# 采用隐式时间步长的神经量子态增强虚时间演化

Maksymilian Kliczkowski，[1,2] Thomas D. Kühne，[1,2,3] Fedor Šimkovic IV，[4]Prachi Sharma，[4] Carlos L. Benavides-Riveros，[4]和 Werner Dobrautz [1,2,5,6]高级系统理解中心（*CASUS*），德国格尔利茨 *02826*[2]亥姆霍兹德累斯顿-罗森多夫中心（*HZDR*），德国德累斯顿 *01328*[3]德累斯顿工业大学人工智能研究所，*Helmholtzstraße 10*，德国德累斯顿 *01069*[4]*IQM*量子计算机公司，*Georg-Brauchle-Ring 23-25*，德国慕尼黑 *80992*[5]德累斯顿/莱比锡可扩展数据分析与人工智能中心（*ScaDS.AI*），德国德累斯顿 *01069* [6]德累斯顿工业大学，德国德累斯顿 *01069*

（日期：2026年11月25日）

寻找量子多体系统的基态仍然是物理学和量子化学中的核心挑战。虚时间演化（ITE）提供了一种自然且广泛使用的途径，但其收敛通常是渐近的。变分方法将这一演化限制在参数化流形上，而神经量子态（NQS）则提供了灵活的表示，通常通过随机重构（SR）进行优化。其性能取决于变分流形的几何结构、ITE 的收敛性质，以及有限步显式（前向）积分器的稳定性。通过利用能量景观的局部几何信息，近期提出的增强虚时间演化（AITE）原则上能够在有限虚时间内达到基态。为了在 NQS 优化中利用 AITE 的加速收敛特性，我们首先将隐式欧拉步与Galerkin投影相结合，推导出一种隐式 SR 方案，该方案允许采用大步长且保持稳定，其性能显著优于传统方法。我们利用该方案构造 AITE 的隐式版本，并表明，在变分投影和有限步积分下，其加速效果依然存在，同时能量误差呈渐近二次收敛。在有限采样条件下，隐式方案仍能带来稳定的性能提升，而隐式AITE 的优势也日益明显。

arXiv:2611.32236v1 [quant-ph] 24 Nov 2026

量子多体系统的基态决定了其零温性质，但希尔伯特空间的指数级增长使精确对角化仅适用于小尺寸系统 [1]。为避免这种维数灾难，变分方法将波函数限制在一个参数化流形中，其规模远小于相应的希尔伯特空间 [2, 3]。张量网络表示变分波函数 [4–6]，而变分蒙特卡罗（VMC）则通过随机采样对其进行优化 [7, 8]。这两种方法都利用了这样一个事实：许多物理系统的低能态都可以用紧凑的表示很好地近似 [9, 10]。神经量子态（NQS）使用神经网络（NN）进行参数化 [11–15]。结合 VMC，NQS 已成功应用于各种相互作用的量子多体系统，相关综述见文献 [16, 17]。然而，对于大型且复杂的系统，尤其是在高于一维的情况下，精确的基态优化仍然是一项挑战 [11, 15, 18, 19]。

给定一个变分流形，目标是高效地使状态演化至基态。一种自然的方法是虚时间演化 (ITE)，它通过 Wick旋转 $t \to -i\tau$ 与实时间演化形式上相关；该旋转以非幺正传播子 [20–23]取代幺正动力学。虚时间演化是投影量子蒙特卡罗 [24–30] 的基础，也被广泛用于优化神经量子态，使其趋向基态 [11, 31]。对于哈密顿量 $\hat{H}$ 和一个与基态有非零重叠的初态 $|\psi(0)\rangle$，演化 $|\psi(\tau)\rangle = e^{-\tau H}|\psi(0)\rangle$ 会指数级抑制激发态贡献，使状态渐近趋于基态 $\tau \to \infty$。对于归一化态 $|\phi(\tau)\rangle = |\psi(\tau)\rangle / \||\psi(\tau)\|$，其瞬时期望值为 $\langle A\rangle \equiv \langle\phi(\tau)|A|\phi(\tau)\rangle$，相应的动力学可写为

$$\partial_\tau |\phi(\tau)\rangle = -\hat{G}|\phi(\tau)\rangle \ , \tag{1}$$

其中，依赖于状态的厄米生成元 $\hat{G} = \hat{G}^\dagger$ 满足 $\langle G\rangle = 0$。对于（归一化的）虚时间演化，式（1）给出 $\hat{G}_{\mathrm{ITE}} = H$，其中 $H = H - E(\tau)$ 且 $E(\tau) \equiv \langle H\rangle$。能量衰减速率完全由其方差决定。$\dot{E}_{\mathrm{ITE}} = -2\mu_2$，其中 $\mu_m \equiv \langle H^m\rangle$ 表示瞬时能量分布的第 $m$ 阶中心矩。

最近，Benavides-Riveros 等人提出了增强虚时间演化（AITE） [32]，该方法通过二次的、依赖于状态的生成元利用能量分布的高阶矩

$$\hat{G}_{\mathrm{AITE}} = a\bar{H} + b(\bar{H}^2 - \mu_2) \ , \tag{2}$$

系数为 $a = \cos(2\varphi)$ 和 $b = \sin(2\varphi)/(2\sqrt{\mu_2})$，其中 $\tan(2\varphi) = \kappa/2$，$\kappa = \mu_3/\mu_2^{3/2}$ 为费希尔-皮尔逊偏度系数。由此得到的局域能量流是双重括号公式的一种推广 [33, 34]，并且给出

$$\dot{E}_{\mathrm{AITE}} = -2\mu_2\sqrt{1 + \kappa^2/4} \le \dot{E}_{\mathrm{ITE}} \ , \tag{3}$$

在严格不等式成立的情况下 $\kappa \neq 0$。我们通过绝对能量误差 $\epsilon = E - E_0$ 和相对能量误差 $\epsilon_{\rm rel} = |\epsilon|/|E_0|$来量化收敛性，其中 $E_0$ 表示基态能量。值得注意的是，与仅在有能隙系统中渐近收敛的虚时间演化（ITE）不同，AITE 能在有限虚时间 $\epsilon = 0$达到 $\tau_*$ ，前提是能量降至低于第一激发态能量 [32]。

AITE 的加速收敛能否经受变分 NQS 投影、有限采样和有限步长离散化的考验？传统的 NQS 优化依赖于显式（前向）欧拉步进：在当前状态处计算时间导数，并受到 $G$ 的谱带宽 $W_G$ 所设定的稳定性约束。对于 AITE，该带宽随 $\epsilon \to 0$ 发散。为克服这一限制，我们将隐式（后向）欧拉步进——其改为在更新后的状态处计算时间导数——投影到 NQS 流形上，并比较由此得到的隐式 虚时间演化（iITE）和隐式 AITE（iAITE）方案。两种隐式 方案都显著加快了收敛；随着试探态的表达能力增强，iAITE 始终优于 iITE，表明 AITE 的加速效果能够经受变分投影的考验。

模型哈密顿量。 下文研究正方晶格上的 (i) 横场伊辛（TFI）模型 [11],和 (ii) $J_1$-$J_2$ 海森堡反铁磁体 [18], 其包含 $N = L_x \times L_y$ 个格点。在这两种情况下，我们都采用 $\sigma^z$ 直积基，其中构型为$\mathbf{s} = (s_1, \ldots, s_N)$; $s_\ell = \pm 1$，并施加周期性边界条件（PBC）。此外，对于 $J_1$-$J_2$ 模型，我们施加 Marshall 符号旋转 [35]。我们定义 $\hat{\boldsymbol{S}}_\ell = \hat{\boldsymbol{\sigma}}_\ell/2$，其中 $\hat{\boldsymbol{\sigma}}_\ell = (\hat{\sigma}^x_\ell, \hat{\sigma}^y_\ell, \hat{\sigma}^z_\ell)$。关于晶格、哈密顿量、对称性扇区和变分试探态的详细信息，见补充材料 （SM） [36]。

我们首先考虑较简单、无受挫的横场伊辛模型，

$$\hat{H}_{\rm I} = -J \sum_{\langle \ell, \ell' \rangle} \hat{\sigma}^z_\ell \hat{\sigma}^z_{\ell'} - h \sum_{\ell} \hat{\sigma}^x_\ell \,, \tag{4}$$

其中 $\langle \ell, \ \ell' \rangle$ 表示最近邻键。该模型无法精确求解，并且在热力学极限下于 $h_c/J \simeq 3.04$处存在量子临界点 [37]。我们固定 $J = 1$ ，并考察 $h = 4 > h_c$ 顺磁相中的情形（图 1–3）以及 $h = 2 < h_c$ 铁磁相中的情形（图 4）。为探索更具挑战性的情形，我们选择 $J_1$-$J_2$ 海森堡模型 [38],

$$\hat{H}_{J_1 J_2} = J_1 \sum_{\langle \ell, \ell' \rangle} \hat{\boldsymbol{S}}_\ell \cdot \hat{\boldsymbol{S}}_{\ell'} + J_2 \sum_{\langle\langle \ell, \ell' \rangle\rangle} \hat{\boldsymbol{S}}_\ell \cdot \hat{\boldsymbol{S}}_{\ell'} \,, \tag{5}$$

其中 $\langle\langle \ell, \ \ell' \rangle\rangle$ 表示次近邻键。我们固定 $J_1 = 1$ 并考察 $J_2 = 0.2$ 奈尔有序相中的情形（图 3 和 4) 和 $J_2 = 0.5$ 在强挫折区域（图1和 2） [18, 39]。

变分虚时间演化。 我们将多体波函数表示为 $\Psi_{\boldsymbol{\theta}}(\mathbf{s}) = \langle \mathbf{s} | \Psi_{\boldsymbol{\theta}} \rangle$,

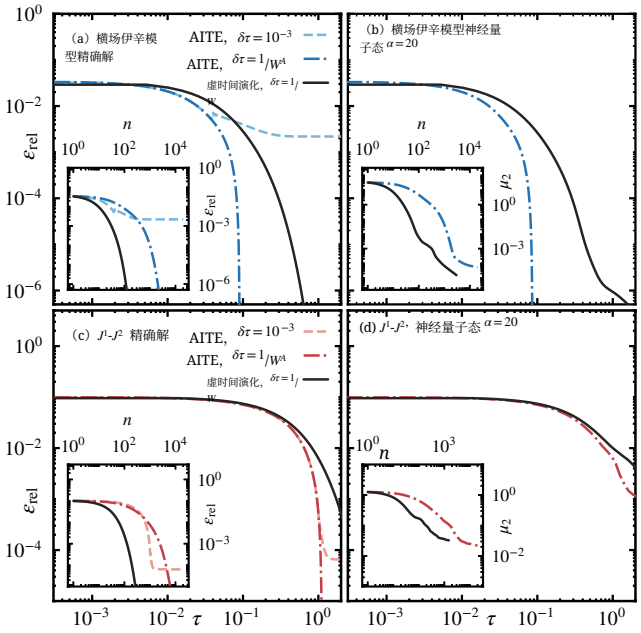


图1. 正方晶格上横场伊辛模型的有限步显式 AITE 和虚时间演化的相对能量误差 $\epsilon_{\rm rel}$ ，对应 (a)、(b)；以及 $J_1$-$J_2$ 模型在 $J_2 = 0.5$ 下的结果，对应 (c)、(d)，随虚时间 $\tau$的变化。(a)、(c) 对横场伊辛模型和 $J_1$-$J_2$ 模型，分别对显式 [式 (8)] 虚时间演化（实线）、固定 $\delta\tau = 10^{-3}$ 的 AITE（虚线）以及步数受限的 $\delta\tau = 1/W_A$（点划线）进行精确计算。(b)、(d) 分别展示两个模型使用对称性适配受限玻尔兹曼机 ($\alpha = 20$) 进行变分演化 [式 (6)] 的相同结果，其中 AITE 仅显示 $\delta\tau = 1/W_A$ 。采用全求和。插图：(a)、(c) 为 $\epsilon_{\rm rel}$；(b)、(d) 为能量方差 $\mu_2$ 随离散时间步 $n$的变化。

由通常为复数的参数 $\boldsymbol{\theta}$参数化，其中$\mathbf{s}$ 表示所选计算基中的构型$\{|\mathbf{s}\rangle\}$ ，其由 $N$ 个自旋 1 2组成。随后，瑞利-里兹变分原理给出基态能量的上界$E_0 \leq E_{\rm var} = \min_{\boldsymbol{\theta}} \langle \Psi_{\boldsymbol{\theta}} | H | \Psi_{\boldsymbol{\theta}} \rangle / \langle \Psi_{\boldsymbol{\theta}} | \Psi_{\boldsymbol{\theta}} \rangle$ [8, 40, 41]。下文中，我们用大写希腊字母表示参数化态，用小写字母表示任意希尔伯特空间态。

当将其限制在变分流形 $\mathcal{M}$ 上时，精确的虚时间演化方向会通过含时变分原理 [2, 22, 42], 投影到其切空间上，从而得到适用于神经量子态的随机重构（SR） [11, 43, 44]。SR 使用量子几何张量（QGT） $S$ 对变分力进行预条件处理，这等价于在 $\mathcal{M}$ [45–49] 上进行自然梯度优化。近期研究改进了其可扩展性和条件数 [50–52]，或对其进行了修改 $\mathcal{M}$ [53–55], ， 同时仍将投影更新作为一种显式 离散化。SR 与虚时间演化之间的联系表明，还存在一条互补且非排斥的途径——加速底层演化本身。

神经量子态使用神经网络对波函数进行参数化，其中包含大量 $N_p$ 变分参数。

变分态为$|\Psi_{\boldsymbol{\theta}}\rangle = \sum_{\mathbf{s}} \Psi_{\boldsymbol{\theta}}(\mathbf{s})\,|\mathbf{s}\rangle$，而变分蒙特卡洛通过从$p_{\boldsymbol{\theta}}(\mathbf{s}) \propto |\Psi_{\boldsymbol{\theta}}(\mathbf{s})|^2$中采样来估计期望值。为单独考察虚时间动力学的作用，我们首先采用具有隐单元密度[11, 56, 57] 的对称性适配复数受限玻尔兹曼机（RBM） $\alpha$，并让参数在 $N$晶格平移下共享。除非另有说明，我们使用精确希尔伯特空间求和 $N \le 16$，以将变分投影与随机效应区分开来；对于较大的系统，则使用马尔可夫链蒙特卡洛（MCMC）采样。我们介绍采用精确求和的受限玻尔兹曼机，以及采用 MCMC 采样的视觉变换器（ViT） [58] 。数值实现使用 NetKet [59, 60], 、JAX [61],和 Flax [62], ，并采用我们自行实现的演化算法。

显式随机重构。将式（1）投影到 NQS 流形的切空间上，可得到随机重构方程 [42, 44], 此处针对全纯试探态书写。对于 ViT 等采用实值参数的情况，取 $S$ 和$\mathbf{F}$的实部，

$$(S + \lambda I)\mathbf{d} = -\mathbf{F}[\hat{G}]\,, \quad S_{kl} = \mathrm{Cov}(O_k^*, O_l)\,, \qquad (6)$$

其中 $\mathrm{Cov}(X,Y) \equiv \langle XY\rangle - \langle X\rangle\langle Y\rangle$，$\mathbf{d} \equiv \partial_\tau \boldsymbol{\theta}$ 表示变分更新方向，$O_k = \partial_{\theta_k} \log \Psi_{\boldsymbol{\theta}}$给出对数导数，$S$ 是 QGT [45], ，而$\lambda \in \mathbb{R}_+$ 用于对线性方程组 [13, 44]进行正则化。变分力为 $F_j[G] = \mathrm{Cov}(O_j^*, G_{\mathrm{loc}})$，其中 $\hat{}G_{\mathrm{loc}}$ 是 $G$ 的局域估计量。

在 $\boldsymbol{\theta}$固定时，矩阵 $S$ 不依赖于生成元，因此显式 虚时间演化和 AITE 的区别仅在于变分力 $\mathbf{F}$。对于 AITE，有

$$\mathbf{F}_{\mathrm{AITE}} = a\,\mathbf{F}[\hat{H}] + b\left[\mathbf{F}[\hat{H}^2] - 2E\,\mathbf{F}[\hat{H}]\right]\,. \qquad (7)$$

进入 $a$ 和 $b$ 的矩由同一组蒙特卡洛样本估计，而二次项通常需要计算 $\hat{}H^2$。局部估计量见 SM [36]。

有限时间步长。 连续时间随机重构方程源自 McLachlan变分原理，该原理通过将式（1）的残差范数投影到变分切空间 [21, 42]上来对其进行最小化。尽管还存在其他变分原理和投影策略 [63], 我们关注有限步长离散化（见文末材料），在引入虚时间网格 $\tau_{n+1} = \tau_n + \delta\tau$后得到离散随机重构方程，其中 $\tau_0 = 0$ 且 $n = 0,\ 1,\ 2,\ \ldots\ldots$，其中步长 $\delta\tau$可能随$n$变化。对于归一化态 $|\phi_n\rangle \equiv |\phi\,(\tau_n)\,\rangle$，显式欧拉近似随后给出 [31, 64],

$$|\phi_{n+1}\rangle \simeq \left(1 - \delta\tau \hat{G}_n\right)|\phi_n\rangle\,, \quad \hat{G}_n \equiv \hat{G}[\phi_n]\,, \qquad (8)$$

其中，生成元和更新方向均在当前时刻求值 $|\phi_n\rangle$. 将这一离散化的演化投影到变分切空间上，可得到显式

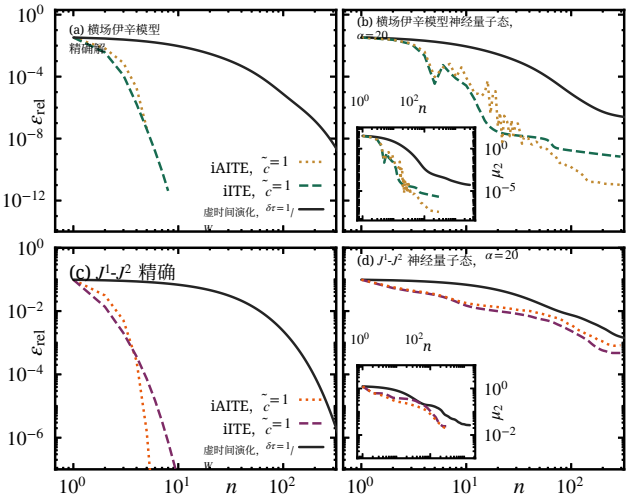


图 2. 正方晶格上有限步长 iITE、iAITE 和显式虚时间演化的相对能量误差 $\epsilon_{\mathrm{rel}}$ ： $L_x = L_y = 4$ （a）、（b）横场伊辛模型在 $h = 4.0$ 处，以及（c）、（d） $J_1$-$J_2$ 模型在 $J_2 = 0.5$ 处，作为离散时间步 $n$的函数。（a）、（c）在 $\delta\tau = 1/W$ 处的精确显式虚时间演化 [式（8）] （实线），以及在 $\delta\tau_0 = 1$ 处、采用方差封顶步长 $\delta\tau_n = \min(\delta\tau_0, \tilde{c}/\sqrt{2\mathrm{Var}(G_n)})$ 和 $\tilde{c} = 1$ 的精确隐式 [式（10）] 虚时间演化（虚线）和 AITE（点线）。（b）、（d）对称性适配受限玻尔兹曼机（$\alpha = 20$）的对应结果。采用全求和。插图：能量方差 $\mu_2$ 随 $n$ 的变化。

