Irreversible Markov Samplers In Higher Dimensions

Eoghan Lemoine. SN: 21010005
Supervisor: Professor Edina Rosta
Department of Physics and Astronomy
University College London
Course: Physics MSci

April 2025

Declaration form.

UCL DEPARTMENT OF PHYSICS & ASTROMOMY

Submission of coursework for Physics and Astronomy course PHAS0097/PHAS0096/PHAS0048

Please sign, date and return this form with your coursework by the specified deadline.

DECLARATION OF OWNERSHIP

I confirm that I have read and understood the guidelines on plagiarism, that I understand the meaning of plagiarism and that I may be penalised for submitting work that has been plagiarised.

I confirm that all work will also be submitted electronically and that this can be checked using the JISC detection service, Turnitin®.

I declare that all material presented in the accompanying work is entirely my own work except where explicitly and individually indicated and that all sources used in its preparation and all quotations are clearly cited.

Should this statement prove to be untrue, I recognise the right of the Board of Examiners to recommend what action should be taken in line with UCL's regulations.

Signed eoghan lemoine

PrintName Eoghan Lemoine

Dated 09/04/2025

Title	Date Received	Examiner	Examiner's Signature	Mark

Abstract

Markov Chain Monte Carlo (MCMC) methods are widely used to sample from complex probability distributions. Traditional MCMC algorithms, such as the Metropolis–Hastings algorithm, rely on the detailed balance condition to guarantee convergence to a target distribution. However, detailed balance often limits sampling efficiency, especially in high-dimensional spaces. Irreversible MCMC methods, which relax this condition, have recently attracted attention due to their potential to accelerate sampling. In this investigation, we extend ideas behind the skewed detailed balance framework to develop and benchmark generalised methods for skewness implementation in higher-dimensional sampling problems.

We demonstrate that irreversible samplers incorporating directional bias in 2D significantly outperform reversible samplers in terms of relaxation time and mean first passage time (MFPT). Even simple, unoptimised implementations of skewness achieve substantial efficiency gains, which scale with the number of states. Additionally, we introduce a Hessian-guided skewness method leveraging local curvature information, which significantly accelerates sampling initially but suffers numerical instabilities at longer simulation times, leading to deviations from the correct stationary distribution.

Contents

1 Introduction

Markov Chain Monte Carlo (MCMC) methods are powerful computational techniques used to understand and explore complicated probability distributions. They are particularly useful when dealing with problems that involve many variables or complicated relationships, making direct calculations difficult or even impossible. Instead of evaluating every possible outcome, MCMC methods efficiently generate representative samples from the distribution, allowing researchers to approximate and analyse complex systems.

These methods underpin critical applications in molecular dynamics, Bayesian inference, and statistical physics. Traditional reversible MCMC algorithms, such as the Metropolis-Hastings method, rely fundamentally on the detailed balance condition to ensure convergence towards the desired stationary distribution [1]. The detailed balance condition ensures that, for every possible pair of states, the probability of transitioning from one to the other is exactly balanced by the reverse transition, keeping the overall distribution stable. However, adherence to detailed balance can substantially limit sampling efficiency, particularly in higher-dimensional parameter spaces, frequently resulting in impossibly long convergence times - often known as topological freezing [2]. Consequently, reversible MCMC methods have become increasingly inadequate for the computational demands of contemporary scientific and statistical challenges. Numerous recent studies have highlighted these limitations explicitly [3-7]. For instance, Schaefer et al. [5] employed reversible MCMC to investigate Quantum Chromodynamics (QCD) effects within lattice simulations, encountering severe topological freezing and exceedingly large autocorrelation times at fine lattice spacings. Similarly, Bonati et al. [7] analysed a quantum mechanical particle subject to various constraints, observing critical slowdown and eventual loss of ergodicity as the continuum limit was approached. This emerging consensus underscores the need for innovative sampling strategies capable of overcoming the intrinsic deficiencies associated with reversible sampling. Recent research [8-13] has increasingly focused on breaking detailed balance through the development of irreversible MCMC methods. By deliberately employing asymmetric proposal distributions, these irreversible algorithms enable directional and persistent exploration of the state space, significantly enhancing convergence rates, exploration efficiency, and computational economy. Bernard et al. [9] demonstrated such an approach through their development of event-chains Monte Carlo algorithms for hard-sphere systems, which, unlike Metropolis-Hastings, are rejection-free and employ deterministic particle movements up to a predefined displacement. Their irreversible algorithms notably outperformed traditional reversible methods and even surpassed molecular dynamics simulations in equilibration speed. Similarly, Suwa and Todo [10] introduced a "weight landfill" approach that minimises rejections by inducing directed flows within configuration space. This technique reduced random-walk behaviours, significantly shortening autocorrelation times and achieving convergence rates up to six times faster than the Metropolis algorithm in complex physical models such as the Potts model and quantum Heisenberg spin chain. Irreversible samplers offer distinct advantages for tackling high-dimensional, multimodal problems frequently encountered in areas like molecular dynamics simulations, high-dimensional Bayesian inference, and large-scale network analysis. Despite their potential, irreversible samplers remain relatively unexplored in higher-dimensional contexts, reflecting a noticeable gap in current research. The inherent complexity in designing and implementing irreversible MCMC methods presents significant challenges with the primary challenge being the design and implementation of a framework to optimises the biases within the transition dyamics. Ensuring correct stationary distributions without relying on detailed balance presents challenges which require new frameworks controlling transition dynamics, with complexity rapidly escalating in higher dimensions. Although other literature has explored auxiliary techniques for optimising reversible MCMC algorithms—such as Geyer et al.'s [14] simulated annealing method, which in-

troduces temperature parameters to escape local minima— the direct optimisation of irreversible sampling remains relatively unexplored. The primary aim of this research is to address this gap by developing a robust, practical framework for breaking detailed balance using the skewness optimisation approach introduced by Turitsyn et al [15] in two or more dimensions. This model is particularly appealing for extension into higher dimensions due to its straightforward implementation. This approach enables asymmetric Markov chain transitions that inherently break reversibility while reliably preserving the stationary distribution. By extending this practical framework for skewness optimisation to higher-dimensional spaces, this study seeks to advance irreversible sampling methodologies, providing quantitative in depth performance comparisons with traditional reversible MCMC methods across a range of applications. In the subsequent sections, the necessary background theory underpinning the methods is outlined, followed by a description of the methods employed and their implementation. Finally, benchmark results and analysis of the approaches are presented, quantifying performance improvements over traditional reversible methods.

2 Background and theory

Markov chains are a class of stochastic processes where the future state of a system depends only on its present state [16], not its past history. This property of 'memorylessness' is known as the Markov property, and it allows us to model a wide range of phenomena, from weather patterns to financial markets and statistical physics [17]. Monte Carlo methods, on the other hand, are a class of computational techniques relying on repeated random sampling to obtain a target distribution [18]. By generating a large number of random samples and analysing their statistical properties, these methods can approximate a 'target' distribution that is otherwise difficult or impossible to solve analytically. Markov Chain Monte Carlo (MCMC) methods amalgamate the concepts of Markov chains with algorithmic techniques, such as the Metropolis-Hastings algorithm [19], which generates a sequence of samples by proposing a candidate sample based on the current state of the Markov chain and then accepting or rejecting the candidate based on an acceptance criterion derived from the target distribution.
In essence we construct a Markov chain whose stationary distribution converges to a desired target distribution. By simulating this chain for an adequate duration - the mixing time, one can recover the underlying distribution being studied.
A Markov matrix P is a stochastic matrix satisfying two important conditions. For a given transition matrix P over a discrete state space $\Omega = \{1, \ldots, N\}$:

1. **Non-Negativity:**

$$P_{ij} \geq 0, \quad \forall i, j \in \Omega, \tag{1}$$

   Every entry represents a probability and must be non-negative.

2. **Row-Stochasticity:**

$$\sum_j P_{ij} = 1, \quad \forall i, j \in \Omega, \tag{2}$$

   The sum of transition probabilities for any state i equals 1.

The conditions and assumptions allowed, and the strictness with which they are imposed can significantly influence the success or failure of a sampling method across different contexts. Aperiodicity removes the possibility of cyclical behaviour (see Figure 1). Ergodicity or irreducibility ensures that the chain can eventually reach any state from any other state in a finite time [16].

$$P^n(i, j) > 0 \tag{3}$$

Equation (3) expresses the condition of ergodicity, ensuring that it is possible to reach any state j from any state i in a finite number of steps n, with non-zero probability.

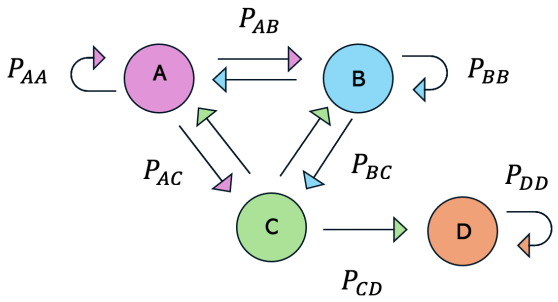


Figure 1: Diagram showing an example Markov chain which is inherently periodic. Once state D is reached, there is zero probability of escape. $P_{DD} = 1$, where P_{DD} is the transition probability from state D to state D.

