seen in plots (a) and (b) of Figure 8.



Figure 8: Particle Trajectories for various Friction Coefficient values of the Base Case

7.2 Base Case Simulation (100,000 Steps)

As seen by the figures in section 7.1, the base case described is a reasonable option to investigate further. The particle explores the PES sufficiently without unpredictable behaviour; the bias region does not cause over-deposition and the shape of the plot is closest to the original landscape. Therefore, those parameters are selected to run the simulation for N = 100,000 steps. This gives the following plots for bias deposition and particle motion:



(a) Bias Deposition of Base Case





Figure 9: Simulation of Base Case for 100,000 steps

Figure 9 (a) shows the largest values of energy levels on the right. Also 9 (b) shows that the particle explored the PES well and extensively.

Then, plots of the position displacements were made to see the particle movements in relation to the minima:



Figure 10: Position Displacement of the Particle

In the above Figure 10, the yellow and purple lines denote the x and y values of the minima. These are the same as those mentioned in section 3.1.2. The yellow lines correspond to the point (-1.87991, 0.78404) and the purple lines are for (1.78624, -0.83124).

Similarly, the changes in velocities v_x and v_y are plotted below in Figure 11. It can be seen that as time progresses, the velocities in the y-direction increase up to 5 times the original. The velocities in the x-direction stay roughly in the range of -4 to 3 m/s. This can be attributed to the fact that the force from the potential energy function, the second-order derivatives and the terms with the bias potential in the y component have a higher magnitude of change compared to x.



Figure 11: Velocity Changes of the Particle

7.3 Reconstruction of the PES

The free energy surface (FES) obtained from the reconstruction in metadynamics represents the free energy landscape of a system. It offers details regarding the relative stability and probability distribution of various system states or configurations. The likelihood of discovering a physical system in a specific state depends on the state's free energy. The lower the free energy, the more stable the state, and the higher the probability of the system occupying that state. The FES allows for the identification of free energy disparities across various states and captures the underlying energy landscape.

The initial potential energy surface (PES), which was modified by the deposition of bias potentials, is reflected in the reconstructed FES generated via metadynamics. By successfully overcoming the energy biases and barriers contained in the original PES, it offers an improved representation of the system's free energy landscape.

Let the free energy surface be represented by F(S(R)). As explained in the paper by Giovanni Bussi and Davide Branduard [4], after a long simulation time, the basic assumption of metadynamics is the following:

$$V_G(S(R)) \sim -F(S(R)) + C(t) \tag{26}$$

where C(t) is a constant dependent on time. For simplification in this report, the above relationship will be modified to define the FES as follows:

$$F(S(R)) \sim -V_G(S(R)). \tag{27}$$

For the base case simulation of 100,000 steps in section 7.2, the reconstructed FES is plotted below in Figure 12. The height of the surface corresponds to the free energy at each point in the system's configuration space.



Figure 12: Contour Plot of the Reconstructed FES

This reconstructed FES, plotted against S_1 gives the Figure 13. The shape is comparable to Figure 1 where the PES is plotted against S_1 .



Figure 13: Reconstructed FES with respect to S_1

In the free energy surface (FES), the regions with the most negative values correspond to the minima of the original potential energy surface (PES). These negative values show reduced free energy and greater stability, which increases the likelihood that the system will occupy those zones. The system can explore and inhabit states that could be energetically unfavourable in the original PES due to the bias potentials that were deposited throughout the simulation.

Additionally, upon the reconstruction, the two minima are successfully identified using metadynamics as seen in the plot 13. The basins of attraction or the most advantageous states that the system can inhabit are therefore represented by the minima in the FES. The relative stability and likelihood of the system being discovered in such states are shown by the respective depths of the minima in the FES. In fact, it appears that the minimum on the plot's left side is favoured because it has more depth. This is the global minimum.

8 Discussion and Conclusion

In this section, the objectives will be revisited and evaluated on the level they have been achieved.

The first goal was to reconstruct the free energy surface through metadynamics and use it to identify equilibrium configurations. This objective evaluated how well metadynamics captured and represented the underlying potential energy surface. Figure 13, which was mentioned in section 7.3, explicitly shows the presence of two minima in the potential energy surface upon reconstruction. Additionally, the reconstruction is effective at precisely differentiating between the local and global minimum. These minima are roughly located in the same regions as the minima of the original PES in Figure 1. Thus, the reconstruction was correctly able to identify the existence of the two minima.