SR 更新 $\boldsymbol{\theta}_{n+1} = \boldsymbol{\theta}_n + \delta\tau\mathbf{d}_n$，其中 $\mathbf{d}_n$ 由方程（6）确定。

对于刚性问题，显式欧拉方法仅具有条件稳定性 [65]。也就是说，将$\hat{}$的谱带宽记为 $G_n$，则 $W_G$是渐近显式欧拉稳定性约束的特征值；该约束是离散性而非物理学的性质，在 $\delta\tau < 2/W_G$接近基态时成立（见文末材料）。对于虚时间演化，$W_G \to W \equiv E_{\max} - E_0$ 是哈密顿量的带宽。相比之下，对于 AITE， $W_G \to W_A$ 会随$n$ 变化，并在接近基态时发散 $\epsilon \to 0$,

$$W_A \sim \frac{W^2}{2\sqrt{\Delta_*\epsilon}}\,, \quad \delta\tau_{\max} \sim \frac{4\sqrt{\Delta_*\epsilon}}{W^2}\,, \qquad (9)$$

其中， $\Delta_* = E_* - E_0$ 是可达能隙，即与该态具有非零重叠的最低激发态对应的能隙（见补充材料 [36]），而 $W_A$ 将低能标度 $\Delta_*$与总带宽 $W$联系起来。因此，AITE 流会变得越来越刚性，典型的稳定显式步长随 $\delta\tau_{\max} \propto \sqrt{\epsilon}$趋于零。在极限情况下，显式AITE 每一步的推进幅度因此小于显式虚时间演化，差一个 $\Delta_*/W$ 因子（见文末材料）。这促使我们引入下文的隐式方案，该方案在收敛附近仍能保持有限的步长。

为展示不同生成元$\hat{}$所引发的动力学和刚性 $G_n$，图1比较了显式虚时间演化和 AITE 在两种模型中的表现：精确希尔伯特空间演化对应[图1(a)、(c)] ，而对称性适配受限玻尔兹曼机 [对应式（6）]

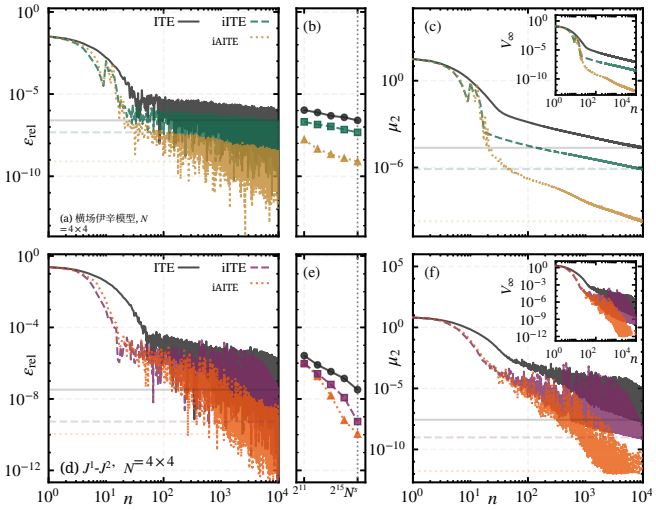


图3. ViT 的变分虚时间演化与离散步数的关系 $n$， [58]采用显式ITE（实线）、iITE（虚线）和 iAITE（点线），使用 $N_s$ MCMC 样本。相对能量误差 $\epsilon_{\rm rel}$ 和能量方差 $\mu_2$ 采用正方晶格 $L_x = L_y = 4$ （a）–（c）横场伊辛模型，参数为 $h = 4.0$；（d）–（f） $J_1$-$J_2$ 模型，参数为 $J_2 = 0.2$。（a）、（d）相对能量误差与离散步数的关系 $n$，使用 $N_s = 32768$ 个样本。（b）、（e）使用 $\bar{\epsilon}_{\rm rel}$ 最后 5% 步数的平均值与 $N_s$的关系。（c）、（f）能量方差 $\mu_2$。插图：V 分数 $V_\infty$ [3]。

采用全求和 [图 1(b)、(d)]。使用适配步长 $\delta\tau = 1/W_A$ 时，精确 AITE 在虚时间 $\tau$ 中比虚时间演化收敛更快，这与式 (3) [32]一致。然而，步长会随 $W_A$缩小，因此 AITE 可能需要比虚时间演化多得多的迭代次数 $n$，才能达到给定的 $\epsilon_{\rm rel}$[（见图 1(a)、(c) 的插图）]。在固定的 $\delta\tau = 10^{-3}$ 下，逐渐增大的 $W_A$ 最终会违反稳定性界限 [（式 (9)）]，导致激发态振幅过冲，使得 $\epsilon_{\rm rel}$ 停滞在有限值（见补充材料 [36]）。在所有运行中，步长中的 $W$ 和 $W_A$ 都被可计算的上界取代（见补充材料 [36]）；变分投影还会增加可表示性约束，进一步限制收敛 [（图 1(b)、(d)）]。

隐式随机重构。克服显式ODE 求解器条件稳定性的一种常见方法是采用隐式时间步进 [65–67]。我们通过将方程（显式欧拉更新（8）替换为其隐式对应形式来推导该方法。与显式欧拉格式不同，该方法在ˆ未来状态处计算时间导数， $|\phi_{n+1}\rangle - |\phi_n\rangle \simeq -\delta\tau \hat{G}_{n+1} |\phi_{n+1}\rangle$，这意味着

$$\left(1 + \delta\tau \hat{G}_{n+1}\right) |\phi_{n+1}\rangle = |\phi_n\rangle \ . \qquad (10)$$

为避免与依赖于状态的生成元相关的非线性隐式求解，我们通过固定ˆ ˆ$G_{n+1} \simeq G_n$来对其进行近似，该生成元作用于更新后的状态 $|\phi_{n+1}\rangle$，而其系数仍取自 $|\phi_n\rangle$。我们将这一线性隐式（后向）欧拉步 [65] 称为隐式，下文均如此称呼。将 $|\phi_{n+1}\rangle$ 展开至一阶

在关于 $|\phi_n\rangle$ 的切空间中，然后对所得关系进行投影，可得到隐式SR方程，

$$(S + \lambda I + \delta\tau K)\,\mathbf{d} = -\mathbf{F} \ , \quad K_{jk} = \langle T_j | \hat{G}_n | T_k \rangle \ , \qquad (11)$$

其中， $|T_k\rangle$ 表示与当前态正交的切向量；数值细节见 SM [36] 。与式（6）不同，式（ $\delta\tau \to 0$）恢复后，式（11）要求方程两侧都包含生成元，通过变分力$\mathbf{F}$ 和切空间矩阵 $K$。我们在 $\tau$处保留局部近似，避免有限步变分投影所需的非线性优化；这类投影基于重叠优化 [68]；见文末材料。

显式欧拉方法需要依赖 $W_G$ 的限制条件，而对于固定的半正定生成元 $G$，式（10）无条件稳定。然而，对于以态为中心且依赖于态的 $G$，这一结论并不能自动推广至投影隐式格式 [式（11）]，。该格式固定生成元，采用线性切空间近似，并根据样本进行估计 [69]。尽管如此，在无噪声双能级极限下，固定 $\delta\tau > 0$时，iAITE 步满足 $\epsilon_{n+1} \simeq 4\epsilon_n^2/(\delta\tau^2\Delta_*^3)$；而 iITE 线性收敛， $\epsilon_{n+1} \simeq \epsilon_n/(1 + \delta\tau\Delta_*)^2$；见 SM [36]。显式方案中导致 $\delta\tau \to 0$的 $W_A$ 发散，正是隐式方案实现二次收敛的原因，因为刚性特征值出现在隐式因子（$1 + \delta\tau$ˆ$G_n$）$^{-1}$的分母中。构造 $K$ 以求解式（11）中的线性方程组所带来的开销，以及增加的采样需求，对于 iITE 仍低于两倍；但对于 iAITE，由于需要计算 ˆ$H^2$，开销可能更大。对于 ITE，式（11）与 Rayleigh–Gauss–Newton 优化 [70],有关，其中惩罚参数取为 $1/\delta\tau$；它也与投影逆迭代 [71, 72], 有关，我们在开发这一方法时注意到了后者，其位移为 $s_n = E_n - 1/\delta\tau$，并需进行正则化。与此同时，我们的虚时间表述将这一方法推广至其他生成元 $G$，包括 AITE 的 $G_A$，并采用基于方差的 $\delta\tau$；见 SM [36]。

在图 2 中，我们展示了隐式方案如何克服显式格式的局限：iITE 和 iAITE 均在基于方差的步长上限中使用 $\delta\tau_0 = 1$ 和 $\tilde{c} = 1$ （见补充材料 [36]），而显式虚时间演化使用 $\delta\tau = 1/W$。对于精确演化， [图 2(a)、(c)], 均以少两个数量级的步数达到给定的$\epsilon_{\rm rel}$ $n$ ，优于显式虚时间演化。投影到对称性适配的 RBM 流形（$\alpha = 20$）并进行全求和会限制可达到的精度，但在两个模型中都保留了相对于显式虚时间演化的优势， [见图 2(b)、(d)]。隐式方案的能量方差 $\mu_2$（插图）始终更低。

在随机采样下，这一优势依然存在（图3），其中采用 ViT 的 MCMC 采样 [58] 取代了全求和，适用于 $h = 4$ [图 3(a)–(c)] 中的横场伊辛模型和 $J_1$-$J_2$ 模型 $J_2 = 0.2$ [图 3(d)–(f)]。在 $N_s = 2^{15}$样本 [图 3(a),(d)], 中，两种 隐式 方案均表现出

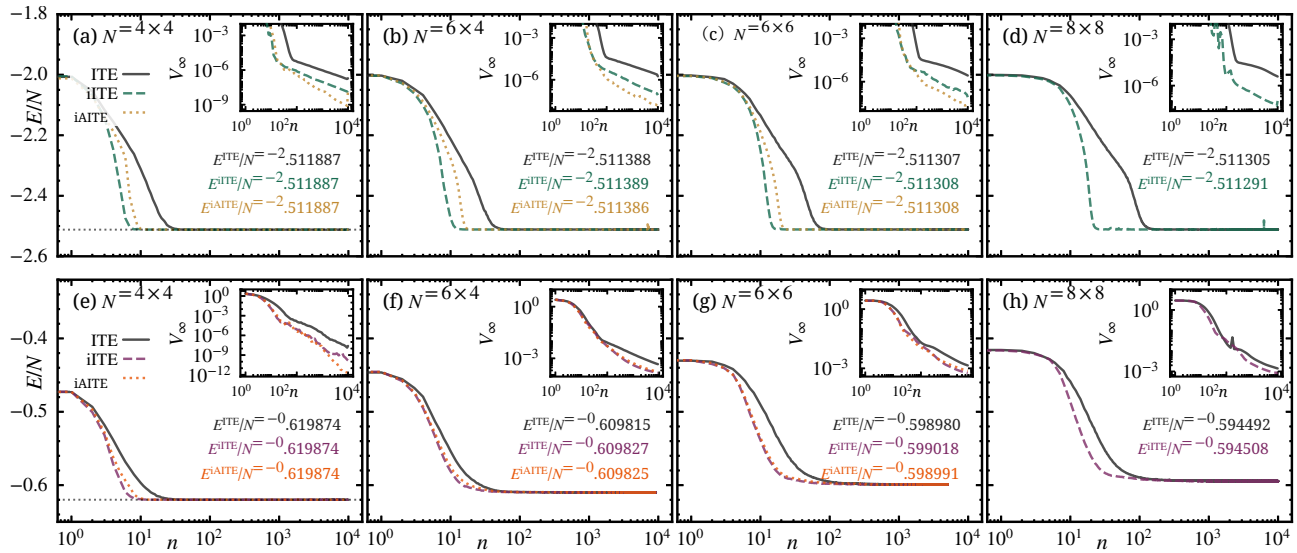


图 4. 每格点能量 $E/N$ 随 ViT [58] 的离散步数 $n$ 变化；ViT 分别采用显式虚时间演化（实线）、iITE（虚线）和 iAITE（点线，对于 $N \le 6\times 6$），并展示不同系统尺寸下的结果：(a)、(e) $N = 4\times 4$；(b)、(f) $6\times 4$；(c)、(g) $6\times 6$；(d)、(h) $8\times 8$，系统位于正方晶格上。面板 (a)–(d) 展示横场伊辛模型在 $h = 2.0$ 下的结果，(e)–(h) 展示 $J_1$-$J_2$ 模型在 $J_2 = 0.2$下的结果；估计的 $E_{(\dots)}/N$为使用每步 $N_s = 16\,384$ 个马尔可夫链蒙特卡罗样本，对最后 $n_{\mathrm{E}} = 500$ 步取平均所得，$\delta\tau_0 = 1$、$c = 1.5$用于显式虚时间演化，$\tilde{c} = 1$用于 iITE 和 iAITE。点线表示可获得的精确 $E_0/N$。插图：V 分数 $V_\infty$ [3]。

显式虚时间演化，其中 iAITE 最终在横场伊辛模型 [图 3(a)–(c)] 中实现了最低的 $\epsilon_{\mathrm{rel}}$。这一优势在不同采样预算下仍然存在 $N_s$ [图 3(b)、(e)]。尽管统计涨落和有限的试探态容量再次限制了可达精度，尤其是对于更具挑战性的 $J_1$-$J_2$ 模型 [图 3(d)–(f)]。对于 iAITE，采样噪声通过设定 $a$ 和 $b$ 的矩 $\mu_2$ 和 $\mu_3$、$\hat{}H^2$ 的估计量，以及 $K$ 的估计值进入；在有限 $N_s$ 下，后者通常不是厄米的（见 SM [36]）。能量方差 [图 3(c)、(f)] 以及插图中所示的相应 V-分数，$V_\infty = N\mu_2/(E - E_\infty)^2$,，其中$E_\infty = \mathrm{Tr}(H)/\dim\mathcal{H}$ 是模拟希尔伯特空间的无限温度能量（对于 $J_1$-$J_2$，为 $S^z = 0$ 扇区）[3],，进一步刻画了收敛变分态的质量。

最后，图4将对比从 $4\times 4$扩展至 $8\times 8$，对应于 $h = 2$ [处的横场伊辛模型，图4(a)–(d)] 以及 $J_1$-$J_2$模型，对应于 $J_2 = 0.2$ [图4(e)–(h)],，其中 iAITE 展示至 $6\times 6$。对于每种尺寸，两种隐式方案都能以比 $E/N$ 少数倍的步数达到平台 $n$，而显式虚时间演化更早且更大幅度地降低 V 分数（见插图）。对于较大的晶格，不同方案得到的能量在每个格点上相差不超过 $\sim 10^{-5}$，且 iAITE 的演化轨迹与 iITE 紧密吻合。随着隐式方案的持续提升$N$，其收敛优势依然存在；而更强的试探态表达能力则有望带来进一步改进，尤其是对 iAITE 而言。

讨论。显式 欧拉步无法利用 AITE 更快的演化速度，因为其刚性迫使 $\delta\tau \to 0$ 取为 $\epsilon \to 0$。我们将隐式 欧拉步与伽辽金投影相结合——以 [73–75] 消除这一限制，从而为一般厄米 生成的流 构造隐式 SR 方案$\hat{}G$ [式（1）]。我们证明，在所测试的模型中，AITE [32]相较于虚时间演化的加速效果，在有限时间步长、变分投影和 MCMC 采样下基本得以保留。在精确 演化中，两种隐式 方案达到给定 $\epsilon_{\mathrm{rel}}$ 所需的步数约比显式虚时间演化少两个数量级；对于采用全求和的投影神经量子态，隐式方案相较于显式 虚时间演化的优势也得以保留。在有限 $N_s$条件下，这些优势在 MCMC 采样中依然存在，尽管要实现加速，试探态 $\Psi_{\boldsymbol{\theta}}$ 的切向量必须能够准确捕捉由$\hat{}G$生成的流的方向。

尽管 iAITE 的表现可能优于 iITE，但计算 $\hat{}H^2$ 及更高阶能量矩 $\mu_2$ 和 $\mu_3$ 需要额外的计算开销，并且需要谨慎进行随机估计。未来改进重要性采样 [76, 77] 或可在一定程度上缓解这一问题。与此同时，iITE 相较于传统 SR 仅带来适度的计算开销，这与相关的 RGN 结果 [70] 一致。

我们的结果表明，隐式变分虚时间演化是改进基态优化的一条通用且前景广阔的路径。在采样、线性求解器和变分试探态方面的进展 $\Psi_{\boldsymbol{\theta}}$，可能会使其应用范围扩展到规模大得多、挑战性更高的系统，包括小能隙区域 [71]。未来有前景的研究方向包括凝聚态物理学和量子化学中的费米子 [19, 78]及非厄米 [79]系统，例如固体和晶格模型 [80–83]以及分子 [84–92]的变相关哈密顿量。

数据和代码可用性。用于生成图的模拟代码和脚本 [93] 将在发表后公开提供，而训练好的网络参数和更大的数值数据集可根据合理请求提供。

***AI*** 辅助工具的使用。AI 工具，包括 Claude Opus 5.5（Anthropic）和 ChatGPT（OpenAI），用于协助英文润色、数值测试开发以及线性求解器优化。作者对所有 AI 辅助贡献均进行了审查和验证。

致谢。我们感谢德国联邦研究、技术与航天部（BMFTR）通过研究计划“量子系统”及资助项目 Quantum Futur 3 提供的资助，项目编号 13N17229；同时感谢亥姆霍兹协会通过倡议与网络基金为亥姆霍兹青年研究者计划下的项目编号 VH-NG-21-08，以及亥姆霍兹量子应用征集下的项目 KA-QUS-02（qFLOW）和 KA-QUS-03（QT-Batt）提供资助。我们感谢帕德博恩大学并行计算中心（PC2）（项目 hpc-prf-qehpc 和 hpc-prf-nniap）的 NHR 中心高性能计算机，以及亥姆霍兹德累斯顿-罗森多夫中心（HZDR）的 Rosi高性能计算集群为我们提供计算时间。PC2 由德国联邦研究、技术与航天部和参与 NHR 的州政府共同支持（www.nhr-verein.de/unsere-partner）。