Our first three assumptions (equation(1), equation(2), equation(3)) allow us to assume that the normalised long term probabilities across all states converge to the stationary distribution as the number of steps taken approaches infinity.
Overall we find an invariant target probability distribution $\pi = (\pi_1, \ldots, \pi_N)$ with $\pi_i > 0$ and $\sum_{i=1}^{N} \pi_i = 1$
The global balance condition is the most basic and fundamental of conditions, ensuring that, in a stationary state, the rate at which probability mass flows into a state i is equal to the rate at which it flows out. For a static probability distribution [20] we find that:

$$\pi_i = \sum_{j=1}^{N} \pi_j P_{ji}, \quad \forall i, j \in \Omega, \tag{4}$$

where:

- π_i: The stationary probability of being in state i.
- P_{ji}: The transition probability from state j to state i.

Samplers in the past have long relied on satisfaction of the detailed balance condition (DBC) to ensure convergence to the target distribution. The DBC is expressed as:

$$\pi_i P_{ij} = \pi_j P_{ji} \tag{5}$$

Detailed balance (equation(5)) is a more stringent condition [19], ensuring the probability flux between any two given states is equal. While always guaranteeing convergence, adhering to DBC will often limit the efficiency of sampling. In practice these conditions lead to computationally prohibitive convergence times. While the chain is still irreducible the time and computational cost associated with sheer number of calculations required make convergence to the static target distribution impossible in practice.

As previously stated, recently several algorithms have been proposed which break the detailed balance condition [8-13].

We focus on the model proposed by Turitsyn et al [15], which introduces a lifting framework combined with a skewed detailed balance condition (SDBC) to break the symmetry of the system and create an irreversible Markov chain. The asymmetry introduced is controlled, as shown in Figure 2. By introducing an additive variable $\epsilon \in \{+, -\}$ into the state space $I = \{1, 2, \ldots, S\}$, and lifting this space, the system's state space is effectively doubled (or expanded further, depending on the number of replicas). Under these conditions, a new balance condition can be derived from the global balance condition (equation (4)):

$$\pi_i P_{ij}^{+\varepsilon} = \pi_j P_{ji}^{-\varepsilon}, \tag{6}$$

The skewed detailed balance (equation (6)) ensures the overall conservation of probability flux within the system. The inter-replica transition probabilities Λ are adjusted to maintain the stochasticity condition allowing convergence to a steady-state distribution. The proposal algorithm must be changed to reflect these new transitions. An irreversible variation of the metropolis hastings algorithm must be used, handling the creation of random variables and the rejection/acceptance of new states.

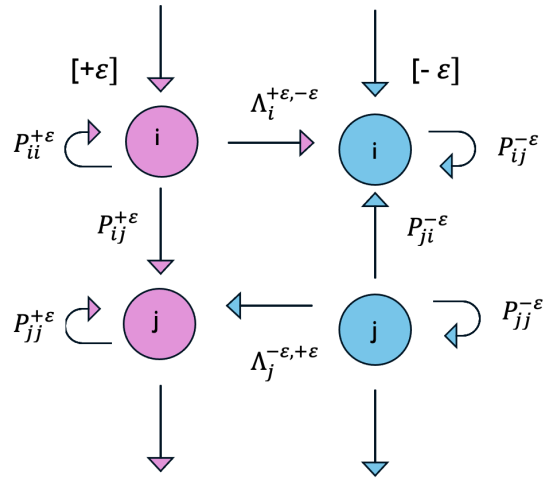


Figure 2: Diagram showing lifting variable and extension of the phase space

The newly allowed transitions between replicas, represented by $\Lambda^{(+,-)}$ and $\Lambda^{(-,+)}$, facilitate faster exploration of the state space. This allows the system to switch between different replicas, helping it avoid getting trapped in loops and local minima. The updated transition matrix, incorporating these inter-replica transitions, is given by:

$$\hat{P} = \begin{pmatrix} P^{(+)} & \Lambda^{(+,-)} \\ \Lambda^{(-,+)} & P^{(-)} \end{pmatrix}$$

Relaxation time is one of the primary metrics by which the performance of a Markov sampler can be quantified. Relaxation time describes the characteristic time it takes for the system to converge from an arbitrary initial distribution to its stationary distribution π [21]. A shorter relaxation time indicates faster convergence and thus a more efficient sampler, assessing the ability of a Markov chain to explore its own state space.

Relaxation time is often associated with mixing time, which measures how quickly the Markov chain reaches a state that is close enough to its stationary distribution, typically quantified in terms of the total variation distance (TVD). The TVD measures the difference between the current distribution of the Markov chain and the stationary distribution [22]:

$$\text{TVD}(P_t, \pi) = \frac{1}{2} \sum_i |P_t(i) - \pi(i)| \tag{7}$$

The mixing time depends on the threshold chosen to determine when the total variation distance between the current and stationary distributions is sufficiently small (often below a certain threshold, e.g., $\epsilon = 1/2$). At this point, the matrix is deemed to be 'mixed'. This threshold can be adjusted depending on the system and the desired accuracy.

$$t_{\text{mix}}(\epsilon) = \min \left\{ t : \max_x \left| P^t(x, \cdot) - \pi \right|_{\text{TV}} \leq \epsilon \right\} \tag{8}$$

The relaxation time is influenced by the spectral properties of the transition matrix, particularly the second-largest eigenvalue of the transition matrix λ_2. The larger the spectral gap, defined as $1 - |\lambda_2|$, the faster the chain mixes and thus the shorter the relaxation time [23]. The mixing time, T_{mix}, is typically proportional to

$$T_{\text{mix}} \sim \frac{1}{1 - |\lambda_2|} \tag{9}$$

3 Methodology

3.1 Summary

The core aim of this project is to develop a generalised framework extending the skewness optimisation method proposed by Turitsyn et al. [15] into higher-dimensional sampling settings.

It is important to note that in this study we explicitly assume knowledge of the target distribution for the purpose of developing, benchmarking, and testing the irreversible sampling optimisation methodology. This approach, while idealised, is essential for understanding theoretical properties, quantifying efficiency gains, and verifying the correctness of the optimisation algorithms. The primary objective is to establish methodological groundwork, providing a clear pathway for future work where these methods may be adapted to more realistic scenarios where the target distribution is only partially known or completely unknown.

While conceptually straightforward in one dimension, implementing skewness optimisation in two or more dimensions poses significant methodological challenges. These arise primarily due to the intricate geometry and anisotropic features of high-dimensional potential landscapes. Unlike the straightforward energy barriers of 1D systems, higher dimensions introduce complex ridges, competing transition pathways, and varying directional dependencies.

Such complexities make it significantly more difficult to implement a generalised optimisation approach that remains numerically stable, efficient, and free from biases that could distort the equilibrium distribution. Moreover, high-dimensional systems frequently feature metastable regions where samplers risk becoming trapped (topological freezing) making directional optimisation even more critical.

To systematically address these methodological challenges, our approach involves carefully selecting and constructing potential surfaces of increasing complexity.

This progressive increase in complexity enables systematic validation and clear attribution of performance improvements or degradations. By comparing irreversible and reversible sampling methods across these different systems, we quantify efficiency gains in key performance metrics such as mixing time, relaxation time, and autocorrelation. This structured, incremental approach ensures interpretable results and provides the groundwork necessary for the practical implementation of irreversible sampling methods in increasingly complex, realistic settings.

These conceptual insights into the difficulties and theoretical requirements directly inform our practical implementation approach, described in detail in Section 4. These insights will also inform the development of more advanced Skewed unoptimised and Hessian-based optimisation techniques, described subsequently in Section 4.1 and 4.2 respectively. In particular, the Hessian-based approach incorporates local geometric information to optimise directional sampling bias dynamically, thus providing a pathway to robustly addressing higher-dimensional sampling complexities.

3.2 Skewness Functions and Skewness Optimisation

Skewness optimisation [15] involves modifying transition rates to introduce a controlled directional bias. Given a transition probability matrix T_{ij}, a skewness function $\Theta_{ij}(\epsilon)$ is introduced to enforce irreversibility while maintaining normalised probability:

$$P_{ij}(\epsilon) = \Theta_{ij}(\epsilon) P_{ij} \tag{10}$$

In one dimension, the skewness function is straightforward since transitions occur along a single axis. The function Θ_{ij} is commonly defined to enhance transitions in one direction while suppressing the reverse:

$$\Theta_{ij}(\epsilon) = \begin{cases} 1, & i \geq j \\ 0, & \text{otherwise} \end{cases} \tag{11}$$

This unidirectional bias is effective in reducing the sampling time between barriers. This framework allows for controlled breaking of detailed balance

(equation (5)) while adhering to the skewed detailed balance condition (equation (6)) and therefore meeting the ergodicity requirements which ensured convergence in the reversible case.

