1   Introduction

In our $T = 0$ Lattice Path Integral Monte Carlo code (insert hyperlink), the acceptance ratio each update usually involves an exponential factor. This factor comes in three forms:

- Case 1: $\frac{W'}{W} \propto e^{-c(\tau - a)}$ $(a < \tau)$

- Case 2: $\frac{W'}{W} \propto e^{-c(b - \tau)}$ $(b > \tau)$

- Case 3: $\frac{W'}{W} \propto e^{-c(\tau_2 - \tau_1)}$ $(a < \tau_1 < \tau_2 < b)$

where τ is a previously sampled imaginary time, a, b are lower and upper bounds of the sampling interval, respectively, and c is a constant, that in practice turns out to be related to the energy difference between pre and post-update configurations. In the third case, two imaginary times τ_1, τ_2 are sampled subject to the constraint: $(a < \tau_1 < \tau_2 < b)$.

Before delving into the specifics of what the symbols will represent in the simulation, I would like to first establish the goal: cancelling these exponentials from the acceptance ratio.

Recall that the acceptance ratio is given by:

$$R = \frac{W'}{W} \frac{P}{P'} \tag{1}$$

where W', W are the weights of worldline configurations, which for our algorithm is obtained from the path-integral formulation of the ground state weight, and P, P' are probabilities of proposing a certain update and can vary depending on the decision process we follow to propose the update. Primed quantities refer to post-update configurations, and unprimed ones refer to the configuration before the update. Since most updates involve the sampling of an imaginary time, we can choose to sample these times from distributions that look similar to the exponentials in the weight ratios. Strategically choosing from what distribution to

sample from, can change the proposal probability ratios P/P', such that the exponentials in the weight ratios cancel.

We will discuss some examples that illustrate how to sample in each of the cases above.

2   Case: $P(\tau) \propto e^{-c(\tau-a)}$

Some of the updates in out Lattice PIMC algorithm will have an acceptance ratio of the form:

$$R = \frac{W'}{W}\frac{P}{P'} = Ae^{-c(\tau-a)}\frac{P}{P'} \tag{2}$$

where $a < \tau$. A is just a constant (or actually product of constants that is obtained from the weight ratio calculation). In order to cancel the exponential factor coming from the weights, the proposal probability ratio needs to be of the form:

$$\frac{P}{P'} = \frac{B}{\frac{c}{\mathcal{Z}}e^{-c(\tau-a)}} \tag{3}$$

where B is in general a product of probabilities dependent on how the update is proposed, and the prefactor $1/\mathcal{Z}$a normalizes the distribution. As part of most updates, an imaginary time τ needs to be proposed or sampled, and we will sample it from what is known as a truncated exponential distribution:

$$P' \equiv P(\tau) = \frac{c}{\mathcal{Z}}e^{-c(\tau-a)} \tag{4}$$

From now on, we can be quite methodic. A random variate τ can be obtained from $P(\tau)$ by: 1) Computing the normalization constant, 2) Computing the cumulutaive distribution function as a function of the random variate, and 3) inverting the cdf.

To normalize $P(\tau)$, integrate over all possible values of τ and equate to unity:

$$1 = \frac{c}{\mathcal{Z}} \int_a^b e^{-c(\tau-a)} d\tau \tag{5}$$

$$= \frac{c}{\mathcal{Z}} e^{ca} \int_a^b e^{-c\tau} d\tau \tag{6}$$

$$= \frac{c}{\mathcal{Z}} e^{ca} [-\frac{1}{c}][e^{-cb} - e^{-ca}] \tag{7}$$

$$= -\frac{1}{\mathcal{Z}} [e^{-c(b-a)} - 1] \tag{8}$$

$$= \frac{1}{\mathcal{Z}} [1 - e^{-c(b-a)}] \tag{9}$$

$$\tag{10}$$

Solving for $\mathcal{Z}$, the normalization constant becomes:

$$\mathcal{Z} = 1 - e^{-c(b-a)} \tag{11}$$

The next to step is to compute the cumulative distribution function (CDF) and equate it to a uniformly distributed random number $x \in [0, 1)$:

$$x \equiv \frac{c}{\mathcal{Z}} \int_a^\tau e^{-c(\tau'-a)} d\tau\prime = \frac{1}{\mathcal{Z}} [1 - e^{-c(\tau-a)}] \tag{12}$$