## 文末材料

### A：变分投影与显式SR

我们通过在时间网格 $\tau_n$上采用显式欧拉步 [方程（8）]，得到显式SR 更新。对于步 $n$处的归一化变分态 $|\Phi_n\rangle \equiv |\Phi(\boldsymbol{\theta}_n)\rangle = |\Psi_n\rangle / \|\Psi_n\|$，欧拉目标态 $|\phi_{n+1}\rangle$ 通常不属于该流形 $\mathcal{M}$ [42]。与此同时，参数偏移$\delta\boldsymbol{\theta} = (\delta\theta_1, \ldots, \delta\theta_{N_p})$ 给出了未归一化变分态的一阶展开

$$|\Psi_{n+1}\rangle \simeq |\Psi_n\rangle + \sum_k \delta\theta_k \, |\partial_k \Psi_n\rangle \ , \tag{A1}$$

其中 $|\partial_k \Psi\rangle \equiv \partial_{\theta_k} |\Psi\rangle$ 表示在步 $n$处求得的偏导数。令变分速度为 $\partial_\tau \boldsymbol{\theta}|_{\tau_n}$，且 $\delta\boldsymbol{\theta} = \delta\tau \mathbf{d}_n$，方程（A1）变为

$$|\Psi_{n+1}\rangle \simeq |\Psi_n\rangle + \delta\tau \sum_k d_{n,k} \, |\partial_k \Psi_n\rangle \ . \tag{A2}$$

这些导数包含与 $|\Phi_n\rangle$平行的分量，它们只影响归一化和全局相位 [2, 94]。去除这些分量，并通过 $\Pi_\perp^{(n)} = I - |\Phi_n\rangle\langle\Phi_n|$ 处理归一化后，得到切向量

$$\begin{aligned} |T_k^{(n)}\rangle &\equiv \|\Psi_n\|^{-1} \Pi_\perp^{(n)} \, |\partial_k \Psi_n\rangle \\ &= \|\Psi_n\|^{-1} \left( |\partial_k \Psi_n\rangle - |\Phi_n\rangle \langle\Phi_n|\partial_k \Psi_n\rangle \right) \ . \end{aligned} \tag{A3}$$

对数导数 $O_k^{(n)}(\mathbf{s}) = \partial_{\theta_k} \log \Psi_{\boldsymbol{\theta}_n}(\mathbf{s})$ 定义对角算符 $\hat{O}_k^{(n)}|\mathbf{s}\rangle \equiv O_k^{(n)}(\mathbf{s})\,|\mathbf{s}\rangle$，使得 $|\partial_k \Psi_n\rangle = \hat{O}_k^{(n)} |\Psi_n\rangle$ 对于复参数和实参数均成立，且式（A3）变为

$$|T_k^{(n)}\rangle = \left( \hat{O}_k^{(n)} - \left\langle O_k^{(n)} \right\rangle \right) |\Phi_n\rangle \ . \tag{A4}$$

根据McLachlan变分原理 [21,42]，$\mathbf{d}_n$ 可使残差平方范数最小，其中 $|R^{(n)}\rangle = \sum_k d_{n,\,k} |T_k^{(n)}\rangle + \hat{G}_n |\Phi_n\rangle$，用于衡量与虚时间演化的偏差，

$$\mathcal{R}(\mathbf{d}_n) \equiv \|R^{(n)}\|^2 = \left\| \sum_k d_{n,k} |T_k^{(n)}\rangle + \hat{G}_n |\Phi_n\rangle \right\|^2 \ . \tag{A5}$$

接下来定义变分力

$$F_k^{(n)} \equiv \langle T_k^{(n)} | \hat{G}_n |\Phi_n\rangle = \mathrm{Cov}(O_k^{(n)*}, G_{\mathrm{loc}}^{(n)}) \ , \tag{A6}$$

它们与局域估计量 $G_{\mathrm{loc}}^{(n)}(\mathbf{s}) = \langle \mathbf{s}|\hat{G}_n|\Phi_n\rangle / \langle \mathbf{s}|\Phi_n\rangle$ 以及量子几何张量 $S_{kl}^{(n)} \equiv \langle T_k^{(n)} | T_l^{(n)} \rangle = \mathrm{Cov}(O_k^{(n)*}, O_l^{(n)})$ 一起，得到平方残差，其展开式为

$$\begin{aligned} \mathcal{R}(\mathbf{d}_n) = &\sum_{k,l} d_{n,k}^* \, d_{n,l} \, S_{kl}^{(n)} \\ &+ \sum_k \left( d_{n,k}^* F_k^{(n)} + d_{n,k} F_k^{(n)*} \right) + \langle \hat{G}_n^2 \rangle \ . \end{aligned} \tag{A7}$$

假设$d_{n,\,k}$ 和 $d_{n,\,k}^*$ 相互独立，并施加驻定性条件，对后者求得 $\partial_{d_{n,\,k}^*} \mathcal{R} = 0$，可得

$$\sum_l S_{kl}^{(n)} \, d_{n,l} + F_k^{(n)} = 0 \ , \tag{A8}$$

这就得到了未正则化的随机重配置关系 $S^{(n)}\mathbf{d}_n = -\mathbf{F}^{(n)}[\hat{G}_n]$。实际上，当 $\lambda I$ 奇异或条件数较差时，会引入对角位移 $S^{(n)}$ 来稳定解，从而得到方程（6）。当 $\boldsymbol{\theta}$ 被限制为实数时，驻定条件则给出 实部$\{S^{(n)}\}\mathbf{d}_n = -$实部$\{\mathbf{F}^{(n)}[\hat{G}_n]\}$。

对于全纯试探态，McLachlan方法、Dirac-Frenkel方法和Galerkin方法在复线性切空间上是一致的 [42, 95]。特别是，对于显式欧拉法，

变分态与欧拉目标之差除以 $\delta\tau$，恰好就是残差 $|R^{(n)}\rangle$[式（A5）]。将其投影到瞬时切空间上的Galerkin投影为 [73]

$$0 = \langle T_j^{(n)} | R^{(n)} \rangle = \sum_k S_{jk}^{(n)} d_{n,k} + F_j^{(n)} \ , \tag{A9}$$

这与方程（A8）中的驻定条件一致。

### B：谱刚性与不同的欧拉格式

显式欧拉格式。我们比较方程（8）中引入的厄米生成元 $G_n$ 的显式欧拉离散化的谱稳定性。对于虚时间演化和 AITE 生成元，它们都是 $H$ 的函数，因此共享其本征态 $G_n|E_j\rangle = g_{j,n}|E_j\rangle$。将归一化后的态展开为 $|\phi_n\rangle = \sum_j c_{j,n}|E_j\rangle$，其中 $c_{j,n} = \langle E_j|\phi_n\rangle$，则系数按如下方式演化：

$$c_{j,n+1} = \|(1-\delta\tau\hat{G}_n)|\phi_n\rangle\|^{-1}(1-\delta\tau g_{j,n})c_{j,n} \ . \tag{B1}$$

归一化因子在比值中相消，因此得到

$$\frac{c_{j,n+1}}{c_{0,n+1}} = \frac{c_{j,n}}{c_{0,n}} \frac{1-\delta\tau g_{j,n}}{1-\delta\tau g_{0,n}} \ , \tag{B2}$$

当且仅当$|(1-\delta\tau g_{j,n})/(1-\delta\tau g_{0,n})| < 1$时，每个激发态分量相对于基态的抑制都会减弱，因此 $0 < \delta\tau < 2/(g_{j,n} + g_{0,n})$，对每个 $j > 0$均成立。定义 $W_G \equiv \max_j g_{j,n} - g_{0,n}$，只要分母为正，中心步长在$0 < \delta\tau < 2/(W_G + 2g_{0,n})$时仍保持稳定。接近基态时，$g_{0,n} \to 0$，由此得到

$$\delta\tau_{\max} \simeq 2/W_G \ . \tag{B3}$$

对于虚时间演化，$W_G = W$；而对于 AITE，$W_G = W_A$ 由瞬时能量矩决定，并且渐近发散，这意味着 $\delta\tau_{\max} \to 0$。

为估计渐近时间步长，我们采用双能级近似（参见 SM [36]），此时剩余的激发态布居主要由最低的可达激发构成，可达能隙为 $\Delta_*$。由于 $\epsilon \to 0$，相应的偏度按 $\kappa \simeq \sqrt{\Delta_*/\epsilon}$发散，最优 AITE 参数满足

$$a \simeq 2\sqrt{\frac{\epsilon}{\Delta_*}} \ , \quad b \simeq \frac{1}{2\sqrt{\Delta_*\epsilon}} \ . \tag{B4}$$

在这一状态下，增长的 AITE带宽变为

$$W_A = W[a + b(W-2\epsilon)] \simeq \frac{W^2}{2\sqrt{\Delta_*\epsilon}} \ , \tag{B5}$$

这意味着 $\delta\tau_{\rm A}^{\max} \simeq 4\sqrt{\Delta_*\epsilon}/W^2$，参见公式（9）。因此，尽管连续时间AITE加速随着 $\mathcal{S}_A = \sqrt{1+\kappa^2/4} \simeq \frac{1}{2}\sqrt{\Delta_*/\epsilon}$[公式（3）],最大稳定显式欧拉步长趋于零 $\sqrt{\epsilon}$，且它们的乘积保持有限，

$$\mathcal{S}_A \delta\tau_{\rm A}^{\max} \simeq 2\Delta_*/W^2 \ , \tag{B6}$$

其比虚时间演化的取值 $2/W$ 小$\Delta_*/W$倍。

隐式欧拉格式。相比之下，隐式欧拉是 A稳定且 L稳定的，因此刚性高能模会受到阻尼，这与 Crank–Nicolson 格式或隐式中点格式不同 [65]。接近基态时，两个生成元满足 $g_{0,n} \to 0^-$ 以及$g_{j,n} > 0$，其中 $j > 0$。对于虚时间演化，$g_{0,n} = -\epsilon$；而对于 AITE，有$g_{0,n} \simeq -\sqrt{\Delta_*\epsilon}/2$。因此，对于任意固定的 $\delta\tau > 0$ 并且充分接近收敛时，我们有 $0 < 1+\delta\tau g_{0,n} < 1+\delta\tau g_{j,n}$，这保证了 $|(1+\delta\tau g_{0,n})/(1+\delta\tau g_{j,n})| < 1$。因此，与显式方法不同，隐式欧拉不需要 $\delta\tau \to 0$，因为 $\epsilon \to 0$。特别地，理想化的双能级隐式 AITE 在 $\epsilon$中表现出二次收敛；参见 SM [36]。

### C：隐式随机重配置

我们将变分投影扩展到隐式欧拉离散化方程（10）。全隐式格式要求自洽求解新态$\hat{}$ $\hat{}$及其依赖于状态的生成元 $G_{n+1} = G[\Phi_{n+1}]$ [67]。此处，我们采用正文中的线性隐式步，$\hat{}$ $\hat{}$并将生成元固定在当前态，$G_{n+1} \simeq G_n \equiv \hat{}G[\Phi_n]$。冻结生成元不会改变不动点，因为在不动点处 $\Phi_{n+1} = \Phi_n$，因此$\hat{}$ $\hat{}G_{n+1} = G_n$。

用关于 $|\Phi_n\rangle$的一阶切空间展开近似新态，并将归一化和相位固定为当前 $n$ 态，得到

$$\left(1+\delta\tau\hat{G}_n\right) \underbrace{\left(|\Phi_n\rangle + \delta\tau \sum_k d_{n,k}|T_k^{(n)}\rangle\right)}_{\simeq|\Phi_{n+1}\rangle} \simeq |\Phi_n\rangle \ . \tag{C1}$$

减去 $|\Phi_n\rangle$ 并除以 $\delta\tau$，得到隐式欧拉残差

$$|i^{(n)}\rangle = \hat{G}_n|\Phi_n\rangle + \sum_k d_{n,k}\left(1+\delta\tau\hat{G}_n\right)|T_k^{(n)}\rangle \ . \tag{C2}$$

由于切空间涵盖所有一阶变分位移，我们施加伽辽金正交条件 [73–75]，$\langle T_j^{(n)}|i^{(n)}\rangle = 0$，对于 $j = 1, \ldots\ldots, N_p$。由此得到

$$0 = \langle T_j^{(n)}|\hat{G}_n|\Phi_n\rangle + \sum_k d_{n,k}\left[\langle T_j^{(n)}|T_k^{(n)}\rangle + \delta\tau\langle T_j^{(n)}|\hat{G}_n|T_k^{(n)}\rangle\right] . \tag{C3}$$

这等价地可表示为

$$\sum_k \left(S^{(n)}_{jk} + \delta\tau K^{(n)}_{jk}\right) d_{n,k} = -F^{(n)}_j \,, \qquad \text{(C4)}$$

其中 $K^{(n)}_{jk} = \langle T^{(n)}_j | \hat{} G_n | T_{(n)\,k} \rangle$ 表示切空间矩阵。加入对角正则化 $\lambda I$后，可得回式（11）。

当 $\boldsymbol{\theta}$ 为实数时（如 ViT），对该式进行投影可得 (C2)

$$\mathrm{Re}\left\{S^{(n)} + \delta\tau K^{(n)}\right\} \mathbf{d}_n = -\,\mathrm{Re}\left\{\mathbf{F}^{(n)}\right\} \,, \qquad \text{(C5)}$$

当 $\delta\tau \to 0$时，该式可化简为 $\mathrm{Re}\{S^{(n)}\}\mathbf{d}_n = -\mathrm{Re}\{\mathbf{F}^{(n)}\}$，这与显式 SR 的实数形式一致。

投影后的隐式 欧拉格式可恢复 显式SR，即 $\delta\tau \to 0$。与残差范数最小化$^{\partial_{d^*_{n\text{相比},\ k}} \|i^{(n)}\|^2 = 0}$Galerkin投影避免了引入涉及 $\hat{} G_n^2$的额外矩阵元，从而降低了计算和统计成本。然而，在时间步长较大时，有效矩阵 $S^{(n)} + \delta\tau K^{(n)}$ 未必仍保持正定。补充材料 [36]中讨论了残差范数最小化、正则化和数值稳定性。

该方法与 VMC线性方法 共用切空间构造 [96, 97]。不同于通过广义本征值问题来最小化能量的线性方法，我们的方法对隐式 Euler 方程进行投影，从而得到一个依赖于时间步长的线性系统，也适用于 AITE 等依赖于状态的生成元。

[1]H. Lin、J. Gubernatis、H. Gould 和 J. Tobochnik，《量子系统的精确对角化方法》，Computers in Physics **7**, 400 (1993)。[2] L. Hackl、T. Guaita、T. Shi、J. Haegeman、E. Demler 和 I. Cirac，《变分方法的几何：封闭量子系统的动力学》，SciPost Physics **9**，048 (2020)。[3] D. Wu、R. Rossi、F. Vicentini、N. Astrakhantsev、F. Becca、X. Cao、J. Carrasquilla、F. Ferrari、A. Georges、M. Hibat-Allah、M. Imada、A. M. Läuchli、G. Mazzola、A. Mezzacapo、A. Millis、J. R. Moreno、T. Neupert、Y. Nomura、J. Nys、O. Parcollet、R. Pohle、I. Romero、M. Schmid、J. M. Silvester、S. Sorella、L. F. Tocchio、L. Wang、S. R. White、A. Wietek、Q. Yang、Y. Yang、S. Zhang 和 G. Carleo，《量子多体问题的变分基准》，Science **386**，296 2024 2302 ( )，arXiv：。04919 [量子物理]。[4] S. R. White，《量子重整化群的密度矩阵表述》，Physical Review Letters **69**，2863 1992( )。[5] U. Schollwöck，《矩阵乘积态时代的密度矩阵重整化群》，Annals of Physics**326**，96 (2011)。

[6] R. Orús，《复杂量子系统的张量网络》，Nature Reviews Physics 1, 538 (2019). [7] W. M. C. Foulkes、L. Mitas、R. J. Needs 和 G. Rajagopal，《固体的量子蒙特卡罗模拟》，Reviews of Modern Physics 73, 33 (2001). [8] F. Becca 和 S. Sorella，《关联系统的量子蒙特卡罗方法》（剑桥大学出版社，剑桥，2017）。[9] F. Verstraete 和 J. I. Cirac，《矩阵乘积态能够忠实地表示基态》，Physical Review B 73, 094423 2006 0505140 ( ), arXiv:cond-mat/ 。[10] M. B. Hastings，《一维量子系统的面积定律》，Journal of Statistical Mechanics: Theory and Experiment 2007, P08024 (2007), arXiv:0705.2024 [quant-ph]。[11] G. Carleo 和 M. Troyer，《用人工神经网络求解量子多体问题》，Science 355, 602 (2017). [12] G. Torlai、G. Mazzola、J. Carrasquilla、M. Troyer、R. Melko 和 G. Carleo，《神经网络量子态层析》，Nature Physics 14, 447 (2018). [13] M. Schmitt 和 M. Heyl，《利用人工神经网络研究二维量子多体动力学》，Physical Review Letters 125, 100503 (2020). [14] M. Hibat-Allah、M. Ganahl、L. E. Hayward、R. G. Melko 和 J. Carrasquilla，《循环神经网络波函数》，Physical Review Research 2, 023358 (2020). [15] M. Reh、M. Schmitt 和 M. Gärttner，《优化神经量子态的设计选择》，Physical Review B 107, 195115 (2023). [16] M. Medvidović 和 J. R. Moreno，《用于多体物理的神经网络量子态》，The European Physical Journal Plus 139, 631 (2024). [17] H. Lange、A. Van De Walle、A. Abedinnia 和 A. Bohrdt，《从架构到应用：神经量子态综述》，Quantum Science and Technology 9, 040501 (2024). [18] K. Choo、T. Neupert 和 G. Carleo，《利用神经网络量子态研究二维受挫 $J_1$-$J_2$ 模型》，Physical Review B 100, 125124 (2019), arXiv:1903.06713 [cond-mat.str-el]。[19] D. Pfau、J. S. Spencer、A. G. d. G. Matthews 和 W. M. C. Foulkes，《利用深度神经网络对多电子薛定谔方程进行第一性原理求解》，Physical Review Research 2, 033429 (2020), arXiv:1909.02487 [physics.chem-ph]。[20] G. C. Wick，《贝特-萨尔皮特波函数的性质》，Physical Revie w 96, 1124 (1954). [21] A. McLachlan，《含时薛定谔方程的变分解》，Molecular Physics 8, 39 1964( ). [22] S. McArdle、T. Jones、S. Endo、Y. Li、S. C. Benjamin 和 X. Yuan，《基于变分试探态的虚时间演化量子模拟》，npj Quantum Information 5, 75 (2019). [23] F. Verstraete、J. J. García-Ripoll 和 J. I. Cirac，《矩阵乘积密度算符：有限温度和耗散系统的模拟》，Phys. Rev. Lett. 93, 207204 2004( ). [24] S. Zhang 、 J. Carlson 和 J. E. Gubernatis，《费米子基态的约束路径蒙特卡罗方法》，Physical Review B 55, 7464 (1997). [25] A. W. Sandvik 和 H. G. Evertz，《价键基中的变分与投影量子蒙特卡罗模拟的环更新》