3.3 Hessian-Based Skewness Optimisation

To address the insufficiency of a unidirectional skewness in higher dimensions, we employ an adaptive skewness function that dynamically adjusts based on the key points in the landscape.
The Hessian matrix of the potential function provides insight into the local curvature at a given point (x, y), helping to identify favourable escape directions. For a 2 dimensional system Hessian is defined as [24]:

$$H = \begin{bmatrix} \frac{\partial^2 V}{\partial x^2} & \frac{\partial^2 V}{\partial x \partial y} \\ \frac{\partial^2 V}{\partial y \partial x} & \frac{\partial^2 V}{\partial y^2} \end{bmatrix} \tag{12}$$

- V: The potential function.
- x, y: State space.

The $eigenvalues$ of H describe the curvature of the potential:

- $\lambda > 0$: Locally stable direction (resists movement)
- $\lambda < 0$: Unstable direction (preferred escape path)

The corresponding $eigenvectors$ indicate the principal axes of curvature. The eigenvector associated with the most negative eigenvalue represents the steepest escape direction from saddle points. See Figure 3.
By incorporating this information into the skewness function, we define a modified $\Theta_{ij}(\epsilon)$ that dynamically weights the skewness function based on the unstable eigenvector v_{unstable}:

$$\Theta_{ij}(\epsilon) = f(v_{\text{unstable}} \cdot r_{ij}) \tag{13}$$

where:

- r_{ij}: is the displacement vector between states i and j
- $f(\cdot)$ is a function that enhances transitions along the unstable eigenvector direction

This formulation ensures that transitions are biased along the most efficient escape routes with the aim of improving overall sampling efficiency.

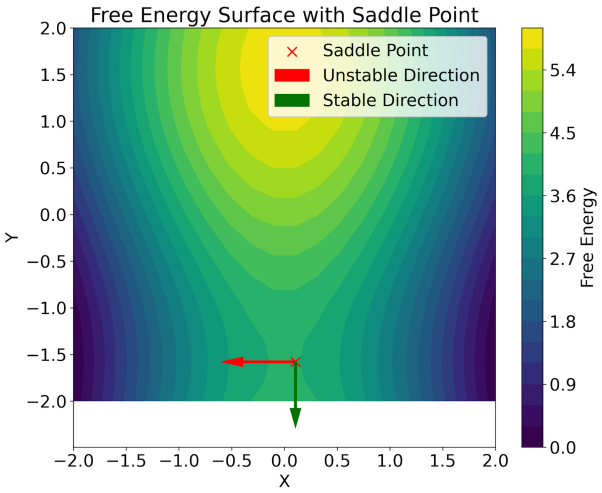


Figure 3: A free energy surface showing the stable and unstable directions at the saddle point, computed from the Hessian matrix.

4 Experimental Setup and Implementation

Selection of Potential Surfaces:
We choose a simple one-dimensional potential as a starting point for the analysis, investigating the basic effects and benefits of skewness optimisation and irreversibility in 1D. To allow for the possibility of future comparison between one- and two-dimensional analyses, we employ a double-well potential in both cases.
Subsequently, we extended this analysis to a simplified two-dimensional potential that varies only in one direction, allowing us to isolate and clearly understand the directional effects introduced by skewness. Finally, we explore a more complex two-dimensional potential landscape featuring multiple minima and saddle points, representative of realistic, challenging scenarios encountered in practical applications.
This structured incremental approach provides a controlled testing environment and allows for better interpretability of results.

Discretisation of Potentials and Rate Matrix Construction:
The double well potential used for 1D analysis is shown in Figure 4. States are created by discretizing the space along x. In the case of Figure 4 the system was discretised into n=40 states.

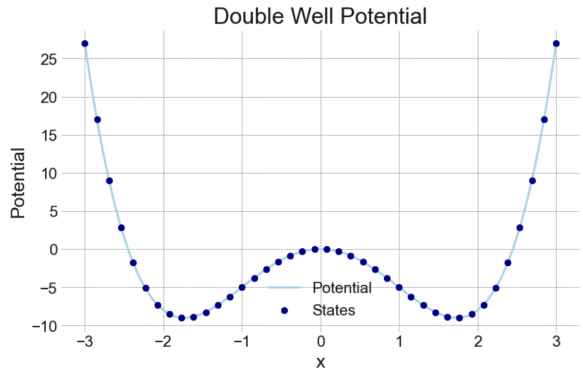


Figure 4: The 1D double well potential varying along x discretised into n = 40 states.

In 2D, potential surfaces were discretised into a finite number of states by averaging the potential within a chosen number of discretised grid cells. Initially, an 8×8 discretisation scheme is employed, as illustrated in Figure 5. This relatively coarse discretisation offers two primary advantages: simplifying the analysis by making the resulting rate matrix more interpretable and significantly reducing computational complexity during both sampling and optimisation processes.

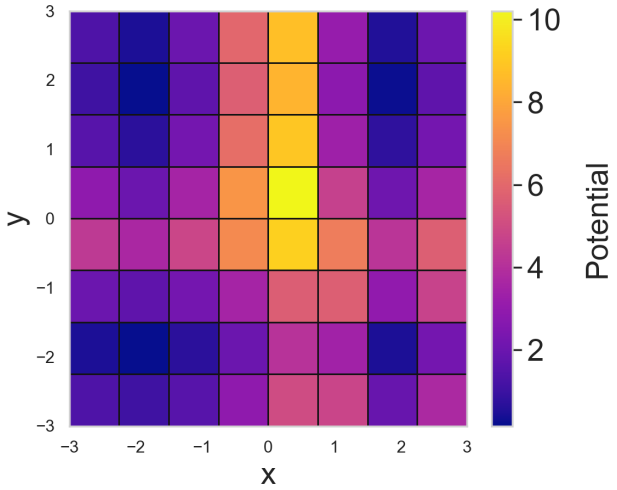


Figure 5: An 8×8 discretisation scheme for a representative 2D potential surface. The z-axis and colour bar indicate the potential values.

From these discretised potentials, reversible rate matrices are constructed according to the Metropolis–Hastings criterion. The rate R_{ij} represents how strongly the system wants to move from state i to j based on their energy difference, before adjusting these values to ensure they form valid probabilities that sum to 1.

$$R_{ij} = A \cdot \exp\left(\frac{V_i - V_j}{k_B T}\right) \tag{14}$$

where:

- R_{ij} is the rate from state i to j.
- A is a scaling pre-factor.
- V_i and V_j represent the potential energy at discrete states i and j, respectively.
- k_B is the Boltzmann constant.
- T is the system temperature.

These reversible rate matrices (built from equation (14)) inherently satisfy the detailed balance condition (equation(5)) and therefore guarantee correct equilibrium distributions. The reversible sampler provides reliable baselines for performance comparison in terms of both performance and accuracy metrics (more on this in section 5).

Efficient Implementation of Irreversibility and Skewness Optimisation: Introducing irreversibility by directly modifying all possible state transitions in the rate matrix would be computationally prohibitive due to the exponential scaling with the number of states. To overcome this, a computationally efficient structured skewness optimisation approach is implemented. Instead of assigning independent skewness parameters to each possible transition, a reduced set of parameters selectively perturbs only off-diagonal elements of the rate matrix corresponding to nearest-neighbour transitions. Fewer parameters are required, significantly reducing computational overhead. A binary encoding scheme of length $2N$, where N represents the number of discrete states, efficiently captures the skewness modifications. Each binary entry explicitly encodes whether a transition is biased forward or suppressed. Crucially, these adjustments are constrained to nearest-neighbour transitions, ensuring that the skewness modifications remain computationally feasible and scale more efficiently as discretisation becomes finer. This adjustment also simplifies any further optimisation, which is done with respect to this reduced set of parameters.

Computational Tools and Implementation Details: All computational work—including construction of rate matrices, implementation of sampling algorithms, and skewness optimisation—was conducted in Python. The codebase leverages several core scientific libraries for numerical computation, optimisation, and statistical analysis:

- `numpy.correlate` — for time-series correlation analysis
- `scipy.optimize` — for parameter fitting and optimization
- `statsmodels.api.tsa.acf` — for empirical autocorrelation computation

Additional tools from `scipy.linalg`, `matplotlib`, and `statsmodels` are used for matrix operations, trajectory simulation, and convergence diagnostics. A complete list of the modules used, along with important functions are included in the appendix under sections 1 and 2.

4.1 Unoptimised Skewness Implementation in 2D

In this implementation, the irreversible sampler is constructed by breaking detailed balance through the introduction of a controlled skewness into the transition rates. Recall that the core idea is to add a directional push to the dynamics so that the Markov chain can more efficiently overcome energy barriers and escape metastable regions.

Directional Bias via Grid Partitioning. The implementation leverages the 2D grid structure by applying coordinate-based transition biases:

- **Horizontal Bias:** States in the left half receive enhanced rightward rates and suppressed leftward rates, while the right half experiences the opposite bias.
- **Vertical Bias:** States in the top half are biased downward, while bottom half states are biased upward.

This fixed, cyclic bias is encoded in a rate modification matrix with skewness parameter δ:

$$\Theta_{ij} = \begin{cases} 1 \pm \delta, & \text{for horizontal transitions} \\ 1 \pm \delta, & \text{for vertical transitions} \end{cases} \tag{15}$$

where the sign depends on position relative to grid centre.

