arXiv:2302.04240v1 [physics.comp-ph] 8 Feb 2023

通过直接采样截断指数分布缩短格点路径积分量子蒙特卡罗中的自相关时间

Emanuel Casiano-Diaz[a,b,*]，Kipton Barros[c]，Ying Wai Li[b]，Adrian Del Maestro[a,d,e]

[a]田纳西大学物理与天文学系，美国田纳西州诺克斯维尔，*TN 37996*[b]洛斯阿拉莫斯国家实验室计算、计算科学与统计科学部，美国新墨西哥州洛斯阿拉莫斯，*NM 87545*[c]洛斯阿拉莫斯国家实验室理论部与非线性研究中心，美国新墨西哥州洛斯阿拉莫斯，*NM 87545*[d]田纳西大学明·*H*·高电气工程与计算机科学系，美国田纳西州诺克斯维尔，*TN 37996*[e]田纳西大学先进材料与制造研究所，美国田纳西州诺克斯维尔，*TN 37996*

摘要

在蒙特卡罗模拟中，根据接受率接受或拒绝提议构型；接受率取决于底层概率分布以及先验抽样概率。通过谨慎选择随机变量的抽样概率分布，可以缩短自相关时间，从而提高模拟效率。本文展示了如何直接从二维截断指数分布中抽取随机变量。我们证明，与拒绝抽样相比，直接抽样方法能够更快地收敛至目标分布。随后，我们将一维和二维截断指数分布的直接采样应用于近期提出的一种路径积分蒙特卡罗（PIMC）算法，用于模拟零温下的玻色-哈伯德格点模型。新的抽样方法提高了接受率，并缩短了估计量的自相关时间，从而有效加快了模拟。

关键词： 自相关时间；蒙特卡罗；截断指数分布；直接采样 g.

1. 引言

马尔可夫链蒙特卡罗（MCMC）模拟中的随机游走会产生一系列样本，这些样本通常存在统计相关性。统计估计的质量与所获得的有效非相关样本数量直接相关。因此，MCMC 方法开发中的一项关键挑战，是缩短生成充分去相关样本所需的计算时间。

最常见的 MCMC 方法之一是 Metropolis算法 [1, 2, 3, 4], ，它从一种往往不易采样的概率分布中获取样本。在该算法中，细致平衡原理会导出一个非负接受率R，用于判定随机提出的构型是接受还是拒绝。只有当接受率大于从均匀分布$r \sim \mathcal{U}(0, 1)$中抽取的随机数，即$r < R$时，才保留提出的构型；否则予以拒绝。

在大多数应用中，接受率较低的情况很常见，这会导致马尔可夫链效率低下，因为大多数提出的构型都会被拒绝。通过谨慎选择用于抽取蒙特卡罗更新中随机变量的底层概率分布，可以提高接受率，甚至使其达到 1，从而接受每个新构型，进而改善随机游走的动力学，并降低连续样本之间的相关性。

在近期开发的一种路径积分蒙特卡罗（PIMC）算法中， [5], 某些更新的接受率取决于从一维或二维截断指数分布中抽取随机变量；该分布是限制在有限域内的指数分布。本文介绍如何从二维截断指数分布中直接采样两个随机变量，并将这一抽样策略应用于 PIMC。由此得到的方法能够缩短自相关时间，从而加快统计平均值向其精确值收敛。

本文结构如下：第2节回顾如何从一维

*通讯作者。电子邮箱：ecasiano@vols.utk.edu

预印本已提交至 *Computer Physics Communications* *2023*年*2*月*9*日

概率分布中抽取随机变量，方法是对概率密度函数（PDF）的累积分布函数（CDF）求反函数。我们将在一维截断指数分布的背景下介绍这一方法。第3节将此方法推广到一个非平凡的情形：直接从二维截断指数分布中抽取两个随机变量。随后，我们将从一维和二维截断指数分布中直接采样随机变量的方法应用于文献 [5] 中的QMC算法，以模拟零温下的玻色子晶格，并证明在不牺牲性能的情况下，直接采样能够缩短动能和势能基态能量的自相关时间。

2. 一维 截断指数分布的直接采样

一维（1D）截断指数分布定义如下：

$$P_1(x) = \frac{1}{\mathcal{Z}} e^{-c(x-a)} = \frac{c e^{-c(x-a)}}{1 - e^{-c(b-a)}}, \qquad (1)$$

其中，a 和b 分别是有限域的下界和上界，c 是尺度参数，x是截断区间内满足$a \le x \le b$的随机变量。因子 $\mathcal{Z}$ 的选取确保了该分布归一化：$\int_{b\ a} \mathrm{d}x P_1(x) = 1$。$P_1(x)$的累积分布函数（CDF）为：

$$F_1(x) \equiv \int_a^x \mathrm{d}x'\, P_1(x') = \frac{1 - e^{-c(x-a)}}{c\mathcal{Z}}\,. \qquad (2)$$

逆变换抽样法通过反转函数关系$y = F_1(x)$，从式（1）的目标分布$x \sim P_1(x)$中获取样本。第一步是在 0 和 1 之间均匀抽取一个随机变量y 。我们将该随机变量记为$y \sim U(0, 1)$。然后，$x = F_1^{-1}(y)$得到一个服从所需目标分布$x \sim P_1(x)$的随机变量。对式（2）中的累积分布函数求逆，可得：

$$x(y) = a - \frac{\ln(1 - c\mathcal{Z}y)}{c}\,. \qquad (3)$$

当累积分布函数无法通过解析方式求逆时，一种常见的实用方法是归因于 von Neumann 的拒绝抽样 [6, 7], ，它通过依次比较两个独立采样的随机数，对有限定义域上的$P_1(x)$进行暴力采样，其中$P_{1,\max} \equiv \max_{a \le x \le b} P_1(x)$。

拒绝抽样

1. 从均匀分布$x \sim U(a,b)$中采样一个随机数。2. 再独立地从均匀分布 $\chi \sim U(0, P_{1,\max})$中采样另一个随机数。3. 如果 $\chi < P_1(x)$，则接受x。否则，拒绝该提议，并返回步骤1。

