HW5 - Jacob Shkrob

Jacob Shkrob

November 30, 2022

1 Exercise 64

In this exercise, we consider the 2-dimensional XY model of statistical physics, where we consider $\sigma_i \in \mathbb{R}^2$, where $i \in \mathbb{Z}_L$ and $\|\sigma_i\| = 1$. The XY model assigns to each σ_i the probability

$$\pi(\sigma) = \frac{e^{\beta \sum_{i \leftrightarrow j} \sigma_i \sigma_j}}{\mathcal{Z}}.$$

Rewriting in terms of angles $\theta_i \in [-\pi, \pi)$ gives that the density is

$$\pi(\theta) = \frac{e^{\beta \sum_{i \leftrightarrow j} \cos(\theta_i - \theta_j)}}{\mathcal{Z}}.$$

Taking the log gradient of the density π with respect to θ gives

$$\nabla_\theta \log \pi(\theta) = \begin{pmatrix} -\beta \sin(\theta_1 - \theta_2) \\ \beta \sin(\theta_1 - \theta_2) - \beta \sin(\theta_2 - \theta_3) \\ \beta \sin(\theta_2 - \theta_3) - \beta \sin(\theta_3 - \theta_4) \\ \vdots \\ \beta \sin(\theta_{L-1} - \theta_L). \end{pmatrix}$$

Therefore, we can write that our (unadjusted) Langevin algorithm for fixed step parameter $h > 0$ as

$$X_{t+1,h} = X_{t,h} + h \nabla \log(\pi(X_t))) + \sqrt{2h}\xi_t, \quad \xi_t \sim \mathcal{N}_L(\mathbf{0}, I)\text{i.i.d.}.$$

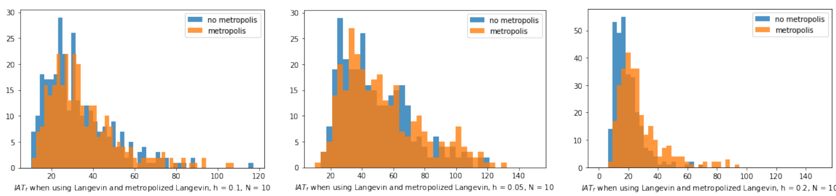
In addition to unadjusted Langevin, we included Metropolis-adjusted Langevin dynamics in the model. In this scheme, the proposal distribution $q(x|y)$ is based on the sampled point in the Langevin scheme. In other words,

$$q(x|y) \sim \mathcal{N}_L(y + h \nabla \log(\pi(y)), 2hI),$$

and therefore, if we write $X'_{t+1,h}$ is the newly sampled point according to $q(x|X_{t,k})$ our acceptance probability p_{acc} can be written as

$$p_{acc} = \min\left\{1, \frac{\pi(X'_{t+1,h}) q(X_{t,k}|X'_{t+1,h})}{\pi(X_{t,k}) q(X'_{t+1,h}|X_{t,k})}\right\}.$$

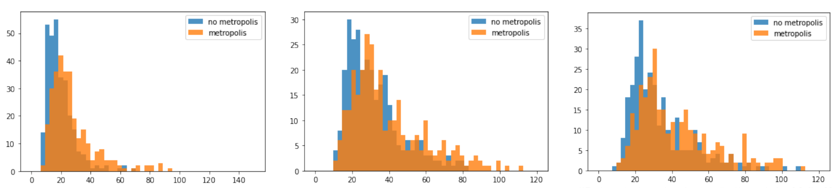
Figure 1 details the effect of decreasing and increasing step size h in both the unadjusted Langevin and Metropolis-adjusted Langevin schemes. When the step size increases, metropolis does not perform as well, most likely because the step sizes are too large for adequate exploration and mixing. When the step sizes are smaller, Metropolis-adjusted Langevin does much better.



(a) $\beta = 1,\ h = 0.1$ (b) $\beta = 1,\ h = 0.05$ (c) $\beta = 1,\ h = 0.2$

Figure 1: Histogram of IAT_f for different step legnths $h \in \{0.05, 0.1, 0.2\}$, $\beta = 1$, $N = 10$, after 1000 iterations

Another important aspect of the models to investigate is their robustness towards increased dimensionality (specficially, increased lattice size L in $\mathbb{Z}_L$), which is detailed in Figure 2. It appears that increased the benefit that unadjusted Langevin had with the larger step size $h = 0.2$ decreases as the lattice number N increases rapidly, which is to be expected.