与投影量子蒙特卡罗模拟的环更新，Physical Review B 82, 024407 (2010). [26] G. H. Booth、A. J. W. Thom 和 A. Alavi，《无固定节点的费米子蒙特卡罗：斯莱特行列式空间中的生死消长》，《化学物理学杂志》131, 054106 (2009). [27] D. Cleland、G. H. Booth 和 A. Alavi，《通讯：适者生存：加速全组态相互作用量子蒙特卡罗的收敛》，《化学物理学杂志》132, 041103 (2010). [28] J. S. Spencer、N. S. Blunt、S. Choi、J. Etrych、M.-A. Filip、W. M. C. Foulkes、R. S. T. Franklin、W. J. Handley、F. D. Malone、V. A. Neufeld、R. Di Remigio、T. W. Rogers、C. J. C. Scott、J. J. Shepherd、W. A. Vigor、J. Weston、R. Xu 和 A. J. W. Thom，《HANDE-QMC 项目：从基态出发的开源随机量子化学》，《化学理论与计算杂志》15, 1728 (2019). [29] K. Guther、R. J. Anderson、N. S. Blunt、N. A. Bogdanov、D. Cleland、N. Dattani、W. Dobrautz、K. Ghanem、P. Jeszenszki、N. Liebermann、G. L. Manni、A. Y. Lozovoi、H. Luo、D. Ma、F. Merz、C. Overy、M. Rampp、P. K. Samanta、L. R. Schwarz、J. J. Shepherd、S. D. Smart、E. Vitale、O. Weser、G. H. Booth 和 A. Alavi，《NECI：N 电子组态相互作用，侧重于最先进的随机方法》，《化学物理学杂志》153, 034107 (2020). [30] H. Shi 和 S. Zhang，《真实材料辅助场量子蒙特卡罗的一些最新进展》，《化学物理学杂志》154, 024107 (2021). [31] E. Ledinauskas 和 E. Anisimovas，《神经网络量子态的可扩展虚时间演化》，SciPost Physics **15**, 229 2023 2307 15521 ( ), [ ] arXiv: 。quant-ph 。[32] C. L. Benavides-Riveros、P. Sharma、P. Frisoni 和 F. Šimkovic，《利用局部几何信息增强虚时间演化》(2026), arXiv:2606.23934 [quant-ph]。[33] L. Xiaoyue、M. Robbiati、A. Pasquale、E. Pedicillo、A. Wright、S. Carrazza 和 M. Gluza，《优化双重括号量子算法的策略》(2024), arXiv:2408.07431 [quant-ph]。[34] M. Gluza、J. Son、B. H. Tiang、R. Zander、R. Seidel、Y. Suzuki、Z. Holmes 和 N. H. Y. Ng，《用于量子虚时间演化的双重括号量子算法》，Physical Review Letters 136, 020601 (2026). [35] W. Marshall，《反铁磁性》，《伦敦皇家学会会刊：A辑，数学与物理科学》232, 48 (1955). [36] 参见补充材料 [URL...]，了解有关虚时间演化、随机重构和数值方法的更多详情，其中包括参考文献 [98–106]。[37] H. W. J. Blöte 和 Y. Deng，《横场伊辛模型的团簇蒙特卡罗模拟》，《物理评论 E》66, 066110 (2002). [38] P. Chandra 和 B. Doucot，《受挫方形海森堡晶格在大 S 下可能存在的自旋液体态》，Physical Review B 38, 9335 (1988). [39] S.-S. Gong、W. Zhu、D. N. Sheng、O. I. Motrunich 和 M. P. A. Fisher，《方格有序相与量子自旋-1/2 J1-J2模型方格海森堡模型的相图，Physical Review Letters 113, 027201 (2014), arXiv:1311.5962 [凝聚态物理：强关联电子系统]。[40] W. Ritz，《一种求解数学物理中某些变分问题的新方法》，Jour- nal für die reine und angewandte Mathematik 135, 1 (1909)。[41] J. K. L. MacDonald，《Rayleigh-Ritz 变分法的逐次近似》，Physical Review 43, 830 1933( )。[42] J. Haegeman、J. I. Cirac、T. J. Osborne、I. Pizorn、H. Verschelde 和 F. Verstraete，《量子晶格的含时变分原理》，Physical Review Letters 107, 070601 (2011), arXiv:1103.0936 [凝- 聚态物理：强关联电子系统]。[43] S. Sorella，《用于变分量子蒙特卡罗的广义 Lanczos 算法》，Physical Review B 64, 024512 (2001)。[44] S. Sorella，《变分蒙特卡罗方法中的波函数优化》，Physical Review B 71, 241103 (2005)。[45] J. P. Provost 和 G. Vallee，《量子态流形上的黎曼结构》，Communications in Math- ematical Physics 76, 289 (1980)。[46] S.-i. Amari，《自然梯度在学习中的高效应用》，Neural Computation 10, 251 (1998)。[47] J. Stokes、J. Izaac、N. Killoran 和 G. Carleo，《量子自然梯度》，Quantum 4, 269 (2020)。[48] J. Gacon、C. Zoufal、G. Carleo 和 S. Woerner，《量子费舍尔信息的同步扰动随机近似》，Quantum 5, 567 (2021)。[49] D. Fitzek、R. S. Jonsson、W. Dobrautz 和 C. Schäfer，《使用 qBang 优化变分量子算法：交织度量与动量，高效探索平坦能量景观》，Quantum 8, 1313 (2024)。[50] A. Chen 和 M. Heyl，《通过高效优化增强深度神经量子态》，Nature Physics 20, 1476 (2024)。[51] R. Rende、L. L. Viteritti、L. Bardone、F. Becca 和 S. Goldt，《用于优化大规模神经网络量子态的简单线性代数恒等式》，Communi- cations Physics 7, 260 (2024)。[52] C.-Y. Park 和 M. J. Kastoryano，《神经量子态学习的几何》，Physical Review Research 2, 023232 2020( )。[53] R. Rende、S. Goldt、F. Becca 和 L. L. Viteritti，《神经网络量子态的微调》，Physi- cal Review Research **6**, 043280 2024 2403 07795 ( ), [ ] arXiv: . 凝聚态物理：强关联电子系统 . [54] S. B. Kožić、V. Zlatić、F. Franchini 和 S. M. Gi- ampaolo，《探索基旋转对 NQS 性能的影响》(2025)。[55] S. Cantori、L. Brodoloni、E. Recchi、E. Costa、B. Juliá-Díaz 和 S. Pilati，《用于量子多体系统的自适应基于采样的神经对角化》，Sci- Post Physics 20, 158 (2026)。[56] Y. Nomura，《通过恢复对称性，助力受限玻尔兹曼机表示量子态》，Journal of Physics: Condensed Matter 33, 174003 (2021)。[57] K. Choo、G. Carleo、N. Regnault 和 T. Neupert，《利用神经网络量子态研究对称性与多体激发态》，Physical Review Letters 121,

167204 (2018), arXiv:1807.03325 [凝聚态物理：强关联电子系统]。[58] L. L. Viteritti、R. Rende 和 F. Becca，《受挫量子自旋系统的 Transformer 变分波函数》，Physical Review Letters 130, 236401 (2023), arXiv:2211.05504 [凝聚态物理：强关联电子系统]。[59] F. Vicentini、D. Hofmann、A. Szabó、D. Wu、C. Roth、C. Giuliani、G. Pescia、J. Nys、V. Vargas-Calderón、N. Astrakhantsev 和 G. Carleo，《Netket 3：用于多体量子系统的机器学习工具箱》，Sci- Post Phys. Codebases , 7 (2022)。[60] G. Carleo、K. Choo、D. Hofmann、J. E. T. Smith、T. Westerhout、F. Alet、E. J. Davis、S. Efthymiou、I. Glasser、S.-H. Lin、M. Mauri、G. Mazzola、C. B. Mendl、E. van Nieuwenburg、O. O'Reilly、H. Théveni- aut、G. Torlai、F. Vicentini 和 A. Wietek，《Netket：用于多体量子系统的机器学习工具包》，SoftwareX 10, 100311 (2019)。[61] J. Bradbury、R. Frostig、P. Hawkins、M. J. John- son、Y. Katariya、C. Leary、D. Maclaurin、G. Nec- ula、A. Paszke、J. VanderPlas、S. Wanderman-Milne 和 Q. Zhang，《JAX：Python+NumPy 程序的可组合变换》(2018)。[62] J. Heek、A. Levskaya、A. Oliver、M. Ritter、B. Ron- depierre、A. Steiner 和 M. van Zee，《Flax：适用于 JAX 的神经网络库与生态系统》(2024)。[63] X. Yuan、S. Endo、Q. Zhao、Y. Li 和 S. Benjamin，《变分量子模拟理论》，Quantum 3, 191 2019 1812 ( )，arXiv: . 08767 [量子物理]。[64] S. Sorella，《基于随机重构的格林函数蒙特卡罗》，Physical Review Letters 80, 4558 (1998), arXiv:cond-mat/9803107。[65] E. Hairer 和 G. Wanner，《常微分方程求解 II》，Springer 计算数学丛书，第 14 卷（Springer Berlin Heidelberg，柏林，海德堡，1991）。[66] J. C. Butcher，《常微分方程数值方法》（J. Wiley，奇切斯特，西萨塞克斯，英格兰；霍博肯，新泽西州，2003）。[67] F. Y. Leong、W.-B. Ewe 和 D. E. Koh，《变分量子演化方程求解器》，Scientific Reports 12, 10817 (2022), arXiv:2204.02912 [量子物理]。[68] A. Sinibaldi、C. Giuliani、G. Carleo 和 F. Vicentini，《通过投影量子演化消除含时变分蒙特卡罗的偏差》，Quantum 7, 1131 (2023), arXiv:2305.14294 [量子物理]。[69] J.-F. Chassagneux 和 A. Richou，《倒向随机微分方程欧拉格式的数值稳定性分析》(2014), arXiv:1407.0887 [math.PR]。[70] R. J. Webber 和 M. Lindsey，《利用增强采样进行变分蒙特卡罗的 Rayleigh-Gauss-Newton 优化》，Physical Review Research 4, 033099 (2022), arXiv:2106.10558 [stat.ML]。[71] H. Zhang、V. Armegioiu、J. Carrasquilla、S. Mishra、J. Müller、J. Nys 和 M. Zeinhofer，《投影逆迭代：用神经量子态进行基态计算的特征值方法》(2026), arXiv:2606.07825 [量子物理]。[72] M. Medvidović、A. Orfi、J. Carrasquilla 和 D. Sels，《神经网络量子态的绝热输运》，PRX Intelligence **1**，013020 2026 2510 15030 ( )，[ ] arXiv: . 量子物理 .[73] O. Koch 和 C. Lubich，《动力学低秩近似》，SIAM Journal on Matrix Analysis and Appli- cations 29, 434 (2007)。

[74] C. Lubich，《从量子到经典分子动力学：约化模型与数值分析》，苏黎世高等数学讲座（欧洲数学会，瑞士苏黎世，2008）。[75] E. Kieri、C. Lubich 和 H. Walach，《存在小奇异值时的离散化动力低秩近似》，《SIAM 数值分析杂志》54，1020（2016）。[76] J. R. Trail，《量子蒙特卡罗中的重尾随机误差》，《物理评论 E》77，016703（2008），arXiv:0909.5505 [物理学：计算物理]。[77] A. Misery、L. Gravina、A. Santini 和 F. Vicentini，《另辟蹊径：通过重要性采样改进变分蒙特卡罗梯度》，《机器学习：科学与技术》7，015035（2026），arXiv:2507.05352 [量子物理]。[78] K. Choo、A. Mezzacapo 和 G. Carleo，《用于从头算电子结构的费米子神经网络态》，《自然·通讯》11，2368（2020），arXiv:1909.12852 [物理学：计算物理]。[79] M. Solinas、B. Barton、Y. Zhang、J. Nys 和 J. Carrasquilla，《二维非厄米系统的双正交神经网络方法》，《Physical Review Letters》**136**，126501 2026 2508 01072 （），[ ] arXiv: 。量子物理 。[80] S. Tsuneyuki，《变相关方法：通向固体电子结构计算的另一种可能途径》，《理论物理进展增刊》176，134（2008）。[81] W. Dobrautz、H. Luo 和 A. Alavi，《通过非幺正相似变换获得的二维排斥型哈伯德模型的紧凑数值解》，Physical Review B 99，075119（2019）。[82] A. Baiardi 和 M. Reiher，《变相关密度矩阵重整化群》，《化学物理学杂志》153，164115（2020）。[83] I. O. Sokolov、W. Dobrautz、H. Luo、A. Alavi 和 I. Tavernelli，《通过精确变相关方法将量子计算机上量子多体问题的精度提高多个数量级》，Physical Review Research 5，023174（2023）。[84] S. F. Boys 和 N. C. Handy，《确定具有完全电子关联的能量和波函数》，《伦敦皇家学会会刊 A：数学与物理科学》310，43 （1969 ）。[85] A. J. Cohen、H. Luo、K. Guther、W. Dobrautz、D. P. Tew 和 A. Alavi，《通过 Jastrow 因子分解对电子薛定谔方程进行相似变换》，《化学物理学杂志》151，061101（2019）。[86] M. Motta、T. P. Gujarati、J. E. Rice、A. Kumar、C. Masteran、J. A. Latone、E. Lee、E. F. Valeev 和 T. Y. Takeshita，《采用变相关哈密顿量的电子结构量子模拟：以更小的量子计算机资源获得更高精度》，《物理化学化学物理》22，24270（2020）。[87] J. P. Haupt、S. M. Hosseini、P. López Ríos、W. Dobrautz、A. Cohen 和 A. Alavi，《优化变相关方法的 Jastrow 因子》，《化学物理学杂志》158，224105（2023）。[88] E. M. Christlmaier、T. Sc hraivogel、P. L. Ríos、A. Alavi 和 D. Kats，《xTC：变相关方法中三体相互作用的高效处理》，《化学物理学杂志》**159**, 014113 2023 2304 07006 （），[ ] arXiv: 。物理学：化学物理 。

[89] W. Dobrautz、I. O. Sokolov、K. Liao、P. L. Ríos、M. Rahm、A. Alavi 和 I. Tavernelli，《通过变相关方法在当前量子硬件上实现真正的化学精度》，《化学理论与计算杂志》20，4146（2024）。[90] E. Magnusson、A. Fitzpatrick、S. Knecht、M. Rahm 和 W. Dobrautz，《面向量子化学的高效量子计算：利用变相关方法和自适应试探态技术降低电路复杂度》，《法拉第讨论》254，402（2024）。[91] K. Liao、Y. Cheng、W. Dobrautz、T. Zhu 和 A. Alavi，《用于变相关哈密顿量的插值可分离密度拟合》（ 2026 2607 17314），[ ] arXiv: 。物理学：化学物理 。[92] C.-L. Hong、E. Koridon、P. K. Faehrmann、T. D. Kühne、S. Polla 和 W. Dobrautz，《从波函数正则性到特征向量条件性：非厄米变相关哈密顿量量子算法的精度–成本权衡》（2026），arXiv:2609.39452 [量子-物理]。[93] M. Kliczkowski、T. D. Kühne、F. Šimkovic、P. Sharma、C. L. Benavides-Riveros 和 W. Dobrautz，《“具有隐式时间步的神经量子态增强虚时间演化”的数据》（2026）。[94] P. Kramer 和 M. Saraceno 编，《量子力学中含时变分原理的几何学》，《物理学讲义》第 140 卷（Springer Berlin Heidelberg，柏林、海德堡，1981）。[95] J. Broeckhove、L. Lathouwers、E. Kesteloot 和 P. Van Leuven，《论含时变分原理的等价性》，《化学物理快报》149，547（1988）。[96] J. Toulouse 和 C. J. Umrigar，《通过能量最小化优化量子蒙特卡罗波函数》，《化学物理学杂志》126，084102（2007），arXiv:physics/0701039。[97] J. T. Frank 和 M. J. Kastoryano，《利用线性方法学习神经网络量子态》（2021），arXiv:2104.11011 [量子-物理]。[98] K. Pearson，《X. 关于演化的数学理论的贡献——II. 均质材料中的偏斜变异》，《伦敦皇家学会哲学汇刊》（A）343，1895。[99] R. Brockett，《对列表进行排序、对矩阵进行对角化并求解线性规划问题的动力系统》，《线性代数及其应用》146，79（1991）。[100] A. M. Bloch、R. W. Brockett 和 T. S. Ratiu，《完全可积梯度流》，《数学物理通讯》147，57（1992）。[101] M. Alghadeer、K. U. Giang、S. Cao、S. D. Fasciati、M. Piscitelli、N. Ng、P. J. Leek、M. Gluza 和 M. Bakr，《双括号算法冷却》（2025），arXiv:2510.00302 [量子-物理]。[102] J. Richter、N. B. Ivanov 和 K. Retzlaff，《受挫 $J_{1}-J_{2}$ 海森堡反铁磁体中马歇尔–佩尔斯符号规则的违反》，《欧洲物理快报》（EPL）25，545（1994），arXiv:cond-mat/9407041。[103] T. Reinhart 和 G. De les Coves，《深入探索、触及更远：玻尔兹曼机能够高效表示什么？》（2026），arXiv:2610.39731 [cs.LG]。[104] N. J. Higham，《数值算法的精度与稳定性》，第 2 版（工业与应用数学学会，2002）。

[105]A. Frommer 和 U. Glässner，《移位线性方程组的重启 GMRES》，《SIAM 科学计算杂志》**19，15（1998）**。[106] D. W. Marquardt，《非线性参数最小二乘估计的一种算法》，《工业与应用数学学会杂志》**11，431（1963）**。[107] R. Zander、R. Seidel、L. Xiaoyue 和 M. Gluza，《黎曼几何在双括号量子虚时间演化中的作用》，载于《几何信息科学（*GSI 2025*）》，《计算机科学讲义》，第 16035 卷（施普林格·自然瑞士，查姆，2025），第 105–114 页，arXiv:2504.01065 [量子物理]。

# 具有隐式时间步的神经量子态增强虚时间演化的补充材料

下文将给出推导和数值细节，以补充正文内容。第S1 节回顾归一化虚时间演化（ITE），给出收敛界，并分析双能级渐近行为。第S2 节详细介绍增强虚时间演化（AITE） [32]，第S3 节介绍一种可计算的带宽（刚性）界。第S4节介绍变分试探态和采样，第S5节阐明隐式随机重构方程，第S6 节说明欧拉时间步，第S7节介绍估计量和线性求解器，第S8节分析隐式演化在双能级极限下的局部收敛性。补充数值结果见第S9节。

除非另有说明，未归一化态记为 $|\psi\rangle$，而归一化态记为 $|\phi\rangle = |\psi\rangle / \sqrt{\langle\psi|\psi\rangle}$。参数化态使用大写希腊字母 $\{|\Psi\rangle, |\Phi\rangle\}$，任意态使用小写字母表示。只要上下文明确，虚时间 $\tau$ 的依赖关系将被省略，例如 $|\phi\rangle \equiv |\phi(\tau)\rangle$ 和 $E \equiv E(\tau)$。