该过程返回的值 x 将是来自分布 P_1的良好样本。不过请注意，如果P_1 与$U(a,b)$差异很大，那么在返回一个良好样本之前，可能会经历多次拒绝。

由于存在逆函数，这里无需使用拒绝抽样。为了开展对二维D 情形的分析，我们通过分别采用直接采样和拒绝抽样生成随机样本的直方图$x \sim P_1$，对这两种方法进行比较。 8×10^5 个样本的结果如图 1 所示。两个直方图都与公式 (1) 中的预期结果一致，但拒绝抽样的结果波动更大。这可归因于直接采样数据集中良好样本数量较多，因为根据公式 (1) 得到的随机变量总会被接受。相比之下，拒绝抽样方法中的相当一部分迭代并未得到良好样本。

对于一维概率分布，例如$P_1(x)$，可以使用柯尔莫哥洛夫-斯米尔诺夫（KS）检验来量化随机变量数据集对目标分布的拟合程度。图 2 展示了 KS距离（或KS统计量）随数据集样本数的变化，该指标衡量经验累积分布函数与理论累积分布函数之间的最大差异。对于直接采样数据集，KS距离的衰减速度远快于拒绝抽样数据集。这是预期之中的，因为拒绝抽样数据集包含的来自公式 (1) 的统计样本更少。

在回顾了 1D 截断指数分布的标准直接反演法和拒绝抽样法之后，我们现在将其推广到 2D 情形，这与路径积分量子蒙特卡罗模拟相关。

3. 2D 截断指数分布的直接采样

概率分布式 (1) 可推广至两个随机变量 x_1,x_2 ，如下：

$$P_2(x_1, x_2) = \frac{1}{\mathcal{Z}_J} e^{-c(x_2 - x_1)}, \qquad (4)$$

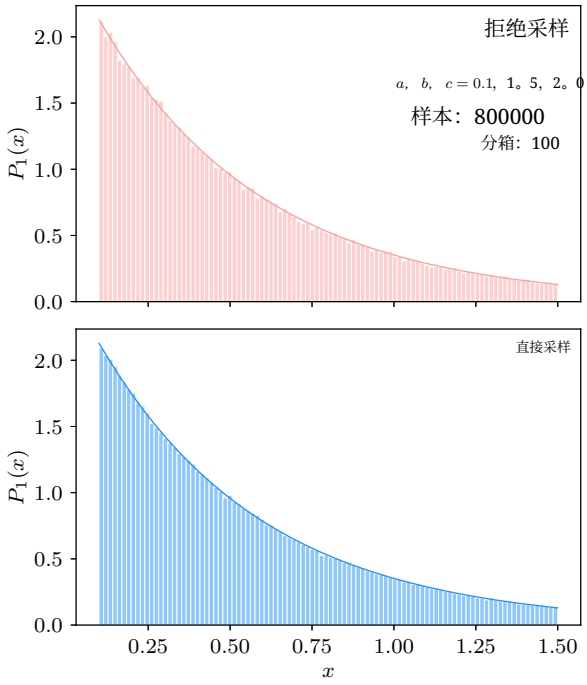


图1：通过拒绝抽样（上图）和直接采样（下图）生成的一维截断指数分布，见公式（1）。对于所示的样本数量（800,000），两种方法都成功生成了所需的一维截断指数分布随机变量x。直接采样的数据更接近精确分布（实线曲线），因为其中包含了更多有效样本。

其中 $a \le x_1 < x_2 \le b$，且该分布的归一化条件为：

$$\mathcal{Z}_J = \frac{e^{-c(b-a)} - ac + bc - 1}{c^2}. \tag{5}$$

随机变量x_1 和x_2 可通过将公式（4）分解为边缘概率与条件概率的乘积，依次进行采样

$$P_2(x_1, x_2) = P_2(x_1)P_2(x_2|x_1), \tag{6}$$

其中

$$P_2(x_1) = \int_{x_1}^{b} \mathrm{d}x_2\, P_2(x_1, x_2) = \frac{1 - e^{-c(b-x_1)}}{c\mathcal{Z}_J} \tag{7}$$

且 $P_2(x_2|x_1)$ 是变量 $x_2 \in [x_1,b]$ 上的一维截断指数分布。

边缘分布的累积分布函数为

$$\begin{aligned} F_2(x_1) &= \int_a^{x_1} dx_1' P_2(x_1') \\ &= \frac{1}{c^2\mathcal{Z}_J}\left[e^{-c(b-a)} - e^{-c(b-x_1)} - c(a - x_1)\right]. \end{aligned} \tag{8}$$

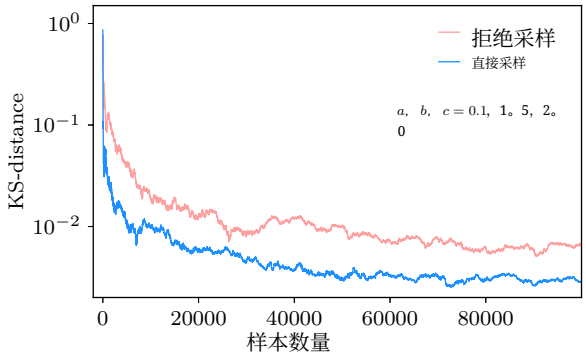


图2：柯尔莫哥洛夫-斯米尔诺夫（KS）检验结果。图中绘出了两种采样方法的 KS距离随样本数量的变化情况。直接采样数据集的 KS距离较小，这表明它对真实的一维截断指数分布（式（1））的估计更准确。

记$y = F_2(x_1)$。要从边缘分布$x_1 \sim P_2(x_1)$中生成样本，可以从$y \sim U(0,1)$中均匀采样，然后通过求F_2的逆来计算x_1：

$$x_1 = F_2^{-1}(y). \tag{9}$$

可以用朗伯（乘积对数）函数 [8], 给出x_1 的解析解，该函数定义为对函数关系 $f(\alpha) = \alpha e^{\alpha}$求逆。由于该映射不是单射，其反函数

$$W_k(\alpha e^{\alpha}) = \alpha, \tag{10}$$

有多个解分支k。当 α 为实数时，存在两个解分支；按照惯例，分别标记为$k = 0$（对应 $\alpha \ge -1$）和$k = -1$（对应 $\alpha \le -1$）。