Finally, solving for the random time τ, we obtain:

$$\tau = a - \frac{1}{c} \ln(1 - \mathcal{Z}x) \tag{13}$$

Sampling τ from this distribution leads to an acceptance ratio that has the form:

$$R = \frac{W'}{W}\frac{P}{P'} \tag{14}$$

$$= Ae^{-c(\tau-a)}\frac{B}{\frac{c}{\mathcal{Z}}e^{-c(\tau-a)}} \tag{15}$$

$$R = \frac{AB\mathcal{Z}}{c} \tag{16}$$

The original exponential coming from the weight ratio has been cancelled. The idea is that since the acceptance ratio R only involves constants now, and no random variates, simulation parameters, such as the fugacity η, could be smartly equilibrated to obtain a desired acceptance ratio, that should be roughly constant, independent of the random time sampled. This might in turn enhance the dynamics of the worldline configuration and reduce autocorrelation and numbers of samples needed to achieve convergence of an estimator.

3   Case: $P(\tau) \propto e^{-c(b-\tau)}$

Some other updates will have an acceptance ratio of the form

$$R = \frac{W'}{W}\frac{P}{P'} = Ae^{-c(b-\tau)}\frac{P}{P'} \tag{17}$$

where $b > \tau$. Note that exponential coming from the weight ratio now depends on b, which is the upper bound of the interval, rather than a, which is the lower bound. Additionally, there is a sign of the random variate τ, which is now negative, and b is positive. The process follows the same steps as case 1, but we now sample τ instead from the following distribution:

$$P' \equiv P(\tau) = \frac{c}{\mathcal{Z}}e^{-c(b-\tau)} \tag{18}$$

where the normalization constant is left unchanged. The CDF is obtained almost the same way as before, but the limits of integration go from τ, the random time, to b, the upper

bound:

$$x \equiv \frac{c}{\mathcal{Z}} \int_{\tau}^{b} e^{-c(b-\tau')} d\tau\prime = \frac{1}{\mathcal{Z}} [e^{-c(b-\tau)}] \tag{19}$$

and again, solving for τ gives us the random imaginary time sampled from a truncated exponential distribution:

$$\tau = b + \frac{1}{c} \ln(1 - \mathcal{Z}x) \tag{20}$$

which cancels the exponential coming from the weight ratio, when computing the acceptance ratio R:

$$R = \frac{W'}{W} \frac{P}{P'} \tag{21}$$

$$= A e^{-c(b-\tau)} \frac{B}{\frac{c}{\mathcal{Z}} e^{-c(b-\tau)}} \tag{22}$$

$$R = \frac{AB\mathcal{Z}}{c} \tag{23}$$

4   Case: $P(\tau) \propto e^{-c(\tau_2 - \tau_1)}$

The final case is that in which the weight ratio depends on not one, but two random variates τ_1, τ_2. In the discussion that follows, we will assume that the random times follow are subject to the constraint: $a < \tau_1 < \tau_2 < b$. A prototypical acceptance ratio for this case has the following form:

$$R = \frac{W'}{W} \frac{P}{P'} = A e^{-c(\tau_2 - \tau_1)} \frac{P}{P'} \tag{24}$$

As such, as part of the decision process followed to build the proposal probability ratio, it will be wise to sample the imaginary times from the following joint probability distribution:

$$P' \equiv P(\tau_1, \tau_2) = \frac{c}{\mathcal{Z}} e^{-c(\tau_2 - \tau_1)} \tag{25}$$

In practice, we don't sample the two times simultaneously, but one at a time, rather. To do this, we will employ the following property of joint probability distributions:

$$P(\tau_1, \tau_2) = P(\tau_1) P(\tau_2 | \tau_1) \tag{26}$$

where $P(\tau_1)$ is the marginalized probability of τ_1, obtained by integrating out the τ_2 dependence. The factor $P(\tau_2 | \tau_1)$ is the probability of sampling τ_2, conditional on the value of τ_1. The strategy we will follow is to first sample τ_1 from $P(\tau_1)$ and then τ_2 from $P(\tau_2 | \tau_1)$.

