arXiv:2305.01799v1 [quant-ph] 2 May 2023

玻色变分量子电路中的能量依赖型贫瘠高原

Bingzhi Zhang[1] 和 Quntao Zhuang[1,2,*]1南加州大学 *Ming Hsieh*电气与计算机工程系，美国加利福尼亚州洛杉矶 *90089*2南加州大学物理与天文学系，美国加利福尼亚州洛杉矶 *90089*

玻色连续变量变分量子电路（VQC）对于腔量子电动力学和光学系统中的信息处理至关重要，并广泛应用于量子通信、传感和纠错。由于相关物理系统具有无限维特性，缺乏诸如 t-设计等理论工具，因此这类变分量子电路的可训练性尚未得到充分理解。我们克服了这一困难，揭示了这类变分量子电路中存在能量依赖型贫瘠高原。梯度方差按 $1/E^{M\nu}$衰减，随模式数 M 呈指数衰减，而随（每模式）电路能量 E呈多项式衰减。对于浅层电路，指数为 $\nu=1$；对于深层电路，指数为 $\nu=2$。我们针对一般高斯态和数态的态制备证明了这些结果。我们还提供了数值证据，表明这些结果可推广至一般态制备任务。由于电路能量是可控参数，我们提出了一种缓解连续变量变分量子电路中贫瘠高原的策略。

I. 引言

变分量子电路（VQCs） [1] 是在含噪中等规模量子（NISQ）时代实现实用量子优势的候选方案 [2]，此时可扩展的纠错量子计算机尚不可用。变分量子电路利用经典控制来优化量子线路，以解决计算问题，包括优化 [3]，本征系统问题 [4–10]，偏微分方程 [11]，量子模拟 [12–14] 以及机器学习 [15–23]。作为一种通用的量子线路设计方法，它也已应用于量子态的近似 [24]，制备 [25, 26]，分类 [27–31] 和层析成像 [32]。

关于变分量子电路的早期研究聚焦于离散变量（DV）量子比特有限维系统，这类系统天然适用于计算；而连续变量（CV）玻色量子模式系统则较少受到研究。然而，许多重要的量子系统都以量子模式建模。例如，量子通信与网络 [33–37] 依赖光子——这是唯一的飞行量子信息载体。在这方面，连续变量变分量子电路已被证明能够增强量子转导和纠缠蒸馏 [38]；光子量子计算机 [39, 40]也依赖猫码和 Gottesman-Kitaev-Preskill（GKP）码等玻色编码 [41]，，后者展现出巨大潜力 [42, 43]。此类编码态的工程实现因连续变量变分量子电路而得到极大推动 [44–47]；最后，分布式纠缠传感器网络普遍依赖连续变量变分量子电路，以在传感 [48–51] 和数据分类 [52, 53]中实现量子优势。

与传统算法不同，变分量子电路的运行时间取决于训练变分参数以优化代价函数所需的时间。因此，代价函数的景观决定了变分量子电路的可训练性——整个景观中较大的梯度将确保平滑的训练过程。对于离散变量系统，得益于成熟的随机酉算符工具箱和 t-设计 [54–57]，，严谨的结果揭示了贫瘠高原现象 [58–61]，该现象表明，当电路深度并非过浅时，代价函数梯度通常会随着量子比特数量的增加而呈指数衰减。然而，连续变量变分量子电路的可训练性尚未得到探索，这或许是由于在 $t\geq 2$ 无限维 [62–64]中不存在 t-设计这一根本问题。

在本研究中，我们通过提出能量受限随机量子线路的概念，克服了这一问题，并证明这类线路存在一种独特的能量依赖型贫瘠高原现象。对于一个 M 模式的态制备任务，梯度方差渐近地按 $\sim 1/E^{M\nu}$ 衰减：相对于模式数 M 呈指数衰减，而相对于线路能量 E 呈幂律衰减。对于浅层电路，贫瘠高原指数为 $\nu=1$；对于深层电路，则为 $\nu=2$。我们针对一般高斯态和数态的态制备任务证明了这些结果。由于变分量子电路是随机初始化的，我们预期能量依赖型贫瘠高原适用于所有态制备任务，并提供了支持这一预期的数值证据。此外，由于随机初始化时线路能量是一个可控参数，我们提出了一种缓解贫瘠高原的策略：针对每个训练问题，启发式地选择初始电路能量。

II. 结果

一般而言，VQC $U(\boldsymbol{x})$ 的目标是最小化代价函数

$$\mathcal{C}(\boldsymbol{x})=\mathrm{Tr}\left[OU(\boldsymbol{x})\rho_0U^\dagger(\boldsymbol{x})\right] \tag{1}$$

关于可调参数 $\boldsymbol{x}$，其中 ρ_0 是初始态，O 是厄米观测量。变分量子电路的性能

* qzhuang@usc.edu

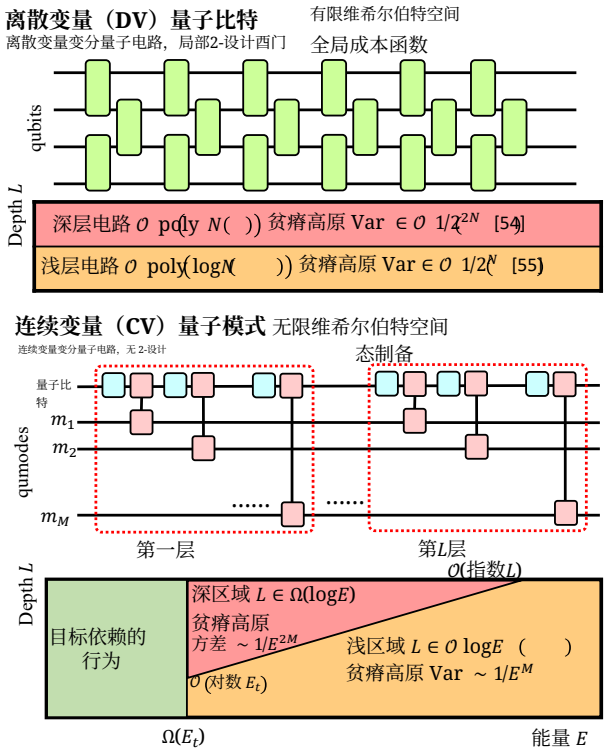


图 1. 离散变量和连续变量系统中变分量子电路可训练性的总结。离散变量系统中的希尔伯特空间是有限维的，而连续变量系统中的希尔伯特空间是无限维的。通用离散变量变分量子电路由局部2设计酉门（黄绿色）构成，通用连续变量变分量子电路则由单量子比特旋转（青色）和回波条件位移（ECD）门（粉色）构成 [47]。在离散变量变分量子电路中，梯度方差随量子比特数呈指数衰减 N，无论是在浅层 [59] 还是深层 [58] 、优化全局成本函数的离散变量变分量子电路中都是如此。本文表明，在连续变量变分量子电路的浅层和深层区域，梯度方差随模式数呈指数衰减 M，但随电路能量呈多项式衰减 E。

取决于表达能力与可训练性之间的平衡。表达能力关注变分量子电路的 ansatz 所能覆盖的解空间大小，可通过充分近似该 ansatz 所需的酉算符集合的基数来量化 [65]；而可训练性关注变分量子电路在解空间内收敛到最优配置的速度。

对离散变量变分量子电路的研究揭示了表达能力与可训练性之间的权衡 [66, 67]。为确保表达能力，离散变量变分量子电路由单量子比特旋转以及双量子比特或多量子比特门组成（见图 1），例如硬件高效拟设 [5]。当层数与量子比特数呈线性关系 N时，变分量子电路可以逼近被分类为t-设计 [54–57]的复杂酉算符系综。然而，提高表达能力会使变分量子电路的训练更具挑战性。随机初始化后，代价函数 $\mathcal{C}$ 关于任意参数的梯度会随着量子比特数的增加而呈指数级减小。更准确地说，参考文献 [58, 59]表明，虽然梯度的均值为零，但梯度方差对于浅层电路按 $\mathcal{O}(1/2^N)$衰减，对于深层电路则按 $\mathcal{O}(1/2^{2N})$衰减（总结见图 1 上图）。

对于连续变量情形，为确保在 M 个振子上的表达能力，我们考虑通过与量子比特进行弱色散耦合来实现回波条件位移（ECD）门的方法 [47] （见图 1）。变分量子电路的每一层都由 M 个 ECD 门组成，它们位于单量子比特旋转之间：量子比特通过 ECD 门控制每个量子模式的位移 $U_{\mathrm{ECD}}(\beta) = D(\beta) \otimes |1\rangle\langle 0| + D(-\beta)$，在一次单量子比特旋转 $U_{\mathrm{R}}(\theta, \phi) = \exp[-i\theta/2(\cos\phi\sigma^x + \sin\phi\sigma^y)]$,之后，其中 σ^x 和 σ^y 分别是泡利 X 和 Y 算符，$D(\beta) \equiv \exp(\beta m^\dagger - \beta^* m)$ 是模式 m上的位移算符。这里，尽管由于维数无限而缺少 Haar 随机酉算符和 t设计（$t \geq 2$） [62–64]，我们仍建立了能量正则化的线路系综，以表示“典型”的量子比特—量子模式线路，从而开展分析。

我们的主要结果表明，在训练此类连续变量变分量子电路以进行态制备时，代价函数会呈现能量依赖型贫瘠高原。具体而言，当制备一个具有 M 个模式、每模能量为 E_t 的态时，如果线路的每模能量 $E \in \Omega(E_t)$，梯度方差会随线路能量呈多项式衰减。当线路具有浅层对数深度（$L \in \mathcal{O}(\log E)$）时，梯度方差按 $1/E^M$ 衰减；在深电路区域（$L \in \Omega(\log E)$），则按 $1/E^{2M}$ 衰减。或者，对于固定电路深度 $L \gtrsim \log(E_t)$，随着线路能量增加，梯度方差首先呈现目标依赖行为，随后迅速 $1/E^{2M}$ 衰减，并在临界能量 $E \sim \exp(L)$处转变为 $1/E^M$标度。我们证明，这些结果适用于一类相当一般的态制备，这类态称为高斯态 [68]，，其中包括对量子计算 [69]和分布式量子传感 [48, 49, 52, 53]有用的非经典多体纠缠态。为将结果推广到高斯态以外的情形，我们还针对福克数态证明了这些结果，并通过数值方法对一般随机态进行了验证。

此外，由于电路能量 E 是一个连续可调的实参数，能量依赖型贫瘠高原为缓解训练中的挑战提供了机会。

A. 能量正则化线路系综

为确保连续变量变分量子电路的表达能力，我们采用适用于量子比特和 M 量子模式的通用门集。虽然已知 ECD 门与单量子比特幺正变换可以在一个量子比特和一个模式上实现通用控制 [47]，，我们将证明扩展到多个量子比特和量子模式的系统（见附录 B）。我们陈述以下引理：

引理 1 通过任意量子模式与任意量子比特之间的所有 *ECD* 门以及所有单量子比特旋转，可以实现对 $M \geq 1$ 量子模式和$N \geq 1$ 量子比特的通用控制。

此外，也可以通过一种

ECD 门的一种变体——条件非位移门 [70]来实现通用控制。在这种情况下，我们论文的结果仍然成立。

尽管 ECD 门的复振幅可以任意大，但现实系统总是受到能量约束，因此我们引入以下定义来刻画连续变量变分量子电路中“典型”的电路布局：

定义 2 能量正则化的 L 层量子比特-量子模式电路 $\mathcal{U}_{E,\,L,\,M}$ 是以下酉算子系综

$$U=\prod_{\ell=1}^{L}\prod_{j=1}^{M}U_{\mathrm{ECD}}^{(j)}\left(\beta_{\ell}^{(j)}\right)U_{\mathrm{R}}\left(\theta_{M(\ell-1)+j},\phi_{M(\ell-1)+j}\right), \tag{2}$$

其中，每个 *ECD* 门的复位移 $\beta_{\ell}^{(j)}\sim$ 服从均值为零、方差为 E/L 的高斯分布，所有量子比特旋转角 $\{\theta_k,\phi_k\}_{k=1}^{ML}\sim\mathbf{U}[0,2\pi)$ 均服从均匀分布。$\mathcal{N}_{E/L}^{\mathrm{C}}$

为简便起见，我们将实部和虚部的方差均为 $\sigma^2/2$ 的零均值复高斯分布记为 $\mathcal{N}_{\sigma^2}^{\mathrm{C}}$。值得注意的是，由于中心极限定理，选择高斯分布来描述 $\beta_{\ell}^{(j)}$ 虽方便，但并非必需。尽管该量子电路系综源自物理正则化，我们稍后将看到，它还使得变分量子电路多种性质的解析解成为可能，并为连续变量量子信息处理提供了绝佳的研究平台。

变分量子电路的表达能力在$\mathcal{U}_{E,L,M}$ 中会随深度 L增加。根据引理 1 中的普适性，$\mathcal{U}_{E,L,M}$ 中具有 $L\gg1$ 的变分量子电路包含与能量尺度 E相关的所有酉算符，而更大的电路能量 E 则能够在更大的希尔伯特空间体积上实现表达能力。同时，$\mathcal{U}_{E,L,M}$中变分量子电路的可训练性尚不明确，这正是我们在本文中希望解决的问题。由于我们关注的是态制备任务，下面考虑将随机变分量子电路作用于平凡初态后生成的态。这类似于酉设计与态设计之间的关系，但遗憾的是，在 CV 情况下二者都不存在。

我们将能量正则化态系综$\Psi_{E,\,L,\,M}$定义为 $|\psi(\boldsymbol{x})\rangle_{q,\,\boldsymbol{m}}=U(\boldsymbol{x})\,|0\rangle_q\otimes_{j=1}^{M}|0\rangle_{m_j}$，其中我们在自旋向上量子比特和真空量子模式上随机施加 $U(\boldsymbol{x})\in\mathcal{U}_{E,\,L,\,M}$。这里，我们将整体参数记为 $\boldsymbol{x}$，包括所有 $\beta_{\ell}^{(j)}$、θ_k 和 ϕ_k。在我们的记号中，$|a\rangle_q$ 表示量子比特“q”的 $a=0$，1 态，而 $|\alpha\rangle_{m_j}$表示模式 m_j 的复振幅为 α 的相干态（见附录 A），$\boldsymbol{m}=(m_1,\ \cdots,\ m_M)$ 表示所有模式。

在式 (2) 中采用随机量子比特旋转 U_{R} 后，对每个模式施加随后的 ECD 门 $U^{(j)}{}_{\mathrm{ECD}}(\beta_{\ell}^{(j)})$，将导致对所有模式 $j=1,\cdots,${v14> 进行复位移 $+\beta_{\ell}^{(j)}$ 和 $-\beta_{\ell}^{(j)}$ 的随机叠加。因此，每一层带有随机量子比特旋转的 ECD 门都对应于 $2M$ 维相空间中随机游走步所有可能选择的叠加。位移的累积会使最终输出态成为随机游走轨迹的叠加。例如，在单模情形 ($M=1$) 中，我们有如下叠加

$$|\psi(\boldsymbol{x})\rangle_{q,m}=\sum_{a=0}^{1}\sum_{\boldsymbol{s}}v_{\boldsymbol{s},a}(\boldsymbol{x})\,|a\rangle_q\,|(-1)^a\boldsymbol{s}\cdot\boldsymbol{\beta}\rangle_m\,, \tag{3}$$

其中，符号向量 $\boldsymbol{s}$ 在满足约束 $\boldsymbol{s}_L=-1$ 和 $\boldsymbol{\beta}=\{\beta_1,\cdots,\beta_L\}$ 为位移振幅的条件下，对 $\{-1,+1\}^L$ 求和。系数 $v_{\boldsymbol{s},\,a}(\boldsymbol{x})$ 的表达式较为冗长，我们将在附录 C 中明确给出这些系数，并提供该态表示的详细证明。

可以直接验证，在经过能量正则化的系综中，对于符号向量 $\boldsymbol{s}$ 的任意选择，位移 $\boldsymbol{s}\cdot\boldsymbol{\beta}$ 均服从高斯分布，并且由于能量正则化，$\Psi_{E,L,1}$ 中态的系综平均能量为 $\mathbb{E}\langle m^{\dagger}m\rangle=E$。此外，由于求和 $\boldsymbol{s}\cdot\boldsymbol{\beta}=\sum_{\ell}(\pm)\beta_{\ell}$ 中位移的累积，中心极限定理表明，最终位移的振幅普遍服从高斯分布，而与每个 β_{ℓ} 的分布无关。

M 模式情形是公式（3）的直接推广，只需将叠加中的相干态扩展为相干态的乘积即可（见附录 F）。与此同时，对应的每模能量仍为 E。

B. 态制备中的贫瘠高原

在能量正则化系综$\mathcal{U}_{E,L,M}$ 中初始化的线路 $U(\boldsymbol{x})$ 的可训练性由式 (1) 中代价函数 $\mathcal{C}$ 的典型梯度表征。这里，我们关注量子模式的态工程，而量子比特充当辅助量子比特。为了制备一般的 M 模式态 $|\psi\rangle_{\boldsymbol{m}}$，可将可观测量选为 $O=|0\rangle\langle0|_q\otimes|\psi\rangle\langle\psi|_{\boldsymbol{m}}$，并且不失一般性地将初始态设为真空态。

考虑关于第 k 个量子比特旋转角 θ_k 的梯度，可以验证参数移位规则 [71] 对连续变量线路中的量子比特旋转角成立，因此 $\partial_{\theta_k}\mathcal{C}=(\langle O\rangle_{k^{(+)}}-\langle O\rangle_{k^{(-)}})/2$，其中 $\langle O\rangle_{k^{(\pm)}}$对应于在旋转角 θ_k 移位 $\pm\pi/2$ 后，通过线路 $U(\boldsymbol{x})$ 制备的态 $\rho_{k^{(\pm)}}$ 对 O 的期望值。按照文献 [58, 59], 的思路，我们考虑随机初始条件——当从能量正则化系综 $\mathcal{U}_{E,\,L,\,M}$ 中随机采样 $U(\boldsymbol{x})$ 时，梯度的均值

$$\mathbb{E}\left[\partial_{\theta_k}\mathcal{C}\right]=\frac{1}{2}\left(\mathbb{E}\left[\langle O\rangle_{k^{(+)}}\right]-\mathbb{E}\left[\langle O\rangle_{k^{(-)}}\right]\right)=0, \tag{4}$$