现在，我们将对式（9）进行一系列代数运算。首先定义 $B = -e^{-cb}$，并将因变量y变换为一个新的变量，

$$u = yc^2Z_J + Be^{ca} + ca. \tag{11}$$

根据式（8），可得简化后的约束方程，

$$u = Be^{cx_1} + cx_1, \tag{12}$$

或者等价地，

$$(u - cx_1)e^{(u-cx_1)} = Be^{u}. \tag{13}$$

接下来，将 W_k 应用于等式两边，并结合 $\alpha = u - cx_1$使用公式（10）。结果为，

$$x_1 = \frac{1}{c}\left[u - W_k(Be^{u})\right]. \tag{14}$$

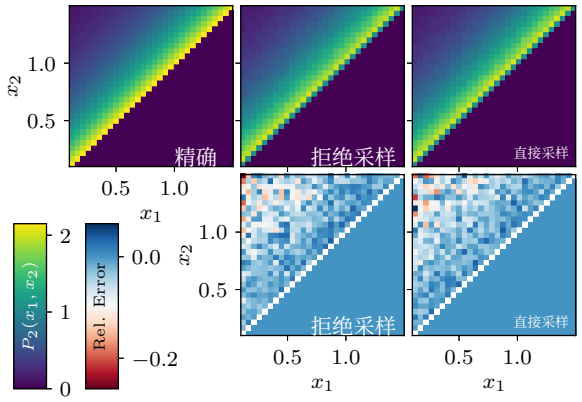


图3：通过拒绝采样和直接采样生成的式（4）中的二维截断指数分布。顶行热图显示了参数$a = 0.1$、$b = 1.5$ 和 $c = 2.0$ 对应的精确分布、拒绝采样分布和直接采样分布，底行则显示了各方法相对于精确分布的相对误差。使用了 8×10^5 个x_1 和x_2 样本。

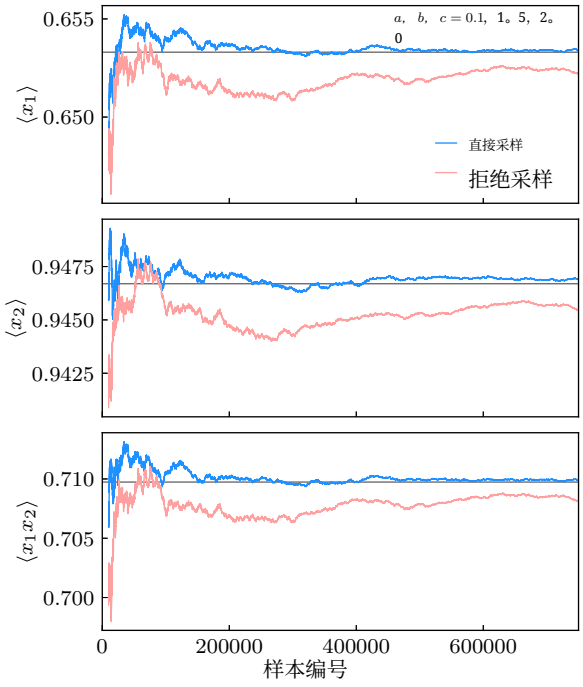


图 4：$\langle x_1 \rangle$、$\langle x_2 \rangle$ 和 $\langle x_1 x_2 \rangle$ 的累积平均值随样本数量的变化。x_1 和 x_2 的取值通过拒绝抽样和直接采样，从式 (4) 的二维截断指数分布中采样得到。对于这三种平均值，通过直接采样获得的数据集比精确结果（水平线）收敛得更快。

利用公式（12），可写为 $\alpha = -\exp[-c(b - x_1)]$。由于 $b - x_1 > 0$，条件 $\alpha \gtrless -1$ 与$c \gtrless 0$一致。因此，我们应选择：

$$k = 0 \ \text{ if } \ c \geq 0, \quad \text{or} \quad k = -1 \ \text{ if } \ c \leq 0. \tag{15}$$

将式（11）代入式（14）右侧，即可得到$x_1(y)$的最终闭式解。注意，$c \to 0$ 的极限是一个可去奇点，对应于$x_1 \to b - (b - a)\sqrt{1 - y}$。

现在可以将同时从联合分布 $P_2(x_1,x_2)$ 中对 (x_1, x_2) 进行采样的最终步骤概括如下。首先，生成一个均匀随机样本 $y \sim U(0,1)$。接下来，结合式 (11)，利用式 (14) 的解析解生成样本 $x_1 \sim P_2(x_1)$，其中已对 x_2 积分消去。此处可以使用现有数值子程序，高效计算朗伯函数 W_0 或 W_{-1} [9]。固定 x_1 后，第二个随机变量 $x_2 \sim P_2(x_2|x_1)$ 可直接从式 (1) 的一维截断指数分布中采样，下界设为 $a \to x_1$，上界仍设为 b。

图 3 展示了从二维概率分布 $P_2(x_1,x_2)$（式 (4)）中抽取样本的结果，参数 a、b 和 c 固定。最左侧的热图显示精确概率分布$P_2(x_1, x_2)$，以便与拒绝采样和直接采样分别获得 8×10^5 个x_1 和x_2 随机样本的结果进行比较。对于这么大的数据集，两种方法都能很好地抽样得到精确分布。观察对应于每种抽样方法的相对误差热图可见，大多数区域的误差都在精确分布的 10% 以内；左上方的部分区域误差较大，这是因为该区域的抽样概率较低。

由于图 3 中拒绝采样与直接采样的结果相似，因此无法直接判断哪种方法在重现二维截断指数分布（式 (4)）方面更高效。已有研究提出了高维柯尔莫哥洛夫–斯米尔诺夫检验 [10, 11], , 但技术问题使其实现起来非常复杂，因此我们选择计算三个量 $\langle x_1 \rangle$、$\langle x_2 \rangle$和 $\langle x_1 x_2 \rangle$ 随样本数量变化的运行平均值，结果如图 4 所示。对于这三个量，通过直接采样获得的样本的运行平均值均比拒绝抽样数据集更快收敛到精确值（使用式 (4) 获得，以水平线表示）。