Integrating out the τ_2 dependence, the marginalized probability distribution of τ_1 can be obtained:

$$P(\tau_1) = \int_{\tau_1}^{b} P(\tau_1, \tau_2) d\tau_2 \tag{27}$$

$$= \frac{c}{\mathcal{Z}_1} \int_{\tau_1}^{b} e^{-c(\tau_2 - \tau_1)} d\tau_2 \tag{28}$$

$$= \frac{c}{\mathcal{Z}_1} e^{c\tau_1} \int_{\tau_1}^{b} e^{-c\tau_2} d\tau_2 \tag{29}$$

$$= -\frac{1}{\mathcal{Z}_1} e^{c\tau_1} [e^{-cb} - e^{-c\tau_1}] \tag{30}$$

$$= -\frac{1}{\mathcal{Z}_1} e^{c\tau_1} [e^{-cb} - e^{-c\tau_1}] \tag{31}$$

$$P(\tau_1) = \frac{1}{\mathcal{Z}_1} [1 - e^{-c(b - \tau_1)}] \tag{32}$$

$$\tag{33}$$

Where $\mathcal{Z}_1$ is the normalization constant of the marginalized distribution $P(\tau_1)$, and the original constant $\mathcal{Z}$, corresponding to the joint distribution has been absorbed by it. Since now both the joint and the marginalized distributions are known, the conditional distribution

can be solved for from the property shown above:

$$P(\tau_2|\tau_1) = \frac{P(\tau_1, \tau_2)}{P(\tau_1)} \tag{34}$$

$$\propto c\frac{\mathcal{Z}_1}{\mathcal{Z}}\frac{e^{-c(\tau_2-\tau_1)}}{1 - e^{-c(b-\tau_1)}} \tag{35}$$

$$P(\tau_2|\tau_1) = \frac{c}{\mathcal{Z}_2}e^{-c(\tau_2-\tau_1)} \tag{36}$$

Since in the expression above, τ_1 is assumed to have already been sampled, it can be treated as a constant. Moreover, all other constants have been absorbed into the normalization constant Z_2, corresponding to the conditional distribution, except c since the benefit of hindsight tells us that we can leave it there to be cancelled after the Z_2 is explicitly computed. Note that the conditional probability is just a simple truncated exponential distribution like the one discussed in case 1.

We are now almost there. Distributions have been computed that will allow for the sampling of τ_1 and then τ_2, such that they effectively are sampled from the joint distribution. The rest from now on follows the same process as the previous cases: computing the normalization constant, the cdf, and the form of the random variates.

For the marginalized distribution, the normalization constant gives:

$$1 = \frac{1}{\mathcal{Z}_1}\int_a^b [1 - e^{-c(b-\tau_1)}]d\tau_1 \tag{37}$$

$$1 = \frac{1}{c\mathcal{Z}_1}[e^{-c(b-a)} + c(b-a) - 1] \tag{38}$$

and the normalization constant becomes:

$$\mathcal{Z}_1 = \frac{1}{c}[e^{-c(b-a)} + c(b-a) - 1] \tag{39}$$

The cdf is:

$$x = \frac{1}{\mathcal{Z}_1} \int_a^{\tau_1} [1 - e^{-c(b-\tau_1')}] d\tau_1' \tag{40}$$

$$x = \frac{1}{\mathcal{Z}_1} [\tau_1 - a + \frac{1}{c}(e^{-c(b-a)} - e^{-c(b-\tau_1)})] \tag{41}$$

$$\tag{42}$$

Solving for τ_1 can get slightly tricky here. First, try leaving all τ_1 dependence to one side of the equation:

$$\mathcal{Z}_1 x + a - \frac{1}{c} e^{-c(b-a)} = \tau_1 - \frac{1}{c} e^{-c(b-\tau_1)} \tag{43}$$

perform the following substitutions for simplicity:

$$y = \mathcal{Z}_1 x + a - \frac{1}{c} e^{-c(b-a)} \tag{44}$$

$$A = \frac{1}{c} e^{-cb} \tag{45}$$

and the equation becomes:

$$y = \tau_1 - A e^{c\tau_1} \tag{46}$$

which, solving for τ_1, has the solution:

$$\tau_1 = y - \frac{1}{c} W_k(-Ace^{cy}) \tag{47}$$

where W_k is the Lambert-W or Product Log function and $k \in \mathbb{Z}$.