Doubled State Space Approach. To maintain the correct equilibrium distribution while introducing irreversibility, a doubled state space technique is employed as described in section 2. The rate matrix is expanded to twice its original size, creating mirror copies of each state. Transitions

within each copy preserve the skewed rates, while additional transitions between copies balance the net probability flows. This ensures that despite breaking detailed balance locally, the overall dynamics still converges to the target equilibrium distribution, while benefiting from the enhanced sampling efficiency of non-reversible dynamics. The results from this model are available in section 6.3.

4.2 Hessian-Guided Skewness Implementation in 2D

To further enhance sampling efficiency in high-dimensional potential landscapes, we implement a Hessian-guided adaptive skewness [24] strategy that leverages the local geometric structure of the potential surface—specifically the curvature information captured by the Hessian matrix—to introduce directional sampling bias that accelerates escape from metastable regions.

Saddle points are first located via numerical optimisation, using a gradient-norm minimisation routine to identify regions where the potential exhibits both stable and unstable curvature. At each identified saddle point, the local Hessian matrix is evaluated, and its eigen-decomposition is computed. The eigenvector corresponding to the most negative eigenvalue is then designated as the unstable mode, representing the energetically favourable escape direction across that saddle.

To construct the skewness field, we define a continuous, position-dependent weighting function that aligns each possible transition vector $\mathbf{r}_{ij}$ with the nearest saddle's unstable direction. Transitions aligned with this direction are preferentially enhanced, while opposing transitions are correspondingly suppressed. The influence of a given saddle on each transition is modulated by a Gaussian-like proximity weighting, ensuring that only nearby saddles influence the bias, thereby maintaining spatial locality and computational efficiency.

Irreversibility is introduced by applying these directional weights multiplicatively to the off-diagonal elements of a base reversible rate matrix. Importantly, the total outflow rate from each state is preserved by re-normalising the diagonal elements accordingly. This ensures that while detailed balance is locally broken, the global balance condition is retained.

The resulting skewed transition matrix therefore introduces irreversible flows that accelerate mixing without distorting equilibrium behaviour. This method is therefore specifically designed for potentials featuring saddle regions, where sampling efficiency is otherwise limited by high energy barriers. Performance of this sampler is evaluated in Section 6.4.

5 Metrics and Quantifying Sampling Performance

To systematically assess and compare the efficiency and accuracy of different sampling methods, several quantitative metrics will be employed. Each metric addresses distinct aspects of sampling performance and will be evaluated separately to provide comprehensive insights into sampler efficiency and accuracy.

Mean First Passage Time (MFPT) measures the expected number of steps required for a sampler to transition from a specified initial state to a defined target state. The MFPT will be calculated numerically using np.linalg.solve, and practically by direct mean sampling measurement. Lower MFPT values indicate improved sampler performance in efficiently traversing energy barriers and metastable states.

Accurately reproducing the stationary distribution is a fundamental requirement for any sampling method. To verify this, we compare the stationary distribution obtained from a sampling run (numerically computed) against the known analytical target distribution derived from the potential landscape. To quantify the discrepancy between these two distributions, we use the Kullback–Leibler (KL) divergence — a standard measure in statistics that captures how one probability distribution differs from another. Conceptually, it can be understood as the average difference in log-likelihood between the two distributions. A smaller KL divergence value indicates that the

sampled distribution closely matches the expected theoretical distribution, confirming that the sampling scheme has converged accurately without introducing significant bias.

The frequency of transitions between metastable states or key regions will be quantified through transition counts. These counts provide insights into the sampler's ability to escape local minima and explore different regions of the potential landscape. Additionally, heatmaps visualising the probability of visiting each state across sampled trajectories will be produced. These heatmaps enable direct visual assessment of exploration patterns, potential trapping scenarios, and overall sampling coverage.

Autocorrelation time measures how quickly a sampled trajectory becomes statistically independent from its past states. It will be quantified by computing the autocorrelation function (ACF) for relevant state variables (e.g., spatial coordinates). A shorter autocorrelation time indicates more efficient decorrelation between successive samples.

Mixing time will be estimated by identifying the lag at which the ACF first falls below a predefined threshold (e.g., 0.1), signifying effective decorrelation. Autocorrelation, relaxation times, MFPT and trajectory analysis will be compared across reversible and irreversible samplers to demonstrate improvements in sampling efficiency.

6 Results and Discussion

6.1 1D results

We begin our investigation with one-dimensional (1D) systems. The following results highlight the advantages of irreversible sampling and review the scaling behaviour of performance.

The relaxation time was measured for four separate samplers on the double well potential in table 1 for n=40 states.

Sampler	T_{relax} **(s)**	**MFPT (s)**
Reversible	337.6	1524
Irreversible	141.9	654
Optimised (RT)	55.43	544.4
Optimised (MFPT)	66.2	356.3

Table 1: Comparison of Relaxation Time and MFPT for Different Sampling Methods for n=40 states

The three irreversible models differ primarily in how the underlying transition rates are skewed and optimised. The reversible sampler is built directly from the potential, with the rate matrix constructed to satisfy detailed balance. In contrast, the Default Irreversible sampler takes this reversible rate matrix and applies a preset modification to break detailed balance, by scaling the forward rates by a factor (and adjusting the backward rates to maintain a form of skewed detailed balance) without any further tuning. This default modification introduces a basic level of asymmetry, or "skewness," in the transition rates.

The Optimised 1 (RT) and Optimised 2 (MFPT) further refine this asymmetry by optimising the skewness parameters with respect to a chosen loss function. In these cases, a set of adjustable reduced parameters defines the rate-modification matrix, and optimisation (via scipy) can minimse either the second eigenvalue (thereby reducing the relaxation time) or the mean first passage time (MFPT) of the system. The optimised models fine-tune this skewness to achieve specific dynamical objectives—resulting in irreversible processes that are better tailored to either rapid equilibration (optimised for relaxation time) or faster state transitions (optimised for MFPT). In practice one can optimise the skewness parameters with respect to any chosen metric, or loss function.

Table 1 compares the performance of four samplers on a one-dimensional double-well potential with n=40 discrete states. The samplers include a standard reversible implementation, a basic irreversible variant, and two optimised irreversible samplers—each targeting different objectives: relaxation time (RT) and mean first passage time (MFPT), respectively.

The reversible sampler exhibits the slowest performance, with a relaxation time of 337.6 seconds and a mean first passage time (MFPT) of 1524 seconds. Introducing irreversibility via a simple, unoptimised skewness function reduces both metrics substantially—to 141.9 seconds and 654 seconds,

respectively—demonstrating the immediate benefits of relaxing detailed balance.

Further improvements are achieved by explicitly optimising the skewness parameters. The sampler optimised for relaxation time achieves a fivefold reduction in relaxation time, while the MFPT-optimised sampler reduces the first passage time by nearly 77 percent compared to the reversible case. Interestingly, both optimised samplers outperform the unoptimised irreversible version across both metrics, despite being tuned for a single one—suggesting that skewness optimisation introduces robustness and general efficiency gains.

Figure 6 shows a plot for autocorrelation against time step, providing some insight into the rate at which these samplers are able to decorrelate from their previous states.

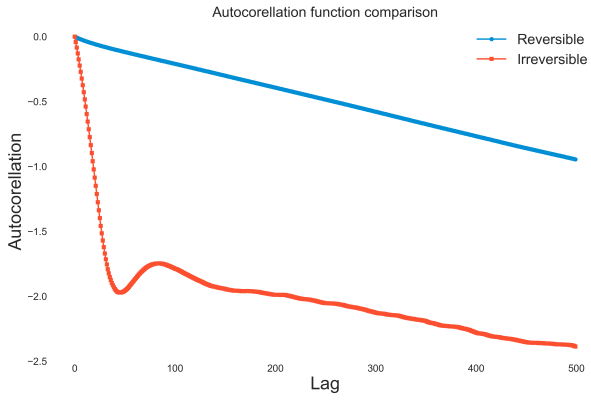


Figure 6: Plot showing falloff of Autocorrelation function with time step for n=40, for reversible and non optimised irreversible model

The skewness function significantly reduces the autocorrelation in the Markov chain, indicating a faster loss of memory and improved mixing efficiency. This effect is particularly evident in the relative accelerated decay of the autocorrelation function seen in Figure 6. One possible explanation for this behaviour is that the unidirectional nature of the skewness function reduces the likelihood of backtracking, encouraging more directed movement through the state space. By limiting returns to recently visited states, the chain more effectively explores new regions, resulting in a greater number of approximately independent samples over fewer iterations.

Figure 7 shows how the relaxation time of the reversible and irreversible samplers scales with the number of states (n).

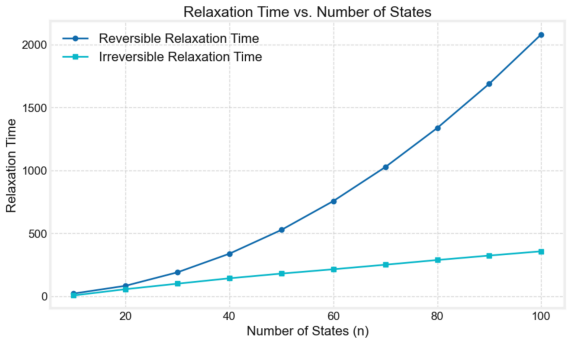


Figure 7: Plot showing how the relaxation time of the reversible and irreversible samplers scales with the number of states (n) for 1D double well potential.