## S1：虚时间演化

我们推导出传统的归一化虚时间演化 [22] 及其等价的密度矩阵（双括号）公式 [33, 34, 107]。我们证明，在方差等于哈密顿量方差的生成元中，该流使瞬时能量下降最大。$\hat{G}$

### S1.1. 归一化虚时间演化

设 $\hat{H} = \hat{H}^\dagger$ 为与时间无关的厄米哈密顿量，并始终假设 $\hbar = 1$。解析延拓至虚时间 [20–22] $t = -i\tau$后，未归一化态

$$|\psi(\tau)\rangle = e^{-\tau\hat{H}} |\psi(0)\rangle \ , \tag{S1}$$

满足虚时间薛定谔方程 $\partial_\tau |\psi\rangle = -\hat{H}|\psi\rangle$。其范数 $\mathcal{N}(\tau) = \langle\psi(\tau)|\psi(\tau)\rangle$ 随时间演化为

$$\partial_\tau \mathcal{N} = -2\langle\psi|\hat{H}|\psi\rangle = -2E(\tau)\mathcal{N}(\tau) \ , \tag{S2}$$

其中，$E(\tau) = \langle\psi(\tau)|\hat{H}|\psi(\tau)\rangle / \mathcal{N}(\tau)$为瞬时能量。由此可得 $\partial_\tau \sqrt{\mathcal{N}} = -E(\tau)\sqrt{\mathcal{N}}$，因此归一化态 $|\phi\rangle = |\psi\rangle / \sqrt{\mathcal{N}}$ 满足

$$\begin{aligned} \partial_\tau |\phi(\tau)\rangle &= \frac{1}{\sqrt{\mathcal{N}}} \partial_\tau |\psi(\tau)\rangle + \left(\partial_\tau \frac{1}{\sqrt{\mathcal{N}}}\right) |\psi(\tau)\rangle \\ &= -\left[\hat{H} - E(\tau)\right] |\phi(\tau)\rangle \ , \end{aligned} \tag{S3}$$

对于中心化哈密顿量 $\bar{H} = \hat{H} - E$，有 $\langle\phi|\bar{H}|\phi\rangle = 0$，从而确保范数保持。对其求导 $E = \langle\phi|\hat{H}|\phi\rangle$ 沿流求导可得

$$\begin{aligned} \dot{E}_{\mathrm{ITE}} &= \langle\partial_\tau\phi|\hat{H}|\phi\rangle + \langle\phi|\hat{H}|\partial_\tau\phi\rangle \\ &= -\langle\phi|\{\bar{H}, \hat{H}\}|\phi\rangle = -2\mu_2 \ , \end{aligned} \tag{S4}$$

其中中心矩 $\mu_m = \langle H^m \rangle$ 与正文中的定义一致。因此，虚时间演化以与瞬时能量方差成正比的速率单调降低能量，$\mu_2 = \mathrm{Var}(H)$。

更一般地，在本文中，我们考虑由厄米算符 $\hat{G} = \hat{G}^\dagger$ 生成的归一化演化，

$$\partial_\tau |\phi\rangle = -\left(\hat{G} - \langle\hat{G}\rangle\right) |\phi\rangle \equiv -\bar{G} |\phi\rangle \ , \tag{S5}$$

以及正文中中心化生成元 $\langle G\rangle = 0$ [1] 式（）。相应的能量变化率为

$$\begin{aligned} \dot{E}[\hat{G}] &= -\langle\phi| \left\{\hat{H} - E, \hat{G} - \langle\hat{G}\rangle\right\} |\phi\rangle \\ &= -2\mathrm{Re}\langle\phi|\bar{H}\bar{G}|\phi\rangle \ , \end{aligned} \tag{S6}$$

根据柯西–施瓦茨不等式，其满足

$$-\dot{E}[\hat{G}] \le 2\left|\langle\bar{H}\phi|\bar{G}\phi\rangle\right| \le 2\sqrt{\mathrm{Var}(\hat{H})\,\mathrm{Var}(\hat{G})} \ . \tag{S7}$$

事实上，具有 $G = H$ 的虚时间演化在固定生成元方差下，以瞬时能量最陡下降的方式达到这一界限 $\mu_2$。AITE [32], 在第S2节中讨论，它通过放宽这一约束，并进一步考虑三阶矩 $\mu_3$，实现了更快的瞬时下降。

### S1.2. 基态收敛

设 $\hat{H}|E_j\rangle = E_j|E_j\rangle$ 具有唯一基态 $E_0 < E_1 \le \cdots \le E_{\mathcal{D}-1} \equiv E_{\max}$，且希尔伯特空间维数为$\mathcal{D} \equiv \dim\mathcal{H}$。我们将激发能记为 $\Delta_j = E_j - E_0$，使得能隙为 $\Delta = \Delta_1$，带宽为$W = \Delta_{\mathcal{D}-1}$，能量误差为 $\epsilon(\tau) = E(\tau) - E_0$，初始值为 $\epsilon_0 = \epsilon(0)$。
对于任意归一化态$|\phi\rangle = \sum_j c_j |E_j\rangle$，其中占据概率为 $p_j = |c_j|^2 \in [0,1]$，$c_j \equiv \langle E_j|\phi\rangle$ 为本征基重叠，能量误差可写为

$$\epsilon = \sum_{j=0}^{\mathcal{D}-1} p_j \Delta_j \ . \tag{S8}$$

将未归一化态从初始 $|\psi(0)\rangle = |\phi\rangle$ 演化得到

$$|\psi(\tau)\rangle = \sum_j c_j e^{-E_j\tau} |E_j\rangle \ , \tag{S9}$$

因此，每个激发态贡献相对于 $|E_0\rangle$都会呈指数衰减，

$$\frac{\langle E_j|\psi(\tau)\rangle}{\langle E_0|\psi(\tau)\rangle} = \frac{c_j}{c_0}\, e^{-\Delta_j \tau}\,, \quad j>0\,, \tag{S10}$$

给定 $c_0 \neq 0$。渐近速率由初始重叠非零的最小激发能决定，

$$\Delta_* = \min_{j>0:\, c_j\neq 0} \Delta_j \geq \Delta\,, \tag{S11}$$

称为可达能隙。对称性限制可能排除能量最低的激发态，从而实际得到$\Delta_* > \Delta$。

最后，利用 $\Delta_j^2 \geq \Delta_*\Delta_j$ for $j>0$，可将能量方差 $\mu_2$ 与能量误差 $\epsilon$联系起来

$$\mu_2 = \sum_{j>0} p_j \Delta_j^2 - \epsilon^2 \geq \epsilon(\Delta_* - \epsilon)\,, \tag{S12}$$

结合式（S4）可得，

$$\dot{\epsilon}_{\rm ITE} \leq -2\epsilon(\Delta_* - \epsilon)\,, \tag{S13}$$

这保证了至少以 $2(\Delta_* - \epsilon)$ 的速率衰减；根据式（S10，渐近速率恰为 $2\Delta_*$，因此可达能隙较小时，虚时间演化的收敛会变慢。

### S1.3. 双能级渐近行为

考虑双能级渐近极限，此时动力学由基态 $|E_0\rangle$和最低的可达激发态 $|E_*\rangle$支配，其中 $E_* = E_0 + \Delta_*$。除相对相位外，

$$|\phi^{(2)}\rangle \simeq \sqrt{1-p}\,|E_0\rangle + \sqrt{p}\,|E_*\rangle\,, \quad p = \frac{\epsilon}{\Delta_*}\,, \tag{S14}$$

以及 $p \ll 1$。不过值得注意的是，在双能级子空间内，以下关系对于任意 $p \in [0,1]$都严格成立。由于 $E = (1-p)E_0 + pE_* = E_0 + p\Delta_*$，中心化哈密顿量 $H$，即虚时间演化生成元，其本征值为$g_0^{\rm ITE} = -p\Delta_*$、$g_*^{\rm ITE} = (1-p)\Delta_*$。中心矩$\mu_m$，

$$\mu_m = (1-p)(-p\Delta_*)^m + p[(1-p)\Delta_*]^m\,, \tag{S15}$$

只取决于当前态。特别地，对于 $m$ =2、3，然后代入 $p = \epsilon/\Delta_*$ [式（S14）]，相关矩为

$$\begin{aligned} \mu_2 &= \Delta_*^2 p(1-p) = \Delta_*\epsilon - \epsilon^2\,, \\ \mu_3 &= \Delta_*^3 p(1-p)(1-2p) = \Delta_*^2\epsilon - 3\Delta_*\epsilon^2 + 2\epsilon^3\,, \end{aligned} \tag{S16}$$

得到能量偏度 [98]正文的

$$\kappa = \frac{\mu_3}{\mu_2^{3/2}} = \frac{1-2p}{\sqrt{p(1-p)}} \simeq \sqrt{\frac{\Delta_*}{\epsilon}} \quad (\epsilon \to 0)\,. \tag{S17}$$

引入激发态与基态振幅比

$$r(\tau) \equiv \left|\frac{\langle\phi^{(2)}|E_*\rangle}{\langle\phi^{(2)}|E_0\rangle}\right| = \sqrt{\frac{p}{1-p}}\,, \tag{S18}$$

并求解 $p$，可得

$$p(\tau) = \frac{r^2}{1+r^2}\,, \quad \epsilon(\tau) = \Delta_* p = \frac{\Delta_* r^2}{1+r^2} \simeq \Delta_* r^2\,, \tag{S19}$$

这对任何演化都成立。

对于虚时间演化，根据方程（S10），$r(\tau) = r(0)\exp(-\Delta_*\tau)$，并对方程（S19）求导可得

$$\dot{\epsilon}_{\rm ITE} = -\frac{2\Delta_*^2 r^2}{(1+r^2)^2} = -2\epsilon(\Delta_* - \epsilon) = -2\mu_2\,, \tag{S20}$$

其中我们使用了 $\mu_2 = \epsilon(\Delta_* - \epsilon) = \Delta_*^2 p(1-p)$ [方程 (S16)]和 $p(1-p) = r^2/(1+r^2)^2$。因此，方程 (S13) 的界达到饱和，即，虚时间演化的速度慢到了能隙所允许的极限。对于 $\epsilon \ll \Delta_*$，速率简化为 $\dot{\epsilon}_{\rm ITE} \simeq -2\Delta_*\epsilon$，因此 $\epsilon_{\rm ITE} \simeq \Delta_* r(0)^2 \exp(-2\Delta_*\tau)$。

### S1.4. 双括号表述

对于纯态密度矩阵 $\hat{\rho} = |\phi\rangle\langle\phi|$，式（S3）等价于双括号流 [33, 34, 99, 100,107]，

$$\partial_\tau \hat{\rho} = -\{\bar{H}, \hat{\rho}\} = -[\hat{\rho}, [\hat{\rho}, \hat{H}]] = -i[\hat{\Omega}_{\rm ITE}, \hat{\rho}]\,, \tag{S21}$$

取厄米切向算符 $\hat{\Omega}_{\rm ITE} \equiv -i[\hat{H}, \hat{\rho}]$。$\hat{}$对于式（1）的任意厄米生成元 $G$，$\hat{}\,\hat{}\,\hat{}^{-}$其中 $\Omega_G \equiv -i[G, \hat{\rho}]$ 且 $\partial_\tau|\phi\rangle = -i\Omega_G|\phi\rangle = -G|\phi\rangle$。$\hat{}$能量变化率以及 $\Omega_G$ 的范数为

$$\dot{E}[\hat{G}] = -\langle\hat{\Omega}_G, \hat{\Omega}_{\rm ITE}\rangle_{\rm HS}\,, \quad \|\hat{\Omega}_G\|_{\rm HS}^2 = 2\,{\rm Var}(\hat{G})\,, \tag{S22}$$

取希尔伯特–施密特内积 $\langle A, B\rangle_{\rm HS} = {\rm Tr}(A^\dagger B)$，其中前者可还原式（S6）。固定 $\hat{}$Ω的希尔伯特–施密特范数 $_G$，等价于固定 $\hat{}$其方差 $G$。根据柯西–施瓦茨不等式，虚时间演化（ITE）便是瞬时能量下降最快的方向 [101]，，如第 S1.1 节所述。

## S2：增强虚时间演化

AITE 演化将纯态密度矩阵 $\hat{}\rho$ 替换为秩一投影算符 $\mathcal{O} \in \mathcal{P}_1(\mathcal{H})$，其中

$$\mathcal{P}_1(\mathcal{H}) \equiv \{\hat{\mathcal{O}} = \hat{\mathcal{O}}^\dagger : \hat{\mathcal{O}}^2 = \hat{\mathcal{O}},\ {\rm Tr}\left[\hat{\mathcal{O}}\right] = 1\}\,, \tag{S1}$$

表示希尔伯特空间上所有纯态密度算符的集合 $\mathcal{H}$，从而推广了双括号表述 [32]。相应地，式（$\hat{}\Omega_{\rm ITE}$ 中的虚时间演化切向算符S21）被替换为$\hat{}\Omega_{\mathcal{O}} = -i[\hat{}H, \hat{}\mathcal{O}]$得到

$$\partial_\tau|\phi\rangle = -i\hat{\Omega}_{\mathcal{O}}|\phi\rangle\,, \tag{S2}$$

ˆ，能量变化率关于投影算符 $\mathcal{O}$ 呈线性，

$$\dot{E} = \mathrm{Tr}\left[\hat{\chi}\,\hat{\mathcal{O}}\right] , \quad \hat{\chi} = -i[\hat{H}, \hat{\Omega}_{\mathrm{ITE}}] . \tag{S3}$$

AITE 搜索满足以下条件的增广投影算符 $\hat{\rho}_A$：该算符使式（S3）在 $\mathcal{P}_1$ [32]上达到最小值，并限制在二阶 Krylov 子空间 $\mathcal{S}_H = \mathrm{span}\{|\phi\rangle, |v\rangle\}$内，其中 $|v\rangle = H|\phi\rangle/\sqrt{\mu_2}$。在该子空间中，

$$\chi\big|_{\mathcal{S}_H} = -2\mu_2 \begin{pmatrix} 1 & \kappa/2 \\ \kappa/2 & -1 \end{pmatrix} , \tag{S4}$$

其本征值为 $\lambda_\pm = \pm 2\mu_2\sqrt{1+\kappa^2/4}$。使用最低本征值对应的本征向量

$$|\lambda_-\rangle = \cos\varphi\,|\phi\rangle + \sin\varphi\,|v\rangle , \tag{S5}$$

其中 $\tan(2\varphi) = \kappa/2$，且 $2\varphi \in (-\pi/2, \pi/2)$，因此增广投影算符变为 $\hat{\rho}_A = |\lambda_-\rangle\langle\lambda_-|$。由于 $\mathrm{Tr}[\hat{\chi}\hat{\rho}_A] = \lambda_-$，式（S3）中的能量变化率给出正文中的式（3）：

$$\dot{E}_{\mathrm{AITE}} = -2\mu_2\sqrt{1+\kappa^2/4} \;\le\; -2\mu_2 = \dot{E}_{\mathrm{ITE}} , \tag{S6}$$

对于 $\kappa \neq 0$，不等式严格成立；即，对于偏斜的瞬时能量分布 [32]。

将 $\hat{\mathcal{O}} = \hat{\rho}_A$ 代入式（S2）得到 $-i\hat{\Omega}_{\rho_A}|\phi\rangle = -\hat{G}_{\mathrm{AITE}}|\phi\rangle$，即，由 $\hat{G}_{\mathrm{AITE}}$[式（S7）],生成的演化

$$\hat{G}_{\mathrm{AITE}} = a\bar{H} + b\,(\bar{H}^2 - \mu_2) , \tag{S7}$$

其中 $a = \cos(2\varphi)$ 和 $b = \sin(2\varphi)/(2\sqrt{\mu_2})$,一常数 $-b\mu_2$ 使生成元 $\langle\hat{G}_{\mathrm{AITE}}\rangle = 0$ 居中。

尽管式（S6）对 $\kappa$ 的任一符号都成立，但我们仅在 $\hat{G}_{\mathrm{AITE}}$当 $\mu_2 > 0$ 且 $\mu_3 > 0$（即，$b > 0$）时才使用，因为随机演化在其他情况下稳定性较差。所有其他情况下，我们都采用虚时间演化步进，（$a$，$b$）=（1，0）。对于所考虑的模型，这种回退很少触发。

### S2.1. 双能级 AITE 渐近行为

回到双能级图像，式（S16）和（S17）给出 $1+\kappa^2/4 = [4p(1-p)]^{-1}$，而式（S6）中的能量变化率变为

$$\dot{\epsilon}_{\mathrm{AITE}} = -\Delta_*\sqrt{\epsilon(\Delta_* - \epsilon)} \;\simeq\; -\Delta_*^{3/2}\sqrt{\epsilon} , \tag{S8}$$

与……相比 $\dot{\epsilon}_{\mathrm{ITE}}$ 它关于……呈线性 $\epsilon$，即，$\epsilon_{\mathrm{ITE}} \propto e^{-2\Delta_*\tau}$。对式（S8）以 $\epsilon = \Delta_*\sin^2\vartheta$ 进行积分，得到 $\dot{\vartheta} = -\Delta_*/2$并且

$$\epsilon_{\mathrm{AITE}}(\tau) = \Delta_*\sin^2\left[\arcsin\sqrt{\frac{\epsilon_0}{\Delta_*}} - \frac{\Delta_*\tau}{2}\right] , \tag{S9}$$

而且值得注意的是，误差在有限时间$\tau_* = (2/\Delta_*)\arcsin\sqrt{\epsilon_0/\Delta_*} \simeq 2\sqrt{\epsilon_0/\Delta_*^3}$($\epsilon_0 \ll \Delta_*$)时消失，其中 $\epsilon_{\mathrm{AITE}} \simeq (\Delta_*^3/4)(\tau_* - \tau)^2$ 为 $\tau \to \tau_*$，而虚时间演化仅渐近收敛。这涉及双能级极限下的连续流，此时忽略了高阶激发。离散更新步骤将在第 S8 节中分析。

根据 $\mu_2 = \Delta_*^2 p(1-p)$ [式（S16）] 以及 $\kappa = (1-2p)/\sqrt{p(1-p)}$ [式（S17）], $\tan 2\varphi = \kappa/2$，可得 $a = 2\sqrt{p(1-p)}$ 和 $b = (1-2p)/(2\Delta_*\sqrt{p(1-p)})$，因此，ˆ 中心化生成元 $G_{\mathrm{AITE}}$ 的二能级本征值为

$$g_0^{\mathrm{AITE}} = -\frac{\Delta_* r}{2} , \quad g_*^{\mathrm{AITE}} = \frac{\Delta_*}{2r} . \tag{S10}$$