(a) $N = 25,\ h = 0.2$ (b) $N = 50,\ h = 0.2$ (c) $N = 100,\ h = 0.2$

Figure 2: Histogram of IAT_f for different step legnths $h \in \{0.05, 0.1, 0.2\}$, $\beta = 1$, $N = 10$, after 1000 iterations

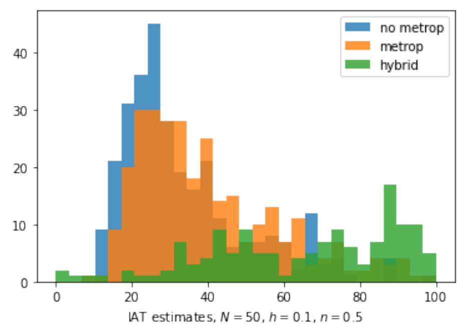
2 Exercise 65

To implement the Hybrid Monte Algorithm, we first need to select function $K(\tilde{x})$ and J used in the velocity verlet scheme discretization. We run the velocity verlet scheme for $n = \lfloor \frac{s}{h} \rfloor$ iteractions, where s is a time point and h is the step parameters (we used $s = 0.5$ and $h = 0.1$) A simple choice to make, and one that fits nicely with the description of the Hybrid MC algorithm, is to simply make $K(\tilde{x})$ a quadratic $(\tilde{x}^\top \tilde{x})$ and consider sampling from $\mathcal{N}_N(\mathbf{0}, I)$, where N is the lattice number. In particular, computing $p_{acc}(X_k, Y_{k+1})$ becomes a ratio of Gaussian densities which is easy to compute. For simplicity, we also chose J

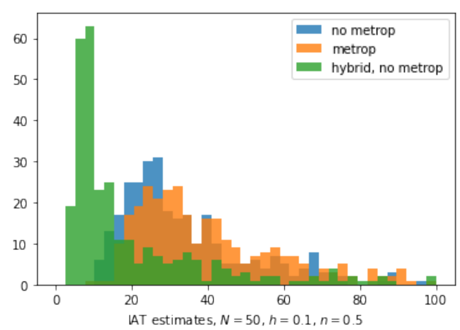
to be the matrix

$$J(x) := \begin{bmatrix} \mathbf{0} & -\hat{J}(\hat{x}) \\ \hat{J}^{\top}(\hat{x}) & \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{0} & -I \\ I & \mathbf{0} \end{bmatrix}.$$

Finally, we compute the magnetization angle by only looking at the projected coordinate in the first N variables. Figure ?? shows the results of using a metropolized Hybrid Monte Carlo scheme for fixed time point $n = 0.5$ and $h = 0.1$. From these results, it appears that hybrid monte carlo is on average worse than the unadjusted and adjusted metropolis langevin dynamics, when using metropolization, but outperforms the two methods when metropolization is turned off. This suggests that perhaps either the flow parameter $n = 0.5$ or step size $h = 0.1$ need to be fine tuned.



(a) IAT estimates for metropolized and unadjusted Langevin and Hybrid MC methods

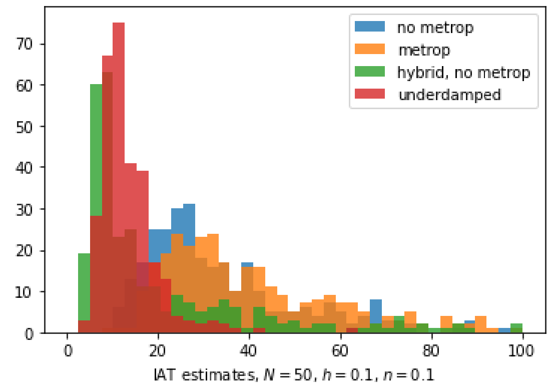


(b) IAT estimates for metropolized and unadjusted Langevin and Hybrid MC methods (without metropolization)