由于对满足 $\theta_k\pm\pi/2$ 的参数进行系综平均后结果相等。在这种情况下，梯度的典型值由方差刻画

$$\mathrm{Var}\left[\partial_{\theta_k}\mathcal{C}\right]=\frac{1}{2}\mathbb{E}\left[\langle O\rangle_{k^{(+)}}^2\right]-\frac{1}{2}\mathbb{E}\Big[\langle O\rangle_{k^{(+)}}\langle O\rangle_{k^{(-)}}\Big]. \tag{5}$$

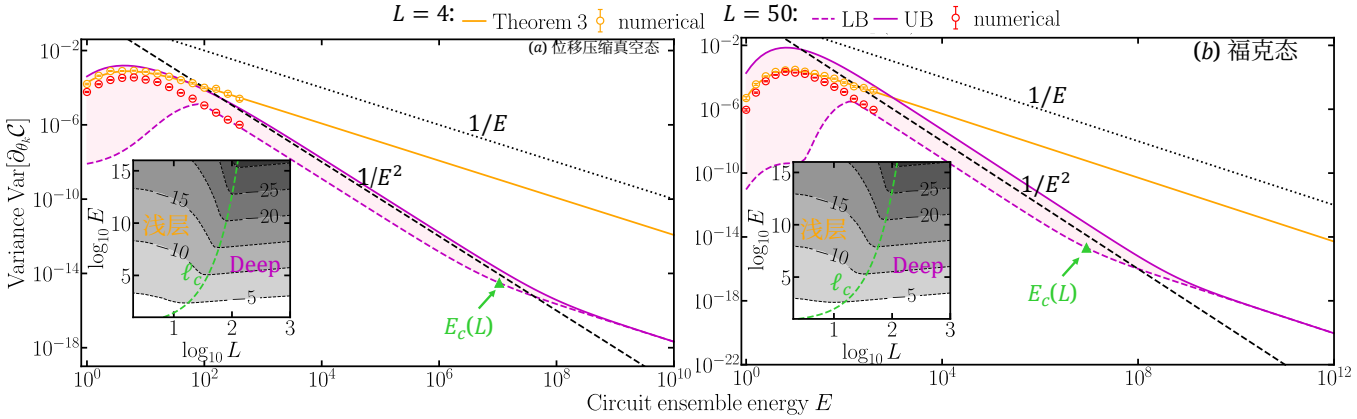


图 2. 在制备（a）位移压缩真空态（DSV）时， $k = L/2$ 处梯度方差 $\mathrm{Var}[\partial_{\theta_k}\mathcal{C}]$，其中 $\gamma = 2$，$\zeta = \sinh^{-1}(2)$；以及（b）福克态时的梯度方差，其中 $E_t = 8$。带有误差棒的橙色和红色圆点分别表示浅层和深层量子电路中的方差数值结果。橙色实线表示定理 3 中的解析方差；洋红色虚线和实线分别表示不等式 (6)、(7) 中的下界和上界。黑色点线和虚线分别表示 $1/E$ 和 $1/E^2$ 的标度。在（a）（b）的插图中，我们绘制了不等式 (7) 中上界的以 10 为底的对数随电路深度和能量的变化。绿色三角形（主图）和线条（插图）分别表示 E_c（L）和 ℓ_c（E）处方差的相应边界。

由于缺乏无限维中的 t-设计等理论工具，$\mathrm{Var}\,[\partial_{\theta_k}\mathcal{C}]$ 的精确计算通常颇具挑战。相反，我们提出的能量正则化系综使我们能够求得一组渐近下界和上界，当 $E \gg 1$ 时成立（忽略高阶项，见附录 F）

$$\mathrm{Var}\,[\partial_{\theta_k}\mathcal{C}] \geq \frac{1}{2}\left[\frac{3^{ML-1}}{4^{ML}}C_1 + \left(\frac{1}{4} - \frac{3^{ML}}{4^{ML}}\right)\min_{\boldsymbol{\ell}} C_2\left(\frac{\boldsymbol{\ell}}{L}\right)\right], \tag{6}$$

$$\mathrm{Var}\,[\partial_{\theta_k}\mathcal{C}] \leq \frac{1}{2}\left[\frac{3^{ML-1}}{4^{ML}}C_1 + \left(\frac{1}{4} + \frac{2^{ML-1}}{4^{ML}}\right)\max_{\boldsymbol{\ell}} C_2\left(\frac{\boldsymbol{\ell}}{L}\right)\right], \tag{7}$$

其中，最小化和最大化是针对向量 $\boldsymbol{\ell} = (\ell^{(1)}, \ldots\ldots, \ell^{(M)})^T$ 进行的，其每个元素均为整数 $\ell^{(j)} \in [1, L-1] \cap \mathbb{N}$。态依赖关联函数为

$$C_1 = \mathbb{E}_{\boldsymbol{\alpha}}\left[|_{\boldsymbol{m}}\langle\psi|\boldsymbol{\alpha}\rangle_{\boldsymbol{m}}|^4\right], \tag{8}$$

$$C_2(\boldsymbol{z}) = \mathbb{E}_{\boldsymbol{\alpha_z},\boldsymbol{\alpha_{1-z}}}\left[\prod_{h=0}^{1}|_{\boldsymbol{m}}\langle\psi|\boldsymbol{\alpha_z} + (-1)^h\boldsymbol{\alpha_{1-z}}\rangle_{\boldsymbol{m}}|^2\right]. \tag{9}$$

这里我们定义向量记法 $|\boldsymbol{\alpha}\rangle_{\boldsymbol{m}} = \bigotimes_{j=1}^{M}|\alpha^{(j)}\rangle_{m_j}$，其中每个 $|\alpha^{(j)}\rangle_{m_j}$ 都是位移为 $\alpha^{(j)}$ 的相干态，对应模式 m_j。根据定义 2，式 (8) 中的系综平均是对每个分量 $\alpha^{(j)}$ 从 $\mathcal{N}_E^{\mathrm{C}}$ 中采样得到的。类似地，$|\boldsymbol{\alpha_z} + (-1)^h\boldsymbol{\alpha_{1-z}}\rangle_{\boldsymbol{m}} = \bigotimes_{j=1}^{M}|\alpha_{z_j} + (-1)^h\alpha_{1-z_j}\rangle_{m_j}$，其中每个分量 $\alpha_{z_j} \sim \mathcal{N}_{\mathrm{C}}^{z_j E}$ 和 $\alpha_{1-z_j} \sim \mathcal{N}_{(1-z_j)E}^{\mathrm{C}}$。由于上述两个关联函数仅取决于态保真度，数值计算通常很高效，有时也可以进行解析计算。请注意，对于 $M \geq 2$，式 (9) 已忽略了 L中指数级小的项，这不会影响我们的结果（见附录 F）。接下来，我们给出一般高斯态和数态的解析结果，以及随机生成态上下界的数值结果。尽管高斯态可以通过由线性光学元件和压缩器构成的高斯电路高效制备 [68]，由于随机初始化变分量子电路没有利用任何启发式方法，我们预计在那里发现的可训练性具有普适性，这一点也得到了其他结果的支持。我们先讨论单模态制备，然后将其推广到多模情形。

1. 单模态制备

对于单模态制备，我们考虑高斯态、数态和随机态。单模（纯）高斯态可表示为位移压缩真空态 $D(\gamma)S(\zeta)|0\rangle_m$，其中 $S(\zeta) = \exp\left[\zeta(m^2 - m^{\dagger 2})\right]$ 是压缩算符（见附录 A）。这是一类重要的态，因为相干态为激光器提供了量子模型，而压缩真空是量子传感的重要资源，例如激光干涉引力波天文台 [72–74] 和暗物质搜索 [75]。这里我们忽略任何可能的相位旋转，相关内容在附录 E

$$|\gamma|^2 + \qquad ^2 \quad \frac{1}{\zeta} \qquad E_t = \text{中}$$

讨论。这种态的能量为 sinh ( )。我们可以解析求出单模态高斯态的两个关联函数：

$$C_1^{\mathrm{Gauss}} = \frac{\mathrm{sech}^2(\zeta)e^{-R(E)/G_1(E)}}{\sqrt{G_1(E)}}, \tag{10}$$

$$C_2^{\mathrm{Gauss}}(z) = \frac{\mathrm{sech}^2(\zeta)e^{-R(zE)/G_1(zE)}}{\sqrt{G_1(E - zE)G_1(zE)}}, \tag{11}$$

其中，我们定义 $G_1(x) = 1+4x+4\mathrm{sech}^2(\zeta)x$ 和 $R(x) = 2|\gamma|^2 + 2\tanh(\zeta)(\mathrm{Re}\{\gamma\}^2 - \mathrm{Im}\{\gamma\}^2) + 4\mathrm{sech}^2(\zeta)|\gamma|^2 x$。

为了超越高斯态，我们考虑含有 $E_t \in \mathbb{N}$ 个光子的福克数态，其关联函数可以用闭式求解为

$$C_1^{\mathrm{Fock}} = \frac{(2E_t)!}{(2^{E_t}E_t!)^2}\frac{(1+1/2E)^{-2E_t}}{1+2E}, \tag{12}$$

$$C_2^{\mathrm{Fock}}(z) = \eta\frac{(1+2E_t)(2E_t)!}{(2^{E_t}E_t!)^2}\frac{(1+1/2Ez)^{-2E_t}}{(1+2Ez)[1+2E(1-z)]}. \tag{13}$$

在第二行中，右侧表示 $C_2^{\mathrm{Fock}}(z)$ 的上下界，因为精确求值变得困难：上界为 $\eta = 1$，下界为 $\eta = {}_2F_1(1/2, -E_t, 1, 1)$，其中 ${}_2F_1$ 是原始超几何函数。通过获得不等式 (6)、(7) 中的这些关联函数，我们可以得到制备相干态和福克态时梯度方差的相应界。

有了上述关联函数，我们发现，当电路深度 L 较浅时，不等式 (6) 和 (7) 中的上下界在 E 较大时关于 $\sim 1/E$ 的主导阶上重合。当深度 L 较大时，$\sim 1/E^2$ 项将在上下界中占主导地位。定量地说，标度的交叉发生在深度 $\ell_c(E) \in \mathcal{O}(\log E)$ 处。关于 $\ell_c(E)$ 的完整公式，请参阅附录 E。等价地，对于深度固定为 L 的线路，标度从 $1/E^2$ 到 $1/E$ 的转变发生在 $E_c(L) \in \Omega(\exp(L))$ 处。总体而言，我们有以下定理：

定理 3 （浅层深度下的贫瘠高原。）对于从能量正则化系综 $\mathcal{U}_{E,L,1}$ 中随机初始化的单模（$M = 1$）连续变量变分量子电路，在浅层深度 $L \le \ell_c(E) \in \mathcal{O}(\log E)$ 下，制备高斯态或福克态时，相对于量子比特旋转角的梯度均值为零，且方差为

$$\mathrm{Var}\left[\partial_{\theta_k}\mathcal{C}\right] = \frac{1}{6}\left(\frac{3}{4}\right)^L C_1 + \mathcal{O}\left(\frac{1}{E^2}\right). \tag{14}$$

其中，C_1 是公式(10) 或(12)中的关联函数。特别地，当 E 趋于无穷大时，$[\partial_{\theta_k}\mathcal{C}] \sim 1/E$ 的方差 为。

定理 4 （深层深度下的贫瘠高原。）对于从能量正则化系综 $\mathcal{U}_{E,L,1}$中随机初始化的单模（$M = 1$）连续变量变分量子电路，在层数为 $L \ge \ell_c(E) \in \Omega(\log E)$时，制备高斯态或福克态时，相对于量子比特旋转角的梯度均值为零，且渐近方差为方差 $\sim 1/E^2$.

就渐近区域而言，我们发现实际中只要 $E \in \Omega(E_t)$，标度 $1/E^2$ 就成立。

在图 2(a)、(b)中，我们展示了浅层和深层量子电路中梯度方差 $\mathrm{Var}\ [\partial_{\theta_k}\mathcal{C}]$ 随系综能量 E 的变化，并将数值结果与各界和定理进行比较。我们考虑制备位移

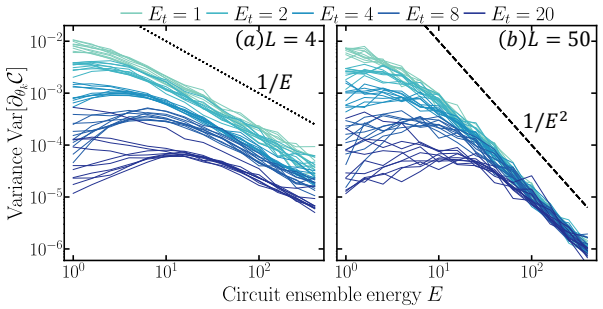


图 3. 在使用 $L = 4$ （a）和 $L = 50$ （b）线路制备随机连续变量态 $|\psi\rangle_m = \sum_n b_n |n\rangle_m$ 时，$k = L/2$ 处梯度方差 Var $[\partial_{\theta_k}\mathcal{C}]$。相同颜色的曲线表示不同样本目标态的方差。（a）和（b）中的黑色点线和虚线分别表示 $1/E$ 和 $1/E^2$的标度。在我们的计算中，截断值选为 $n_c \sim 2E_t$ 和 $\epsilon = 0.1$。

压缩真空 (DSV) 态，其中 $\gamma = 2$，$\zeta = \sinh^{-1}(2)$，以及分别具有 $E_t = 8$ 的福克态。我们看到，浅层量子电路中的数值方差 $L = 4$ （橙色圆点）与定理 3 中给出的式 (14)（橙色线）吻合良好，这表明在 E 的渐近区域内，标度为 $1/E$。对于深层量子电路 $L = 50$，数值结果（红色）位于下界与上界之间，且确实遵循 $1/E^2$ 标度，这支持了定理 4。令我们惊讶的是，在两种情况下，不等式 (6) 中的下界在 E 的渐近区域内都变得极为紧。同时，尽管电路深度很大 $L = 50$，当系综能量远高于 E_c（L）时，与 ℓ_c（E）相比，变分量子电路再次表现为浅层电路，且梯度方差遵循浅层量子电路的 $1/E$ 标度。为理解浅层电路与深层电路之间的转变，我们在图 2 的插图中展示了上界等高线随电路深度 L 和能量 E 的变化。这里，我们发现浅层与深层量子电路在电路深度和系综能量方面存在明显差异，其中表示 $\ell_c \in \mathcal{O}$ (log E) 的边界以绿色曲线标出。除位移压缩真空态外，我们还在附录 G 中考虑了相干态和单模压缩真空态 (SMSV) 的特殊情形，结果再次验证了我们的界和定理。

对于一般的态制备，不等式 (6) 和 (7) 中相关函数 C_1, C_2 的计算可能很困难。不过，依据上述结果，我们猜想定理 3、4 对任意单模态制备都成立。为支持这一猜想，我们给出了制备随机生成的连续变量态的数值证据。这些态是数态基的随机叠加，形式为 $|\psi\rangle_m \propto \sum_{n=0}^{n_c} b_n |n\rangle_m$，其中每个 $b_n \sim \mathcal{N}_2^{\mathbb{C}}$都是随机选取的。我们选取能量窗口 $[E_t-\epsilon, E_t+\epsilon]$内的态。生成目标态后，我们评估制备每个固定态时，梯度方差随电路能量的变化，并考察了不同 E_t 选择下的情况（以颜色表示），如图 3 所示。尽管在低能量区域存在依赖于态的行为，

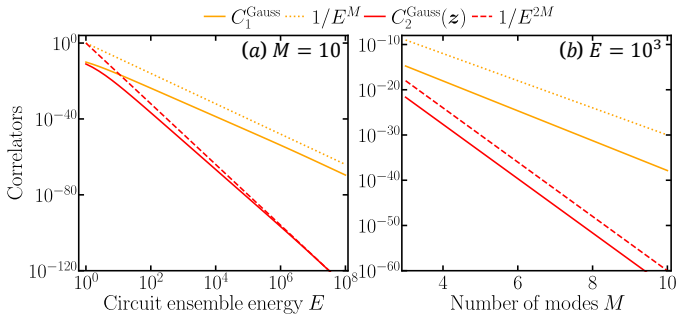


图4. 方程（）、（）中的关联函数 C_1^{Gauss}和 C_2^{Gauss}（$\boldsymbol{z} = \{1/2, \cdots, 1/2\}$ 17 18 ）与 (a) 系综能量 E和 (b) 模式数 M 的关系。目标态 $|\psi\rangle_{\boldsymbol{m}}$ 由全局随机被动高斯幺正变换生成，该变换作用于一个压缩强度为 $r = 8$ 的单模压缩器之后。