Both reversible and irreversible relaxation times increase as the number of states grows. This increase is expected due to finer discretisation of the state space.

The reversible model consistently exhibits a significantly longer relaxation time compared to its irreversible counterpart for all n. The relaxation time of the reversible model appears to scale at least quadratically, consistent with $\mathcal{O}(n^2)$ growth, and may follow a higher-order polynomial trend over the measured range. In contrast, the irreversible relaxation time displays approximately linear scaling with $\mathcal{O}(n)$ behavior, indicating more efficient convergence as the state space resolution increases.

The observed disparity in the scaling behaviour of relaxation times highlights a substantial efficiency advantage of incorporating irreversibility into Markov state model (MSM)-based sampling schemes. This distinction is particularly relevant in applications where a high-resolution discretisation of the state space is required, such as in applications in statistical mechanics and molecular dynamics, where fine-grained sampling resolution is necessary to accurately capture rare events and transition pathways.

Extrapolating the observed trends to $n = 1000$, a regime relevant for such high-resolution models, suggests that 1D irreversible MSMs can achieve equivalent convergence rates with approximately 45–60 times greater efficiency compared to their reversible counterparts.

6.2 2D results - Unidirectional skewness

To evaluate the effects of directional skewness in a controlled setting, we first consider a simple two-dimensional potential that varies in only one direction while remaining constant in the orthogonal direction. This scenario provides a minimal yet informative test case, allowing clear quantification of performance gains attributable solely to directional variation of skewness. It is appropriate to think of directional skewness as applying a push in that direction in order to aid with sampling.

The following results are obtained from sampling on a Gaussian wall, shown in Figure 8 in its continuous (left) and discretised (right) forms. Note that this potential varies in y but is constant along the x direction.

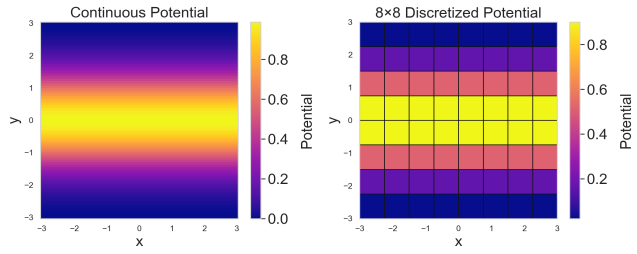


Figure 8: Diagram showing a continuous (left) and the 8x8 discretised scheme (right) for the Gaussian wall potential, where the z-axis (colour bar) shows the value of the potential.

In our implementation, the skew is parametrised by the angle θ via the projection of each displacement vector $\mathbf{d}$ onto the unit vector $\mathbf{u}(\theta) = (\cos\theta, \sin\theta)$, where θ is angle made between the y axis and direction of skewness. The skew factor is then defined by

$$s(\theta) = \text{skew}^{(\mathbf{d}\cdot\mathbf{u}(\theta))},$$

so that transitions aligned with $\mathbf{u}(\theta)$ (yielding a positive dot product) are enhanced, while those opposed (yielding a negative dot product) are suppressed.

By applying skewness in various directions, improvements in the relaxation time were measured. The relaxation time observed from a simple unidirectional skew are compared with the reversible relaxation time in table 2.

Sampler	Relaxation Time / seconds
Reversible	$T_{\text{rev}} = 0.468549$
Y-Skew Irreversible	$T_{\text{y-skew}} = 0.218169$
X-Skew Irreversible	$T_{\text{x-skew}} = 0.468549$
Bidirectional Skew	$T_{\text{bi-skew}} = 0.218112$

Table 2: Effect of Directional Skewness on Relaxation Time for 2D Gaussian Wall Potential

As seen from the results in table 2, even for a trivial potential, when the skewness function is applied along the directional perpendicular to the potential wall which coincides with the direction of greatest change sampling, is improved, and states are explored faster.

Otherwise, when the skewness is applied along a direction of constant potential, there is no improvement in performance, and the relaxation time mirrors that of the reversible sampler. Figure 9 shows how the relaxation time, which has been parametrised in terms of theta, varies with the directional skewness imposed.

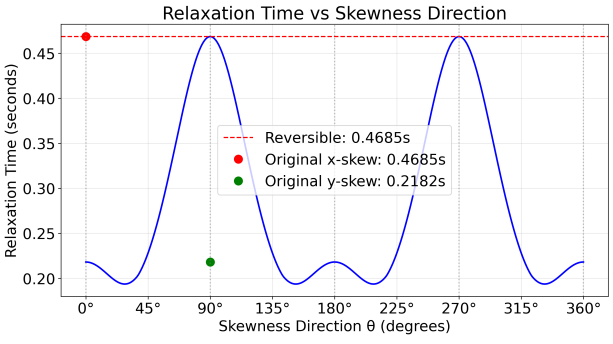


Figure 9: Plot showing the variation of relaxation time with skewness direction angle . The minima at 0° and 180° correspond to alignment with the x-axis.

While, in theory, the relaxation time is expected to exhibit symmetric minima at both $0°$ and $180°$, our simulations reveal a slight asymmetry. This discrepancy arises from discretisation artifacts inherent in the finite grid resolution, which subtly break the rotational symmetry of the potential

These observations concur with how we would expect the 1D skewness function to generalise to higher dimensions. Note importantly that in this the case of this simple potential the sampling performance depends directly on the direction in which the skewness is applied relative the direction of greatest change in the potential.

Crucially, these results offer both theoretical and empirical justification for applying the skewness function in a directionally selective manner in higher-dimensional systems. They provide a strong foundation for more sophisticated applications, where skewing can be guided by local geometric information, such as the Hessian or principal curvature directions.

6.3 2D results - Unoptimised

In this section, we analyse the performance of the unoptimised irreversible sampler applied to a two-dimensional double-well potential, as described in Section 4.1. Specifically, we evaluate both the relaxation time and the mean first passage time (MFPT) between wells, serving as key metrics for sampler efficiency.

We denote the grid resolution by N, such that the total number of discrete states is $n = N^2$. The discretisation begins at $N = 10$ (i.e., $n = 100$), which was empirically determined to be the minimum resolution necessary to capture the relevant features of the double-well landscape.

The *skewness factor* is a tunable parameter within the simulation that modulates the strength of the irreversible perturbation. This parameter, set to a default value of 0.7 arbitrarily for the following experiments, was introduced to systematically evaluate the impact of skew-induced irreversibility on sampling dynamics.

Sampler	T_{relax} **(s)**	**MFPT (s)**
Reversible	41.6565	1220.7532
Irreversible	20.4025	976.6026

Table 3: Relaxation time and MFPT for reversible and unoptimised irreversible samplers on a 2D double-well potential with $n = 400$ states and skewness factor = 0.7.

Sampler	T_{relax} **(s)**	**MFPT (s)**
Reversible	162.9139	4831.7974
Irreversible	62.9264	3865.4379

Table 4: Relaxation time and MFPT for reversible and unoptimised irreversible samplers on a 2D double-well potential with $n = 1600$ states and skewness factor = 0.7.

To evaluate scalability, we examine how relaxation time varies with the number of discrete states n.

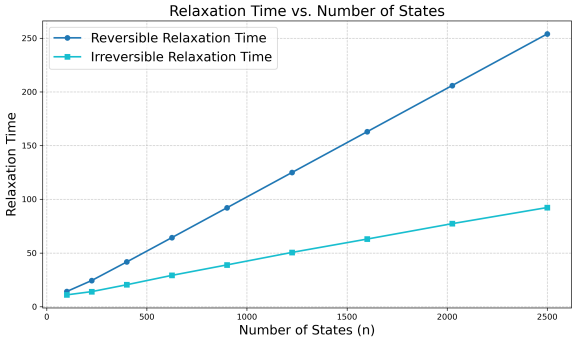


Figure 10: Plot showing how the relaxation time of the reversible and irreversible samplers scales with the number of states (n) for the 2D double-well potential.

From Figure 10 we observe that, in contrast to the 1D case, the reversible and irreversible samplers exhibit the same asymptotic scaling order with respect to n, but with different gradients. This suggests that while both samplers scale linearly with the number of states, the irreversible sampler achieves a consistently lower relaxation time.

A similar trend is observed when examining the mean first passage time (MFPT) across the well as shown in Figure 11.

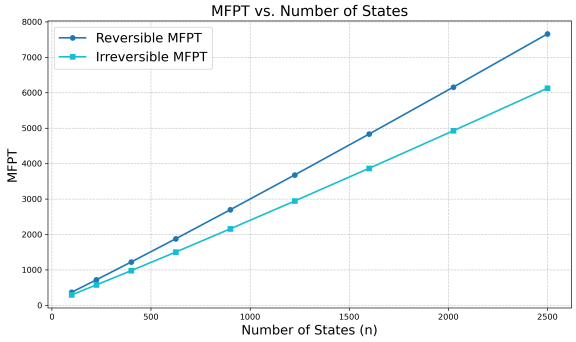


Figure 11: Mean first passage time (MFPT) across the well as a function of the number of states (n) for both reversible and irreversible samplers for the 2D double-well potential.