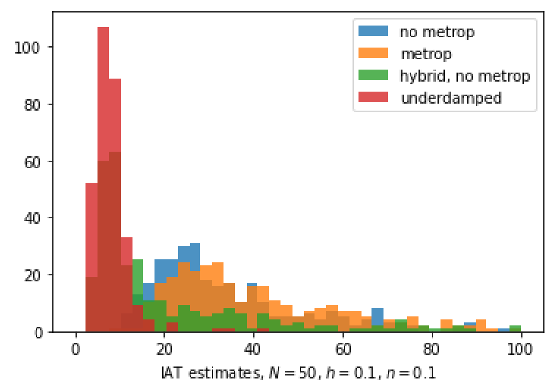
Figure 3: Hybrid MC IAT results for magnetization

3 Exercise 66

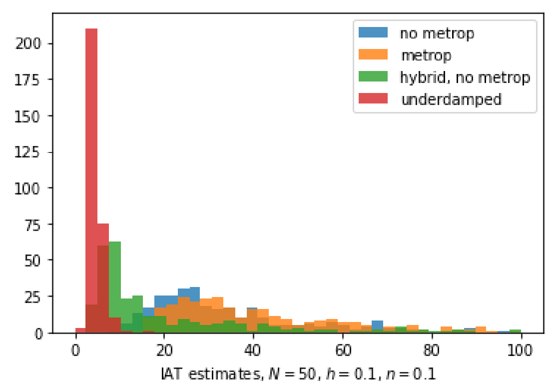
In this exercise, we are using an underdamped Langevin algorithm which uses $\gamma > 0$ and $h > 0$ to control the descent. In the figures below, we ploted the performance of underdamped Langevin as a function of γ and step size $h > 0$, which appear to suggest larger gamma values decrease IAT significantly, as well as larger step size (these graphs are for underdamped Langevin without metropolization). Hence, a relatively large damping parameter γ and small step parameter h would be effective for underdamped Langevin dynamics. In Figure ??, we show the IAT performance of underdamped Langevin with metropolization included, against all the previous models in the exercises (i.e. hybrid, metropolis adjusted Langevin, and regular Langevin) at fixed parameters $N = 50$, $h = 0.1$, $n = 0.1$, and $\gamma \in \{0.1, 0.2\}$. Our findings show that underdamped obtains significantly lower IAT estimates than the other models, suggesting that underdamped is more suitable (increasing γ also tended to lower IAT estimates).



(a) IAT for $\gamma = 0.1$

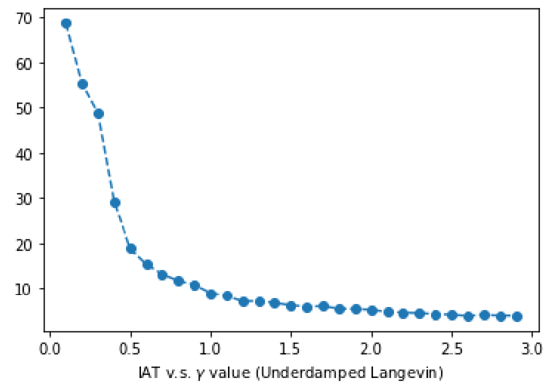


(b) IAT for $\gamma = 0.2$

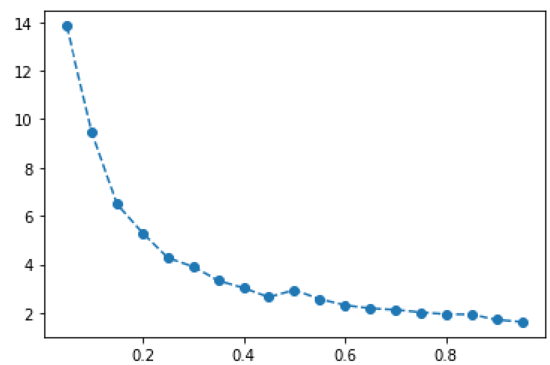


(c) IAT for $\gamma = 0.2$ and no metropolization for underdamped Langevin

Figure 4: IAT_f including Underdamped Langevin



(a) IAT after varying $\gamma \in [0.0, 3.0]$



(b) IAT after varying h step size

Figure 5: IAT_f for Underdamped Langevin as γ and h parameter vary