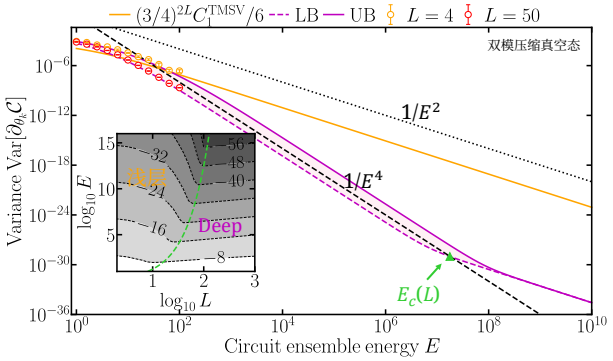


图 5. 制备具有 $\zeta = \sinh^{-1}(2)$ 的 TMSV 态 $|\zeta\rangle_{\mathrm{TMSV}}$ 时，梯度方差 $\mathrm{Var}[\partial_{\theta_k}\mathcal{C}]$ 在 $k = ML/2$ 下的变化。橙色和红色带误差棒的点分别表示浅层和深层电路中的方差数值结果。橙色实线表示作为参考的 $(3/4)^{2L}C_1^{\mathrm{TMSV}}/6$；洋红色虚线和实线分别表示式 (6)、(7) 中的下界和上界。黑色点线和虚线表示 $1/E^2$ 和 $1/E^4$ 的缩放关系。插图展示了以十为底的上界对数随电路深度和能量的变化。绿色三角形（主图）和线条（插图）分别表示 $E_c(L)$ 和 $\ell_c(E)$ 处对应的方差边界。

但在 $E \gtrsim E_t$ 区域可以观察到梯度随能量普遍衰减。浅层电路（子图 (a)）中的衰减呈现 $\sim 1/E$ 的标度，而深层电路（子图 (b)）中的衰减呈现 $\sim 1/E^2$ 的标度。

2. 多模态制备

现在，我们将能量依赖的贫瘠高原结果推广到多模情形，包括关于一般高斯态和数态的解析结果，以及关于随机态的数值结果。我们从乘积态 $|\psi\rangle_{\boldsymbol{m}} = \otimes_{j=1}^{M} |\psi_j\rangle_{m_j}$ 这一简单情形开始，其方程（8）、（9）中的关联函数形式为

$$C_1^{\mathrm{Prod}} = \prod_{j=1}^{M} \left(\mathbb{E}_{\alpha^{(j)} \sim \mathcal{N}_E^{\mathbb{C}}} \left[\left| \langle \psi_j | \alpha^{(j)} \rangle_{m_j} \right|^4 \right] \right), \tag{15}$$

$$C_2^{\mathrm{Prod}}(\boldsymbol{z}) = \prod_{j=1}^{M} \mathbb{E}_{\alpha_y \sim \mathcal{N}_{yE}^{\mathbb{C}}} \left[\prod_{h=0}^{1} \left| \langle \psi_j | \alpha_{z_j} + (-1)^h \alpha_{1-z_j} \rangle_{m_j} \right|^2 \right], \tag{16}$$

它们可化为单模关联函数的乘积。因此，定理 3 和定理 4 可直接推广至单模高斯态乘积和数态乘积的制备：浅层电路 $L \in \mathcal{O}(\log E)$ 中梯度方差的标度为 $1/E^M$，而深层电路 $L \in \Omega(\log E)$ 中则为 $1/E^{2M}$（详细证明见附录 F 1）。

为了超越乘积态，我们考虑一个任意的M-模式高斯态，它通常具有高度纠缠性 [76]。一般的高斯态 $|\psi\rangle_{\boldsymbol{m}}$ 可以由其维格纳函数的均值和协方差矩阵 $\boldsymbol{V_m}$ 来描述（见附录 A）。为简单起见，正文中我们展示均值为零的结果，非零均值的情况见附录 F2。关联函数 C_1 以及 C_2 可以解析地求解为（见附录 F 2）

$$C_1^{\mathrm{Gauss}} = \frac{4^M \det(\boldsymbol{K})}{\sqrt{\det(4\boldsymbol{K} + \boldsymbol{I}/E)} E^M} \tag{17}$$

$$C_2^{\mathrm{Gauss}}(\boldsymbol{z}) = \frac{4^M \det(\boldsymbol{K})}{\sqrt{\det(4\boldsymbol{K} + \boldsymbol{S_z}) \det(4\boldsymbol{K} + \boldsymbol{S_{1-z}})}} \times \frac{1}{\left[\prod_{j=1}^{M} z_j(1 - z_j) \right] E^{2M}}, \tag{18}$$

其中，$\boldsymbol{I}$ 是 $2M \times 2M$ 单位矩阵。这里我们定义了 $\boldsymbol{K} = (\boldsymbol{V_m} + \boldsymbol{I})^{-1}$ 和 $\boldsymbol{S_z} = \oplus_{j=1}^{M} \boldsymbol{I}_2/(z_j E)$，其中 $\boldsymbol{I}_2$ 是 2×2 单位矩阵。在 $E \gg 1$ 的渐近极限下，可以直接看出 $C_1^{\mathrm{Gauss}} \sim 1/E^M$，而 $C_2^{\mathrm{Gauss}}(\boldsymbol{z}) \sim 1/E^{2M}$（非零均值情形的证明见附录 F），由此得到以下定理。

定理5（多模高斯态的贫瘠高原）对于从能量正则化系综 $\mathcal{U}_{E,L,M}$ 中随机初始化的 M 模连续变量变分量子电路，当制备 M 模一般高斯态时，相对于量子比特旋转角的梯度均值为零；在浅层深度下，其渐近方差为 $\mathrm{Var} \sim 1/E^M$，深度为$L \in \mathcal{O}(\log E)$，而深层深度下，$\mathrm{Var} \sim 1/E^{2M}$，深度为$L \in \Omega(\log E)$。

例如，我们考虑一种多部分纠缠的分布式压缩态，它由单模压缩真空态经过随机分束器阵列（全局随机被动高斯幺正变换）生成；在连续变量量子信息研究中，这类态通常具有高度纠缠。

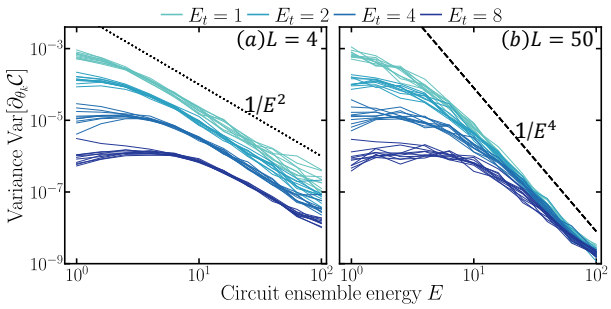


图6。在制备随机双模连续变量态 $|\psi\rangle_{\boldsymbol{m}} = \sum_{n_1,n_2} b_{n_1,n_2} |n_1\rangle_{m_1} |n_2\rangle_{m_2}$ 时，$L = 4$（a）和 $L = 50$（b）量子电路中 $k = ML/2$处梯度的方差 $\mathrm{Var}[\partial_{\theta_k}\mathcal{C}]$ 。相同颜色的曲线表示不同样本目标态的梯度方差。（a）和（b）中的黑色点线和虚线分别表示 $1/E^2$ 和 $1/E^4$的标度。在计算中，我们选择了 $\epsilon = 0.1$ 和 $n_c \sim 2E_t$。

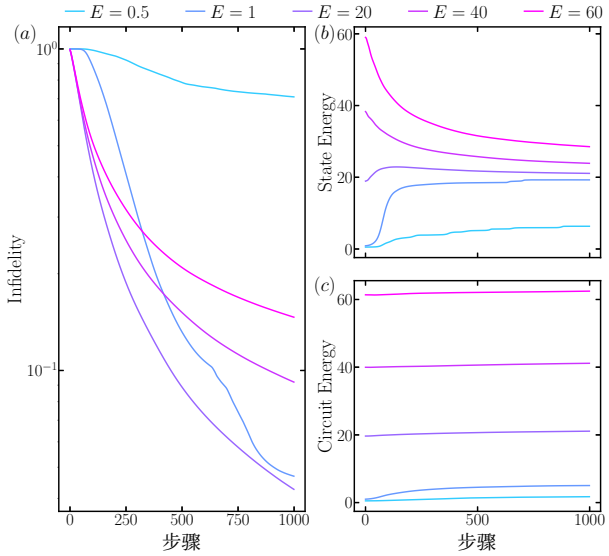


图 7。使用具有不同系综能量 E的 $L = 50$ 连续变量变分量子电路训练福克态 $|\psi\rangle_m = |20\rangle_m$ 。我们展示了 (a) 输出态的平均非保真度、(b) 输出态的平均能量以及 (c) 平均电路能量 $\sum_{j=1}^{L} |\beta_j|^2$ 随训练步数的变化。

搅乱 [77]。它们也是赋能分布式量子传感应用的关键纠缠形式 [48–53]。对于固定的模式数 $M = 10$，从图 4(a) 可以看到，关联函数 C_1^{Gauss}（橙色）和 $C_2^{\mathrm{Gauss}}(\boldsymbol{z})$ 在 $\boldsymbol{z} = \{1/2, \cdots, 1/2\}$（红色）条件下分别趋近于 $1/E^M$（橙色虚点线）和 $1/E^{2M}$（红色虚线）的标度。另一方面，当渐近能量为 $E = 10^3$时，我们看到两个关联函数随模式数 M 呈指数衰减，如图 4(b) 所示。然而，在这种情况下，由于模式数量庞大，直接计算梯度颇具挑战。 $M = 10$ 模式。

为了与定理5进行直接比较，以评估用于高斯态制备的连续变量变分量子电路的梯度，我们考虑双模压缩真空态（TMSV），即贝尔态的连续变量对应形式。TMSV 态 $|\zeta\rangle_{\mathrm{TMSV}}$ 通过对真空态$|0\rangle_{m_1} |0\rangle_{m_2}$施加双模压缩算符 $S_2(\zeta) = \exp\left[\zeta(m_1 m_2 - m_1^\dagger m_2^\dagger)/2\right]$ 生成，其每模能量为 $E_t = \sinh^2(\zeta)$。利用公式（17）、（18）可以求得关联函数，如下所示。

$$C_1^{\mathrm{TMSV}} = \frac{\mathrm{sech}^4(\zeta)}{G_1(E)}, \tag{19}$$

$$C_2^{\mathrm{TMSV}}(z_1, z_2) = \frac{\mathrm{sech}^4(\zeta)}{G_2(z_1, z_2) G_2(1 - z_1, 1 - z_2)}, \tag{20}$$

其中 $G_2(z_1, z_2) = 1 + 2(z_1 + z_2)E + 4\,\mathrm{sech}^2(\zeta) z_1 z_2 E^2$。在渐近区域中，关联函数分别呈现出 $1/E^2$ 和 $1/E^4$ 的标度，因此对应于梯度方差的标度。我们将上述结果与图 5 中的数值模拟进行比较，发现对于浅层和深层量子电路，在 E 的渐近区域内两者吻合良好；而有限能量为 E 的浅层量子电路，其方差可能偏离渐近预测。

与单模情形类似，非高斯态的关联函数的计算通常具有一定难度。我们猜想，对于一般的非高斯态，贫瘠高原现象同样成立，我们通过数值结果加以支持：这些结果来自对随机生成目标态的制备——这是对图 3 所研究的单模情形的一种自然推广，$|\psi\rangle_{\boldsymbol{m}} \propto \sum_{n_1, n_2=0}^{n_c} b_{n_1,\ n_2} |n_1\rangle_{m_1} |n_2\rangle_{m_2}$，其中每个 $b_{n_1,\ n_2} \sim \mathcal{N}_2^{\mathrm{C}}$均为随机选取。我们对能量窗口 $[E_t - \epsilon, E_t + \epsilon]$内的态进行筛选。生成目标态后，我们评估不同目标态能量 E_t（以颜色表示）下，梯度方差随电路能量的变化情况，结果见图 6。我们再次验证了浅层电路（左图）中 $\sim 1/E^2$ 以及深层电路（右图）中 $\sim 1/E^4$ 的标度。然而，由于对计算资源的需求极其庞大，对更多模式的非高斯态进行数值模拟仍然颇具挑战。

C. 规避训练问题的策略

总的来说，贫瘠高原给训练变分量子电路以解决问题带来了挑战。尽管由于复杂度方面的论证，似乎难以通过通用方法彻底解决训练问题，但针对具体问题的启发式方法有望缓解贫瘠高原。对于离散变量变分量子电路，人们已提出多种方法 [78–83] 来缓解训练中的贫瘠高原。

对于连续变量变分量子电路，得益于无限维希尔伯特空间，随机系综除了电路深度之外，还有一个独特的可调参数——电路能量$\mathcal{U}_{E,\ L,\ M}$。因此，可以通过改变电路能量 E采用不同的随机初始化策略。如图 7(a) 所示，在制备福克数态时，当采用不同的

初始电路能量 E时，代价函数（态不保真度）的训练过程会有显著差异。当电路能量 E 与目标态能量 $E_t = 20$ 大致匹配时，收敛效果最佳。这是因为在福克态制备中，梯度方差在图 2(b) 所示位置达到峰值。正如预期，不保真度的下降也反映在系综态能量上：它逐渐趋近于目标态能量，如图 7(b) 所示。相比之下，我们还在图 7(c) 中绘出了定义为 $\sum_{\ell=1}^{L} |\beta_\ell|^2$的电路能量；训练过程中，它们都只发生了轻微变化。类似地，我们也验证了这一策略适用于单模压缩真空态的制备（见附录 G）：采用适当的能量进行初始化同样能够改善收敛。由于在这种情况下梯度在零能量时达到最大值，因此低能量初始化能带来最佳的训练收敛效果。

总而言之，将线路初始化在低能量处，可以缓解贫瘠高原，从而加快收敛。不过，最佳初始能量取决于目标态。例如，要制备高斯态，目标态的正交分量均值会影响 E 的非渐近区域内梯度方差的单调性（见式（10）和（11））：对于均值为零的情况，随着电路能量增加，梯度持续减小；如图 5 中的 TMSV 所示，单模压缩真空态在附录 G 中也有类似结果。对于均值非零的情况，当线路能量 E 与目标态能量 E_t 相匹配时，梯度达到最大值，如图 2(a) 所示。对于一般态，图 3 中的随机连续变量态既有单调递减的情况，也有出现峰值的情况；而数态的梯度方差则呈现峰值，如图 2(b) 所示。在实践中，当训练线路以制备特定态时，可以花费一些计算资源，在低能量区域（$E \lesssim E_t$）评估梯度，以便在正式训练前选出最佳初始点，从而大幅加快收敛。

III. 讨论

本文探讨了基于 ECD 门通用控制实现的连续变量变分量子电路的可训练性。通过制备 M 模一般高斯态和福克态的示例，我们从解析上识别出贫瘠高原现象：梯度方差会随浅层电路的电路能量 Var $\sim 1/E^M$ 而在深层电路中则随 Var $\sim 1/E^{2M}$ 变化。我们还通过数值方法将贫瘠高原现象推广到任意非高斯单模和双模连续变量态。为缓解贫瘠高原，我们提出一种策略：在初始化时调节系综能量，使其与梯度方差的变化规律相匹配。结果表明，该策略能够提升性能。

最后，我们指出几个尚待解决的问题。我们重点研究了相对于量子比特旋转的梯度，以考察连续变量变分量子电路的可训练性（因为部分参数的可训练性较低，就足以证明存在贫瘠高原）；然而，关于 ECD 门中相对于位移参数的梯度如何随能量和模式数增加而衰减，目前尚不清楚。将我们的结果推广到态制备以外的一般任务，也是一项重要工作。对此，我们认为，能量依赖型贫瘠高原可以推广到涉及任意有界多模连续变量可观测量的训练。由于作用于量子模式上的任意有界算符都有一个相干态基底下的 Glauber–Sudarshan P 表示 [84, 85]，因此，有界算符的期望值可以通过对相干态期望值积分得到。由此，任意有界代价函数总可以解释为相干态代价函数的加权平均（权重可能为复数），而态制备中的能量依赖型贫瘠高原通常应当代表涉及任意有界算符的可训练性。除了贫瘠高原之外，探索其他挑战也很有意义，例如连续变量变分量子电路训练中的陷阱 [86]。

致谢

本项目获得美国国家科学基金会CAREER奖（CCF-2142882）的资助。QZ还感谢美国国防高级研究计划局（DARPA）青年教师奖（YFA）项目（资助编号：N660012014029）、美国海军研究办公室（ONR）（资助编号：N000142312296）以及思科系统公司的支持。在QZ参加KITP DYNISQ22研讨会期间，本研究还获得了美国国家科学基金会的部分资助（资助编号：NSF PHY-1748958）。

QZ 提出了这项研究。BZ 在 QZ 的指导下进行了分析并生成了图表。两位作者共同撰写了稿件。

[1] M. Cerezo、A. Arrasmith、R. Babbush、S. C. Benjamin、S. Endo、K. Fujii、J. R. McClean、K. Mitarai、X. Yuan、L. Cincio，等，《变分量子算法》，《自然综述：物理》**3**，625（2021）。[2] J. Preskill，《NISQ 时代及其之后的量子计算》，《Quantum》**2**，79（2018）。[3] E. Farhi、J. Goldstone 和 S. Gutmann，《一种量子近似优化算法》，arXiv:1411.4028 (2014). [4] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik 和 J. L. O'brien，《光子量子处理器上的变分本征值求解器》，《自然·通讯》**5**，4213（2014）。[5] A. Kandala、A. Mezzacapo、K. Temme、M. Takita、M. Brink、J. M. Chow 和 J. M. Gambetta，《适用于小分子和量