### S3：哈密顿量带宽界

显式欧拉稳定性条件 [式（B3）] 用生成元带宽 $W_G = \max_j g_j - g_0$ 表示，其中 $g_j$ 是本征态 $|E_j\rangle$ 对应的 $G$ 本征值，$H$。对于虚时间演化，$W_G = W$ 是 $H$ 的带宽。对于 AITE，$g_j = a(E_j - E) + b[(E_j - E)^2 - \mu_2]$ 和$W_G = W_A = W[a + b(W - 2\epsilon)]$ [式（B5）]。

对于超出精确对角化能力的系统规模，通常无法获得 $W$ 和 $W_A$。对于自旋-1/2 系统，将 $H$展开为泡利字符串 $P_\gamma = \bigotimes_{\ell=1}^N \sigma_\ell^{\gamma_\ell}$后，可直接从哈密顿量得到一个保守界，其中 $\gamma_\ell \in \{0, x, y, z\}$，$\sigma^0 = I$。泡利字符串是厄米的，满足 $P_\gamma^2 = I$，$\mathrm{Tr}\ P_\gamma P_{\gamma'} = 2^N\delta_{\gamma\gamma'}$，并且

$$\hat{H} = h_0\hat{I} + \sum_{\gamma\neq 0} h_\gamma\hat{P}_\gamma , \quad h_\gamma = 2^{-N}\,\mathrm{Tr}\left(\hat{P}_\gamma\hat{H}\right) . \tag{S1}$$

单位算符贡献 $P_0 = I$ 不会影响谱差异，因此可以舍弃。令 $\|\cdot\|_{\mathrm{op}}$ 为算符范数，则 $\|P_\gamma\|_{\mathrm{op}} = 1$，由三角不等式可得

$$W \le 2\left\|\hat{H} - h_0\hat{I}\right\|_{\mathrm{op}} \le 2\sum_{\gamma\neq 0}|h_\gamma| \equiv \overline{W} , \tag{S2}$$

具体取决于模型，这个界可能较宽松。将$W \to W \ge W$ 代入式（B3） [65] 即可得到可计算步长$\delta\tau_{\mathrm{ITE}} < 2/W$。

对于具有 $a, b \ge 0$ 的 AITE，式 (B5) 给出 $W_A = W[a + b(W - 2\epsilon)] \le aW + bW^2 \equiv W_A$。实际应用中，我们使用 $f(x) = ax + b(x^2 - \mu_2)$ 在 $x \in [h_0 - W/2 - E, h_0 + W/2 - E]$ 上的较紧范围 $\max f - \min f$，该范围从不超过 $W_A$。请注意，$W$ 仅取决于 $H$，而 $a$ 和 $b$ 则来自 AITE 已计算的能量矩，因此不需要额外的谱信息。

### S4：变分蒙特卡罗与神经量子态

神经量子态（NQS） [11]为多体波函数提供了强大的神经网络表示——

在变分蒙特卡罗（VMC）方法中 [7, 8]。我们介绍了本文所用的神经量子态试探波函数 [第S4 节。1]以及用于 [第S4 节。2] 随机演化的蒙特卡洛估计量。我们使用 NetKet [59, 60] 进行希尔伯特空间表示和采样，使用 Flax [62] 构建神经网络架构，并使用 JAX [61] 执行所有其他数值运算。

我们采用正文中定义在包含 个格点的 [方形晶格上的横场伊辛模型, ] 以及 $J_1$-$J_2$ [横场伊辛模型] ，并施加周期性边界条件（PBC），即，$\hat{\sigma}^{\nu}_{(L_x+1,\ell_y)} = \hat{\sigma}^{\nu}_{(1,\ell_y)}$ 和 $\hat{\sigma}^{\nu}_{(\ell_x,L_y+1)} = \hat{\sigma}^{\nu}_{(\ell_x,1)}$ ，其中 $\nu = x,\ y,\ z$，其中 $\ell = (\ell_x,\ \ell_y)$。对于对称性适配的 ansatz，我们将 $\Psi_{\boldsymbol{\theta}}$ 限制在总准动量为$\mathbf{k} = \mathbf{0}$的平移不变扇区，其维数为$\mathcal{D} \sim 2^N/N$。

横场伊辛模型还具有额外的 $\mathbb{Z}_2$ 自旋宇称对称性 $\hat{\Pi} = \prod_\ell \hat{\sigma}^x_\ell$，该对称性与晶格平移对易。对于对称性适配受限玻尔兹曼机，我们将 $\Psi_{\boldsymbol{\theta}}$ 限制在$\hat{}$偶宇称扇区 $\Pi = +1$，即，$\Psi_{\boldsymbol{\theta}}(-\mathbf{s}) = \Psi_{\boldsymbol{\theta}}(\mathbf{s})$。结合$\mathbf{k} = \mathbf{0}$，得到 $\mathcal{D}_{\mathrm{TFI}} \sim 2^N/(2N)$。正文中的 ViT 运行未进行宇称投影。没有施加其他对称性。

$J_1$-$J_2$ 模型具有全局 $U(1)$对称性，即守恒总磁化强度$\hat{}$ $S^z_{\mathrm{tot}} = \sum_\ell \hat{S}^z_\ell$。对于所有试探态，我们将 $\Psi_{\boldsymbol{\theta}}$ 限制在 $S^z_{\mathrm{tot}} = 0$ 扇区。结合 $\mathbf{k} = \mathbf{0}$，可得 $\mathcal{D}_{J_1J_2} \sim \binom{N}{N/2}/N$。当 $L_x$和 $L_y$为偶数时，正方晶格是二分的，其子晶格为 $A$ 和 $B$。我们施加马歇尔符号变换 [35], $\hat{U}_{\mathrm{M}} = \prod_{\ell\in B} \hat{\sigma}^z_\ell$。在 $J_2 = 0$处，变换后的基态在 $S^z_{\mathrm{tot}} = 0$ 扇区中可以取为非负。当$J_2 > 0$时，这种符号结构不再严格成立 [102]。

### S4.1. 变分试探态

我们使用正文中的波函数 $\Psi_{\boldsymbol{\theta}}(\mathbf{s})$、$\hat{\sigma}^z$ 乘积基以及对数导数 $O_k$ 。

对于精确求和和小型系统，我们考虑具有 $M = \alpha N$ 个隐单元且隐单元密度为 $\alpha = M/N \in \mathbb{Z}_{>0}$ 的单层复数受限玻尔兹曼机（RBM） [11, 56] ansatz。对隐变量 $h_m = \pm 1$ 求迹后得到

$$\begin{aligned}\Psi^{\mathrm{RBM}}_{\boldsymbol{\theta}}(\mathbf{s}) &= \sum_{\mathbf{h}} \exp\Big(\sum_\ell a_\ell s_\ell + \sum_m h_m \mathcal{F}_m(\mathbf{s})\Big) \\ &= \exp\Big(\sum_\ell a_\ell s_\ell\Big) \prod_{m=1}^{M} 2\cosh \mathcal{F}_m(\mathbf{s}) \,,\end{aligned} \tag{S1}$$

其中， $\mathcal{F}_m(\mathbf{s}) = b_m + \sum_\ell W_{m\ell} s_\ell$和 $\boldsymbol{\theta} = \{a_\ell, b_m, W_{m\ell}\}$ 表示 $N_p = N + M + NM = N + \alpha N(N+1)$ 个复变分参数。增加 $\alpha$ 可扩大试探态空间容量，而可达到的基态表示精度取决于哈密顿量和施加的对称性 [11, 15, 103]。对数导数

受限玻尔兹曼机试探波函数的解析表达式可写成闭式 [11]。

为计及平移对称性，我们进一步采用对称性适配受限玻尔兹曼机 [56, 57, 59], ，其参数在晶格平移下共享，从而将独立复参数的数量减少至 $N_p = \alpha(N+1) + 1$。对于横场伊辛模型，我们还通过将所有偏置设为零来强制施加偶自旋宇称，即 $\Psi_{\boldsymbol{\theta}}(-\mathbf{s}) = \Psi_{\boldsymbol{\theta}}(\mathbf{s})$ ，$a_\ell = b_m = 0$，因为双曲余弦函数是偶函数 [式（S1）], ，这使得$N_p = \alpha N$。在 $S^z_{\mathrm{tot}} = 0$ 扇区的 $J_1$-$J_2$ 模型中，该约束在蒙特卡洛采样器 [59] 中直接施加。由于 $\sum_\ell s_\ell = 0$ 在该扇区中，均匀可见偏置是冗余的，因此将其省略，得到$N_p = \alpha(N+1)$ 。

对于较大的系统或 MCMC 采样，我们采用具有实参数的视觉 Transformer（ViT）试探态 [58] ，其中晶格被划分为图块，再经嵌入后由一组具有多头平移不变注意力的编码器层处理。对数双曲余弦输出给出复对数振幅 $\log\Psi_{\boldsymbol{\theta}}(\mathbf{s})$。所有 ViT 运行均使用 $L = 4$个编码器层、维度为$d = 16$的嵌入以及四个注意力头，除 $4\times4$ 晶格采用 $1\times1$ 个图块外，其余均使用 $2\times2$ 个图块。这对应于 $N_p = 11\,808$（$4\times4$）、$11\,696$（$6\times4$）、$11\,744$（$6\times6$）和 $11\,856$（$8\times8$）个实参数。更大的网络留待未来研究。

### S4.2. 采样与蒙特卡洛估计量

在变分蒙特卡洛中，期望值通过对玻恩分布 $p_{\boldsymbol{\theta}}(\mathbf{s}) = |\Psi_{\boldsymbol{\theta}}(\mathbf{s})|^2/\langle\Psi_{\boldsymbol{\theta}}|\Psi_{\boldsymbol{\theta}}\rangle$取平均来计算。对于算符 $\hat{A}$，局域估计量 [8]为

$$A_{\mathrm{loc}}(\mathbf{s}) = \frac{\langle\mathbf{s}|\,\hat{A}\,|\Psi_{\boldsymbol{\theta}}\rangle}{\Psi_{\boldsymbol{\theta}}(\mathbf{s})} = \sum_{\mathbf{s}'}\langle\mathbf{s}|\hat{A}|\mathbf{s}'\rangle\frac{\Psi_{\boldsymbol{\theta}}(\mathbf{s}')}{\Psi_{\boldsymbol{\theta}}(\mathbf{s})} \,, \tag{S2}$$

从而得到期望值

$$\langle\hat{A}\rangle = \sum_{\mathbf{s}} p_{\boldsymbol{\theta}}(\mathbf{s}) A_{\mathrm{loc}}(\mathbf{s}) \equiv \langle A_{\mathrm{loc}}\rangle_p \,, \tag{S3}$$

其中 $\langle f\rangle_p \equiv \sum_{\mathbf{s}} p_{\boldsymbol{\theta}}(\mathbf{s}) f(\mathbf{s})$。期望值通过对从 $p_{\boldsymbol{\theta}}$ 中采样得到的 $N_s$ 个构型取平均来近似，即， $f = N_s^{-1}\sum_{i=1}^{N_s} f(\mathbf{s}_i)$。中心化对数导数为 $\overline{O_k} = O_k - \overline{O_k}$。

样本通过带有局部移动的 Metropolis–Hastings 链生成 [8]。对于精确求和，方程（S3）在所模拟对称扇区的完备基上求值，使用权重 $w_i = p_{\boldsymbol{\theta}}(\mathbf{s}_i)$ [59]。

### S5：隐式随机重构

我们考察隐式Galerkin投影的变分解释，并将其与直接最小化

以下量隐式欧拉 [65] 残差 $|i^{(n)}\rangle$ 范数[式（C2）]。

### S5.1. Galerkin投影

如正文所示，Galerkin投影 $\langle T_j^{(n)}|i^{(n)}\rangle = 0$会得到隐式随机重构方程 [式（C4）]。与显式随机重构 [11, 44, 64],不同，切空间矩阵 $K^{(n)}$ 会计入对有限步长切向偏移的作用。

$\hat{}$ 对于厄米生成元 $G_n$，投影矩阵 $K^{(n)}$ 也是厄米的，因为 $(K_{jk}^{(n)})^* = \langle T_k^{(n)}|\hat{G}_n|T_j^{(n)}\rangle = K_{kj}^{(n)})$。结合 $S^{(n)}$ 的厄米性，可知投影算符 $\mathcal{K}^{(n)} = S^{(n)} + \delta\tau K^{(n)}$ 是厄米的。因此，Galerkin方程 $\mathcal{K}^{(n)}\mathbf{d}_n = -\mathbf{F}^{(n)}$ 等价于实二次泛函的驻定条件，这与 Ritz–Galerkin方法一致。具体而言

$$\mathcal{Q}_n(\mathbf{d}_n) = \mathbf{d}_n^\dagger \mathcal{K}^{(n)}\mathbf{d}_n + 2\,\mathrm{Re}(\mathbf{d}_n^\dagger \mathbf{F}^{(n)}) , \tag{S1}$$

通过对$\mathbf{d}_n^*$应用 Wirtinger微积分，得到驻定条件

$$\frac{\partial \mathcal{Q}_n}{\partial \mathbf{d}_n^*} = \mathcal{K}^{(n)}\mathbf{d}_n + \mathbf{F}^{(n)} = 0 , \tag{S2}$$

这与正文中的方程（C4）一致。展开平方范数并确定显式McLachlan残差，可得

$$\mathcal{Q}_n = \overbrace{\Big\|\sum_k d_{n,k}\Big|T_k^{(n)}\Big\rangle + \hat{G}_n|\Phi_n\rangle\Big\|^2}^{\mathcal{R}(\mathbf{d}_n)} + \delta\tau\,\mathbf{d}_n^\dagger K^{(n)}\mathbf{d}_n - \langle \hat{G}_n^2\rangle , \tag{S3}$$

使得 $\mathcal{Q}_n$ 与显式McLachlan泛函 [21] $\mathcal{R}$ [方程（A5）] 相差一个有限步长项$\delta\tau\mathbf{d}_n^\dagger K^{(n)}\mathbf{d}_n$，以及一个与$\mathbf{d}_n$无关的加性常数。随着 $\delta\tau \to 0$，有限步长项趋于零，驻定条件$\partial_{\mathbf{d}_n^*}\mathcal{Q}_n = 0$ 恢复显式SR。若 $\mathcal{K}^{(n)}$ 为正定，则驻点是唯一极小值。

### S5.2. 最小化隐式残差

或者，也可以考虑最小化隐式欧拉残差 $\mathcal{I}(\mathbf{d}_n) = \|i^{(n)}\|^2$ 的平方，遵循 McLachlan 最小二乘变分原理 [21, 42]。对于厄米 $\hat{}$ $G_n$，其展开式为

$$\begin{aligned}\mathcal{I}(\mathbf{d}_n) \equiv \|i^{(n)}\|^2 &= \langle\Phi_n|\,\hat{G}_n^2\,|\Phi_n\rangle \\ &+ 2\,\mathrm{Re}\Big[\sum_k d_{n,k}^*\langle T_k^{(n)}|(1+\delta\tau\hat{G}_n)\hat{G}_n\,|\Phi_n\rangle\Big] \\ &+ \sum_{j,k} d_{n,j}^* d_{n,k}\langle T_j^{(n)}|(1+\delta\tau\hat{G}_n)^2|T_k^{(n)}\rangle ,\end{aligned} \tag{S4}$$

并对$d_{n,\,j}^*$ 求驻定条件，得到

$$\begin{aligned}0 = \frac{\partial\mathcal{I}}{\partial d_{n,j}^*} &= \langle T_j^{(n)}|(1+\delta\tau\hat{G}_n)\hat{G}_n\,|\Phi_n\rangle \\ &+ \sum_k d_{n,k}\,\langle T_j^{(n)}|(1+\delta\tau\hat{G}_n)^2|T_k^{(n)}\rangle ,\end{aligned} \tag{S5}$$

展开后可得，

$$\begin{aligned}0 = &\underbrace{\langle T_j^{(n)}|\hat{G}_n\,|\Phi_n\rangle}_{F_j^{(n)}} + \delta\tau\,\underbrace{\langle T_j^{(n)}|\hat{G}_n^2\,|\Phi_n\rangle}_{\Xi_j^{(n)}} \\ &+ \sum_k d_{n,k}\Bigg[\underbrace{\langle T_j^{(n)}|T_k^{(n)}\rangle}_{S_{jk}^{(n)}} + 2\delta\tau\,\underbrace{\langle T_j^{(n)}|\hat{G}_n|T_k^{(n)}\rangle}_{K_{jk}^{(n)}} \\ &+ \delta\tau^2\,\underbrace{\langle T_j^{(n)}|\hat{G}_n^2|T_k^{(n)}\rangle}_{L_{jk}^{(n)}}\Bigg] ,\end{aligned} \tag{S6}$$

因此，与 $\mathcal{U}^{(n)} = S^{(n)} + 2\delta\tau K^{(n)} + \delta\tau^2 L^{(n)}$ 一同

$$\mathcal{U}^{(n)}\mathbf{d}_n = -\mathbf{F}^{(n)} - \delta\tau\boldsymbol{\Xi}^{(n)} , \tag{S7}$$

其中，最小二乘条件包含矩阵元 $L^{(n)}$ 和 $\Xi^{(n)}$ 的 $\hat{}$ $G_n^2$。这些会增加计算成本和统计误差，尤其对于 AITE [方程（2)], 其中 $\hat{}$ $G_n^2$ 包含 $\hat{}H^4$ 项。Galerkin投影将残差与切向量 $|T_j^{(n)}\rangle$进行检验，而非最小二乘方法中的 $(1+\delta\tau\hat{}G_n)|T_{(n)\,j}\rangle$。当 $\delta\tau \to 0$ 时，两种表述相同；通常在有限 $\delta\tau$ 时则有所不同。

### S6：欧拉时间步

我们使用自适应虚时间步 $\delta\tau_n = \tau_{n+1} - \tau_n$ 对虚时间演化进行离散化，使得 $\sum_{k=0}^{n-1}\delta\tau_k$ 。选择这些时间步是为了控制由 $\hat{}$生成元 $G_n$ 驱动的显式 和 隐式 欧拉更新的稳定性。

对于显式方案，按照正文中的[公式（B3）]和第S3节，我们使用

$$\delta\tau_n = \min\big(\delta\tau_0,\; c\,2/\overline{W}_G\big) , \tag{S1}$$

给定猜测步长 $\delta\tau_0$、步长因子 $c$ 和可计算的带宽界 $W_G \geq W_G$。因子 $c > 1$ 超过了界限 $2/W_G$，且只有在 $W_G$ 的范围内才稳定