The linear dependence observed in both relaxation time and MFPT suggests that the computational complexity of both the reversible and unoptimised irreversible samplers scales as $\mathcal{O}(n)$. However, the irreversible sampler consistently demonstrates improved performance, as indicated by a smaller scaling coefficient.

To further investigate the impact of skewness on performance, we generated samplers with varying skewness factors and again plotted the relaxation time against the number of states n.

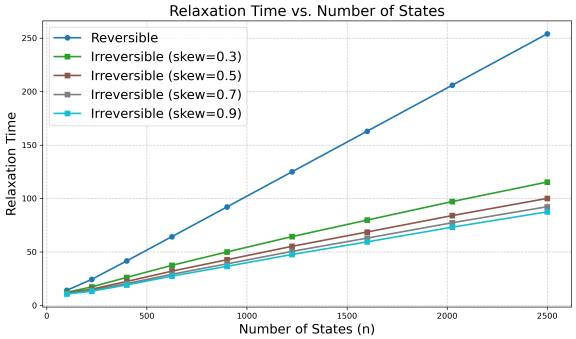


Figure 12: Relaxation time as a function of the number of states (n) for irreversible samplers with varying skewness factors for the 2D double-well potential.

Figure 12 shows that for all tested skewness values, the scaling order remains linear, $\mathcal{O}(n)$, while the gradient decreases with increased skewness. However, the benefit from increasing skewness diminishes beyond a certain point, indicating a regime of diminishing returns.

6.4 2D results - Hessian skewness

To enhance sampling efficiency in more complex two-dimensional (2D) energy landscapes, we employ a Hessian-guided adaptive skewness strategy as introduced in Section 4.2.

Unlike simpler skewing approaches, where asymmetry is applied along fixed grid-aligned directions (e.g., horizontal or vertical), the Hessian-guided skewness imposes directional bias along arbitrary orientations dictated by the local geometry of the potential landscape. As a result, this scheme is inherently better suited to continuous state spaces. When implemented on a discrete lattice, particularly at coarse resolutions (i.e., small n), the limited directional granularity leads to numerical instability and diminished performance gains. Therefore, a sufficiently fine discretisation (large n) is required to faithfully capture the advantages of this method. For the following benchmarks a value of n = 10,000 was chosen.

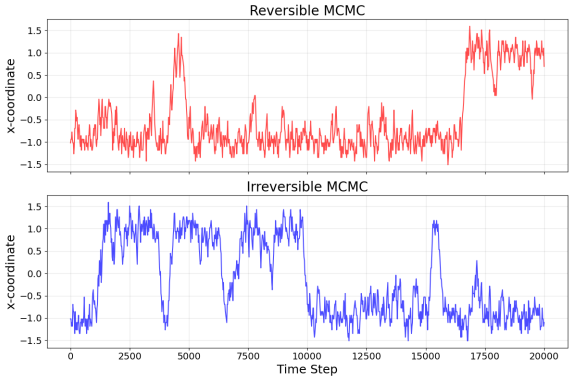


Figure 13: X-coordinate trajectories sampled using reversible (top) and Hessian-guided irreversible (bottom) schemes, showing more frequent barrier crossings in the irreversible case with discretisation : n = 10,000

Figure 13 shows representative trajectories of the x-coordinate under both reversible and Hessian-guided irreversible sampling schemes. The irreversible method enables more persistent and directed motion, facilitating improved exploration of metastable wells and more frequent transitions across energy barriers.

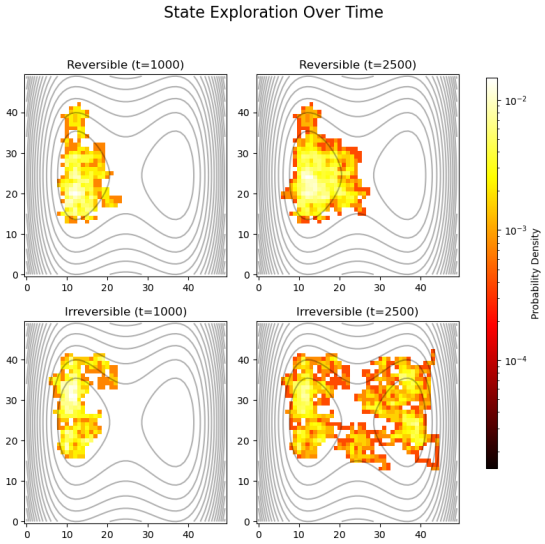


Figure 14: Comparison of state visitation frequencies between reversible (top) and Hessian-guided irreversible (bottom) samplers on the 2D double-well potential with discretisation : n = 10,000

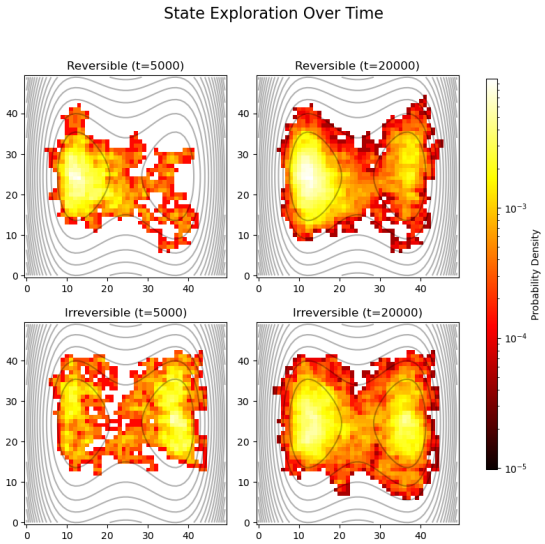


Figure 15: Comparison of state visitation frequencies between reversible (left) and Hessian-guided irreversible (right) samplers on the 2D double-well potential with discretisation : n = 10,000

Figures 14 and 15 show heat maps of state visitation frequencies averaged over numerous trajectories. The irreversible sampler achieves broader and more uniform spatial coverage, indicating faster convergence toward the equilibrium distribution.

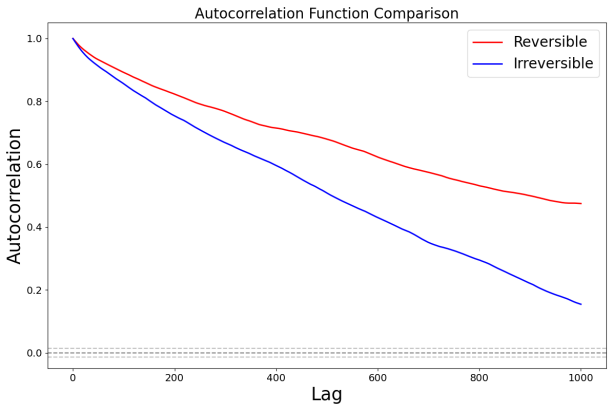


Figure 16: Autocorrelation function (ACF) comparison between reversible and Hessian-guided irreversible sampling, showing faster decorrelation with the irreversible method with discretisation : n = 10,000

Figure 16 presents the autocorrelation function (ACF) along the x-axis, which corresponds to the direction of the principal energy barrier. The Hessian-guided scheme exhibits significantly faster decorrelation, as evidenced by the sharper ACF decay and reduced integrated autocorrelation time, thereby confirming increased sampling efficiency.

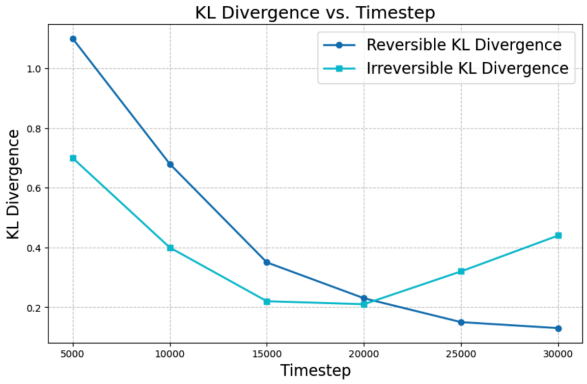


Figure 17: log KL divergence comparison between reversible and Hessian-guided irreversible sampling, showing divergence of stationary distribution with the irreversible method with discretisation : n = 10,000

Figure 17 shows the KL divergence over time for both the reversible and Hessian-guided irreversible samplers applied to a two-dimensional double-well potential. As expected from previous

results, the irreversible sampler initially converges more rapidly, achieving lower KL divergence by around t=15,000-20,000. However, unlike the reversible method—which continues to decrease smoothly—the irreversible sampler exhibits a reversal in performance beyond this point, with KL divergence increasing again over time. This behaviour suggests the onset of numerical instability. A potential contributing factor is the discretisation of the state space: the Hessian-guided skewness relies on aligning transition asymmetry with continuous curvature directions derived from unstable eigenvectors. When implemented on a discrete grid, particularly with misaligned or coarse resolution, this can introduce small but accumulating inconsistencies in the transition probabilities, leading to deviations from the true target distribution. Importantly despite its seemingly faster sampling, the Hessian-optimised irreversible sampler does not converge to the correct target distribution in the large-step limit, rather it demonstrates diverging behaviour. (Plots showing the long term convergence under this scheme are included in the appendix in section 3)

7 Conclusion

To summarise the main findings of this investigation, in 1D even the unoptimised irreversible sampler outperformed the reversible one, reducing both MFPT and relaxation time by over 50%. When skewness was optimised for specific objectives (e.g., relaxation time or MFPT), efficiency gains increased dramatically. These results confirmed the feasibility and flexibility of skewness-based irreversibility and highlighted the further gains possible via further optimisation of skewness parameters.