子磁体的硬件高效变分量子本征求解器》，《自然》**549**，242（2017）。[6] J. R. McClean、J. Romero、R. Babbush 和 A. Aspuru-Guzik，《变分混合量子经典算法理论》，《新物理学杂志》**18**，023023（2016）。[7] P. J. O'Malley、R. Babbush、I. D. Kivlichan、J. Romero、J. R. McClean、R. Barends、J. Kelly、P. Roushan、A. Tranter、N. Ding、等，《分子能量的可扩展量子模拟》，《物理评论X》**6**，031007（2016）。[8] J. I. Colless、V. V. Ramasesh、D. Dahlen、M. S. Blok、M. Kimchi-Schwartz、J. McClean、J. Carter、W. De Jong 和 I. Siddiqi，《利用抗误差算法在量子处理器上计算分子光谱》，《物理评论X》**8**，011021（2018）。[9] C. Bravo-Prieto、J. Lumbreras-Zarapico、L. Tagliacozzo 和 J. I. Latorre，《凝聚态物质系统中变分量子电路深度的标度》，《Quantum》**4**，272（2020）。[10] R. Wiersema、C. Zhou、Y. de Sereville、J. F. Carrasquilla、Y. B. Kim 和 H. Yuen，《在哈密顿量变分拟设中探索纠缠与优化》，PRX Quantum **1**，020319（2020）。[11] M. Lubasch、J. Joo、P. Moinier、M. Kiffner 和 D. Jaksch，《非线性问题的变分量子算法》，Phys. Rev. A **101**，010301（2020）。[12] Y. Li 和 S. C. Benjamin，《融入主动误差最小化的高效变分量子模拟器》，《物理评论X》**7**，021050（2017）。[13] E. F. Dumitrescu、A. J. McCaskey、G. Hagen、G. R. Jansen、T. D. Morris、T. Papenbrock、R. C. Pooser、D. J. Dean 和 P. Lougovski，《原子核的云量子计算》，《物理评论快报》**120**，210501（2018）。[14] S. McArdle、T. Jones、S. Endo、Y. Li、S. C. Benjamin 和 X. Yuan，《基于变分拟设的虚时间演化量子模拟》，《npj量子信息》**5**，75（2019）。[15] M. Schuld、I. Sinayskiy 和 F. Petruccione，《量子机器学习导论》，《当代物理学》**56**，172 2015( )。[16] J. Biamonte、P. Wittek、N. Pancotti、P. Rebentrost、N. Wiebe 和 S. Lloyd，《量子机器学习》，《自然》**549**，195（2017）。[17] V. Dunjko 和 H. J. Briegel，《量子领域中的机器学习与人工智能：近期进展综述》，《物理学进展报告》**81**，074001（2018）。[18] P. Rebentrost、T. R. Bromley、C. Weedbrook 和 S. Lloyd，《量子霍普菲尔德神经网络》，Phys. Rev. A **98**，042308（2018）。[19] N. Killoran、T. R. Bromley、J. M. Arrazola、M. Schuld、N. Quesada 和 S. Lloyd，《连续变量量子神经网络》，《物理评论：研究》**1**，033063（2019）。[20] V. Havlíček、A. D. Córcoles、K. Temme、A. W. Harrow、A. Kandala、J. M. Chow 和 J. M. Gambetta，《利用量子增强特征空间进行监督学习》，《自然》**567**，209（2019）。[21] M. Schuld 和 N. Killoran，《特征希尔伯特空间中的量子机器学习》，《物理评论快报》**122**，040504（2019）。[22] Y. Du、M.-H. Hsieh、T. Liu 和 D. Tao，《参数化量子电路的表达能力》，《物理评论：研究》**2**，033125（2020）。[23] C. Yang、A. Garner、F. Liu、N. Tischler、J. Thompson、M.-H. Yung、M. Gu 和 O. Dahlsten，《机器学习量子模型中可证明的精度优势》，arXiv:2105.14434（2021）。

[24] M. Benedetti、E. Grant、L. Wossnig 和 S. Severini，《用于纯态近似的对抗式量子电路学习》，《新物理学杂志》**21**，043023（2019）。[25] D. Wecker、M. B. Hastings 和 M. Troyer，《实用量子变分算法的进展》，Phys. Rev. A **92**，042303（2015）。[26] M.-C. Chen、M. Gong、X. Xu、X. Yuan、J.-W. Wang、C. Wang、C. Ying、J. Lin、Y. Xu、Y. Wu，等，《利用超导量子协处理器实现绝热变分量子计算》，《物理评论快报》**125**，180501（2020）。[27] A. Patterson、H. Chen、L. Wossnig、S. Severini、D. Browne 和 I. Rungger，《利用含噪量子神经网络进行量子态判别》，《物理评论：研究》**3**，013063（2021）。[28] H. Chen、L. Wossnig、S. Severini、H. Neven 和 M. Mohseni，《通用判别式量子神经网络》，Quantum Mach. Intell. **3**，1（2021）。[29] I. Cong、S. Choi 和 M. D. Lukin，《量子卷积神经网络》，《自然·物理》**15**，1273（2019）。[30] I. MacCormack、C. Delaney、A. Galda、N. Aggarwal 和 P. Narang，《分支量子卷积神经网络》，arXiv:2012.14439（2020）。[31] B. Zhang 和 Q. Zhuang，《变分量子电路中分类误差的快速衰减》，Quantum Sci. Tech-nol.（2022）。[32] Y. Liu、D. Wang、S. Xue、A. Huang、X. Fu、X. Qiang、P. Xu、H.-L. Huang、M. Deng、C. Guo，等，《用于量子态层析成像的变分量子电路》，Phys. Rev. A **101**，052316（2020）。[33] N. Gisin 和 R. Thew，《量子通信》，《自然·光子学》**1**，165（2007）。[34] H. J. Kimble，《量子互联网》，《自然》**453**，1023（2008）。[35] J. Biamonte、M. Faccin 和 M. De Domenico，《从经典到量子的复杂网络》，Commun. Phys. **2**（2019）。[36] S. Wehner、D. Elkouss 和 R. Hanson，《量子互联网：未来发展愿景》，Science **362**（2018）。[37] W. Kozlowski 和 S. Wehner，《迈向大规模量子网络》，载于第六届 *ACM* 纳米尺度计算与通信年度国际会议论文集（2019），第 1–7 页。[38] B. Zhang、J. Wu、L. Fan 和 Q. Zhuang，《机器学习赋能的远程微波量子计算机间混合纠缠分发》，Phys. Rev. Appl. **18**，064016（2022）。[39] B. Q. Baragiola、G. Pantaleoni、R. N. Alexander、A. Karanjai 和 N. C. Menicucci，《利用 Gottesman-Kitaev-Preskill 码实现全高斯通用性和容错》，《物理评论快报》**123**，200502（2019）。[40] M. V. Larsen、C. Chamberland、K. Noh、J. S. Neergaard-Nielsen 和 U. L. Andersen，《容错连续变量测量式量子计算架构》，PRX Quantum **2**，030325（2021）。[41] D. Gottesman、A. Kitaev 和 J. Preskill，《在振荡器中编码量子比特》，Phys. Rev. A **64**，012310（2001）。[42] N. Ofek、A. Petrenko、R. Heeres、P. Reinhold、Z. Leghtas、B. Vlastakis、Y. Liu、L. Frunzio、S. Girvin、L. Jiang，等，《利用超导电路中的纠错延长量子比特寿命》，《自然》**536**，441（2016）。[43] V. Sivak、A. Eickbusch、B. Royer、S. Singh、I. Tsioutsios、S. Ganjam、A. Miano、B. Brock、A. Ding、L. Frunzio，

等，《超越盈亏平衡点的实时量子纠错》（2022），arXiv:quant-ph/2211.09116 [quant-ph]。[44] R. W. Heeres、B. Vlastakis、E. Holland、S. Krastanov、V. V. Albert、L. Frunzio、L. Jiang 和 R. J. Schoelkopf，《利用光子数选择性相位门操控腔态》，《物理评论快报》**115**，137002（2015）。[45] S. Krastanov、V. V. Albert、C. Shen、C.-L. Zou、R. W. Heeres、B. Vlastakis、R. J. Schoelkopf 和 L. Jiang，《利用与量子比特的色散耦合实现振荡器的通用控制》，Phys. Rev. A **92**，040303（2015）。[46] P. Campagne-Ibarcq、A. Eickbusch、S. Touzard、E. Zalys-Geller、N. E. Frattini、V. V. Sivak、P. Reinhold、S. Puri、S. Shankar、R. J. Schoelkopf，等，《对编码在振荡器网格态中的量子比特进行量子纠错》，《自然》**584**，368（2020）。[47] A. Eickbusch、V. Sivak、A. Z. Ding、S. S. Elder、S. R. Jha、J. Venkatraman、B. Royer、S. Girvin、R. J. Schoelkopf 和 M. H. Devoret，《利用与量子比特的弱色散耦合实现振荡器的快速通用控制》，《自然·物理》**18**，1464 1469 2022–（）。[48] Q. Zhuang、Z. Zhang 和 J. H. Shapiro，《利用连续变量多体纠缠进行分布式量子传感》，Phys. Rev. A **97**，032329（2018）。[49] Z. Zhang 和 Q. Zhuang，《分布式量子传感》，Quantum Sci. Technol. **6**，043001（2021）。[50] A. J. Brady、C. Gao、R. Harnik、Z. Liu、Z. Zhang 和 Q. Zhuang，《用于暗物质搜寻的纠缠传感器网络》，PRX Quantum **3**，030333（2022）。[51] Y. Xia、A. R. Agrawal、C. M. Pluchar、A. J. Brady、Z. Liu、Q. Zhuang、D. J. Wilson 和 Z. Zhang，《纠缠增强光机械传感》，Nature Photonics，1（2023）。[52] Q. Zhuang 和 Z. Zhang，《纠缠传感器网络辅助的物理层监督学习》，《物理评论X》**9**，041023（2019）。[53] Y. Xia、W. Li、Q. Zhuang 和 Z. Zhang，《利用变分纠缠传感器网络实现量子增强数据分类》，《物理评论X》**11**，021047（2021）。[54] D. Gross、K. Audenaert 和 J. Eisert，《均匀分布的酉算符：论酉设计的结构》，J. Math. Phys. **48**，052104（2007）。[55] A. Ambainis 和 J. Emerson，《量子 t-设计：量子世界中的 t-路独立性》，载于第二十二届 *IEEE* 计算复杂性年度会议（*CCC'07*）（IEEE，2007），第 129–140 页。[56] D. A. Roberts 和 B. Yoshida，《通过设计实现混沌与复杂度》，J. High Energy Phys. **2017**（4），121。[57] F. G. Brandao、A. W. Harrow 和 M. Horodecki，《局域随机量子电路是近似多项式 t-设计》，Commun. Math. Phys. **346**，397（2016）。[58] J. R. McClean、S. Boixo、V. N. Smelyanskiy、R. Babbush 和 H. Neven，《量子神经网络训练景观中的贫瘠高原》，《自然·通讯》**9**，4812（2018）。[59] M. Cerezo、A. Sone、T. Volkoff、L. Cincio 和 P. J. Coles，《浅层参数化量子电路中依赖代价函数的贫瘠高原》，《自然·通讯》**12**，1791（2021）。[60] S. Wang、E. Fontana、M. Cerezo、K. Sharma、A. Sone、L. Cincio 和 P. J. Coles，《变分量子算法中的噪声诱导贫瘠高原》，《自然·通讯》**12**，6961（2021）。[61] C. Ortiz Marrero、M. Kieferová 和 N. Wiebe，《纠缠诱导的贫瘠高原》，PRX Quantum **2**，040316（2021）。[62] R. Blume-Kohout 和 P. S. Turner，高斯 2-设计何以不存在，Communications in Mathematical Physics **326**, 755 (2014). [63] Q. Zhuang、T. Schuster、B. Yoshida 和 N. Y. Yao，相空间中的搅乱与复杂度，Phys. Rev. A **99**, 062334 (2019). [64] J. T. Iosue、K. Sharma、M. J. Gullans 和 V. V. Albert，连续变量量子态设计：理论与应用，arXiv 预印本 arXiv:2211.05127 (2022). [65] Y. Du、Z. Tu、X. Yuan 和 D. Tao，衡量变分量子算法表达能力的高效方法，Phys. Rev. Lett. **128**, 080506 (2022). [66] M. Larocca、P. Czarnik、K. Sharma、G. Muraleedharan、P. J. Coles 和 M. Cerezo，运用量子最优控制工具诊断贫瘠高原，Quantum **6**, 824 2022( ). [67] B. Zhang、A. Sone 和 Q. Zhuang，组合问题中的量子计算相变，npj Quantum Information **8**, 87 (2022). [68] C. Weedbrook、S. Pirandola、R. García-Patrón、N. J. Cerf、T. C. Ralph、J. H. Shapiro 和 S. Lloyd，高斯量子信息，Rev. Mod. Phys. **84**, 621 (2012). [69] M. Gu、C. Weedbrook、N. C. Menicucci、T. C. Ralph 和 P. van Loock，利用连续变量簇态进行量子计算，Physical Review A **79**, 062318 (2009). [70] A. A. Diringer、E. Blumenthal、A. Grinberg、L. Jiang 和 S. Hacohen-Gourgy，条件非位移：利用单个量子比特实现快速多振子控制，arXiv 预印本 arXiv:2301.09831 (2023). [71] K. Mitarai、M. Negoro、M. Kitagawa 和 K. Fujii，量子线路学习，Phys. Rev. A **98**, 032309 (2018). [72] M. Tse、H. Yu、N. Kijbunchoo、A. Fernandez-Galiana、P. Dupej、L. Barsotti、C. Blair、D. Brown、S. Dwyer、A. Effler、等，引力波天文学时代的量子增强型先进 LIGO 探测器，Phys. Rev. Lett. **123**, 231107 (2019). [73] J. Aasi、J. Abadie、B. Abbott、R. Abbott、T. Abbott、M. Abernathy、C. Adams、T. Adams、P. Addesso、R. Adhikari、等，利用光压缩态提升 LIGO 引力波探测器的灵敏度，Nat. Photonics **7**, 613 (2013). [74] J. Abadie、B. P. Abbott、R. Abbott、T. D. Abbott、M. Abernathy、C. Adams、R. Adhikari、C. Affeldt、B. Allen、G. Allen、等，突破量子散粒噪声极限运行的引力波天文台，Nat. Phys. **7**, 962 (2011). [75] K. Backes、D. Palken、S. Al Kenany、B. Brubaker、S. Cahn、A. Droster、G. C. Hilton、S. Ghosh、H. Jackson、S. Lamoreaux、等，量子增强型暗物质轴子搜索，Nature **590**, 238 (2021). [76] J. T. Iosue、A. Ehrenberg、D. Hangleiter、A. Deshpande 和 A. V. Gorshkov，线性光学中的佩奇曲线与典型纠缠，arXiv:2209.06838 (2022). [77] B. Zhang 和 Q. Zhuang，连续变量随机量子网络中的纠缠生成，npj Quantum Inf. **7**, 33 (2021). [78] E. Grant、L. Wossnig、M. Ostaszewski 和 M. Benedetti，一种用于应对参数化量子电路中贫瘠高原的初始化策略，Quantum **3**, 214 (2019). [79] T. Volkoff 和 P. J. Coles，利用随机参数化量子电路中的关联获得大梯度，Quantum Sci. Technol. **6**, 025008 (2021). [80] B. T. Kiani , G. De Palma , M. Marvian , Z.-W. Liu ,

和 S. Lloyd, 《使用量子地球搬运距离学习量子数据》, Quantum Sci. Technol. **7**, 045002 (2022). [81] S. H. Sack, R. A. Medina, A. A. Michailidis, R. Kueng 和 M. Serbyn, 《利用经典阴影避免贫瘠高原》, PRX Quantum **3**, 020365 (2022). [82] S. Cichy, P. K. Faehrmann, S. Khatri 和 J. Eisert, 《用于延缓变分量子算法中贫瘠高原出现的微扰小工具》, arXiv 预印本 arXiv:2210.03099 (2022). [83] X. Liu, G. Liu, J. Huang 和 X. Wang, 《缓解变分量子本征求解器中的贫瘠高原》, arXiv 预印本 arXiv:2205.13539 (2022). [84] C. Mehta, 《量子算符的对角相干态表示》, 《物理评论快报》**18**, 752 (1967). [85] A. Vourdas, 《量子力学中的解析表示》, J. Phys. A Math. Gen. **39**, R65 (2006). [86] E. R. Anschuetz 和 B. T. Kiani, 《变分量子算法深陷各种陷阱》, 《自然》通讯 **13**, 7760 (2022). [87] P. Marian 和 T. A. Marian, 《双模高斯态之间的 Uhlmann保真度》, Phys. Rev. A **86**, 022340 (2012). [88] G. Spedalieri, C. Weedbrook 和 S. Pirandola, 《量子保真度的极限公式》, J. Phys. A Math. **46**, 025304 2012( ). [89] L. Banchi, S. L. Braunstein 和 S. Pirandola, 《任意高斯态的量子保真度》, 《物理评论快报》**115**, 260501 (2015). [90] S. L. Braunstein 和 P. Van Loock, 《连续变量量子信息》, Rev. Mod. Phys. **77**, 513 (2005). [91] D. d'Alessandro, 量子控制与动力学导论 (Chapman and Hall/CRC, 2021)。

目录

I. 引言 1 II. 结果 1 A. 能量正则化电路系综 2 B. 态制备中的贫瘠高原 3 1. 单模态制备 4 2. 多模态制备 6 C. 规避训练问题的策略 7 III. 讨论 8 致谢 8 参考文献 8 A. 高斯态 11 B. ECD 门集的普适性：引理 1 的证明 12 C. 态的表示：单模情形 13 D. 梯度评估方法 16 1. 预备知识 17

2. 权重四重乘积的系综平均 1

8

E. 单模连续变量态制备中的梯度方差 191. 单模高斯态 222. 福克数态 24

F. 制备多模量子模式连续变量态时梯度的方差 26 1. 乘积态 29 2. 多模高斯态 30

G. 相干态和单模压缩真空态的结果 33

附录 A：高斯态

这里我们简要介绍高斯态。更多详细信息请参见参考文献 [68], 本文采用了其中的约定。

由 M 个模式组成的系统由 M 湮灭算符和产生算符 $\{m_j, m_j^\dagger\}_{j=1}^M$描述，它们满足对易关系 $[m_j, m_{j'}^\dagger] = \delta_{j,\ j'}$。也可以通过相空间中的位置算符和动量算符 $q_j = m_j + m_j^\dagger$ 和 $p_j = i(m_j^\dagger - m_j)$来描述该系统。这些正交分量可以组合成正交分量向量 $\mathcal{X} = (q_1,\ p_1,\ \ldots\ldots,\ q_M,\ p_M)^T$。一阶和二阶矩，也称为平均正交分量和协方差矩阵（CM），为

$$\overline{\mathcal{X}} = \langle \mathcal{X} \rangle \tag{A1}$$

$$\boldsymbol{V}_{ij} = \frac{1}{2} \langle \{ q_i - \langle q_i \rangle, q_j - \langle q_j \rangle \} \rangle, \tag{A2}$$