From here on, the process for sampling τ_2 follows the same steps as Case 1, but replacing $a = \tau_1$ in the exponent and $\tau_2 = \tau$.

By sampling τ_1 and then τ_2 as prescribed, we have effectively sampled from the joint probability distribution $P(\tau_1, \tau_2)$. Finally, let's take a look at the effect of sampling from these joint distribution on the acceptance ratio:

$$R = \frac{W'}{W}\frac{P}{P'} \tag{48}$$

$$= Ae^{-c(\tau_2-\tau_1)}\frac{B}{\frac{c}{\mathcal{Z}}e^{-c(\tau_2-\tau_1)}} \tag{49}$$

$$R = \frac{AB\mathcal{Z}}{c} \tag{50}$$

where the normalization of the joint distribution is:

$$\mathcal{Z} = \frac{1}{c^2}[e^{-c(b-a)} + c(b-a) - 1] \tag{51}$$

In the following section, example distributions will be shown built from sampling random variates in each of the three cases.

5 Sampling examples

5.1 $P(\tau) \propto e^{-c(b-\tau)}$

Parameters: $c = -3.14, a = 0.5, b = 1.4$
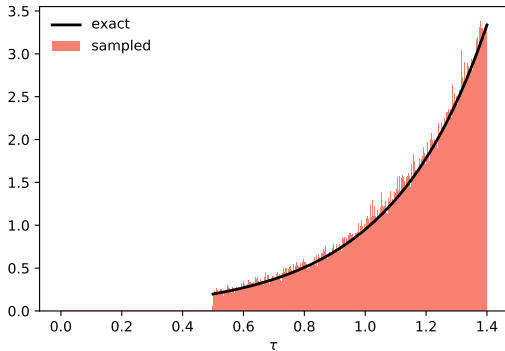
Parameters: $c = +3.14, a = 0.5, b = 1.4$
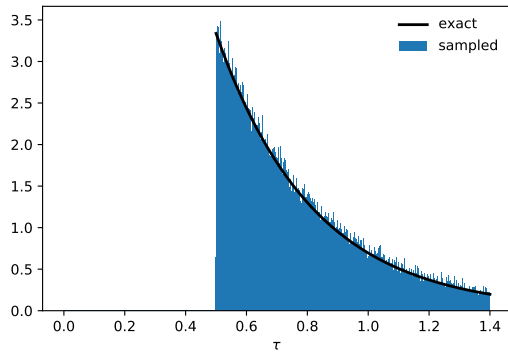
5.2 $P(\tau) \propto e^{-c(b-\tau)}$

Parameters: $c = -3.14, a = 0.5, b = 1.4$
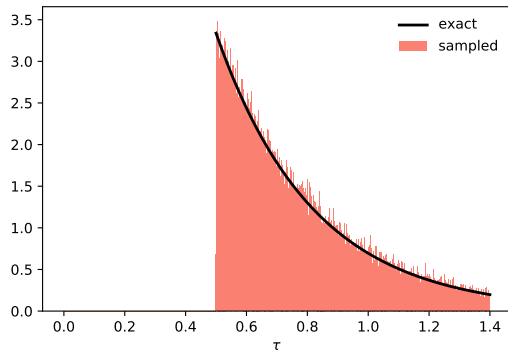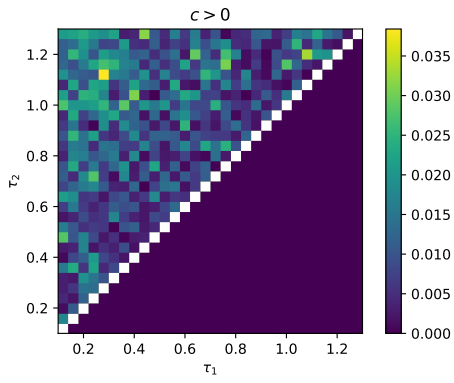