在下一节中，我们将说明，如何从一维和二维截断指数分布中采样随机变量，

能够通过缩短样本之间的自相关时间，提高某些量子蒙特卡罗模拟的效率。

4. 应用：格点路径积分量子蒙特卡罗

基于 Metropolis-Hastings 算法的马尔可夫链蒙特卡罗应用，会根据概率密度函数 $\pi(\nu) = W(\nu)/\mathcal{Z}$抽取构型，从而建立马尔可夫链，其中 ν 表示由问题空间定义的一个构型。随机转移概率$T(\nu \to \nu')$，即从一个构型ν 转移到新构型 ν' 的概率，应独立于随机游走的历史。为此，需要使用满足细致平衡原理的蒙特卡罗更新构成遍历集：$\pi(\nu)T(\nu \to \nu') = \pi(\nu')T(\nu' \to \nu)$。转移概率可分解为选择概率$P(\nu \to \nu')$与接受概率$A(\nu \to \nu')$的乘积。根据细致平衡原理，一般蒙特卡罗更新的接受率可表示为：

$$\frac{A(\nu \to \nu')}{A(\nu' \to \nu)} = \frac{W(\nu')P(\nu' \to \nu)}{W(\nu)P(\nu \to \nu')} \equiv R. \tag{16}$$

通过 MCMC 过程生成的M 个构型可用于近似计算可观测量的期望值：

$$\langle O \rangle = \sum_{\nu} O(\nu)\pi(\nu) \simeq \frac{1}{M}\sum_{i=1}^{M} O_i, \tag{17}$$

其中$O_i = O(\nu_i)$。实际中，构成有限马尔可夫链 $\{\nu_1, \ldots, \nu_M\}$ 的随机样本 $\nu \sim \pi(\nu)$ 可能并非相互独立，因此O_i 和O_j之间存在相关性，即$\langle O_i O_j \rangle \neq \langle O_i \rangle \langle O_j \rangle$。可通过积分自相关时间 $\mathcal{T}_O$来量化可观测量O 的相关性：

$$\mathcal{T}_O = 1 + 2\sum_{\tau=1}^{M} \frac{C_O(\tau)}{C_O(0)}, \tag{18}$$

其中，自相关函数定义为

$$C_O(\tau) = \frac{1}{M-\tau}\sum_{i=1}^{M-\tau} (O_i - \langle O \rangle)(O_{i+\tau} - \langle O \rangle). \tag{19}$$

因此，只有间隔 $\mathcal{T}_O$ 个 MCMC 步的样本才能作为真正独立的测量值，而任何能降低$\mathcal{T}_O$ 的算法改进，都会提高模拟的整体效率。

在近期的一项工作中，[5], 本文的部分作者介绍了一种路径积分蒙特卡罗算法，用于

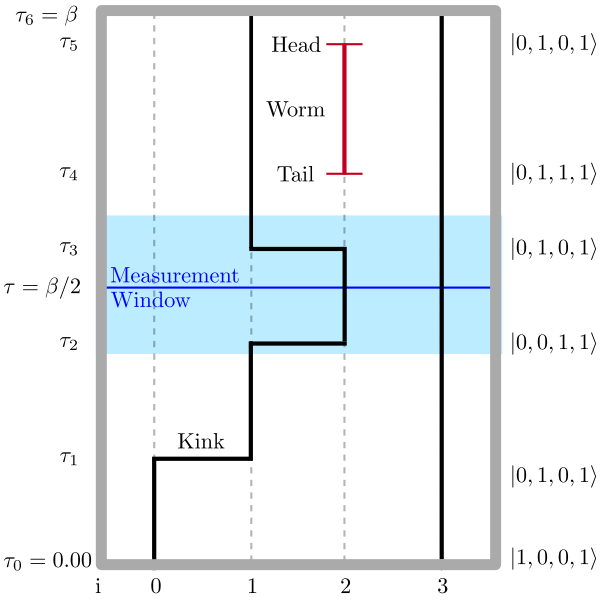


图 5：路径积分蒙特卡罗中，4 个晶格点上的 2 个粒子的世界线构型示例。路径沿虚时间方向（竖直）和空间方向（水平）传播。图中显示了三个拐点，分别位于虚时间 τ_1、τ_2、τ_3 处；它们是粒子在相邻格点间跃迁所致。通过对图中虫子进行更新（尾部位于 τ_4，头部位于 τ_5），可以采样更多构型。一些估计量（如动能）需要对某个量（例如拐点数）在以 $\tau = \beta/2$ 为中心、宽度由用户定义的窗口内取平均。

在零温下（$T = 0$）模拟玻色格点模型，该算法受有限温度 PIMC 虫算法 [12]的启发。关于该算法的完整细节（可用于计算基态期望值），请读者参阅文献 [5]:

$$\langle O \rangle \equiv \frac{\langle \Psi | O | \Psi \rangle}{\langle \Psi | \Psi \rangle} \tag{20}$$

通过对试探态 $|\Psi_T\rangle$施加密度算符的高次幂进行投影：$|\Psi\rangle = \lim_{\beta\to\infty} e^{-\beta H}|\Psi_T\rangle$，其中$H$ 是系统哈密顿量，β 是投影长度。

利用量子力学的路径积分表述，目标构型空间可以表示为一组在虚时间（由 β表征）和空间中传播的路径，这些路径称为世界线。图 5 展示了一维 Bose–Hubbard 模型中含有 $N = 2$ 个玻色子、分布在$L = 4$ 个格点上的世界线构型示例。竖直方向表示虚时间，水平方向表示晶格点。在该图中，线条宽度与格点上的粒子数成正比，虚线表示空格点。图左侧标出的虚时间集合 $\{\tau_i\}$ la-

bel对应福克态（$|n_1,n_2,n_3,n_4\rangle$）发生变化的时刻（右侧），其中n_j表示格点j上的粒子数。占据数的局部变化可以通过两种方式发生：表示相邻格点间粒子跃迁的拐点，或插入、删除一种称为虫子的特殊截断世界线 [13, 14, 15, 16]。形式上，虫尾和虫头分别表示玻色子产生算符和湮灭算符。通过对虫子执行更新，可以对整个构型空间进行采样。

将虫子插入世界线构型的接受率形式如下：

$$R = \text{const} \times \frac{1}{\mathrm{P}(\tau_\mathrm{h}, \tau_\mathrm{t})} \times \mathrm{e}^{-\mathrm{c}(\tau_\mathrm{h}-\tau_\mathrm{t})} \quad (\mathrm{a} \le \tau_\mathrm{t} < \tau_\mathrm{h} \le \mathrm{b}), \tag{21}$$