其中，$\langle \cdot \rangle$ 表示期望值，$\{A, B\}$ 是算符 $A,\ B$的反对易子。任意单模式纯高斯态都可以表示为经过旋转和位移的压缩态

$$|\psi\rangle_m = D(\gamma) R(\tau) S(\zeta) |0\rangle_m, \tag{A3}$$

其中，$R(\tau) = \exp(-i\tau m^\dagger m)$ 是相位旋转，而 $S(\zeta) = \exp\left[\zeta(m^2 - m^{\dagger 2})/2\right]$ 是压缩算符。位移算符 $D(\beta) = \exp\left(\beta m^\dagger - \beta^* m\right)$ 满足编织关系$D(\alpha)D(\beta) = e^{(\alpha\beta^\star - \alpha^\star\beta)/2} D(\alpha+\beta)$。其平均正交分量和协方差矩阵为

$$\overline{\mathcal{X}} = (2\,\mathrm{Re}\{\gamma\}, 2\,\mathrm{Im}\{\gamma\})^T \tag{A4}$$

$$V = \begin{pmatrix} e^{2\zeta}\sin^2(\tau) + e^{-2\zeta}\cos^2(\tau) & \sin(2\tau)\sinh(2\zeta) \\ \sin(2\tau)\sinh(2\zeta) & e^{2\zeta}\cos^2(\tau) + e^{-2\zeta}\sin^2(\tau) \end{pmatrix}, \tag{A5}$$

下面我们给出一些示例中的 CM 和均值；相干态具有 $\tau = \zeta = 0$。因此，CM 简化为 $\boldsymbol{I}_2$单位矩阵，而均值可以是任意值；单模压缩真空态（SMSV）

态具有 $\gamma=\tau=0$。因此，其均值为零 $\mathcal{X}=0$，且 CM 为对角矩阵 $V=\mathrm{diag}(e^{-2\zeta},\ e^{2\zeta})$。

双模压缩真空态（TMSV）是两个模式的最大纠缠高斯态，可通过对真空态的乘积态 $|0\rangle_{m_1}|0\rangle_{m_2}$施加双模压缩算符 $S_2(\zeta)=\exp\left[\zeta(m_1m_2-m_1^\dagger m_2^\dagger)/2\right]$ 生成。它的平均值也为零，其协方差矩阵变为

$$\boldsymbol{V}_{\mathrm{TMSV}}=\begin{pmatrix}\cosh(2\zeta)\boldsymbol{I}_2 & \sinh(2\zeta)\sigma^z\\ \sinh(2\zeta)\sigma^z & \cosh(2\zeta)\boldsymbol{I}_2\end{pmatrix},\tag{A6}$$

其中 $\boldsymbol{I}_2$ 是 2×2 恒等算符，σ^z 是泡利 Z 算符。

两个高斯量子态之间的保真度ρ_A，ρ_B 完全由平均正交分量决定$\mathcal{X}_A$，$\mathcal{X}_B$以及协方差矩阵 V_A，V_B。此外，它可以通过解析方法求解 [87–89]。当其中一个态为纯态时，其形式可简化为 [88]

$$F(\rho_A,\rho_B)=\left(\mathrm{Tr}\left\{\sqrt{\sqrt{\rho_A}\rho_B\sqrt{\rho_A}}\right\}\right)^2\tag{A7}$$

$$=\frac{2^M}{\sqrt{\det(V_A+V_B)}}\exp\left[-\frac{1}{2}\boldsymbol{d}^T(V_A+V_B)^{-1}\boldsymbol{d}\right],\tag{A8}$$

其中 $\boldsymbol{d}\equiv\overline{\mathcal{X}}_A-\overline{\mathcal{X}}_B$。

附录 B：ECD 门集的普适性：引理 1 的证明

对于正文中涉及量子比特与量子模式相互作用的连续变量变分量子电路，文献 [47] 已证明，ECD门和单量子比特旋转构成的门集是通用的，也就是说，门生成元的重复嵌套对易子的线性组合涵盖了量子比特-量子模式系统 [90, 91]的整个李代数。下面我们回顾这一证明。ECD门和单量子比特旋转的生成元为

$$\mathcal{G}=\{q\sigma^z,p\sigma^z,\sigma^x,\sigma^y\},\tag{B1}$$

其中，q,p 分别是量子模式的位置和动量，σ^c和 $c\in\{x,y,z\}$则是量子比特上的泡利算符。首先，$q\sigma^z,p\sigma^z$ 与 σ^x,σ^y之间的对易子会产生形式为 $q\sigma^c,p\sigma^c$ 且 $c\in\{x,y,z\}$ 的算符。为了获得 q,p的高阶多项式算符，可以考虑对易子 $[q\sigma^a,q\sigma^b]\propto\epsilon_{abc}q^2\sigma^c$，其中 ϵ_{abc} 是三维列维-奇维塔符号。重复进行对易运算可以生成满足 $j\geq2$ 的算符$q^j\sigma^a$，对于 $p^j\sigma^a$也可类似处理。为了获得 the

将正交分量耦合在一起的项以及泡利算符，我们考虑对易子

$$[q^{j+1}\sigma^a,p\sigma^b]=2i\epsilon_{abc}q^{j+1}p\sigma^c+(j+1)\epsilon_{abc}q^j\sigma^c,\tag{B2}$$

假设 $a\neq b$。结合 $q^j\sigma^c$，可得 $p^{j+1}\sigma^c$。通过对 $p\sigma^b$反复取对易子，可得到所有多项式正交分量项 $q^jp^k\sigma^c$，其中 $c\in\{x,y,z\}$。最后一步是通过对易子 $[q^{j+1}p^k\sigma^a,p\sigma^a]\propto q^jp^k$消去泡利算符。因此，利用包含ECD门和单量子比特旋转的门集，可以实现生成元形式为 $q^jp^k\sigma^c$ 、且满足 $\sigma^c\in\{\boldsymbol{I}_2,\sigma^x,\sigma^y,\sigma^z\}$ 的所有酉算符。

现在，我们将 ECD 门和单量子比特幺正变换的普适性推广到任意数量的 $M\geq1$ 量子模式和$N\geq1$量子比特。ECD 门（任意量子比特—量子模式对之间）的生成元与单量子比特旋转结合后为

$$\mathcal{G}=\left[\bigcup_{\ell=1}^{M}\bigcup_{r=1}^{N}\{q_\ell\sigma_r^z,p_\ell\sigma_r^z\}\right]\bigcup\left[\bigcup_{r=1}^{N}\{\sigma_r^x,\sigma_r^y\}\right].\tag{B3}$$

对于通过 ECD 门与量子比特 ℓ,ℓ' 耦合的两个模式，$q_\ell^{j_\ell}p_\ell^{k_\ell}\sigma_r^a$ 与 $q_{\ell'}^{j_{\ell'}}p_{\ell'}^{k_{\ell'}}\sigma_r^b$之间的对易子（假设 $\ell\neq\ell'$）为

$$\begin{aligned}\left[q_\ell^{j_\ell}p_\ell^{k_\ell}\sigma_r^a,q_{\ell'}^{j_{\ell'}}p_{\ell'}^{k_{\ell'}}\sigma_r^b\right]&=q_\ell^{j_\ell}p_\ell^{k_\ell}q_{\ell'}^{j_{\ell'}}p_{\ell'}^{k_{\ell'}}\left[\sigma_r^a,\sigma_r^b\right]\\&\propto q_\ell^{j_\ell}p_\ell^{k_\ell}q_{\ell'}^{j_{\ell'}}p_{\ell'}^{k_{\ell'}}\epsilon_{abc}\sigma_r^c.\end{aligned}\tag{B4}$$

最后，对易子 $\left[q_\ell^{j_\ell}p_\ell^{k_\ell}q_{\ell'}^{j_{\ell'}+1}p_{\ell'}^{k_{\ell'}}\sigma_r^x,p_{\ell'}\sigma_r^x\right]\propto q_\ell^{j_\ell}p_\ell^{k_\ell}q_{\ell'}^{j_{\ell'}}p_{\ell'}^{k'_{\ell'}}$可以生成形如$q_\ell^{j_\ell}p_\ell^{k_\ell}q_{\ell'}^{j_{\ell'}}p_{\ell'}^{k_{\ell'}}\sigma_r^c$且满足 $\sigma_r^c\in\{\boldsymbol{I}_2,\sigma_r^x,\sigma_r^y,\sigma_r^z\}$的任意算符。对其他$M-2$ 模式重复上述过程，即可生成由 $\left(\prod_{\ell=1}^{M}q_\ell^{j_\ell}p_\ell^{k_\ell}\right)\sigma_r^c$作为生成元、对 M 个模式和一个量子比特具有通用控制的任意酉变换。接下来，我们考虑连接到一个模式 ℓ的两个量子比特r,r'，$q_\ell^{j_\ell}p_\ell^{k_\ell+1}\sigma_r^z$ 与 $q_\ell\sigma_{r'}^z$ 之间的对易子为

$$\begin{aligned}\left[q_\ell^{j_\ell}p_\ell^{k_\ell+1}\sigma_r^z,q_\ell\sigma_{r'}^z\right]&=q_\ell^{j_\ell}\left[p_\ell^{k_\ell+1},q_\ell\right]\sigma_r^z\sigma_{r'}^z\\&\propto q_\ell^{j_\ell}p_\ell^{k_\ell}\sigma_r^z\sigma_{r'}^z.\end{aligned}\tag{B5}$$

通过 $q_\ell^{j_\ell}p_\ell^{k_\ell}\sigma_r^z\sigma_{r'}^z$ 与 $\{\sigma_r^x,\sigma_{r'}^x,\sigma_r^y,\sigma_{r'}^y\}$ 之间的对易子，我们得到 $q_\ell^{j_\ell}p_\ell^{k_\ell}\sigma_r^c\sigma_{r'}^{c'}$ 与 σ_r^c，$\sigma_{r'}^{c'}\in\{\boldsymbol{I}_2,\sigma_r^x,\sigma_r^y,\sigma_r^z\}$ 的形式。重复上述讨论的过程，我们可以生成所有生成元具有 $q_\ell^{j_\ell}p_\ell^{k_\ell}$（$\prod_r\sigma_r^c$）形式的酉算符。结合 M 模式和一个量子比特的结果，我们最终得到 $\left(\prod_{\ell=1}^{M}q_\ell^{j_\ell}p_\ell^{k_\ell}\right)\left(\prod_{r=1}^{N}\sigma_r^c\right)$ 形式的生成元，其中证明了可在具有 M 模式和 n 个量子比特的系统上实现通用控制。

附录 C：态的表示：单模情形

本节将证明公式（3）中单模能量正则化态系综的表示形式。多模情形的推广见附录 F。我们首先证明该表示对所有 L 层线路均成立，然后分析能量正则化。

为简化记号，我们定义 $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_L)^T$，$\boldsymbol{\theta} = (\theta_1, \ldots, \theta_L)^T$，$\boldsymbol{\phi} = (\phi_1, \ldots, \phi_L)^T$ 及总体而言参数 $\boldsymbol{x} = (\boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{\phi})$。考虑一个由 ECD模块组成的单模量子比特-量子模式变分电路，在量子比特的 $|0\rangle$ 和 $|1\rangle$ 基底下，每个 ECD模块都可以写成矩阵形式，即

$$U_{\mathrm{ECD}}(\beta)U_{\mathrm{R}}(\theta,\phi) = \begin{pmatrix} e^{i\phi}\sin\frac{\theta}{2}D(-\beta) & \cos\frac{\theta}{2}D(-\beta) \\ \cos\frac{\theta}{2}D(\beta) & e^{i(\pi-\phi)}\sin\frac{\theta}{2}D(\beta) \end{pmatrix}, \tag{C1}$$

为方便起见，我们将变量重新标记为 $\phi - \pi/2 \to \phi$，并在下文中采用此定义。酉算符 $U = \prod_{\ell=1}^{L} U_{\mathrm{ECD}}(\beta_\ell)U_{\mathrm{R}}(\theta_\ell, \phi_\ell)$ 作用于输入态 $|0\rangle_q |0\rangle_m$ 所得的输出态是

$$\begin{aligned} |\psi(\boldsymbol{\beta},\boldsymbol{\theta},\boldsymbol{\phi})\rangle_{q,m} &= U(\boldsymbol{\beta},\boldsymbol{\theta},\boldsymbol{\phi})\,|0\rangle_q\,|0\rangle_m \\ &= \sum_{a=0}^{1}\sum_{\boldsymbol{s}} w_{\boldsymbol{s},a}(\boldsymbol{\theta},\boldsymbol{\phi})\,|a\rangle_q\, e^{i\chi_{\boldsymbol{s}}(\boldsymbol{\beta})}\,|(-1)^a\boldsymbol{s}\cdot\boldsymbol{\beta}\rangle_m \equiv \sum_{a=0}^{1}\sum_{\boldsymbol{s}} w_{\boldsymbol{s},a}(\boldsymbol{\theta},\boldsymbol{\phi})\,|a\rangle_q\,|B_{\boldsymbol{s},a}\rangle_m \end{aligned} \tag{C2}$$

其中，长度为 L 的符号向量 $\boldsymbol{s}$ 定义为 $\boldsymbol{s} = (\boldsymbol{s}_{1:L-1}, -1)$，其中 $\boldsymbol{s}_{1:L-1} \in \{-1,1\}^{L-1}$。为方便起见，我们将额外的相位吸收到量子模右矢向量中，表示为 $|B_{\boldsymbol{s},a}\rangle \equiv e^{i\chi_{\boldsymbol{s}}(\boldsymbol{\beta})}|(-1)^a\boldsymbol{s}\cdot\boldsymbol{\beta}\rangle_m$，其中相位定义为

$$\chi_{\boldsymbol{s}}(\boldsymbol{\beta}) \equiv \sum_{\ell=1}^{L-1}\sum_{\ell'=\ell+1}^{L} \boldsymbol{s}_\ell \boldsymbol{s}_{\ell'}(\mathrm{Re}\{\beta_\ell\}\,\mathrm{Im}\{\beta_{\ell'}\} - \mathrm{Im}\{\beta_\ell\}\,\mathrm{Re}\{\beta_{\ell'}\}), \tag{C3}$$

通过定义 $v_{\boldsymbol{s},a}(\boldsymbol{\theta},\boldsymbol{\phi},\boldsymbol{\beta}) \equiv e^{i\chi_{\boldsymbol{s}}(\boldsymbol{\beta})} w_{\boldsymbol{s},a}(\boldsymbol{\theta},\boldsymbol{\phi})$，公式 (C2) 等价于主论文中的公式（3）。在公式 (C2) 中，$|a\rangle_q$ 是计算基中的量子比特态。每个叠加态的权重为

$$w_{\boldsymbol{s},a}(\boldsymbol{\theta},\boldsymbol{\phi}) = e^{i\Phi_{\boldsymbol{s},a}(\boldsymbol{\phi})}T_{\boldsymbol{s},a}(\boldsymbol{\theta}), \tag{C4}$$

其中

$$\Phi_{\boldsymbol{s},a}(\boldsymbol{\phi}) \equiv a(n_{\boldsymbol{s}}\pi + \phi_1) + \mathbb{P}(a)\sum_{\ell=1}^{L}(\mathbf{d}\boldsymbol{s}_\ell - 1)\,(\boldsymbol{s}_\ell\phi_\ell - \delta_{\boldsymbol{s}_\ell,1}\pi)\,, \tag{C5}$$

$$T_{\boldsymbol{s},a}(\boldsymbol{\theta}) \equiv \prod_{\ell=2}^{L}\sin\left(\frac{\theta_\ell + \mathbf{d}\boldsymbol{s}_\ell\pi}{2}\right)\sin\left(\frac{\theta_1}{2} + \frac{(\mathbb{P}(a)\boldsymbol{s}_1 + 1)\pi}{4}\right). \tag{C6}$$

有一些符号需要说明。我们将 $\mathbb{P}(x) \equiv (-1)^x$ 定义为变量 x 的奇偶性，当 x 为偶数/奇数时，其值为 ±1。我们定义差分符号向量 $\mathbf{d}\boldsymbol{s}$，用于表示向量 $\boldsymbol{s}$ 中符号的变化——差分符号向量的第 ℓ 个元素是

$$\mathbf{d}\boldsymbol{s}_\ell \equiv |\boldsymbol{s}_\ell - \boldsymbol{s}_{\ell-1}|/2, \tag{C7}$$

当 $\boldsymbol{s}_{\ell+1} = \boldsymbol{s}_\ell$ 时其值为零，否则为 1。注意，$(-1)^{\mathbf{d}\boldsymbol{s}_\ell} = \boldsymbol{s}_\ell \times \boldsymbol{s}_{\ell-1}$。由于 $\boldsymbol{s}_L = -1$ 始终成立，我们赋予 $\mathbf{d}\boldsymbol{s}_1 = |\boldsymbol{s}_1 - \boldsymbol{s}_L|/2 = (\boldsymbol{s}_1 + 1)/2$，使其反映 $\boldsymbol{s}_1$ 的值。

我们可以发现，$\{\mathbf{d}\boldsymbol{s}_\ell | \mathbf{d}\boldsymbol{s}_\ell = 0\}_{\ell\geq2}$ 的大小为

$$n_{\boldsymbol{s}} = L - 1 - \sum_{\ell=2}^{L}\mathbf{d}\boldsymbol{s}_\ell \tag{C8}$$