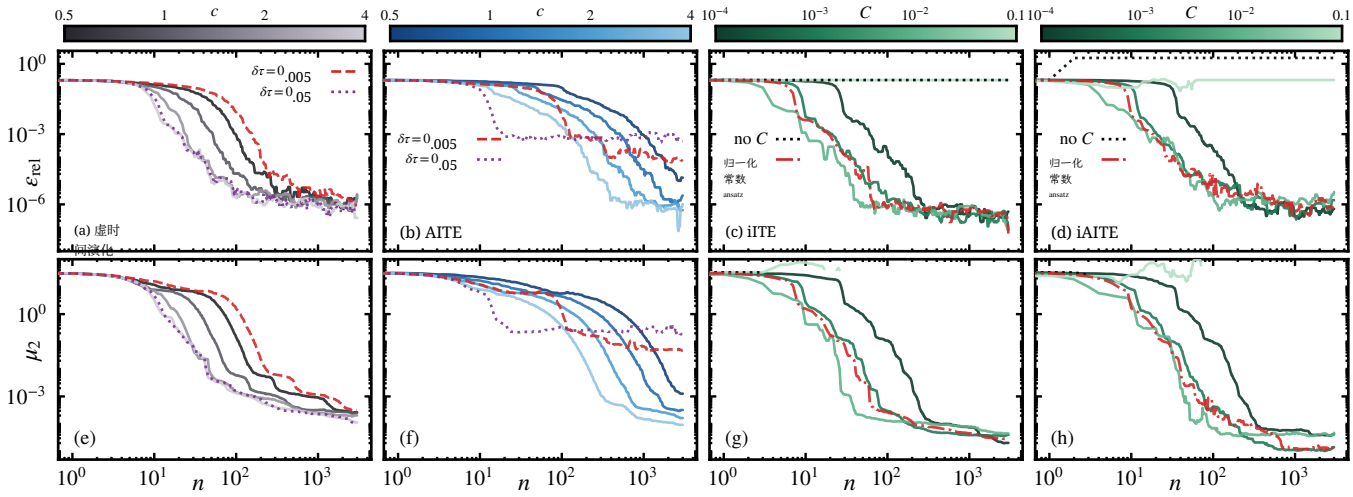


图 S1. (a)-(d) 相对能量误差 $\epsilon_{\rm rel}$ 和 (e)-(h) 能量方差 $\mu_2$ 随步数 $n$ 的变化：在偶宇称扇区中，对于 $h=2$ 处的 $4\times4$ 横场伊辛模型，使用每步 $\alpha=8$ 和 $N_s=4096$ 个样本的对称性适配受限玻尔兹曼机，分别采用显式虚时间演化 (a,e)、显式反绝热虚时间演化 (b,f)、隐式虚时间演化 (c,g) 和隐式 AITE (d,h)。为清晰起见，曲线采用 11 点滑动中位数进行平滑。(a)、(b)、(e)、(f) 中，颜色表示步长因子 $c\in\{0.5,1,2,4\}$ [式 (S1)],，其中 $\delta\tau_0=1$；红色虚线和紫色点线分别表示固定的 $\delta\tau=0.005$ 和 0.05。(c)、(d)、(g)、(h) 中，颜色表示信赖域半径$C\in\{10^{-4},10^{-3},10^{-2},10^{-1}\}$ [式 (S13)],；黑色点线和红色点划线分别表示无约束演化（无$C$）和 $C_{\rm ansatz}$ [式 (S13)],。隐式方案采用 $\delta\tau_0=1$ 和 $\tilde{c}=1$ [式 (S2)],；在 (a)、(b)、(e)、(f) 中，信赖域固定为 $C=C_{\rm ansatz}$。

高估了 $W_G$。对于虚时间演化，$\overline{W_G}=W$ [公式 (S2)], 而对于 AITE，$W_G=\overline{W_A}$ 会在每次更新时根据当前的 $a$、$b$和 $E$重新计算。

在正文中，显式方案使用 $c\in[0.5,1.5]$，而隐式方案使用 $\tilde{c}=1$ [公式（S2）]，其中 $\delta\tau_0$ =1。图1 和2 对虚时间演化使用 $c=1/2$，而图1对 AITE 使用，即，$\delta\tau=1/W_G$，其中 $\delta\tau_0=10^3$，因此截断值$c\,2/W_G$ 始终适用。

此外，图3 对虚时间演化使用 $c=1.5$ 和 $\delta\tau_0=1$。在图4中，我们为每种方案绘出在可用的步长控制参数中最终 $E/N$ 最低的一次运行；虚时间演化的步长控制参数为 $\delta\tau_0=1$ 和 $c=1.5$。

作为比较，在图S1中，我们改为在更宽的范围内扫描 $c$，并调整 $C$ 隐式方案的信赖域半径，仅用于检查它们对运行的影响。

另一方面，对于衰减模态，隐式更新是 A 稳定和 L稳定的（$g\geq0$）[65]。然而，中心化后的生成元包含负本征值 $g_{0,\ n}<0$，参见第 B 节。相应的隐式因子（$1+\delta\tau g_{0,\ n}$）$^{-1}$ 在$\delta\tau=-1/g_{0,\ n}$处变为奇异，之后改变符号。^ 在每个依赖于状态的 $G_n$ 处，我们将步长限制为

$$\delta\tau_n=\min\left(\delta\tau_0,\frac{\tilde{c}}{\sqrt{2\,\mathrm{Var}(G_n)}}\right)\ , \tag{S2}$$

在正文和补充材料的所有图中使用 $\delta\tau_0=1$ 和 $\tilde{c}=1$。此外，$g_{0,\ n}$ 可以使用 Var($G_n$) 进行界定。

接下来是第 S1.2节，

$$\begin{aligned}\mathrm{Var}(G_n)&=\sum_j p_j g_{j,n}^2=p_0g_{0,n}^2+\sum_{j>0}p_jg_{j,n}^2\\&\geq p_0g_{0,n}^2+\frac{1}{1-p_0}\Big(\sum_{j>0}p_jg_{j,n}\Big)^2\\&=p_0g_{0,n}^2+\frac{p_0^2}{1-p_0}\,g_{0,n}^2=\frac{p_0}{1-p_0}\,g_{0,n}^2\ ,\end{aligned} \tag{S3}$$

第二行由柯西–施瓦茨不等式 $(\sum_{j>0}p_jg_{j,\ n})^2\leq\sum_{j>0}p_j\sum_{j>0}p_jg_{j,\ n}^2$ 推得，而最后一个等式则来自中心化条件$\sum_j p_jg_{j,\ n}=0$。在双能级极限下以及对于 $p_0>1/3$，式（S3）给出$|g_{0,\ n}|<[2\mathrm{Var}(G_n)]^{1/2}$，因此式（S2）的上限保证$1+\delta\tau_ng_{0,\ n}>0$。对于 AITE，其他生成元特征值可能低于 $g_{0,\ n}$，因此无法排除其他极点。图 S1 对比了所有方案的步长控制。

此外，在一阶近似下，态的变化量为 $\||\phi_{n+1}\rangle-|\phi_n\rangle\|\simeq\delta\tau_n\|G_n|\phi_n\rangle\|=\delta\tau_n\ G_n^{\ 1/2}$ Var()。因此，式（S2）中的上限将每一步在希尔伯特空间中的移动距离限制为 $\tilde{c}/\sqrt{2}$。对于虚时间演化，$\mathrm{Var}(G_n)=\mu_2\to0$ 在收敛附近，从而恢复 $\delta\tau_n=\delta\tau_0$。然而，该上限并不能保证投影随机演化的稳定性，因为这一论证适用于确定性线性问题，不能直接推广到采样得到的变分更新 [69]。

### S7：样本空间显式和隐式估计量

这些量是在当前态 $|\Phi_n\rangle$下计算的，生成元为$\hat{}$ $G_n$ ，步长为 $\delta\tau_n$ [，使用第S6], 节中给出的估计量，见第S4.2节；为简洁起见，我们省略下标 $n$。

量子几何张量和变分力 [式（A6）] 由同一样本集上的协方差估计量计算得到 [8]

$$S_{jk} = \langle \widetilde{O}_j^* \widetilde{O}_k \rangle_p \,, \quad F_j = \langle \widetilde{O}_j^* G_{\mathrm{loc}} \rangle_p \,, \tag{S1}$$

$\widetilde{}$并且由于中心化，$\langle O_j \rangle_p = 0$, $F_j$ 在任意$\hat{}$$\hat{}$平移 $G \to G + c$下保持不变。

此外，隐式 SR 表述需要估计$\hat{}$切空间矩阵的 $K_{jk} = \langle T_j|G|T_k\rangle$~[式（11）]。由 $|T_k\rangle = O_k|\Phi_n\rangle$ [式（A4）], 可得 [70–72],

$$\begin{aligned} K_{jk} &= \sum_{\mathbf{s}} \widetilde{O}_j^*(\mathbf{s})\, \Phi_n^*(\mathbf{s}) \langle \mathbf{s}|\hat{G}|T_k\rangle \\ &= \sum_{\mathbf{s}} |\Phi_n(\mathbf{s})|^2 \widetilde{O}_j^*(\mathbf{s})\, \frac{\langle \mathbf{s}|\hat{G}|T_k\rangle}{\Phi_n(\mathbf{s})} \\ &= \langle \widetilde{O}_j^* B_k \rangle_p \,, \end{aligned} \tag{S2}$$

其中 $B_k(\mathbf{s}) \equiv \langle \mathbf{s}|\hat{}G|T_k\rangle / \Phi_n(\mathbf{s})$。对式（S2）中的局域估计量求导，得到

$$\frac{\langle \mathbf{s}|\hat{A}|T_k\rangle}{\Phi_n(\mathbf{s})} = \partial_{\theta_k} A_{\mathrm{loc}}(\mathbf{s}) + \widetilde{O}_k(\mathbf{s}) A_{\mathrm{loc}}(\mathbf{s}) \,, \tag{S3}$$

从而可以在不构造 $B_k$ 的情况下计算 $|T_k\rangle$。$\hat{}$ $\hat{}$对于 $A = G_n$，求导时将其矩和系数视为固定值$\hat{}$。对于对角 $A$，第一项消失；一般而言，式（S3）涉及对通过 $\Psi_{\boldsymbol{\theta}}(\mathbf{s}')/\Psi_{\boldsymbol{\theta}}(\mathbf{s})$连接的构型上的振幅比求导。$\hat{}$这可能会带来较大的计算开销，尤其是对于 AITE [式（S7）]。

$\bar{H} = \hat{}H - E$ 和 $\bar{H}^2$ 的局域估计量为

$$\begin{aligned} x(\mathbf{s}) &= H_{\mathrm{loc}}(\mathbf{s}) - E \,, \\ y(\mathbf{s}) &= (H^2)_{\mathrm{loc}}(\mathbf{s}) - 2E H_{\mathrm{loc}}(\mathbf{s}) + E^2 \,, \end{aligned} \tag{S4}$$

其中 $(H^2)_{\mathrm{loc}}$ 使用了 $\hat{}H$ （第 S3节）的 Pauli 字符串展开的乘积。由于$\hat{}H$ 是厄米的，中心矩变为

$$\mu_2 = \langle |x|^2 \rangle_p \,, \quad \mu_3 = \mathrm{Re}\langle x^* y \rangle_p \,, \quad \mu_4 = \langle |y|^2 \rangle_p \,, \tag{S5}$$

这只要求$H$ 的幂最高达到$H^2$。由于 $(H^2)_{\mathrm{loc}} \neq H_{\mathrm{loc}}^2$, $\mu_3$ 并不是采样局部能量的三阶中心矩。

对于 AITE 生成元，由于 $\kappa = \mu_3/\mu_2^{3/2}$ 在$\mu_2 \to 0$时发散，我们将 $\tan(2\varphi) = \kappa/2$ 代入系数中

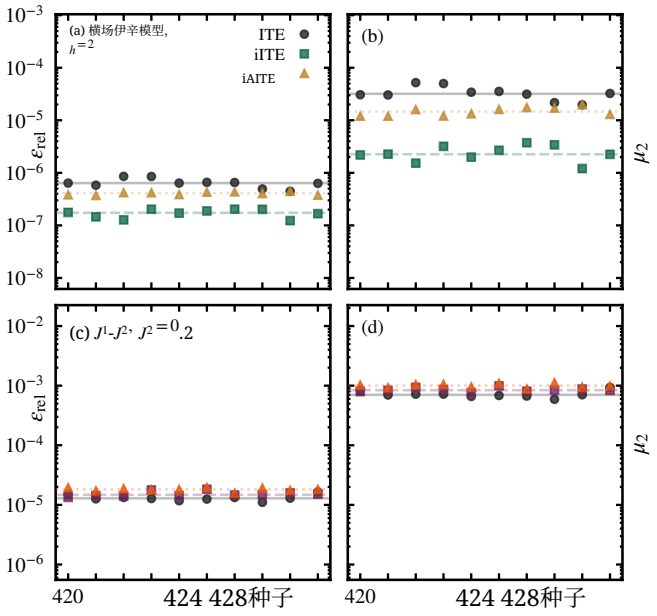


图 S2。（a）、（c）相对能量误差 $\epsilon_{\mathrm{rel}}$ 以及（b）、（d）最终 10% 运行步中的能量方差 $\mu_2$ ，分别与随机种子进行对比，其中（a、b）为 $h = 2$ 处的横场伊辛模型，（c、d）为 $J_2 = 0.2$处的 $J_1$-$J_2$ 模型，且 $\alpha = 20$、$N_s = 16\,384$。标记表示最终 10% 步骤上的中位数。

公式（S7），并直接根据矩对其进行计算，

$$a = \frac{2\mu_2^{3/2}}{\varrho} \,, \quad b = \frac{\mu_3}{2\sqrt{\mu_2}\,\varrho} \,, \tag{S6}$$

与 $\varrho = \sqrt{4\mu_2^3 + \mu_3^2}$对于 $\mu_3 < 0$， $b < 0$ 可能产生冻结 AITE 生成元的负高能本征值。如果 $\mu_2 \leq 0$ 或 $\mu_3 \leq 0$ ，我们则退回到虚时间演化 $(a,\ b) = (1,\ 0)$。

最后，利用式（S4），$\hat{}$ $G_{\mathrm{AITE}}$[式（2）]的局域估计量与 AITE 力给出

$$\begin{aligned} G_{\mathrm{loc}}^{\mathrm{AITE}} &= a x + b(y - \mu_2) \,, \\ F_j^{\mathrm{AITE}} &= a\langle \widetilde{O}_j^* x \rangle_p + b\langle \widetilde{O}_j^* y \rangle_p \,, \end{aligned} \tag{S7}$$

生成元的方差可由此获得

$$\begin{aligned} \mathrm{Var}(G_{\mathrm{AITE}}) &= \langle \hat{G}_{\mathrm{AITE}}^2 \rangle \\ &= a^2\mu_2 + 2ab\mu_3 + b^2(\mu_4 - \mu_2^2) \,, \end{aligned} \tag{S8}$$

对于虚时间演化，这可化简为 $\mu_2$。在数值实现中，我们改用 $G_{\mathrm{loc}}^{\mathrm{AITE}}$ 的样本方差。

对于 $N_s$ 个样本，式（S1）中的估计量是样本矩阵的乘积。对于 $N_s \times N_p$ 矩阵 $X_{ik} =$~$O_k(\mathbf{s}_i)/\sqrt{N_s}$ 和向量 $u_i = [G_{\mathrm{loc}}(\mathbf{s}_i) - \overline{G_{\mathrm{loc}}}]/\sqrt{N_s}$，其中 $G_{\mathrm{loc}}$ 是样本均值， $S = \overline{X^\dagger X}$ 以及 $\mathbf{F} = X^\dagger \mathbf{u}$。

$\bar{}$此外，在隐式更新 Eq. (S3) 应用于 $H$和 $\bar{H}^2$时，结合局域估计量 $x$和 $y$ [Eq.(S4)], 可得

$N_s \times N_p$ 矩阵

$$Y_{ik} = \frac{\langle \mathbf{s}_i|\bar{H}|T_k\rangle}{\sqrt{N_s}\,\Phi_n(\mathbf{s}_i)} = \frac{[\partial_{\theta_k} x + \widetilde{O}_k\, x](\mathbf{s}_i)}{\sqrt{N_s}}\,,$$
$$Z_{ik} = \frac{\langle \mathbf{s}_i|\bar{H}^2|T_k\rangle}{\sqrt{N_s}\,\Phi_n(\mathbf{s}_i)} = \frac{[\partial_{\theta_k} y + \widetilde{O}_k\, y](\mathbf{s}_i)}{\sqrt{N_s}}\,. \tag{S9}$$

将$\hat{G} = a\bar{H} + b(\bar{H}^2 - \mu_2)$代入 Eq. (S2) 得到 $K = X^\dagger(aY + bZ) - b\mu_2 S$。隐式更新[Eq.(11)]的矩阵可因式分解为

$$S + \delta\tau K = X^\dagger V\,,$$
$$V = (1 - \delta\tau\, b\mu_2)X + \delta\tau(aY + bZ)\,. \tag{S10}$$

精确的 $K$ 是厄米的 [第 S5]节。只有在进行全求和时，或对于 $N_s \to \infty$，其估计值才具有同样的性质。$\hat{}$这是因为 $G$ 仅作用于方程（S2）的右矢侧。因此，在有限 $N_s$下，该矩阵 $X^\dagger V$ 不是厄米的。

### S7.1. 求解线性方程组

显式 [方程（6）] 和隐式 [方程（11）] 更新步骤都求解

$$(\tilde{\lambda} I + \mathcal{K})\mathbf{d} = -\mathbf{F}\,, \tag{S11}$$

其中，$\mathcal{K} = S + \delta\tau K = X^\dagger V$对于隐式 更新步骤[方程（S10）]，，而对于 $\mathcal{K}$ 显式 $S = X^\dagger X$ 更新步骤，取代为。对角位移 $\tilde{\lambda} \geq \lambda$，其中 $\lambda = 10^{-3}$，可对系统进行正则化，因为采样得到的 $S$ 对于$\tilde{}$是奇异的。我们使用 $\lambda = \lambda$，除非应用下面的信赖域。

对于显式 更新步骤，方程（S11）的矩阵是厄米且正定的，因此使用 Cholesky 分解。对于隐式 更新步骤，线性方程组不是厄米的，因此我们使用 LU 分解 [104]。对于$N_p > N_s$，则可以通过样本空间中的$N_s \times N_s$ 矩阵 [50, 51], 来求解，这会将稠密求解的成本从 $\mathcal{O}(N_p^3)$ 降至 $\mathcal{O}(N_p N_s^2 + N_s^3)$，使用

$$(\tilde{\lambda} I + X^\dagger V)^{-1} = \tilde{\lambda}^{-1}\left[I - X^\dagger(\tilde{\lambda} I + V X^\dagger)^{-1} V\right]\,, \tag{S12}$$

以 $X$ 替代 $V$，用于显式 更新。在本研究中，我们不采用这种简化。对于大型系统，我们使用 Arnoldi 过程，它只需要与 $X$、$V$ 和 $X^\dagger$ 的乘积，且其 Krylov 子空间不依赖于 $\lambda$ [105]$\tilde{}$。

此外，我们通过信赖域限制每次更新的参数变化，

$$\delta\tau^2 \|\mathbf{d}\|^2 \leq C\,, \quad C_{\text{ansatz}} = 10^{-4}\,\frac{\min(n_\theta, 2N_s)}{N}\,, \tag{S13}$$