其中，τ_t 和 τ_h 分别表示虫尾和虫头的虚时间，并从联合概率分布$P(\tau_h,\tau_t)$中随机抽样。选择二维截断指数分布（式 (4)）作为该分布后，式 (21) 中的指数因子会抵消，使接受率只剩下一个常数因子：

$$R = \text{const} \times \frac{1}{\mathrm{e}^{-\mathrm{c}(\tau_\mathrm{h}-\tau_\mathrm{t})}} \times \mathrm{e}^{-\mathrm{c}(\tau_\mathrm{h}-\tau_\mathrm{t})} = \text{const}, \tag{22}$$

其中，抽取 τ_t、τ_h 所用的二维截断指数分布的归一化常数已并入 const。我们预计，通过实现一个利用迭代方法调整模拟参数的预平衡阶段，可以进一步优化这一恒定接受率；这类似于通过调节化学势 μ 来设定粒子平均数 [17, 18]的方法。其余更新的接受率也可以化为常数，这些更新涉及拐点的插入与删除，或沿虚时间方向移动虫子端点；具体做法是从一维截断指数分布（式 (1)）中抽取虚时间。对于沿虚时间方向移动虫子端点的更新，从该分布中抽样实际上可实现完美直接采样 [12, 5]（$R = 1$）。

在接下来的讨论中，对于从截断指数分布中抽取虚时间、使接受率具有式（22）形式的结果，我们称之为直接采样。传统的拒绝采样方案则是从重新缩放的均匀分布中抽取每个虚时间，$\tau \sim U(a,b)$，其中a、b 分别是区间的下界和上界。因此，这种情况下的联合概率分布为 $P(\tau_h,\tau_t) = 1/(b-a)^2$。然而，两种方案仍都包含梅特罗波利斯采样步骤：通过比较随机数$r \sim U(0,1)$是否满足$r < R$来决定是否接受更新，否则便拒绝。从形式上看，使用这两种方案得到的采样分布相同，只差一个前因子。

我们在巡游玻色子于晶格 [19] 上的 Bose–Hubbard模型基态量子蒙特卡罗模拟中，对所提出的直接采样方法进行了基准测试：

$$H = -t \sum_{\langle i,j \rangle} b_i^\dagger b_j + \frac{U}{2} \sum_i n_i(n_i-1) - \mu \sum_i n_i \,, \tag{23}$$

其中，t 是相邻晶格点 $\langle i,j\rangle$ 之间的隧穿振幅，$U > 0$ 是排斥相互作用势能，μ 是化学势，$b^\dagger{}_i(b_i)$ 是格点 i 上的玻色子产生（湮灭）算符，满足对易关系：$[b_i, b^\dagger{}_j] = \delta_{i,j}$，其中 $n_i = b^\dagger{}_i b_i$是局域数算符。模拟在正则系综中进行，其中 μ 是模拟参数。该模型会发生量子相变：相互作用较弱时处于超流体相，而排斥相互作用较强时则转变为莫特绝缘体，此时玻色子高度局域化。准确确定量子临界点一直是众多研究的动力
[20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35]；下文将报告相互作用强度固定为 $U/t = 3.3$ 时的模拟结果。该参数具有代表性，处于量子临界区，此时空间和时间关联长度均发散，导致众所周知的临界减速问题 [36]，即 MCMC 样本之间可能存在很强的关联。

动能估计量在福克基中是非对角的，其数值由图 5 [5] 中测量窗口内拐点的平均数量确定：

$$\langle \mathrm{K} \rangle = -\frac{\langle \mathrm{N_{kinks}} \rangle}{\Delta\beta} \tag{24}$$

其中 $\Delta\beta$ 是窗口宽度。势能估计量在福克基中是对角的，通过测量

$$\langle V \rangle = \frac{U}{2} \sum_i \langle n_i(n_i-1) \rangle \tag{25}$$

虚时间 $\tau = \beta/2$ 处的值获得。

为了理解直接采样与拒绝采样在我们的量子蒙特卡罗算法中的作用，我们对单位填充的单维 Bose-Hubbard 模型进行了模拟，$L = N$，其中 L 是格点数，

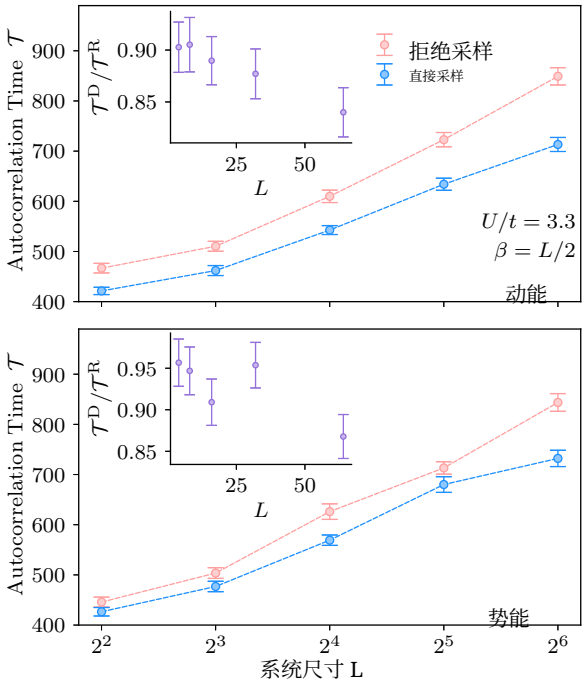


图6：通过路径积分蒙特卡罗得到的一维玻色-哈伯德模型基态动能（上）和势能（下）能量估计量的积分自相关时间随系统尺寸的变化，L 处于单位填充时：$L = N$。数值越低越好。插图显示了针对一维和二维截断指数分布，使用直接采样和拒绝采样（均匀分布）方法进行采样所得自相关时间的比值。