由于 $\mathbf{d}\boldsymbol{s}_\ell = 0, 1$。我们想指出，$n_{\boldsymbol{s}}$ 也等于 $\Phi_{\boldsymbol{s},a}$ 中系数非零的 $\{\phi_\ell\}_{\ell\geq2}$ 的数量，以及 $T_{\boldsymbol{s},a}$ 中出现的 $\{\sin(\theta_\ell/2)\}_{\ell\geq2}$ 的数量。由式 (C5) 可知，$\Phi_{\boldsymbol{s},a}$ 中每个 ϕ_ℓ 在 $\ell \geq 2$ 下的系数为 $\mathbb{P}(a)(\mathbf{d}\boldsymbol{s}_\ell - 1)\boldsymbol{s}_\ell$。由于 $\mathbb{P}(a) = (-1)^a$ 且 $\boldsymbol{s}_\ell = \pm1$，只有 $\mathbf{d}\boldsymbol{s}_\ell = 1$ 能使系数为零，因此 $n_{\boldsymbol{s}}$ 统计的是系数非零的 $\{\phi_\ell\}_{\ell\geq2}$ 的数量。此外，由式 (C6) 还可知，若 $\mathbf{d}\boldsymbol{s}_\ell = 0$ 满足 $\ell \geq 2$，则 $T_{\boldsymbol{s},a}$ 中 θ_ℓ 的函数形式为 $\sin(\theta_\ell/2)$；而若 $\mathbf{d}\boldsymbol{s}_\ell = 1$，其形式则为 $\cos(\theta_\ell/2)$。为了统计

在 $\Phi_{\boldsymbol{s},a}$ 中系数非零的 $\{\phi_\ell\}_{\ell=1}^{L}$ 的总数，记为 $N_{\boldsymbol{s},a}$。由式 (C5) 可知， ϕ_1 的系数为$a+\mathbb{P}(a)(\mathbf{d}\boldsymbol{s}_1-1)\boldsymbol{s}_1 = a+(-1)^a(\mathbf{d}\boldsymbol{s}_1-1)(2\mathbf{d}\boldsymbol{s}_1-1) = 1-a-(-1)^a\mathbf{d}\boldsymbol{s}_1 \geq 0$。其中，第一个等式利用了定义 $\boldsymbol{s}_1 = 2\mathbf{d}\boldsymbol{s}_1 - 1$，第二个等式利用了 $(\mathbf{d}\boldsymbol{s}_1)^2 = \mathbf{d}\boldsymbol{s}_1$ 和 $a+(-1)^a = 1-a$。由于 ϕ_1 的系数只能为零或一，因此总数就是

$$N_{\boldsymbol{s},a} = n_{\boldsymbol{s}} + 1 - a - (-1)^a\mathbf{d}\boldsymbol{s}_1. \tag{C9}$$

注意，这也就是 $T_{\boldsymbol{s},a}$ 中 $\{\sin(\theta_\ell/2)\}_{\ell=1}^{L}$ 的总数，因为只有当 $1-(\mathbb{P}(a)\boldsymbol{s}_1+1)/2 = 1$ 时，$\sin(\theta_1/2)$ 才会出现在 $T_{\boldsymbol{s},a}$ 中。利用恒等式 $\boldsymbol{s}_1 = 2\mathbf{d}\boldsymbol{s}_1 - 1$ 和 $1+(-1)^a = 2a$，我们得到 $1-(-1)^a\mathbf{d}\boldsymbol{s}_1 + a$，这正好就是 $N_{\boldsymbol{s},a}$ 中 $n_{\boldsymbol{s}}$ 后面的各项，即我们上面给出的结果。

为帮助理解这些记号，我们提供如下 $L=2$ 示例

$$\begin{aligned}
&|\psi(\boldsymbol{\beta},\boldsymbol{\theta},\boldsymbol{\phi})\rangle_{q,m} \\
&= |0\rangle_q \left(e^{i(\phi_1+\phi_2+\beta_1^{\mathrm{R}}\beta_2^{\mathrm{I}}-\beta_1^{\mathrm{I}}\beta_2^{\mathrm{R}})} \sin\frac{\theta_1}{2}\sin\frac{\theta_2}{2} |-\beta_1-\beta_2\rangle_m + e^{i(\beta_1^{\mathrm{I}}\beta_2^{\mathrm{R}}-\beta_1^{\mathrm{R}}\beta_2^{\mathrm{I}})} \cos\frac{\theta_1}{2}\cos\frac{\theta_2}{2} |+\beta_1-\beta_2\rangle_m \right) \\
&+ |1\rangle_q \left(e^{i(\pi-\phi_2+\beta_1^{\mathrm{R}}\beta_2^{\mathrm{I}}-\beta_1^{\mathrm{I}}\beta_2^{\mathrm{R}})} \cos\frac{\theta_1}{2}\sin\frac{\theta_2}{2} |+\beta_1+\beta_2\rangle_m + e^{i(\phi_1-\beta_1^{\mathrm{R}}\beta_2^{\mathrm{I}}+\beta_1^{\mathrm{I}}\beta_2^{\mathrm{R}})} \sin\frac{\theta_1}{2}\cos\frac{\theta_2}{2} |-\beta_1+\beta_2\rangle_m \right).
\end{aligned} \tag{C10}$$

其中， $\beta_{1\,2}^{\mathrm{R/I}}$, 表示 $\beta_{1,2}$ 的实部或虚部，以缩短公式。根据公式（C4）中的定义，可以验证其与公式 (C2) 中的态表示一致。

下面我们通过数学归纳法，详细证明公式 (C2)（等价于正文中的公式（3））所表示的态。

证明。首先，我们从 $L=2$ 出发：在这种情况下，公式（C10）已经证明公式 (C2) 成立。然后我们假设，对于一个 L 层线路，它同样具有公式 (C2) 的形式；而对于 $L+1$ 层线路，我们有

$$\begin{aligned}
&|\psi_{L+1}(\boldsymbol{\beta},\boldsymbol{\theta},\boldsymbol{\phi})\rangle = U_{\mathrm{ECD}}(\beta_{L+1})U_{\mathrm{R}}(\theta_{L+1},\phi_{L+1})|\psi_L(\boldsymbol{\beta},\boldsymbol{\theta},\boldsymbol{\phi})\rangle \\
&= \begin{pmatrix} e^{i\phi_{L+1}}\sin\left(\frac{\theta_{L+1}}{2}\right)D(-\beta_{L+1}) & \cos\left(\frac{\theta_{L+1}}{2}\right)D(-\beta_{L+1}) \\ \cos\left(\frac{\theta_{L+1}}{2}\right)D(\beta_{L+1}) & e^{i(\pi-\phi_{L+1})}\sin\left(\frac{\theta_{L+1}}{2}\right)D(\beta_{L+1}) \end{pmatrix} \begin{pmatrix} \sum_{\boldsymbol{s}} w_{\boldsymbol{s},0}(\boldsymbol{\theta},\boldsymbol{\phi})\prod_{\ell=1}^{L}D(\boldsymbol{s}_\ell\beta_\ell)|0\rangle_m \\ \sum_{\boldsymbol{s}} w_{\boldsymbol{s},1}(\boldsymbol{\theta},\boldsymbol{\phi})\prod_{\ell=1}^{L}D(-\boldsymbol{s}_\ell\beta_\ell)|0\rangle_m \end{pmatrix} \\
&= \begin{pmatrix} \sum_{\boldsymbol{s}} e^{i(\Phi_{\boldsymbol{s},0}+\phi_{L+1})}\sin\left(\frac{\theta_{L+1}}{2}\right)T_{\boldsymbol{s},0}D(-\beta_{L+1})\prod_{\ell=1}^{L}D(\boldsymbol{s}_\ell\beta_\ell)|0\rangle_m + \sum_{\boldsymbol{s}} e^{i\Phi_{\boldsymbol{s},1}}\cos\left(\frac{\theta_{L+1}}{2}\right)T_{\boldsymbol{s},1}D(-\beta_{L+1})\prod_{\ell=1}^{L}D(-\boldsymbol{s}_\ell\beta_\ell)|0\rangle_m \\ \sum_{\boldsymbol{s}} e^{i\Phi_{\boldsymbol{s},0}}\cos\left(\frac{\theta_{L+1}}{2}\right)T_{\boldsymbol{s},0}D(\beta_{L+1})\prod_{\ell=1}^{L}D(\boldsymbol{s}_\ell\beta_\ell)|0\rangle_m + \sum_{\boldsymbol{s}} e^{i(\Phi_{\boldsymbol{s},1}+\pi-\phi_{L+1})}\sin\left(\frac{\theta_{L+1}}{2}\right)T_{\boldsymbol{s},1}D(\beta_{L+1})\prod_{\ell=1}^{L}D(-\boldsymbol{s}_\ell\beta_\ell)|0\rangle_m \end{pmatrix}.
\end{aligned} \tag{C11}$$

注意，上述位移算符的乘积是从 β_1 到 β_L进行计算的。为了与公式 (C2) 中的表示进行比较，我们定义 $\boldsymbol{s}'$ 为 $(\pm\boldsymbol{s},-1)$，这实际上涵盖了长度为 $L+1$ 的符号向量定义中的所有可能情况。为了证明这一结果，我们只需证明公式（C11）与公式 (C2) 中的表示一致即可。

我们从量子模式上的位移开始。公式（C11）中的量子模式总位移可直接看出满足 $(-1)^a\boldsymbol{s}'\cdot\boldsymbol{\beta}$，其中 $\boldsymbol{s} = (\pm\boldsymbol{s},-1)$。由于位移算符的编织关系 $D(\alpha) = e^{\alpha m^\dagger - \alpha^* m} = e^{-|\alpha|^2/2}e^{\alpha m^\dagger}e^{-\alpha^* m}$，会产生一个相位；我们可以直接证明它具有公式（C3）的形式，具体如下

$$\begin{aligned}
&D(\alpha_L)\cdots D(\alpha_1) = e^{-\frac{1}{2}\sum_{\ell=1}^{L}|\alpha_\ell|^2}e^{\alpha_L m^\dagger}e^{-\alpha_L^* m}\cdots e^{\alpha_1 m^\dagger}e^{-\alpha_1^* m} \\
&= e^{-\frac{1}{2}\sum_{\ell=1}^{L}|\alpha_\ell|^2}\left(\prod_{\ell=1}^{M}e^{\alpha_\ell m^\dagger}\right)\left(\prod_{\ell=1}^{L}e^{-\alpha_\ell^* m}\right)\left(\prod_{\ell=1}^{L-1}e^{-\alpha_\ell(\sum_{\ell'=\ell+1}^{L}\alpha_{\ell'}^*)}\right) && \text{(C12)} \\
&= e^{-\frac{1}{2}\sum_{\ell=1}^{L}|\alpha_\ell|^2}e^{\sum_{\ell=1}^{L}\alpha_\ell m^\dagger}e^{-\sum_{\ell=1}^{L}\alpha_\ell^* m}e^{-\sum_{\ell=1}^{L-1}\sum_{\ell'=\ell+1}^{L}\alpha_\ell\alpha_{\ell'}^*} && \text{(C13)} \\
&= e^{-\frac{1}{2}\sum_{\ell=1}^{L}|\alpha_\ell|^2}e^{-\sum_{\ell=1}^{L-1}\sum_{\ell'=\ell+1}^{L}\alpha_\ell\alpha_{\ell'}^*}e^{\frac{1}{2}|\sum_{\ell=1}^{L}\alpha_\ell|^2}D\left(\sum_{\ell=1}^{L}\alpha_\ell\right) && \text{(C14)} \\
&= e^{i\sum_{\ell=1}^{L-1}\sum_{\ell'=\ell+1}^{L}(\mathrm{Re}\{\alpha_\ell\}\,\mathrm{Im}\{\alpha_{\ell'}\}-\mathrm{Im}\{\alpha_\ell\}\,\mathrm{Re}\{\alpha_{\ell'}\})}D\left(\sum_{\ell=1}^{L}\alpha_\ell\right), && \text{(C15)}
\end{aligned}$$

其中，第二行中我们对量子模算符重新排序，使所有湮灭算符都位于产生算符之后，并根据 Baker–Campbell–Hausdorff 恒等式引入额外的相位 $e^Ae^B = e^Be^Ae^{[A,B]}$（高阶对易子为零时）。最后一行通过展开每个 $\alpha_\ell = \mathrm{Re}\{\alpha_\ell\} + i\mathrm{Im}\{\alpha_\ell\}$ 得出。令$\alpha_\ell = (-1)^a\boldsymbol{s}_\ell\beta_\ell$，即可得到公式（C3）中的公式。

接下来，我们需要证明权重满足公式（C4）的形式。从公式（C11）第一行的第一项出发，该项对应于 $\boldsymbol{s}'=(\boldsymbol{s},-1)$ 和 $a=0$，符号向量之差为 $\mathbf{d}\boldsymbol{s}'=(\mathbf{d}\boldsymbol{s},\ 0)$，于是权重中的相位和振幅可化简为

$$\Phi_{\boldsymbol{s},0}+\phi_{L+1}=\sum_{\ell=1}^{L}(\mathbf{d}\boldsymbol{s}_\ell-1)\boldsymbol{s}_\ell\phi_\ell-(\mathbf{d}\boldsymbol{s}_\ell-1)\delta_{\boldsymbol{s}_\ell,1}\pi+\phi_{L+1}=\sum_{\ell=1}^{L+1}(\mathbf{d}\boldsymbol{s}'_\ell-1)\boldsymbol{s}'_\ell\phi_\ell-(\mathbf{d}\boldsymbol{s}'_\ell-1)\delta_{\boldsymbol{s}'_\ell,1}\pi=\Phi_{\boldsymbol{s}',0},\quad\text{(C16)}$$

$$\begin{aligned}\sin\left(\frac{\theta_{L+1}}{2}\right)T_{\boldsymbol{s},0}&=\sin\left(\frac{\theta_{L+1}}{2}\right)\prod_{\ell=2}^{L}\sin\left(\frac{\theta_\ell+\mathbf{d}\boldsymbol{s}_\ell\pi}{2}\right)\sin\left(\frac{\theta_1}{2}+\frac{(\boldsymbol{s}_1+1)\pi}{4}\right)\\&=\prod_{\ell=2}^{L+1}\sin\left(\frac{\theta_\ell+\mathbf{d}\boldsymbol{s}'_\ell\pi}{2}\right)\sin\left(\frac{\theta_1}{2}+\frac{(\boldsymbol{s}'_1+1)\pi}{4}\right)=T_{\boldsymbol{s}',0},\end{aligned}\quad\text{(C17)}$$

其中，在公式（C16）和（C17）中，我们利用了以下事实：$\mathbf{d}\boldsymbol{s}'_{L+1}=0$。

公式（C11）第二行的最后一项也对应于 $\boldsymbol{s}'=(\boldsymbol{s},-1)$，但 $a=1$ 由于量子模式中的总位移为 $(-1)^a\boldsymbol{s}\cdot\boldsymbol{\beta}$，因此权重变为

$$\Phi_{\boldsymbol{s},1}+\pi-\phi_{L+1}=n_{\boldsymbol{s}}\pi+\phi_1-\Phi_{\boldsymbol{s},0}+\pi-\phi_{L+1}=(n_{\boldsymbol{s}}+1)\pi+\phi_1-\Phi_{\boldsymbol{s}',0}=n_{\boldsymbol{s}'}\pi+\phi_1-\Phi_{\boldsymbol{s}',0}=\Phi_{\boldsymbol{s}',1},\quad\text{(C18)}$$

$$\begin{aligned}\sin\left(\frac{\theta_{L+1}}{2}\right)T_{\boldsymbol{s},1}&=\sin\frac{\theta_{L+1}}{2}\prod_{\ell=2}^{L}\sin\left(\frac{\theta_\ell+\mathbf{d}\boldsymbol{s}_\ell\pi}{2}\right)\sin\left(\frac{\theta_1}{2}+\frac{(-\boldsymbol{s}_1+1)\pi}{4}\right)\\&=\prod_{\ell=2}^{L+1}\sin\left(\frac{\theta_\ell+\mathbf{d}\boldsymbol{s}'_\ell\pi}{2}\right)\sin\left(\frac{\theta_1}{2}+\frac{(-\boldsymbol{s}'_1+1)\pi}{4}\right)=T_{\boldsymbol{s}',1},\end{aligned}\quad\text{(C19)}$$

其中，在公式（C18）的第二个等式中，我们应用了公式（C16）的结果，而公式（C18）的最后一个等式则是通过直接利用 $n_{\boldsymbol{s}'}=L-\sum_{\ell=2}^{L+1}\mathbf{d}\boldsymbol{s}'_\ell=n_{\boldsymbol{s}}+1$ 在 $\mathbf{d}\boldsymbol{s}'=(\mathbf{d}\boldsymbol{s},0)$ 中给出的定义得到的。