Moving to two-dimensional systems, we sampled a potential which varied only in one direction, isolating directional aspect of skewness. The results demonstrated a clear dependence on alignment between skew direction and the direction of greatest potential change. When correctly aligned, relaxation times were reduced significantly; when misaligned (e.g., skew applied along a direction of constant potential), no benefit was observed. The results from this section give two new key insights: Firstly for a symmetric potential increasing skewness along a direction of no interest (i.e the direction where potential is constant) does not deteriorate sampling performance beyond that of the reversible case. Secondly, the skewness optimisation method proposed by Turitsyn et al. [15] can be applied effectively in 2D. This result is particularly important as it led to the development of the two samplers presented in sections 4.1 and 4.2, for which results are shown in 6.3 and 6.4.

In the 2D double-well system, the unoptimised irreversible sampler again outperformed their reversible counterpart. Importantly, both the MFPT and relaxation time scaled linearly with the number of states n, indicating that the efficiency gains from irreversibility increase as the system size increases. However, in contrast to the 1D case—both samplers in 2D showed similar asymptotic scaling order, differing only in their gradient. Note therefore that for the irreversible regime employed in 6.3 - the relative advantage of irreversibility is more pronounced in one-dimensional systems with 2D yielding a reduction in the pre-factor rather than a change in the asymptotic complexity. Measurements with varying skewness factors in 2D revealed diminishing returns beyond a certain threshold.

A sampler incorporating hessian-guided adaptive skewness method was developed, leveraging local curvature information to dynamically set the skewness direction. This method significantly accelerated exploration around saddle points. However, numerical instability and deviations from the equilibrium distribution were observed in the long-term limit. We suspect these issues arose due to discretisation constraints. This highlighted a key challenge : the extension of these methods to discrete, grid-based implementations in higher dimensions requires careful attention to numerical stability, especially when continuous dynamics are approximated with finite steps.

Future Work: Several promising avenues can be explored to build upon this work. First, implementing continuous state space sampler or interpolation techniques could improve the repre-

sentation of curvature-based (i.e Hessian) dynamics within discretised state spaces. Additionally, incorporating adaptive correction mechanisms or re-weighting schemes would enhance equilibrium accuracy while maintaining efficiency gains. Extending the skewness optimisation framework to higher-dimensional spaces (3D and beyond) would be highly beneficial, possibly necessitating algorithmic advancements to efficiently handle complex potential curvature. Finally, applying these irreversible MCMC methods to real-world scenarios—such as atomic charge distribution simulations - where target distributions are not known in advance, but can be easily verified - offers a critical step towards practical implementation.

Overall Findings and Reflections:

This study extends the skewed detailed balance framework initially proposed by Turitsyn et al. [15] to higher-dimensional settings, addressing a critical gap in the literature. To our knowledge, no previous research has explored the performance and benefits of irreversible Markov Chain Monte Carlo (MCMC) samplers beyond one-dimension.

Our results clearly demonstrate that introducing irreversibility consistently enhances sampling efficiency across both one-dimensional and two-dimensional potential landscapes. Importantly, even simple, unoptimised irreversible methods significantly outperform traditional reversible samplers, underscoring the practical utility and potential for broader adoption of these techniques in multidimensional sampling scenarios.

This project lays a foundation for practical and scalable irreversible samplers in 2D. Future extensions that integrate adaptive control, continuous approximations, and higher-dimensional generalisations will further improve robustness and applicability, helping irreversible MCMC become a viable standard for tackling the most challenging sampling problems in modern computational science.

References

[1] Metropolis, F., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. (1953). Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(6), 1087–1092.

[2] Bonanno, C., Nada, A., & Vadacchino, D. (2024). Mitigating topological freezing using out-of-equilibrium simulations. *Journal of High Energy Physics*, 2024(4).

[3] Alles, B., Boyd, G., D'Elia, M., Di Giacomo, A., & Vicari, E. (1996). Hybrid Monte Carlo and topological modes of full QCD. *Physics Letters B*, 389(1), 107–111.

[4] Del Debbio, L., Manca, G. M., & Vicari, E. (2004). Critical slowing down of topological modes. *Physics Letters B*, 594(3–4), 315–323.

[5] Schaefer, S., Sommer, R., Virotta, F., & the ALPHA Collaboration. (2011). Critical slowing down and error analysis in lattice QCD simulations. *Nuclear Physics B*, 845(1), 93–119.

[6] McGlynn, G., & Mawhinney, R. D. (2014). Diffusion of topological charge in lattice QCD simulations. *Physical Review D*, 90(7), 074502.

[7] Bonati, C., & D'Elia, M. (2018). Topological critical slowing down: Variations on a toy model. *Physical Review E*, 98(1), 013308.

[8] Diaconis, P., Holmes, S., & Neal, R. M. (2000). Analysis of a nonreversible Markov chain sampler. *Annals of Applied Probability*, 10(3), 726–752.

[9] Bernard, E. P., Krauth, W., & Wilson, D. B. (2009). Event-chain Monte Carlo algorithms for hard-sphere systems. *Physical Review E*, 80(5), 056704.

[10] Suwa, H., & Todo, S. (2010). Markov chain Monte Carlo method without detailed balance. *Physical Review Letters*, 105(12), 120603.

[11] Schram, R. D., & Barkema, G. T. (2015). Monte Carlo methods beyond detailed balance. *Physica A: Statistical Mechanics and its Applications*, 418, 88–93.

[12] Kaiser, M., Jack, R. L., & Zimmer, J. (2017). Acceleration of convergence to equilibrium in Markov chains by breaking detailed balance. *Journal of Statistical Physics*, 168(2), 259–287.

[13] Ichiki, A., & Ohzeki, M. (2013). Violation of detailed balance accelerates relaxation. *Physica D: Nonlinear Phenomena*, 88(1), 012124.

[14] Geyer, C. J., & Thompson, E. A. (1995). Annealing Markov chain Monte Carlo with applications to ancestral inference. *Journal of the American Statistical Association*, 90(431), 909–920.

[15] Turitsyn, K. S., Chertkov, M., & Vucelja, M. (2011). Irreversible Monte Carlo algorithms for efficient sampling. *Physica D: Nonlinear Phenomena*, 240(4–5), 410–414.

[16] Kemeny, J. G., & Snell, J. L. (1976). *Finite Markov Chains*. New York: Springer-Verlag.

[17] Brooks, S., Gelman, A., Jones, G. L., & Meng, X.-L. (Eds.). (2011). *Handbook of Markov Chain Monte Carlo*. CRC Press.

[18] Rubinstein, R. Y., & Kroese, D. P. (2017). *Simulation and the Monte Carlo Method* (3rd ed.). Wiley.

[19] Metropolis, F., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. (1953). Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(6), 1087–1092.

[20] Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1), 97–109.

[21] Lovász, L., & Simonovits, M. (1993). Random walks in a convex body and an improved volume algorithm. *Journal of Computer and System Sciences*, 47(3), 384–403.

[22] Levin, D. A., Peres, Y., & Wilmer, E. L. (2009). *Markov Chains and Mixing Times*. Providence, RI: American Mathematical Society.

[23] Sinclair, A. (1992). Improved bounds for mixing rates of Markov chains and multicommodity flow. *SIAM Journal on Computing*, 21(4), 721–746.

[24] Balboni, D., & Bacciu, D. (2024). ADLER – An efficient Hessian-based strategy for adaptive learning rate. In *Proceedings of the 2024 Conference on Machine Learning* (pp. 131–136).