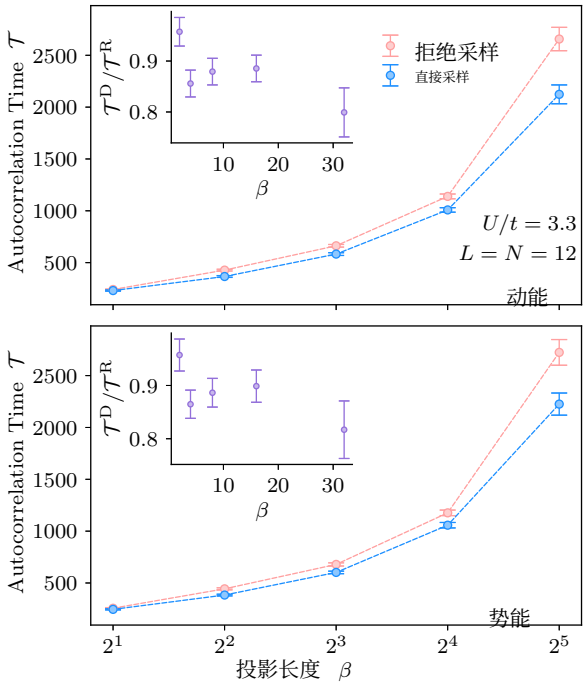


图7：通过路径积分蒙特卡罗得到的一维玻色-哈伯德模型基态动能（上）和势能（下）能量估计量的自相关时间随投影长度 β 的变化。数值越低越好。插图显示了针对一维和二维截断指数分布，使用直接采样和拒绝采样（均匀分布）方法进行采样所得自相关时间的比值。

而 N 是粒子数。对于不同的 L 值，我们使用两种采样方法，在$U/t = 3.3$（超流-绝缘临界点附近）计算了式（18）中的积分自相关时间。结果见图6，自相关时间是通过 emcee Python 库 [37], 计算的，该库基于参考文献 [38, 39]中所述的方法。对于系统尺寸不超过 $L = 2^6 = 64$ 的情形，在执行蠕虫更新时，直接从截断指数分布中抽取虚时间所得的动能和势能自相关时间均更短。插图展示了从截断指数分布（直接采样）中抽样所得的自相关时间与从均匀分布（拒绝采样）中抽样所得的自相关时间之比 $\mathcal{T}^D/\mathcal{T}^R$ 随系统尺寸的变化。比值小于 1 表明样本间相关性降低；在所研究的最大系统尺寸下，改善幅度达到 ~ 15%。

直接采样对动能估计量自相关时间的影响更大，因为模拟——

中平均拐点数的动力学与模拟中虫子动力学的改善直接相关（见图 5）。误差棒表示由 80 次独立模拟计算得出的自相关时间均值标准误。

由于参考文献 [5] 中描述的晶格基态路径积分蒙特卡罗算法是一种投影算法，因此会产生系统误差；随着投影长度 β的增加，该误差会减小。因此，要获得可靠的可观测量估计值，需要针对 β的不同取值进行模拟，并根据β 的指数拟合加上一个加性常数，在误差范围内外推得到精确值：$\langle O(\beta)\rangle = C_1 e^{-\beta C_2} + \langle O\rangle$，其中$C_1$,$C_2$,和 $\langle O\rangle$ 为拟合参数。因此，为了了解直接采样对 β 外推的影响，我们绘制了固定系统尺寸$L = 12$ 且$U/t = 3.3$时动能和势能的β相关自相关时间，见图6。再次可以看到，直接从截断指数分布中采样虚时间能改善自相关时间；插图显示时间约缩短了20%。

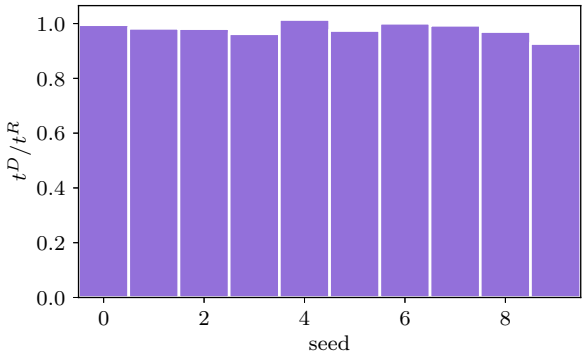


图 8：使用拒绝采样（R）和直接采样（D）从截断指数分布中抽样，对尺寸为 D Bose-Hubbard 晶格的路径积分蒙特卡罗模拟所需的墙钟时间，其中 $L = N = 12$，且 $U/t = 3.3$。每种方法均从不同的随机种子开始进行 10 次独立运行，并收集了 10^6 个动能和势能样本。采用不同抽样方案时，墙钟时间没有显著变化，墙钟时间的平均比值为 $t^{\mathrm{D}}/t^{\mathrm{R}} = 1.02 +/- 0.01$。

对于最大的 β 个值。

上述结果表明，在 [5]所述算法中，直接从截断指数分布采样可以缩短估计量的自相关时间。然而，对于我们对该算法的实现，我们还发现，直接从截断指数分布采样所需的墙钟时间并不比原始版本更长；在原始版本中，虚时间是从均匀分布中采样的。换言之，直接采样方案的实现不会影响实际运行时间。对于一个包含$L = N = 12$个玻色子、满足$U/t = 3.3$且 $\beta = 16$的系统，我们进行了10次模拟，每次均使用不同的随机种子，并发现采用两种采样方案时墙钟时间之比为$t^{\mathrm{D}}/t^{\mathrm{R}} = 1.02 +/- 0.01$。图8展示了每个随机种子下直接采样相对于拒绝法的墙钟时间比$t^{\mathrm{D}}/t^{\mathrm{R}}$。因此，这种新的采样方案成功缩短了自相关时间，同时没有增加墙钟时间，从而有效加速了文中讨论的量子蒙特卡罗应用。

确认本节所述结果所需的所有代码、脚本和数据均可在开源代码仓库中获取 [40, 41, 42]。

5. 结论

本文展示了如何通过两步逆采样法，直接从二维截断指数分布中获取随机变量。对于有限数量的样本，无论是一维还是二维，直接从截断指数分布中获得的随机变量数据集都能更好地再现目标分布。随后，我们将截断指数分布的直接采样应用于晶格路径积分蒙特卡罗；该分布出现在虫子 更新的接受概率中，并使虚时间世界线构型空间得以高效抽样。直接采样方法降低了动能和势能估计量的积分自相关时间，同时使两种抽样方案的模拟运行时间基本相同。对于所考察的系统尺寸，整体效率提升了 15 %。