公式（C11）第一行中的第二项对应于 $\boldsymbol{s}'=(-\boldsymbol{s},-1)$ 和 $a=0$，其中差值为 $\mathbf{d}\boldsymbol{s}'=(1-\mathbf{d}\boldsymbol{s}_1,(\mathbf{d}\boldsymbol{s})_{2:L},1)$。权重中的相位和振幅为

$$\begin{aligned}\Phi_{\boldsymbol{s},1}&=n_{\boldsymbol{s}}\pi+\phi_1-\sum_{\ell=1}^{L}(\mathbf{d}\boldsymbol{s}_\ell-1)(\boldsymbol{s}_\ell\phi_\ell-\delta_{\boldsymbol{s}_\ell,1}\pi)=\left(L-1-\sum_{\ell=2}^{L}\mathbf{d}\boldsymbol{s}_\ell\right)\pi+\phi_1-\sum_{\ell=1}^{L}(\mathbf{d}\boldsymbol{s}_\ell-1)(\boldsymbol{s}_\ell\phi_\ell-\delta_{\boldsymbol{s}_\ell,1}\pi)\\&=-\sum_{\ell=2}^{L}(\mathbf{d}\boldsymbol{s}_\ell-1)\boldsymbol{s}_\ell\phi_\ell+\left[1-(\mathbf{d}\boldsymbol{s}_1-1)\boldsymbol{s}_1\right]\phi_1-\sum_{\ell=2}^{L}(\mathbf{d}\boldsymbol{s}_\ell-1)(1-\delta_{\boldsymbol{s}_\ell,1})\pi+(\mathbf{d}\boldsymbol{s}_1-1)\delta_{\boldsymbol{s}_1,1}\pi\\&=\sum_{\ell=2}^{L}(\mathbf{d}\boldsymbol{s}'_\ell-1)\boldsymbol{s}'_\ell\phi_\ell+(1-\mathbf{d}\boldsymbol{s}'_1\cdot\boldsymbol{s}'_1)\phi_1-\sum_{\ell=2}^{L}(\mathbf{d}\boldsymbol{s}'_\ell-1)\delta_{\boldsymbol{s}'_\ell,1}\pi-\mathbf{d}\boldsymbol{s}'_1(1-\delta_{\boldsymbol{s}'_\ell,1})\pi\\&=\sum_{\ell=1}^{L+1}(\mathbf{d}\boldsymbol{s}'_\ell-1)\boldsymbol{s}'_\ell\phi_\ell-\sum_{\ell=1}^{L+1}(\mathbf{d}\boldsymbol{s}'_\ell-1)\delta_{\boldsymbol{s}'_\ell,1}\pi=\Phi_{\boldsymbol{s}',0},\end{aligned}\quad\text{(C20)}$$

and

$$\begin{aligned}\cos\left(\frac{\theta_{L+1}}{2}\right)T_{\boldsymbol{s},1}&=\cos\left(\frac{\theta_{L+1}}{2}\right)\prod_{\ell=2}^{L}\sin\left(\frac{\theta_\ell+\mathbf{d}\boldsymbol{s}_\ell\pi}{2}\right)\sin\left(\frac{\theta_1}{2}+\frac{(1-\boldsymbol{s}_1)\pi}{4}\right)\\&=\prod_{\ell=2}^{L+1}\sin\left(\frac{\theta_\ell+\mathbf{d}\boldsymbol{s}'_\ell\pi}{2}\right)\sin\left(\frac{\theta_1}{2}+\frac{(\boldsymbol{s}'_1+1)\pi}{4}\right)=T_{\boldsymbol{s}',0},\end{aligned}\quad\text{(C21)}$$

其中，在公式（C20）的最后一行，我们应用恒等式 $1-\mathbf{d}\boldsymbol{s}'_1\cdot\boldsymbol{s}'_1=(\mathbf{d}\boldsymbol{s}'_1-1)\boldsymbol{s}'_1$ 和 $\mathbf{d}\boldsymbol{s}'_1(1-\delta_{\boldsymbol{s}'_1,1})=(\mathbf{d}\boldsymbol{s}'_1-1)\delta_{\boldsymbol{s}'_1,1}$。将 $\boldsymbol{s}'_1=\pm1$ 分别代入即可轻松验证这些恒等式。

类似地，公式（C11）第二行中的第一项，带有 $\boldsymbol{s}'=(-\boldsymbol{s},-1)$ 但没有 $a=1$，其权重为

$$\Phi_{\boldsymbol{s},0}=n_{\boldsymbol{s}}\pi+\phi_1-\Phi_{\boldsymbol{s},1}=n_{\boldsymbol{s}'}\pi+\phi_1-\Phi_{\boldsymbol{s}',0}=\Phi_{\boldsymbol{s}',1} \tag{C22}$$

$$\begin{aligned}\cos\left(\frac{\theta_{L+1}}{2}\right)T_{\boldsymbol{s},0}&=\cos\left(\frac{\theta_{L+1}}{2}\right)\prod_{\ell=2}^{L}\sin\left(\frac{\theta_\ell+\mathbf{d}\boldsymbol{s}_\ell\pi}{2}\right)\sin\left(\frac{\theta_1}{2}+\frac{(\boldsymbol{s}_1+1)\pi}{4}\right)\\&=\prod_{\ell=2}^{L+1}\sin\left(\frac{\theta_\ell+\mathbf{d}\boldsymbol{s}'_\ell\pi}{2}\right)\sin\left(\frac{\theta_1}{2}+\frac{(1-\boldsymbol{s}'_1)\pi}{4}\right)=T_{\boldsymbol{s}',1},\end{aligned} \tag{C23}$$

其中，在公式（C22）的第二个等式中，我们利用了以下事实：$n_{\boldsymbol{s}'}=L-\sum_{\ell=2}^{L+1}\mathbf{d}\boldsymbol{s}_\ell=L-\sum_{\ell=2}^{L}\mathbf{d}\boldsymbol{s}_\ell-1=n_{\boldsymbol{s}}$给定 $\mathbf{d}\boldsymbol{s}'=(1-\mathbf{d}\boldsymbol{s}_1,(\mathbf{d}\boldsymbol{s})_{2:L},1)$。

因此，公式（C4）中的权重得以证明；结合量子模式位移，我们证明了一个（$L+1$）层线路的输出态满足公式（C2）中的形式。

■

现在我们引入能量正则化。不失一般性地，我们考虑每一步的位移均服从复高斯分布，$\beta_\ell\sim\mathcal{N}^{\mathrm{C}}_{E/L}$，或等价地，$\mathrm{Re}\{\beta_\ell\}$，$\mathrm{Im}\{\beta_\ell\}\sim\mathcal{N}_{E/2L}$的均值为零、方差为 $E/2L$，从而输出态的系综平均能量为 E。由公式（C2）可知，输出态的系综平均能量为

$$\begin{aligned}\mathbb{E}\left[\langle\psi|m^\dagger m|\psi\rangle\right]&=\sum_{a,a'=0}^{1}\sum_{\boldsymbol{s},\boldsymbol{s}'}\mathbb{E}\left[w_{\boldsymbol{s},a}w^*_{\boldsymbol{s}',a'}\right]\langle a'|a\rangle\,\mathbb{E}\left[e^{i(\chi_{\boldsymbol{s},a}-\chi_{\boldsymbol{s}',a'})}\langle(-1)^{a'}\boldsymbol{s}'\cdot\boldsymbol{\beta}|m^\dagger m|(-1)^a\boldsymbol{s}\cdot\boldsymbol{\beta}\rangle\right]\\&=\sum_{a=0}^{1}\sum_{\boldsymbol{s}}\mathbb{E}\left[|w_{\boldsymbol{s},a}|^2\right]\mathbb{E}\left[|\boldsymbol{s}\cdot\boldsymbol{\beta}|^2\right]\\&=2\cdot2^{L-1}\cdot\frac{1}{2^L}E=E,\end{aligned} \tag{C24}$$

其中 我们应用了恒等算符$\langle a|a'\rangle=\delta_{a,\ a'}$，$\mathbb{E}[w^*_{\boldsymbol{s}',a}w_{\boldsymbol{s},a}]=\mathbb{E}[|w_{\boldsymbol{s},a}|^2]\delta_{\boldsymbol{s},\ \boldsymbol{s}'}=1/2^L\delta_{\boldsymbol{s},\ \boldsymbol{s}'}$ ⟶ d

$$\mathbb{E}\left[|\boldsymbol{s}\cdot\boldsymbol{\beta}|^2\right]=\mathbb{E}\left[\mathrm{Re}\{\boldsymbol{s}\cdot\boldsymbol{\beta}\}^2\right]+\mathbb{E}\left[\mathrm{Im}\{\boldsymbol{s}\cdot\boldsymbol{\beta}\}^2\right]=2L\mathbb{E}\left[\mathrm{Re}\{\beta_\ell\}^2\right]=E, \tag{C25}$$

利用 了每个 β_ℓ 在 $\boldsymbol{\beta}$ 中的独立性以及 β的实部和虚部的对称性 ℓ。

附录 D：梯度评估方法

本节将为附录E中最终评估梯度方差提供一些准备材料。

如正文所述，代价函数通常可以写为 $\mathcal{C}=\mathrm{Tr}\left[OU\rho_0U^\dagger\right]$，其中 ρ_0是初始态，O 是可观测量。考虑关于第 k个量子比特旋转角 θ_k的梯度，则该梯度可写为

$$\partial_{\theta_k}\mathcal{C}=\partial_{\theta_k}\,\mathrm{Tr}\left[OU\rho_0U^\dagger\right]=\frac{-i}{2}\,\mathrm{Tr}\left[OU_{\mathrm{right}}[G(\phi_k),U_{\mathrm{left}}\rho_0U^\dagger_{\mathrm{left}}]U^\dagger_{\mathrm{right}}\right]=\frac{1}{2}\Big(\langle O\rangle_{k^{(+1)}}-\langle O\rangle_{k^{(-1)}}\Big), \tag{D1}$$

其中，我们将 $G(\phi_k)\equiv\cos\phi_k\sigma^x+\sin\phi_k\sigma^y$、$U_{\mathrm{left}}=\prod_{j=1}^{k-1}U_{\mathrm{ECD}}(\beta_j)U_{\mathrm{R}}(\theta_j,\phi_j)$分别记为第 k层之前的线路，以及从第 k层到第 L层的互补电路 U_{right}。最后一个等式通过应用参数移位规则 [71]、$[G(\phi_k)\ \rho]=i\left[U_{\mathrm{R}}\left(\frac{\pi}{2},\phi_k\right)\rho U^\dagger\mathrm{R}\left(\frac{\pi}{2},\ \phi_k\right)-U_{\mathrm{R}}\left(-\frac{\pi}{2},\ \phi_k\right)\rho U^\dagger\mathrm{R}\left(-\frac{\pi}{2},\ \phi_k\right)\right]$得到，而 $\langle O\rangle_{k^{(\pm1)}}$对应于使用包含 O 的输出态对 L层线路求期望，其中 θ_k 移位为 $\theta_k\to\theta_k\pm\pi/2$。为方便起见，在以下讨论中，我们将 $w_{\boldsymbol{s},\ a,\ k^{(\mu)}}$ 记为公式（C4）中定义的权重，其中 θ_k 通过 $\mu=\pm1$移位了 $\mu\pi/2$。

容易验证，

$$\mathbb{E}\left[\partial_{\theta_k}\mathcal{C}\right]=\frac{1}{2}\left(\mathbb{E}\left[\langle O\rangle_{k^{(+1)}}\right]-\mathbb{E}\left[\langle O\rangle_{k^{(-1)}}\right]\right)=0, \tag{D2}$$

由于系综平均是对 $\theta_j\in[0,2\pi$ 进行的）。根据方差的定义，方差可化简为

$$\mathrm{Var}\left[\partial_{\theta_k}\mathcal{C}\right]=\mathbb{E}\left[(\partial_{\theta_k}\mathcal{C})^2\right]=\frac{1}{2}\left(\mathbb{E}\left[\langle O\rangle^2_{k^{(+1)}}\right]-\mathbb{E}\left[\langle O\rangle_{k^{(+1)}}\langle O\rangle_{k^{(-1)}}\right]\right), \tag{D3}$$

这里我们再次取 $\mathbb{E}[\langle O\rangle_{k^{(+)}}^2]=\mathbb{E}[\langle O\rangle_{k^{(-)}}^2]$。对于本文所考虑的态制备任务，算符$O=|\phi\rangle\langle\phi|_q\otimes|\psi\rangle\langle\psi|_m$, 上述方差中的两项可通过公式 (C2) 中的输出态表示展开为

$$\begin{aligned}&\mathbb{E}\left[\langle O\rangle_{k^{(+1)}}\langle O\rangle_{k^{(\mu)}}\right]\\&=\sum_{\substack{a,a',\\b,b'=0}}^{1}\langle\phi|a\rangle\langle a'|\phi\rangle\langle\phi|b\rangle\langle b'|\phi\rangle\sum_{\substack{\boldsymbol{s},\boldsymbol{s}',\\\boldsymbol{r},\boldsymbol{r}'}}\mathbb{E}\left[w_{\boldsymbol{s},a,k^{(+1)}}w^*_{\boldsymbol{s}',a',k^{(+1)}}w_{\boldsymbol{r},b,k^{(\mu)}}w^*_{\boldsymbol{r}',b',k^{(\mu)}}\right]\mathbb{E}\left[\langle\psi|B_{\boldsymbol{s},a}\rangle\langle B_{\boldsymbol{s}',a'}|\psi\rangle\langle\psi|B_{\boldsymbol{r},b}\rangle\langle B_{\boldsymbol{r}',b'}|\psi\rangle\right].\end{aligned}\tag{D4}$$

上述方差的精确计算较为困难，因为$w_{\boldsymbol{s},\ a,\ k^{(\mu)}}$ 取决于$\mathbf{d}\boldsymbol{s}$ ，而$B_{\boldsymbol{s},\ a}$ 取决于 $\boldsymbol{s}$。因此，我们考虑由以下基本不等式构造的下界和上界。对于两组按递增顺序排列的 N实数 $x_1\le x_2\le\cdots\le x_N$和 $y_1\le y_2\le\cdots\le y_N$，存在一个众所周知的重排不等式

$$\sum_{j=1}^{n}x_jy_{n+1-j}\le\sum_{j=1}^{n}x_{\sigma(j)}y_j\le\sum_{j=1}^{n}x_jy_j,\tag{D5}$$

其中， $\sigma(j)\in\mathcal{S}_n$ 是 n 个元素的任意排列。一般而言， x_i 可以为正也可以为负，因此我们仅在对所有 j均有 $y_j>0$ 这一假设下，考虑界的一个放宽版本；该假设将 x 和 y之间的指标依赖关系解耦。上下界为

$$\sum_{j=1}^{n}x_{\sigma(j)}y_j\ge\left(\sum_{\substack{j=1\\x_j\ge0}}^{n}x_j\right)y_1+\left(\sum_{\substack{j=1\\x_j<0}}^{n}x_j\right)y_n,\tag{D6a}$$

$$\sum_{j=1}^{n}x_{\sigma(j)}y_j\le\left(\sum_{\substack{j=1\\x_j\ge0}}^{n}x_j\right)y_n+\left(\sum_{\substack{j=1\\x_j<0}}^{n}x_j\right)y_1.\tag{D6b}$$

证明。我们仅证明下界，上界则可通过交换最小值y_1 和最大值 y_n自然得到。公式（D6a）中的下界直接由公式（D5）中的第一个不等式得到，即

$$\sum_{j=1}x_{\sigma(j)}y_j\ge\sum_{j=1}^{n}x_jy_{n+1-j}=\sum_{\substack{j=1\\x_j>0}}^{n}x_jy_{n+1-j}+\sum_{\substack{j=1\\x_j<0}}^{n}x_jy_{n+1-j}\ge\sum_{\substack{j=1\\x_j>0}}^{n}x_jy_1+\sum_{\substack{j=1\\x_j<0}}^{n}x_jy_n.\tag{D7}$$

■

1. 预备知识

在本部分中，我们介绍一些在附录 E 中评估梯度方差时所需的预备引理和命题。回顾式 (C7) 中差分符号向量 $\mathbf{d}\boldsymbol{s}_\ell=|\boldsymbol{s}_\ell-\boldsymbol{s}_{\ell-1}|/2$ 的定义，它具有以下性质。

引理 6 差分符号向量中所有元素之和始终为偶数， $\mathbb{P}\left(\sum_{\ell=1}^{L}\mathbf{d}\boldsymbol{s}_\ell\right)=+1$。

证明。由于 $(-1)^{\mathbf{d}\boldsymbol{s}_\ell}=\boldsymbol{s}_\ell\times\boldsymbol{s}_{\ell-1}$ 对于 $\ell\ge2$ 且 $(-1)^{\mathbf{d}\boldsymbol{s}_1}=\boldsymbol{s}_1\times\boldsymbol{s}_L$，我们有

$$\prod_{\ell=1}^{L}(-1)^{\mathbf{d}\boldsymbol{s}_\ell}=(-1)^{\sum_{\ell=1}^{L}\mathbf{d}\boldsymbol{s}_\ell}=\boldsymbol{s}_1\times\boldsymbol{s}_L\times\prod_{\ell=2}^{L}(\boldsymbol{s}_\ell\times\boldsymbol{s}_{\ell-1})=\prod_{\ell=1}^{L}\boldsymbol{s}_\ell^2=1,\tag{D8}$$

因为 $\boldsymbol{s}_\ell=\pm1$。因此， $\sum_{\ell=1}^{L}\mathbf{d}\boldsymbol{s}_\ell$ 为偶数，且 $\mathbb{P}\left(\sum_{\ell=1}^{L}\mathbf{d}\boldsymbol{s}_\ell\right)=+1$

■

引理 6 的一个直接结果涉及 $\Phi_{\boldsymbol{s},a}$ 中 ϕ_js 的数量，该数量满足

$T_{\boldsymbol{s},a}$ \ $T_{\boldsymbol{r},a}$	$\sin(\theta_\ell/2)$	$\cos(\theta_\ell/2)$
$\sin(\theta_\ell/2)$	$n_{\sin}$	$N_{\boldsymbol{s},a}-n_{\sin}$
$\cos(\theta_\ell/2)$	$N_{\boldsymbol{r},a}-n_{\sin}$	$N_T-n_{\sin}$

表 I。θ_ℓ 中$T_{\boldsymbol{s},a}$和$T_{\boldsymbol{r},a}$的数量及其对应形式。

推论 7 在 $\Phi_{\boldsymbol{s},a}$ 中具有非零系数的 ϕ_ℓ**s** 的数量为 $N_{\boldsymbol{s},a}=n_{\boldsymbol{s}}+1-a-(-1)^a\mathbf{d}\boldsymbol{s}_1$，其奇偶性满足 $\mathbb{P}(N_{\boldsymbol{s},a})=\mathbb{P}(a)\mathbb{P}(L)$。