具有 $n_\theta$ 个实参数、$N_s$ 个样本（全求和时的基态数）以及 $N$ 个晶格点。

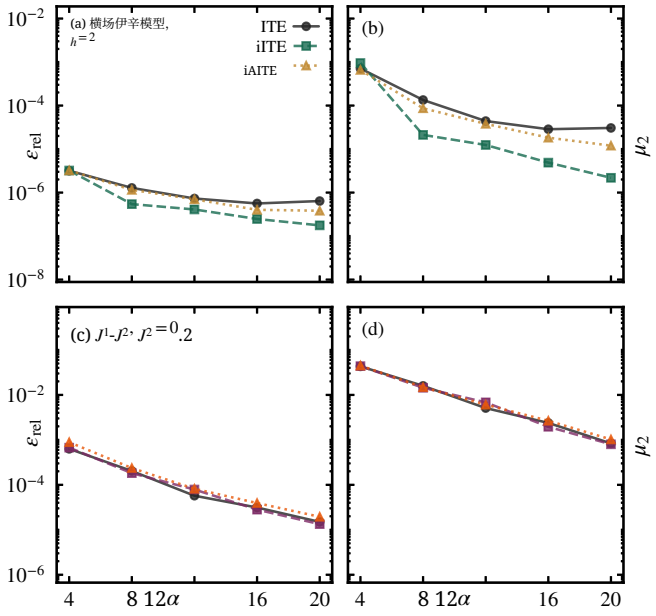


图 S3。（a）、（c）相对能量误差 $\epsilon_{\text{rel}}$ 和（b）、（d）能量方差 $\mu_2$，分别针对每次运行最后 10% 的结果，随受限玻尔兹曼机宽度 $\alpha$ 的变化；其中（a）、（b）为 $h = 2$ 下的横场伊辛模型，（c）、（d）为 $J_2 = 0.2$下的 $J_1$-$J_2$ 模型，且 $N_s = 16\,384$。

该界限适用于所有方案的更新，包括显式和隐式方案。除非另有说明，所有运行中我们均使用 $C = C_{\text{ansatz}}$。如果在 $\lambda = \lambda$ 处的解违反此界限，我们就增大 $\tilde{\lambda}$，直到 $\delta\tau\|\mathbf{d}\| = \sqrt{C}$，这就是一个 Levenberg–Marquardt 步长 [106]。

### S8：双能级欧拉渐近性

考虑S1.3 和S2.1 中$\hat{}$固定步长 $\delta\tau$下的双能级极限，其中生成元 $G_n$ 在当前态 $|\phi_n\rangle$处冻结，其本征值为$g_{0,\ n}$和$g_{*,\ n}$。显式、$r^{(n+1)}$以及隐式、 $r'^{(n+1)}$欧拉更新 [65] 给出的振幅比为

$$r^{(n+1)} = r^{(n)}\left|\frac{1 - \delta\tau g_{*,n}}{1 - \delta\tau g_{0,n}}\right|\,,$$
$$r'^{(n+1)} = r'^{(n)}\left|\frac{1 + \delta\tau g_{0,n}}{1 + \delta\tau g_{*,n}}\right|\,, \tag{S1}$$

这些关系在分母非零处均成立。

虚时间演化。对于虚时间演化，将第 $g_0^{\text{ITE}}$、$_{n} = -\Delta_* p^{(n)}$ 和$g_*^{\text{ITE}}$、$_{n} = \Delta_*(1 - p^{(n)})$ 节 S1.3中在第 $n$ 步、取 $p^{(n)} = (r^{(n)})^2/[1 + (r^{(n)})^2]$ [的本征值代入式（S19）]，，再代入式（S1），并利用 $\mathcal{T} = \delta\tau\Delta_*$可得

$$r^{(n+1)} = r^{(n)}\left|\frac{1 - \mathcal{T} + (r^{(n)})^2}{1 + (1 + \mathcal{T})(r^{(n)})^2}\right|\,,$$
$$r'^{(n+1)} = r'^{(n)}\left|\frac{1 + (1 - \mathcal{T})(r'^{(n)})^2}{1 + \mathcal{T} + (r'^{(n)})^2}\right|\,. \tag{S2}$$

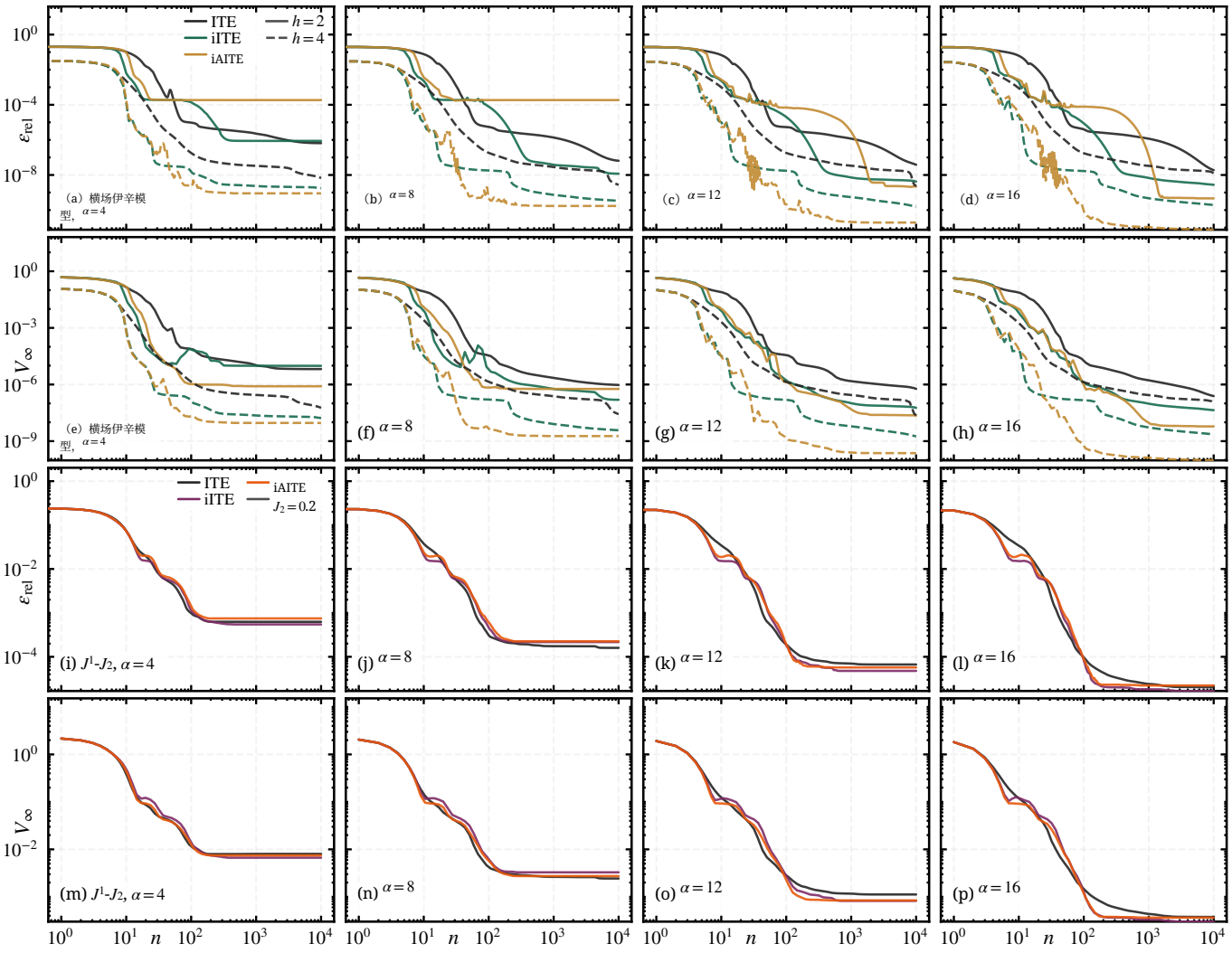


图 S4. (a)–(d) 相对能量误差 $\epsilon_{\text{rel}}$ 和 (e)–(h) V 分数 $V_\infty$ 随迭代次数 $n$ 变化的曲线，针对横场伊辛模型在 $h = 2$（实线）和 $h = 4$（虚线）时的情况；以及 (i)–(l) $\epsilon_{\text{rel}}$ 和 (m)–(p) $V_\infty$ 针对 $J_1$-$J_2$ 模型在 $J_2 = 0.2$时的情况，其中在 $4 \times 4$ 晶格上对所有 $2^{16}$ 构型进行全求和。各列对应的宽度为 $\alpha = 4$，8，12，16。

在固定 $\delta\tau > 0$ 时，将 $\epsilon^{(n)} \simeq \Delta_*(r^{(n)})^2$ 代入 $r^{(n)} \to 0$，显式 $\epsilon^{(n+1)}$ 和隐式 $\epsilon'^{(n+1)}$ 更新步骤的渐近递推关系为

$$\begin{aligned} \epsilon_{\text{ITE}}^{(n+1)} &\sim (1-\mathcal{T})^2 \epsilon_{\text{ITE}}^{(n)} , \\ \epsilon_{\text{ITE}}^{\prime(n+1)} &\sim \frac{\epsilon_{\text{ITE}}^{\prime(n)}}{(1+\mathcal{T})^2} . \end{aligned} \tag{S3}$$

一般而言，两种方案在$\epsilon^{(n)}$上均呈线性收敛。显式虚时间演化需要 $0 < \mathcal{T} < 2$，而隐式虚时间演化原则上对任意固定的$\mathcal{T} > 0$均局部稳定。然而，对于 $\mathcal{T} > 2$，激发态也是隐式更新的稳定不动点，因为 $r'^{(n+1)}/r'^{(n)} \to |1-\mathcal{T}| > 1$ 对于 $r'^{(n)} \to \infty$ [式（S2）]成立。事实上，采用冻结生成元的隐式虚时间演化步是一种位移为 $E_n - 1/\delta\tau$的反幂迭代，它收敛到最接近该位移的本征态 [71]。因此，较大的隐式步可能收敛到激发态，具体取决于初态。

增强虚时间演化。对于 AITE，将式（S10）的本征值代入式（S1），并令$\tilde{\mathcal{T}} \equiv \mathcal{T}/2$，可得显式

$$r^{(n+1)} = r^{(n)} \left| \frac{1-\tilde{\mathcal{T}}/r^{(n)}}{1+\tilde{\mathcal{T}} r^{(n)}} \right| = \frac{|r^{(n)} - \tilde{\mathcal{T}}|}{1+\tilde{\mathcal{T}} r^{(n)}} , \tag{S4}$$

and

$$\begin{aligned} r'^{(n+1)} &= r'^{(n)} \left| \frac{1-\tilde{\mathcal{T}} r'^{(n)}}{1+\tilde{\mathcal{T}}/r'^{(n)}} \right| \\ &= \frac{(r'^{(n)})^2 |1-\tilde{\mathcal{T}} r'^{(n)}|}{r'^{(n)} + \tilde{\mathcal{T}}} . \end{aligned} \tag{S5}$$

随着 $r^{(n)} \to 0$, $g_{*,\ n}$ 发散，在固定 $\tilde{\mathcal{T}} > 0$ 时，显式因子 $1 - \tilde{\mathcal{T}}/r^{(n)}$ 会超过激发态振幅，

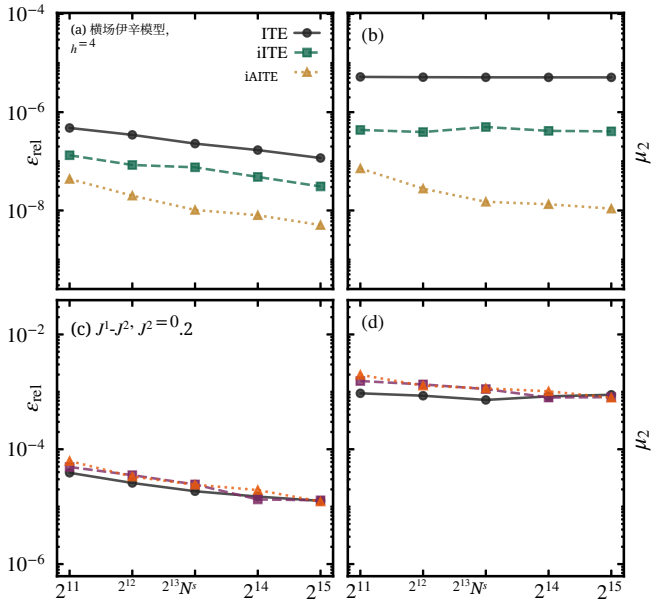


图 S5. (a)、(c) 相对能量误差 $\epsilon_{\rm rel}$ 以及 (b)、(d) 各次运行最后 10% 的能量方差 $\mu_2$ 与每次更新的样本数 $N_s$ 之间的关系，其中 (a)、(b) 为 $h=4$ 下的横场伊辛模型，(c)、(d) 为 $J_2=0.2$ 下的 $J_1$-$J_2$ 模型，且 $\alpha=20$。

给出

$$\lim_{r^{(n)}\to 0^+} r^{(n+1)} = \tilde{\mathcal{T}}\,,$$
$$\lim_{r^{(n)}\to 0^+} \epsilon_{\rm AITE}^{(n+1)} = \frac{\Delta_*\tilde{\mathcal{T}}^2}{1+\tilde{\mathcal{T}}^2} > 0\,, \tag{S6}$$

因此，固定步长的显式AITE 不会局部收敛到基态。相反，选择递减步长 $\tilde{\mathcal{T}}_n = cr^{(n)}$，并令 $0<c<2$，可得

$$r^{(n+1)} = \frac{|1-c|r^{(n)}}{1+c(r^{(n)})^2}\,,$$
$$\epsilon_{\rm AITE}^{(n+1)} \sim (1-c)^2\epsilon_{\rm AITE}^{(n)}\,. \tag{S7}$$

因此，采用满足 $\delta\tau_n \sim (\epsilon^{(n)}_{\rm\ AITE})^{1/2}$ 的通用稳定自适应步长，再次得到线性收敛。在此设置下，经过调优的 $c=1$ 可在一次迭代中消除激发态分量。

˜相比之下，对于固定 $\mathcal{T}>0$ 下的隐式更新，递推关系变为 $r'^{(n+1)} \sim (r'^{(n)})^2/\tilde{\mathcal{T}}$，因此

$$\epsilon_{\rm AITE}^{(n+1)} \sim \frac{4}{\delta\tau^2\Delta_*^3}\left(\epsilon_{\rm AITE}^{(n)}\right)^2\,. \tag{S8}$$

在双能级极限下，隐式AITE 的局部能量误差因此呈现二次收敛，而非线性收敛，且不要求时间步长 $\delta\tau$ 趋于零。

式（S8）的二次收敛在渐近意义上也适用于完整谱。在带有 $a$，$b>0$ 的基态区间中，所有更高能级的生成元特征值都大于 $g_{*,\ n}$，因此衰减得更快，所以最低可达激发决定了渐近行为。

## S9：补充数值结果

本节汇总支持正文中各项论断的结果。除非另有说明，此处的每次运行均使用投影的对称性适配受限玻尔兹曼机，并采用带周期性边界条件（PBC）和对角位移 $\lambda = 10^{-3}$ 的 $4\times4$ 晶格，通过玻恩采样进行采样。

我们选取 $h=2$ 偶数全局$X$字称扇区中的 TFI，以及 $J_1$-$J_2$ 模型在 $J_2=0.2$ 的 $S^z_{\rm tot}$ =0 扇区中进行计算。除非另有说明，所有计算均采用 $\delta\tau_0=1$，其中 $c=1.5$用于 ITE，$\tilde{c}=1$ 用于 iITE 和 iAITE（第 S6节），并采用信赖域 $C = C_{\rm ansatz}$ [式（S13）], 。在图 S2–S5中，相对能量误差 $\epsilon_{\rm rel}$ 和能量方差 $\mu_2$ 取最后 10% 更新步骤的中位数。

### S9.1. 种子依赖性

NQS 优化会通过马尔可夫链和初始参数受到伪随机数生成器种子的影响。为量化这种依赖性，图S2 针对种子 420–429 重复展示了在 $\alpha=20$和 $N_s = 16\,384$下的一种配置。图 S2(a)、(c) 展示了 $\epsilon_{\rm rel}$，图S2(b)、(d) 则展示了 $\mu_2$ 每个种子下横场伊辛模型的虚时间演化、iITE 和 iAITE 结果，分别对应图 S2(a)、(b)；以及 $J_1$-$J_2$ 模型的结果，对应图S2(c)、(d)。对于本文研究的模型，所有种子都得到了相一致的结果。

### S9.2. 拟设空间容量

为区分试探态的局限性与优化的局限性，我们在图 S3 中改变对称性适配 RBM 的宽度，取值为 $\alpha=4$、8、12、16、20，并在 $N_s = 16\,384$下使用相同的种子。图S3(a)、(c) 展示了 $\epsilon_{\rm rel}$，图 S3(b)、(d) $\mu_2$ 分别给出了横场伊辛模型 (a)、(b) 和 $J_1$-$J_2$ 模型 (c)、(d) 的虚时间演化、iITE 和 iAITE 结果。总体而言，在固定其余超参数时，增大 $\alpha$ 会降低最终的 $\epsilon_{\rm rel}$。本文考察的所有模型均呈现这一一致趋势。

在图 S4 中，为了隔离投影到受限玻尔兹曼机流形上的影响，我们进一步对 $4\times4$ 晶格模拟扇区的完备基求和，其中 $\alpha=4$、8、12、16；对于 $h=2$ 和 $h=4$ 处的横场伊辛模型，以及 $J_2=0.2$处的 $J_1$-$J_2$ 模型。图 S4(a)–(d) 和图 S4(i)–(l) 的面板展示了 $\epsilon_{\rm rel}$，而图 S4(e)–(h) 和图 S4(m)–(p) 的面板展示了 ITE、iITE 和 iAITE 的 V 分数 $V_\infty$，其随有限演化步数 $n$ 的变化；前者对应横场伊辛模型 [图 S4(a)–(h)]，后者对应 $J_1$-$J_2$ 模型 [图 S4(i)–(p)]。每一列对应一个宽度 $\alpha$。

### S9.3. 样本数量

最后，为了检验可达精度如何取决于估计量的统计误差，图 S5 将样本数从 $N_s = 2^{11}$、.....、215 变化至 $\alpha$ =20。图 S5(a)、(c) 的面板展示了 ITE、iITE 和 iAITE 的 $\epsilon_{\text{rel}}$ ，而图 S5(b)、(d)$\mu_2$ 展示了它们在 $h$ =4 [图 S5(a)、(b)] 处的横场伊辛模型以及在 $J_2 = 0.2$[图 S5(c)、(d)]处的 $J_1$-$J_2$ 模型中的结果。尽管两个模型均显示出一致的结果，但 $J_1$-$J_2$ 模型从增加的样本数 $N_s$ 中获益较少，并且相比横场伊辛模型，各方案之间的精度差异也更小。