未来的研究方向包括实现迭代方法，以优化当前虫子更新中接受率恒定的非物理算法参数，从而进一步改善动力学并接近理想的抽样极限。

6. 致谢

我们感谢 N. Prokof'ev 和 M. Thamm 进行了富有成效的讨论。

本研究部分得到洛斯阿拉莫斯国家实验室 (LANL) 实验室定向研究与开发经费的支持。LANL 由 Triad National Security, LLC 代表美国能源部国家核安全管理局运营（合同编号：89233218CNA000001）。K. B.、Y. W. L. 和 A. D. 感谢美国能源部科学办公室基础能源科学办公室通过项目编号 DE- SC0022311 提供的支持。

参考文献

[1] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, E. Teller, 《快速计算机进行状态方程计算》, 《化学物理学杂志》21 (6) (1953) 1087–1092. doi:10.1063/1.1699114。URL https://doi.org/10.1063/1.1699114 [2] K. Binder, D. Heermann, 《统计物理学中的蒙特卡罗模拟：导论》, 物理学研究生教材, Springer Berlin Heidelberg, 2010年。[3] W. Krauth, 《统计力学：算法与计算》, 牛津大学出版社, 牛津, 2006年。[4] N. Giordano, H. Nakanishi, 《计算物理学》, Pearson/Prentice Hall, 2006年。[5] E. Casiano-Diaz, C. M. Herdman, A. Del Maestro, 面向晶格玻色子纠缠的路径积分基态蒙特卡罗算法（2022年）。 doi:10.48550/ARXIV.2207.11301。URL https://arxiv.org/abs/2207.11301

[6] G. E. Forsythe, 《冯·诺依曼正态分布及其他分布随机抽样比较法》, 数学计算 26 (120) (1972) 817–826。doi:10.2307/2005864。[7] M. E. J. Newman, G. T. Barkema, 《统计物理学中的蒙特卡罗方法》, 克拉伦登出版社, 牛津, 1999年。[8] R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, D. E. Knuth, 《关于 LambertW 函数》, 计算数学进展 5 (1) (1996年) 329。 doi:10.1007/bf02124750。URL https://link.springer.com/article/10.1007/BF02124750[9] T. Fukushima, 《无需超越函数求值即可精确快速计算 Lambert W 函数》, 计算与应用数学杂志 244 (2013年) 77。doi:10.1016/j.cam.2012.11.021。URL https://linkinghub.elsevier.com/retrieve/pii/S0377042712005213[10] W. H. Press, S. A. Teukolsky, 《二维数据的柯尔莫哥洛夫-斯米尔诺夫检验》, 计算物理 2 (4) (1988) 74–77。doi:10.1063/1.4822753。URL https://aip.scitation.org/doi/abs/10.1063/1.4822753[11] A. Justel, D. Pe˜na, R. Zamar, 《多变量拟合优度柯尔莫哥洛夫-斯米尔诺夫检验》, 统计与概率快报 35 (3) (1997 251 259 doi:https://doi.org/10.1016/S0167-) – 。7152(97)00020-5 URL https://www.sciencedirect.com/science/article/pii/S0167715297000205[12] N. V. Prokof'ev, B. V. Svistunov, I. S. Tupitsyn, 《用于离散量子系统统计的精确、完备且通用的连续时间世界线蒙特卡罗方法》, 实验与理论物理学杂志 87 (1998) 310。 doi:10.1134/1.558661。URL https://link.springer.com/article/10.1134/1.558661[13] M. Boninsegni, N. V. Prokof'ev, B. V. Svistunov, 《虫算法与图示蒙特卡罗: 连续空间路径积分蒙特卡罗模拟的新方法》, 物理评论 E 74 2006 036701 doi:10.1103/PhysRevE.74. ( ) 。036701[14] M. Boninsegni, N. V. Prokof'ev, B. Svistunov, 《连续空间路径积分蒙特卡罗模拟的虫算法》, 《物理评论快报》 96 2006 070601 doi:10.1103/PhysRevLett. ( ) 。[15] N. V. Prokof'ev, B. V. Svistunov, I. S. Tupitsyn, 《量子蒙特卡罗模拟中的“虫”算法》, 物理学快报 A 238 (1998) 253。doi:10.1016/S0375-9601(97)00957-2。URL https://doi.org/10.1016/s0375-9601(97)00957-2[16] N. Prokof'ev, B. Svistunov, 《经典统计模型的虫算法》, 《物理评论快报》87 (2001) 160601。doi:10.1103/PhysRevLett.87.160601。URL https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.87.160601[17] C. M. Herdman, A. Rommal, A. Del Maestro, 《量子蒙特卡罗测量 ^{4}He 的化学势》, 物理评论 B 89 (2014) 224502。 doi:10.1103/PhysRevB.89.224502。URL https://link.aps.org/doi/10.1103/PhysRevB.89.224502[18] C. Miles, B. Cohen-Stead, O. Bradley, S. Johnston, R. Scalettar, K. Barros, 《通过动态调节化学势在巨正则系综蒙特卡罗模拟中实现目标粒子数》, 物理评论 E 105 (2022年) 045311。doi:10.1103/PhysRevE.105.045311。URL https://link.aps.org/doi/10.1103/PhysRevE.105.045311[19] H. A. Gersch, G. C. Knollman, 《量子胞模型》 玻色子, 物理评论 129 1963 959 967 doi: ( ) – .. 10.1103/PhysRev.129.959 URL https://link.aps.org/doi/10.1103/PhysRev.129.959[20] J. Carrasquilla、S. R. Manmana、M. Rigol, 《一维玻色-哈伯德模型中超流体-莫特绝缘体转变过程中的能隙、保真度易感率和布洛赫振荡的标度行为》, 物理评论 A 87 (4) (2013) 043606. doi:10.1103/PhysRevA.87.043606. URL https://journals.aps.org/pra/abstract/10.1103/PhysRevA.87.043606[21] T. D. Kuhner、S. R. White、H. Monien, 《具有最近邻相互作用的一维玻色-哈伯德模型》, 物理评论 B 61 2000 12474 12489 doi: ( ) – ..URL https://link.aps.org/doi/10.1103/PhysRevB.61.12474[22] J. K. Freericks、H. Monien, 《纯净及无序玻色-哈伯德模型的强耦合展开》, 物理评论 B 53 (1996) 2691–2700. doi:10.1103/PhysRevB.53.2691. URL https://link.aps.org/doi/10.1103/PhysRevB.53.2691[23] N. Elstner、H. Monien, 《玻色-哈伯德模型的动力学和热力学》, 物理评论 B 59 (1999) 12184–12187. doi:10.1103/PhysRevB.59.12184. URL https://link.aps.org/doi/10.1103/PhysRevB.59.12184[24] T. D. K¨uhner、H. Monien, 《一维玻色-哈伯德模型的各相》, 物理评论 B 58 (1998) R14741–R14744. doi:10.1103/PhysRevB.58.R14741. URL https://link.aps.org/doi/10.1103/PhysRevB.58.R14741[25] V. A. Kashurnikov、A. V. Krasavin、B. V. Svistunov, 《一维玻色哈伯德模型中的莫特绝缘体-超流液体转变: 量子蒙特卡罗方法》, 实验与理论物理学杂志 64 (2) (1996) 99–104. doi:10.1134/1.567139 URL https://doi.org/10.1134/1.567139[26] V. A. Kashurnikov、B. V. Svistunov, 《精确对角化结合重整化群理论: 研究一维超流体-绝缘体转变的精确方法》, 物理评论 B 53 (1996) 11776–11778. doi:10.1103/PhysRevB.53.11776. URL https://link.aps.org/doi/10.1103/PhysRevB.53.11776[27] S. Ejima、H. Fehske、F. Gebhard, 《一维玻色-哈伯德模型的动力学性质》, 《欧洲物理快报》 (《欧洲物理快报》) 93 (3) (2011) 30002. doi:10.1209/0295-5075/93/30002 URL https://doi.org/10.1209/0295-5075/93/30002[28] J. Zakrzewski、D. Delande, 《精确确定一维玻色-哈伯德模型中的超流体-绝缘体转变》, 《AIP会议论文集》1076 (1) (2008) 292 300 arXiv:https://aip.scitation.org/doi/pdf/ – .,. 10.1063/1.3046265 doi:10.1063/1.3046265 URL https://aip.scitation.org/doi/abs/10.1063/1.3046265[29] S. M. A. Rombouts、K. Van Houcke、L. Pollet, 《正则系综中的量子蒙特卡罗模拟环更新》, 《物理评论快报》96 (2006) 180603. doi:10.1103/PhysRevLett.96.180603. URL https://link.aps.org/doi/10.1103/PhysRevLett.96.180603[30] W. Krauth, 《一维玻色哈伯德模型的贝特拟设》, 物理评论 B 44 (1991) 9772–9775. doi:10.1103/PhysRevB.44.9772. URL https://link.aps.org/doi/10.1103/PhysRevB.