Appendix

1. Modules

- numpy
- matplotlib.pyplot
- statsmodels.api.tsa.acf
- scipy.linalg
- scipy.interpolate
- scipy.constants
- scipy.optimize
- statsmodels.api
- matplotlib.pyplot
- pandas
- seaborne
- time

2. Code - Important Functions

```python
def relaxation_time(rate_matrix):
    """
    Calculate relaxation time from the rate matrix (inverse of the spectral gap).
    """
    eigenvalues = linalg.eigvals(rate_matrix)
    eigenvalues = sorted(eigenvalues, key=lambda x: np.real(x), reverse=True)
    return 1.0 / np.abs(np.real(eigenvalues[1]))
```

A function that calculates the relaxation time of a Markov process by computing the eigenvalues of the rate matrix and returning the inverse of the spectral gap (the difference between the largest and second-largest eigenvalues

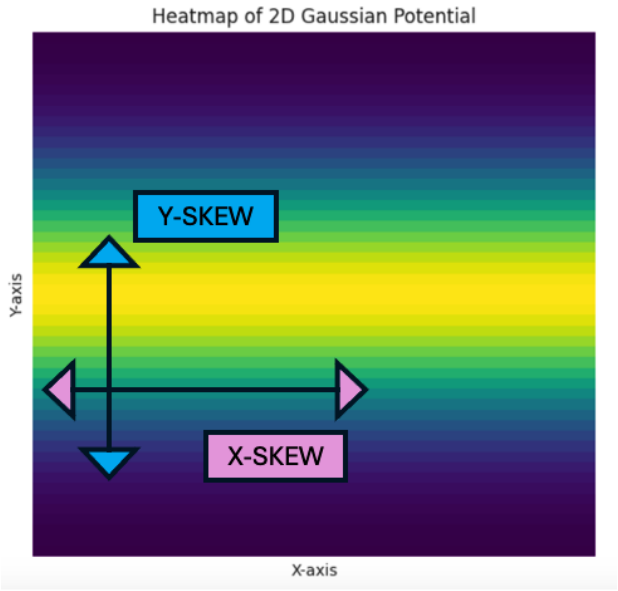
```python
def construct_rate_modification(num_states, skewness=0.5, cyclic_flow=True):
    """
    Construct a rate modification matrix for irreversible dynamics.
    """
    n_states = num_states**2
    rate_mod = np.ones((n_states, n_states), dtype=np.float64)
    for i in range(num_states):
        for j in range(num_states):
            idx = i * num_states + j
            x = -2 + 4 * j / (num_states - 1)
            y = -2 + 4 * i / (num_states - 1)
            if j < num_states - 1:
                right_idx = i * num_states + (j + 1)
                if j < num_states // 2:
                    rate_mod[idx, right_idx] = 1 + skewness
                    rate_mod[right_idx, idx] = 1 - skewness
                else:
                    rate_mod[idx, right_idx] = 1 - skewness
                    rate_mod[right_idx, idx] = 1 + skewness
            if i < num_states - 1:
                down_idx = (i + 1) * num_states + j
                if i < num_states // 2:
                    rate_mod[idx, down_idx] = 1 + skewness
                    rate_mod[down_idx, idx] = 1 - skewness
                else:
                    rate_mod[idx, down_idx] = 1 - skewness
                    rate_mod[down_idx, idx] = 1 + skewness
    return rate_mod
```

A function that constructs a rate modification matrix for introducing directional bias in irreversible dynamics. It creates a matrix that enhances transitions in one direction while suppressing them in the opposite direction based on a skewness parameter and spatial position.

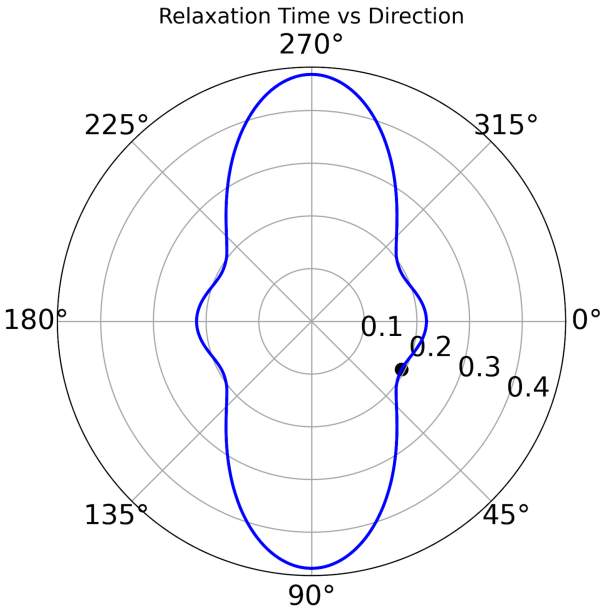
```python
def create_irreversible_rate_matrix(rate, rate_mod, peq):
    """
    Create an irreversible rate matrix (doubled state space) that satisfies skewed detailed balance.
    """
    n_states = rate.shape[0]
    irrev_rate = np.zeros((2 * n_states, 2 * n_states))
    # Upper-left block: copy and modify reversible rate matrix.
    for i in range(n_states):
        for j in range(n_states):
            if i != j:
                irrev_rate[i, j] = rate[i, j] * rate_mod[i, j]
                if peq[j] > 1e-10:
                    irrev_rate[j + n_states, i + n_states] = irrev_rate[i, j] * (peq[i] / peq[j])
    # Connect the two copies to balance net flows.
    for i in range(n_states):
        irrev_rate[i, i] = 0.0
        irrev_rate[i + n_states, i + n_states] = 0.0
        sum_orig = np.sum(irrev_rate[i, :n_states])
        sum_mirror = np.sum(irrev_rate[i + n_states, n_states:])
        imbalance = sum_mirror - sum_orig
        if imbalance > 0:
            irrev_rate[i, i + n_states] = imbalance
        else:
            irrev_rate[i + n_states, i] = -imbalance
    for i in range(2 * n_states):
        irrev_rate[i, i] = -np.sum(irrev_rate[i, :])
    return irrev_rate
```

A function that creates an irreversible rate matrix using the doubled state space approach. It applies the rate modification to the original rate matrix and connects the two replicas of the state space while preserving the skewed detailed balance condition.

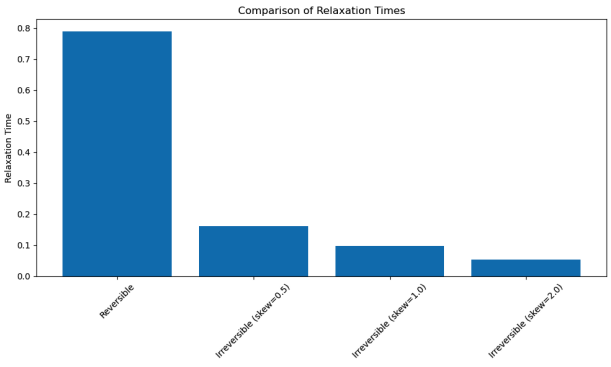
3. Supplementary Figures



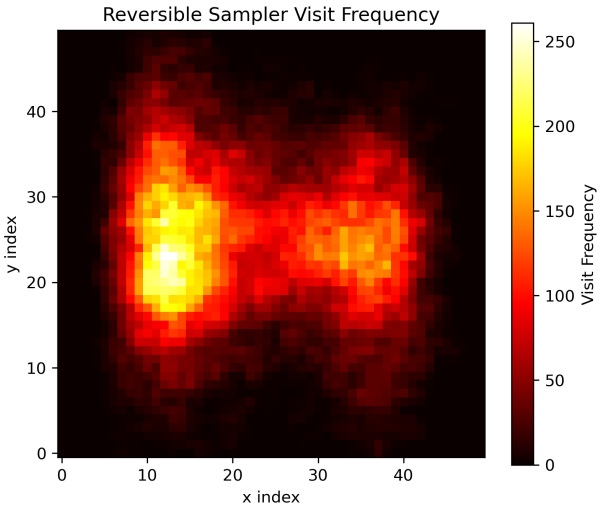
2D double well with direction skewness shown



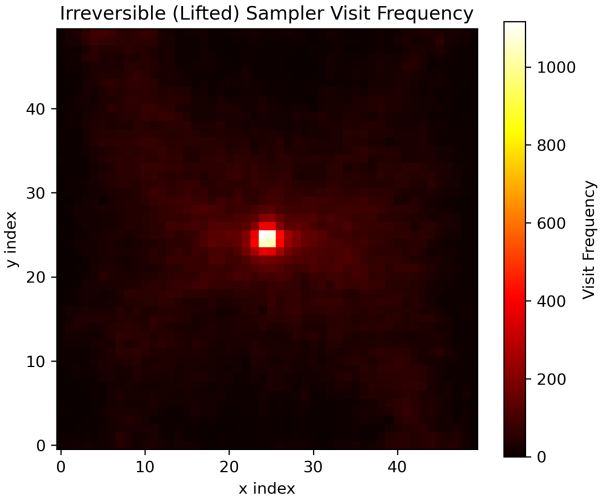
Polar plot showing relaxation time against skewness direction (theta) for 2D Gaussian wall potential.



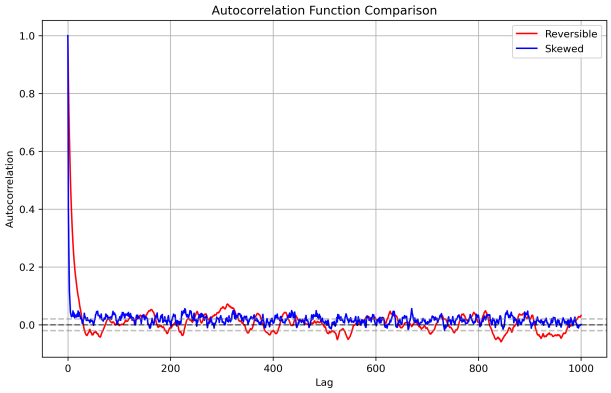
Comparison of relaxation time in 1D for reversible vs irreversible skew values tested on the double well potential for different skew values.



Plot showing long term occupancy heatmap for 2D Gaussian wall potential and reversible sampler for n=1000.



Plot showing long term occupancy heatmap for 2D Gaussian wall potential and Hessian optimised irreversible sampler for n=1000, confirming severe disruption in convergence accuracy



Plot showing decrease in autocorrelation function for 2D double well potential for the unoptimised skewness scheme with n=40. This plot shows both reversible and irreversible samplers.