The second goal then focused on the impact of the bias potential function's parameters. This objective was important for optimizing the metadynamics simulation. To achieve effective sampling, reliable results, and an accurate depiction of energy barriers and transition states, it is important to understand how these characteristics affect the exploration of the PES. The parameter analysis in section 7.1 made clear how crucial it is to scale the variables accurately to create an effective metadynamics investigation. The best suitable configuration for the metadynamics had the lowest Gaussian width ($\sigma = 0.25$), a moderately weighted Gaussian hill ($\omega = 6 \times 10-7$) and a relatively low friction coefficient ($\gamma = 1 \times 10^{-5}$). This was justifiable as long as the bias was being deposited every ten steps ($\tau = 10$). As a result, the PES was sufficiently explored while maintaining a reasonable simulation time.

Thirdly, section 5.2 examined the Milstein scheme as a numerical integration technique. The correct implementation of this objective was crucial for accurately simulating stochastic processes. When dealing with stochastic differential equations, the Milstein scheme provides more reliable results than simpler integration schemes. Researchers may simulate complex systems affected by both deterministic and stochastic influences by comprehending the theoretical foundation and practical application of the Milstein scheme. The theoretical information about the scheme was obtained from reading the works of literature [7] and [1]. Moreover, a way to modify it for this specific investigation was identified. Since the plots of the bias depositions in section 7.1 are similar to the original PES, the implementation can be regarded as accurate. Particle trajectories in the base case remain within the specified PES as well. However, there is a discrepancy in the particle's velocities at each time step; the *y*-direction velocities experience more abrupt changes. The section 7.2 discusses the causes of this.

In addition to these goals, there are other possible routes for a metadynamics study. The PES is capable of involving additional equilibrium configurations, more degrees of freedom for the particle, and can have multiple particles interacting in the system. The list of the research's potential areas of study is by no means exhaustive, thus, there is room for improvement and further research. An example would be to look into other numerical integration techniques to the Milstein scheme.

Furthermore, adjustments can be made during the FES reconstruction. Even though the minima were found, the investigation's accuracy was constrained by the absence of the constant C(t) in the FES approximation in equation 26. It might be feasible to investigate defining this constant function dependent on time. Such constants can accumulate to produce significant errors in higher-dimensional real-world systems. This could result in the inaccurate identification of equilibrium configurations, which would have serious consequences.

Combining the objectives and concepts discussed in this paper, it is clear that using metadynam-

ics to investigate potential energy surfaces with numerous minima has shown to be an effective method for studying complex systems. The energy landscape has been successfully navigated and the dynamics of transitions between two separate minima have been captured. This investigation has provided information on the equilibrium configurations of the particle system and the interaction between deterministic and stochastic forces. A thorough framework for comprehending the behaviour of 2-dimensional systems has been created by combining metadynamics, FES reconstruction, and the Milstein scheme. This framework addresses the difficulties of investigating complex energy landscapes and rare events.

Metadynamics has laid the path for future advances in understanding and manipulating complex systems by unravelling the complicated dynamics of particle systems, promoting innovation and discovery in a variety of domains.

References

- Yousef Alnafisah. "The Implementation of Milstein Scheme in Two-Dimensional SDEs Using the Fourier Method". In: Abstract and Applied Analysis 2018 (2018). ISSN: 16870409. DOI: 10.1155/2018/3805042.
- [2] Alessandro Barducci, Massimiliano Bonomi, and Michele Parrinello. *Metadynamics*. Sept. 2011. DOI: 10.1002/wcms.31.
- [3] Dominic Breit. "Preliminaries". In: Existence Theory for Generalized Newtonian Fluids (2017), pp. 161-171. DOI: 10.1016/B978-0-12-811044-7.00011-2. URL: https://linkinghub. elsevier.com/retrieve/pii/B9780128110447000112.
- [4] Giovanni Bussi and Davide Branduardi. "Free-Energy Calculations with Metadynamics: Theory and Practice". In: *Reviews in Computational Chemistry*. Vol. 28. May 2015, pp. 1–49. DOI: 10.1002/9781118889886.ch1.
- [5] Michele Invernizzi and Michele Parrinello. "Making the Best of a Bad Situation: A Multiscale Approach to Free Energy Calculation". In: *Journal of Chemical Theory and Computation* 15.4 (Apr. 2019), pp. 2187–2194. ISSN: 15499626. DOI: 10.1021/acs.jctc.9b00032.
- Benjamin Lindner. "The diffusion coefficient of nonlinear Brownian motion". In: New Journal of Physics 9 (May 2007). ISSN: 13672630. DOI: 10.1088/1367-2630/9/5/136.
- [7] G.N. Milstein. Numerical Integration of Stochastic Differential Equations. Ed. by M Hazewinkel. Vol. 313. Ekatarinburg: Springer Science+Business Media, B.Y., 1995, pp. 37–41.