44.9772[31] R. V. Pai、R. Pandit、H. R. Krishnamurthy、S. Ramasesha, 《一维无序玻色哈伯德模型：密度矩阵重整化群研究》, 《物理评论快报》76 (1996) 2937–2940. doi:10.1103/PhysRevLett.76.2937. URL https://link.aps.org/doi/10.1103/PhysRevLett.76.2937[32] G. G. Batrouni、R. T. Scalettar、G. T. Zimanyi, 《一维玻色系统中的量子临界现象》, 《物理评论快报》 65 1990 1765 1768 doi:10.1103/PhysRevLett.65.1765. URL https://link.aps.org/doi/10.1103/PhysRevLett.65.1765[33] M. Pino、J. Prior、A. M. Somoza、D. Jaksch、S. R. Clark, 《一维玻色-哈伯德模型中的再入现象与纠缠》, 物理评论 A 86 (2012) 023631. doi:10.1103/PhysRevA.86.023631. URL https://link.aps.org/doi/10.1103/PhysRevA.86.023631[34] A. M. L¨auchli、C. Kollath, 《一维玻色–哈伯德模型量子淬火后的关联扩展与纠缠》, 《统计力学：理论与实验》2008 (05) (2008) P05018. doi:10.1088/1742-5468/2008/05/p05018. URL https://doi.org/10.1088/1742-5468/2008/05/p05018[35] S. Rachel、N. Laflorencie、H. F. Song、K. Le Hur, 《利用双分区涨落检测量子临界点》, 《物理评论快报》 108 2012 116401 doi:10.1103/PhysRevLett.108.116401. URL https://link.aps.org/doi/10.1103/PhysRevLett.108.116401[36] U. Wolff, 《临界减速》, 《核物理 B：会议增刊》 17 1990 93 102 doi:10.1016/0920-5632(90)90224-i. URL https://doi.org/10.1016/0920-5632(90)90224-i[37] D. Foreman-Mackey、D. W. Hogg、D. Lang、J. Goodman, 《emcee：MCMC锤》, 《太平洋天文学会出版物》125 (925) (2013) 306. doi:10.1086/670067. URL https://dx.doi.org/10.1086/670067[38] J. Goodman、J. Weare, 《具有仿射不变性的系综采样器》, 《应用数学与计算科学通讯》 5 1 2010 65 80 doi:10.2140/camcos.2010.5.65.
[39] A. D. Sokal, 《统计力学中的蒙特卡罗方法：基础与新算法——致读者》, 载于：《泛函积分：基础与应用》, 施普林格美国, 1996年, 第131–192页。[40] E. Casiano-Diaz、K. Barros、Y. Wai Li、A. Del Maestro, 《本文所用的所有代码、脚本和数据均收录于 GitHub 仓库》, https://github.com/DelMaestroGroup/papers-code-truncExponSampling 2023 ()。URL https://github.com/DelMaestroGroup/papers-code-truncExponSampling[41] E. Casiano-Diaz、K. Barros、Y. Wai Li、A. Del Maestro, 《截断指数采样数据集》（2023 doi:10.5281/）。. URL https://doi.org/10.5281/zenodo.7566332[42] E. Casiano-Diaz、C. Herdman、A. Del Maestro, 《PIGSFLI——格点实现的路径积分基态方法》, https://github.com/DelMaestroGroup/pigsfli 2022 ()。doi:10.5281/zenodo.6885505. URL https://doi.org/10.5281/zenodo.6885505