Parameters: $c = +3.14, a = 0.5, b = 1.4$
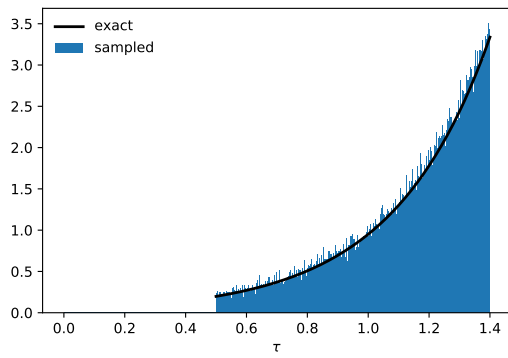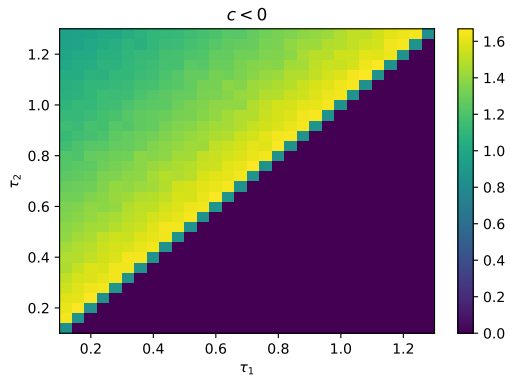
5.3 $P(\tau) \propto e^{-c(b-\tau)}$

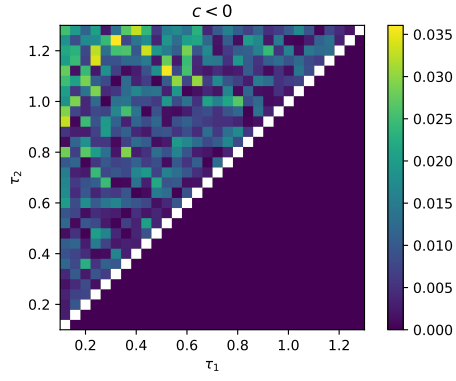$c = -0.5, a = 0.1, b = 1.3$
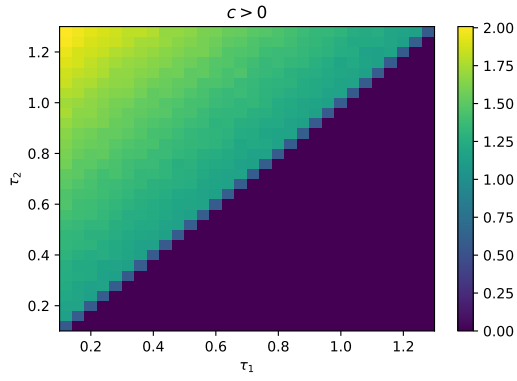
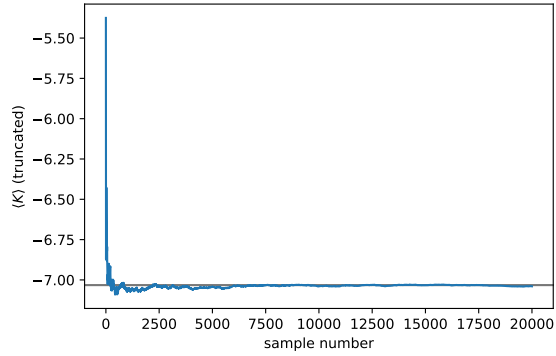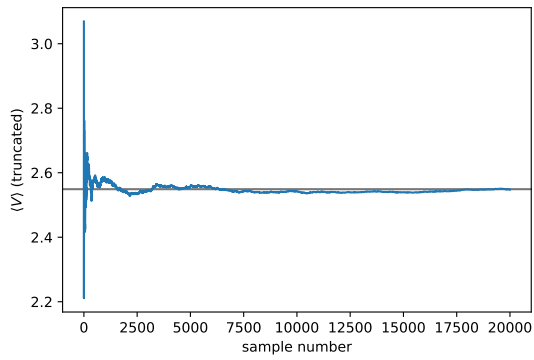Relative error:

$c = 0.5, a = 0.1, b = 1.3$

Relative error:

c < 0

6 Energy benchmark

Energy benchmark for $N = 4$ particles, at unit filling, with $U/t = 3.3$:

Kinetic:

Potential:

Truncated sampling:

$\langle K \rangle$	-7.0322	-7.0378 ± 0.0132	0.4254σ
$\langle V \rangle$	2.5490	2.5482 ± 0.0104	0.0710σ

Uniform sampling:

$\langle K \rangle$	-7.0322	-7.0447 ± 0.0126	0.9953σ
$\langle V \rangle$	2.5490	2.5446 ± 0.0089	0.4919σ

7 Wall clock and Autocorrelation times

In this section, a comparison will be made between the performance of the code using the truncated exponential sampling proposed here, and the original version of the code, which used mostly uniform sampling for the updates. For this comparison, the kinetic energy and the potential energy will be computed for a system using each sampling method. Then, the runtime of the code and the autocorrelation time will be computed for each. The simulation parameters can be read off from the command-line call:

The run times for 20 particles and for each method were:

Truncated: $t = 68.03$ **seconds and** $t = 229.89$ **seconds (4 times more bins)**

Uniform: $t = 67.43$ **seconds and** $t = 230.14$ **seconds (4 times more bins)**

Where the first time corresponds to generating 20000 bins and the second, 80000 bins.

Since the actual runtime was more or less the same for both runs, we now take a look at the autocorrelation time. This quantity gives an estimate of how correlated the samples are. Data with a smaller autocorrelation time becomes decorrelated more quickly and less samples need to be stored to obtain as good statistics as those obtained from more correlated data.
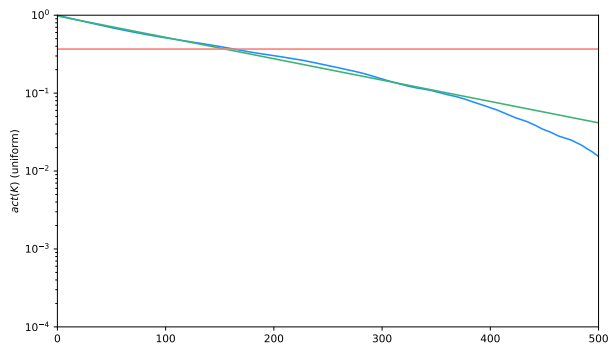
Kinetic energy (truncated):

Kinetic energy (truncated):

Truncated: $act = 106.02$ **Uniform:** $act = 158.35$

Potential energy (uniform):

13
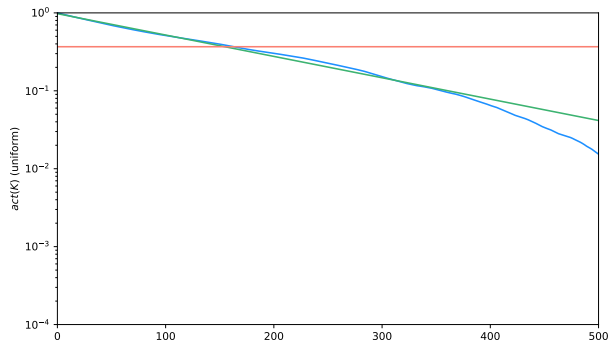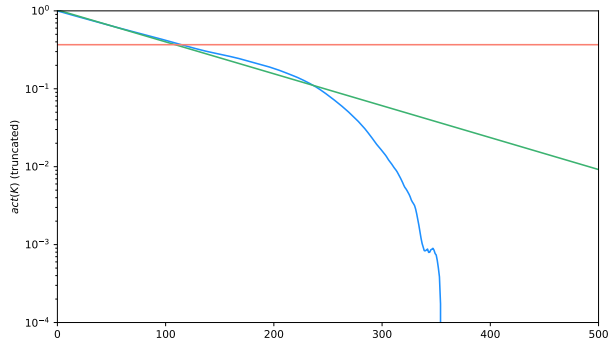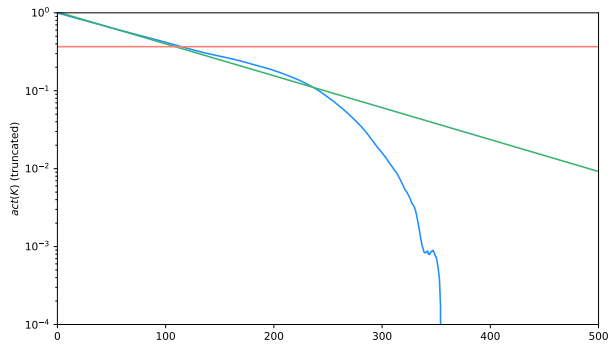
Potential energy (uniform):

Truncated: $act = 151.75$ **Uniform:** $act = 199.69$