证明。由于 $N_{\boldsymbol{s},a}=n_{\boldsymbol{s}}+1-a-(-1)^a\mathbf{d}\boldsymbol{s}_1$，我们可以看到 $N_{\boldsymbol{s},a}$ 的奇偶性满足

$$\mathbb{P}(N_{\boldsymbol{s},a})=(-1)^{n_{\boldsymbol{s}}}(-1)^{(-1)^a\mathbf{d}\boldsymbol{s}_1}(-1)^{1-a}=(-1)^{L-1-\sum_{\ell=2}^L\mathbf{d}\boldsymbol{s}_\ell}(-1)^{-\mathbf{d}\boldsymbol{s}_1}(-1)^{1-a}=(-1)^{L-a}(-1)^{-\sum_{\ell=1}^L\mathbf{d}\boldsymbol{s}_\ell}=\mathbb{P}(a)\mathbb{P}(L), \tag{D9}$$

其中，在第二个等式中，我们将 $(-1)^{(-1)^a\mathbf{d}\boldsymbol{s}_1}=(-1)^{-\mathbf{d}\boldsymbol{s}_1}$ 重写为符号，因为指数的符号不会改变其值，而最后一个等式由引理 6 得出。 ■

命题 8 对于均匀随机抽样的两个任意不同的符号向量 $\boldsymbol{s},\boldsymbol{r}$，$\boldsymbol{s}_\ell=\mathbf{d}\boldsymbol{r}_\ell$ 的元素数量为 $N_T=L-\sum_{\ell=1}^L|\mathbf{d}\boldsymbol{s}_\ell-\mathbf{d}\boldsymbol{r}_\ell|$，且分布概率为 $p(N_T)=\binom{L}{N_T}/(2^{L-1}-1)$，满足约束$0\le N_T\le L-2$ 且 $\mathbb{P}(N_T)=\mathbb{P}(L)$。

证明。对于任意两个符号向量 $\boldsymbol{s}\neq\boldsymbol{r}$，由于引理 6 中的奇偶约束，我们有 $0\le N_T\le L-2$。由于 $\mathbf{d}\boldsymbol{s}_\ell=0$ 具有等先验，因此其分布概率为 $\binom{L}{N_T}/(2^{L-1}-1)$。假设在 $T_{\boldsymbol{s},a}$ 和 $T_{\boldsymbol{r},a}$ 中，均为 $\sin(\theta_\ell/2)$ 形式的 θ_ℓ 的数量为 $n_{\sin}$，而均为 $\cos(\theta_\ell/2)$ 形式的 θ_ℓ 的数量为 $N_{\boldsymbol{s},a}-n_{\sin}$。另一方面，在 $T_{\boldsymbol{s},a}$ 中为 $\sin(\theta_\ell/2)$ 形式、但在 $T_{\boldsymbol{r},a}$ 中为 $\cos(\theta_\ell/2)$ 形式的 θ_ℓ 的数量为 $N_T-n_{\sin}$，而反向对应的数量为 $N_{\boldsymbol{r},a}-n_{\sin}$。上述陈述总结于表 I 中。由于 $T_{\boldsymbol{s},a}$ 和 $T_{\boldsymbol{r},a}$ 的总数 θ_ℓ 均为 L，因此表 I 中各项之和应等于 L，$N_{\boldsymbol{s},a}+N_{\boldsymbol{r},a}+N_T-2n_{\sin}=L$，且奇偶关系为

$$1=\mathbb{P}(L-N_{\boldsymbol{s},a}-N_{\boldsymbol{r},a}-N_T+2n_{\sin})=\mathbb{P}(L)\mathbb{P}(N_{\boldsymbol{s},a})\mathbb{P}(N_{\boldsymbol{r},a})\mathbb{P}(N_T)=\mathbb{P}(L)\mathbb{P}(N_T), \tag{D10}$$

其中，在第二个等式中，我们利用了推论 7 中的 $\mathbb{P}(N_{\boldsymbol{s},a})=\mathbb{P}(N_{\boldsymbol{r},a})$。 ■

命题 9 对于两个任意不同的符号向量 $\boldsymbol{s},\boldsymbol{r}$，在约束条件 $1\le\ell\le L-1$ 下，满足分布 $p(\ell)=\binom{L-1}{\ell-1}/(2^{L-1}-1)$ 的元素数量为 $\boldsymbol{s}_\ell=\boldsymbol{r}_\ell$**is** $\ell=L-\sum_{\ell=1}^L|\boldsymbol{s}_\ell-\boldsymbol{r}_\ell|/2$。

2. 四重权重乘积的系综平均

在本部分中，我们计算公式 (D4) 中出现的四重权重的系综平均，并讨论其性质。一般而言，四重权重乘积为 $w_{\boldsymbol{s},a,k^{(+1)}}w^*_{\boldsymbol{s}',a',k^{(+1)}}w_{\boldsymbol{r},b,k^{(\mu)}}w^*_{\boldsymbol{r}',b',k^{(\mu)}}$，其中 $\mu=\pm1$，对所有 θ 和 ϕ 的系综平均为

$$\begin{aligned}&\mathbb{E}_{\boldsymbol{\theta},\boldsymbol{\phi}}\left[w_{\boldsymbol{s},a,k^{(+1)}}(\boldsymbol{\theta},\boldsymbol{\phi})w^*_{\boldsymbol{s}',a',k^{(+1)}}(\boldsymbol{\theta},\boldsymbol{\phi})w_{\boldsymbol{r},b,k^{(\mu)}}(\boldsymbol{\theta},\boldsymbol{\phi})w^*_{\boldsymbol{r}',b',k^{(\mu)}}(\boldsymbol{\theta},\boldsymbol{\phi})\right]\\&=\mathbb{E}_{\boldsymbol{\phi}}\left[e^{i\left(\Phi_{\boldsymbol{s},a}(\boldsymbol{\phi})-\Phi_{\boldsymbol{s}',a'}(\boldsymbol{\phi})+\Phi_{\boldsymbol{r},b}(\boldsymbol{\phi})-\Phi_{\boldsymbol{r}',b'}(\boldsymbol{\phi})\right)}\right]\mathbb{E}_{\boldsymbol{\theta}}\left[T_{\boldsymbol{s},a,k^{(+1)}}(\boldsymbol{\theta})T_{\boldsymbol{s}',a',k^{(+1)}}(\boldsymbol{\theta})T_{\boldsymbol{r},b,k^{(\mu)}}(\boldsymbol{\theta})T_{\boldsymbol{r}',b',k^{(\mu)}}(\boldsymbol{\theta})\right].\end{aligned} \tag{D11}$$

如果相位 $\Phi_{\boldsymbol{s},a}-\Phi_{\boldsymbol{s}',a'}+\Phi_{\boldsymbol{r},b}-\Phi_{\boldsymbol{r}',b'}$ 中还剩有任何 ϕ_ℓ，那么对 ϕ 求平均的结果显然为零；否则，结果可能为 ±1，这取决于相位中是否存在偶数个 π。关于 $\boldsymbol{\theta}$ 的平均为

命题 10 四重权重乘积中对 θ_ℓ**s** 的系综平均为

$$\begin{aligned}&\mathbb{E}_{\boldsymbol{\theta}}\left[T_{\boldsymbol{s},a,k^{(+1)}}(\boldsymbol{\theta})T_{\boldsymbol{s}',a',k^{(+1)}}(\boldsymbol{\theta})T_{\boldsymbol{r},b,k^{(\mu)}}(\boldsymbol{\theta})T_{\boldsymbol{r}',b',k^{(\mu)}}(\boldsymbol{\theta})\right]\\&=\frac{1}{8^L}\prod_{\substack{\ell=2\\\ell\neq k}}^{L}\left(2\delta_{\mathbf{d}\boldsymbol{s}_\ell,\mathbf{d}\boldsymbol{s}'_\ell}\delta_{\mathbf{d}\boldsymbol{r}_\ell,\mathbf{d}\boldsymbol{r}'_\ell}+\cos\left[\frac{\pi}{2}(\mathbf{d}\boldsymbol{s}_\ell+\mathbf{d}\boldsymbol{s}'_\ell-\mathbf{d}\boldsymbol{r}_\ell-\mathbf{d}\boldsymbol{r}'_\ell)\right]\right)\left(2\delta_{\mathbf{d}\boldsymbol{s}_k,\mathbf{d}\boldsymbol{s}'_k}\delta_{\mathbf{d}\boldsymbol{r}_k,\mathbf{d}\boldsymbol{r}'_k}+\mu\cos\left[\frac{\pi}{2}(\mathbf{d}\boldsymbol{s}_k+\mathbf{d}\boldsymbol{s}'_k-\mathbf{d}\boldsymbol{r}_k-\mathbf{d}\boldsymbol{r}'_k)\right]\right)^{1-\delta_{k,1}}\\&\times\left(2\delta_{\mathbb{P}(a)\boldsymbol{s}_1,\mathbb{P}(a')\boldsymbol{s}'_1}\delta_{\mathbb{P}(b)\boldsymbol{r}_1,\mathbb{P}(b')\boldsymbol{r}'_1}+(1+(\mu-1)\delta_{k,1})\cos\left[\frac{\pi}{4}(\mathbb{P}(a)\boldsymbol{s}_1+\mathbb{P}(a')\boldsymbol{s}'_1-\mathbb{P}(b)\boldsymbol{r}_1-\mathbb{P}(b')\boldsymbol{r}'_1)\right]\right).\end{aligned} \tag{D12}$$

证明。由于各个 θ_ℓ 彼此独立，因此可以分别处理每个 θ_ℓ 的平均值。根据公式 (C6)，我们只需要分别计算 $\ell \geq 2$ 条件下 θ_1 和 θ_ℓ 的平均值。对于 $\mathbb{E}\left[T_{\boldsymbol{s},a,k^{(+1)}} T_{\boldsymbol{s}',a',k^{(+1)}} T_{\boldsymbol{r},b,k^{(\mu)}} T_{\boldsymbol{r}',b',k^{(\mu)}}\right]$ 中满足 $\ell \geq 2$ 条件的 θ_ℓ 的系综平均为

$$
\begin{aligned}
&\mathbb{E}_{\boldsymbol{\theta}}\left[\sin\left(\frac{\theta_\ell+\pi \mathbf{d}\boldsymbol{s}_\ell}{2}+\frac{\pi}{4}\right)\sin\left(\frac{\theta_\ell+\pi \mathbf{d}\boldsymbol{s}'_\ell}{2}+\frac{\pi}{4}\right)\sin\left(\frac{\theta_\ell+\pi \mathbf{d}\boldsymbol{r}_\ell}{2}+\frac{\mu\pi}{4}\right)\sin\left(\frac{\theta_\ell+\pi \mathbf{d}\boldsymbol{r}'_\ell}{2}+\frac{\mu\pi}{4}\right)\right] \\
&=\frac{1}{8}\left(\cos\left[\frac{\pi}{2}(\mathbf{d}\boldsymbol{s}_\ell-\mathbf{d}\boldsymbol{s}'_\ell+\mathbf{d}\boldsymbol{r}_\ell-\mathbf{d}\boldsymbol{r}'_\ell)\right]+\cos\left[\frac{\pi}{2}(\mathbf{d}\boldsymbol{s}_\ell-\mathbf{d}\boldsymbol{s}'_\ell-\mathbf{d}\boldsymbol{r}_\ell+\mathbf{d}\boldsymbol{r}'_\ell)\right]+\mu\cos\left[\frac{\pi}{2}(\mathbf{d}\boldsymbol{s}_\ell+\mathbf{d}\boldsymbol{s}'_\ell-\mathbf{d}\boldsymbol{r}_\ell-\mathbf{d}\boldsymbol{r}'_\ell)\right]\right) \\
&=\frac{1}{8}\left(2\cos\left[\frac{\pi(\mathbf{d}\boldsymbol{s}_\ell-\mathbf{d}\boldsymbol{s}'_\ell)}{2}\right]\cos\left[\frac{\pi(\mathbf{d}\boldsymbol{r}_\ell-\mathbf{d}\boldsymbol{r}'_\ell)}{2}\right]+\mu\cos\left[\frac{\pi}{2}(\mathbf{d}\boldsymbol{s}_\ell+\mathbf{d}\boldsymbol{s}'_\ell-\mathbf{d}\boldsymbol{r}_\ell-\mathbf{d}\boldsymbol{r}'_\ell)\right]\right) \\
&=\frac{1}{8}\left(2\delta_{\mathbf{d}\boldsymbol{s}_\ell,\mathbf{d}\boldsymbol{s}'_\ell}\delta_{\mathbf{d}\boldsymbol{r}_\ell,\mathbf{d}\boldsymbol{r}'_\ell}+\mu\cos\left[\frac{\pi}{2}(\mathbf{d}\boldsymbol{s}_\ell+\mathbf{d}\boldsymbol{s}'_\ell-\mathbf{d}\boldsymbol{r}_\ell-\mathbf{d}\boldsymbol{r}'_\ell)\right]\right),
\end{aligned}
\tag{D13}
$$

其中在最后一行，我们应用恒等式 $\cos(\pi(x-y)/2)=\delta_{x,\,y}$ 当 $x,\ y\in\{0,1\}$ 时。对于 θ_1，平均值为

$$
\begin{aligned}
&\mathbb{E}_{\boldsymbol{\theta}}\left[\sin\left(\frac{\theta_1}{2}+\frac{(\mathbb{P}(a)\boldsymbol{s}_1+2)\pi}{4}\right)\sin\left(\frac{\theta_1}{2}+\frac{(\mathbb{P}(a')\boldsymbol{s}'_1+2)\pi}{4}\right)\sin\left(\frac{\theta_1}{2}+\frac{(\mathbb{P}(b)\boldsymbol{r}_1+\mu+1)\pi}{4}\right)\sin\left(\frac{\theta_1}{2}+\frac{(\mathbb{P}(b')\boldsymbol{r}'_1+\mu+1)\pi}{4}\right)\right] \\
&=\frac{1}{8}\left(\cos\left[\frac{\pi}{4}\left(\mathbb{P}(a)\boldsymbol{s}_1-\mathbb{P}(a')\boldsymbol{s}'_1+\mathbb{P}(b)\boldsymbol{r}_1-\mathbb{P}(b')\boldsymbol{r}'_1\right)\right]+\cos\left[\frac{\pi}{4}\left(\mathbb{P}(a)\boldsymbol{s}_1-\mathbb{P}(a')\boldsymbol{s}'_1-\mathbb{P}(b)\boldsymbol{r}_1+\mathbb{P}(b')\boldsymbol{r}'_1\right)\right]\right. \\
&\quad\left.+\mu\cos\left[\frac{\pi}{4}\left(\mathbb{P}(a)\boldsymbol{s}_1+\mathbb{P}(a')\boldsymbol{s}'_1-\mathbb{P}(b)\boldsymbol{r}_1-\mathbb{P}(b')\boldsymbol{r}'_1\right)\right]\right) \\
&=\frac{1}{8}\left(2\cos\left[\frac{\pi(\mathbb{P}(a)\boldsymbol{s}_1-\mathbb{P}(a')\boldsymbol{s}'_1)}{4}\right]\cos\left[\frac{\pi(\mathbb{P}(b)\boldsymbol{r}_1-\mathbb{P}(b')\boldsymbol{r}'_1)}{4}\right]+\mu\cos\left[\frac{\pi}{4}(\mathbb{P}(a)\boldsymbol{s}_1+\mathbb{P}(a')\boldsymbol{s}'_1-\mathbb{P}(b)\boldsymbol{r}_1-\mathbb{P}(b')\boldsymbol{r}'_1)\right]\right) \\
&=\frac{1}{8}\left(2\delta_{\mathbb{P}(a)\boldsymbol{s}_1,\mathbb{P}(a')\boldsymbol{s}'_1}\delta_{\mathbb{P}(b)\boldsymbol{r}_1,\mathbb{P}(b')\boldsymbol{r}'_1}+\mu\cos\left[\frac{\pi}{4}(\mathbb{P}(a)\boldsymbol{s}_1+\mathbb{P}(a')\boldsymbol{s}'_1-\mathbb{P}(b)\boldsymbol{r}_1-\mathbb{P}(b')\boldsymbol{r}'_1)\right]\right),
\end{aligned}
\tag{D14}
$$

其中在最后一行，我们应用 $\cos(\pi(x-y)/4)=\delta_{x,\,y}$ 当 $x,\ y=\pm1$ 时。请注意，对所有 θ_ℓ 的平均值取决于 k 的选择，更具体地说，取决于是否 $k>1$。可以分别写出这两种情况下的系综平均，并发现对所有 θ_ℓ 的系综平均可以统一写成

$$
\begin{aligned}
&\mathbb{E}_{\boldsymbol{\theta}}\left[T_{\boldsymbol{s},a,k^{(+1)}}(\boldsymbol{\theta})T_{\boldsymbol{s}',a',k^{(+1)}}(\boldsymbol{\theta})T_{\boldsymbol{r},b,k^{(\mu)}}(\boldsymbol{\theta})T_{\boldsymbol{r}',b',k^{(\mu)}}(\boldsymbol{\theta})\right] \\
&=\frac{1}{8^L}\prod_{\substack{\ell=2\\ \ell\neq k}}^{L}\left(2\delta_{\mathbf{d}\boldsymbol{s}_\ell,\mathbf{d}\boldsymbol{s}'_\ell}\delta_{\mathbf{d}\boldsymbol{r}_\ell,\mathbf{d}\boldsymbol{r}'_\ell}+\cos\left[\frac{\pi}{2}(\mathbf{d}\boldsymbol{s}_\ell+\mathbf{d}\boldsymbol{s}'_\ell-\mathbf{d}\boldsymbol{r}_\ell-\mathbf{d}\boldsymbol{r}'_\ell)\right]\right)\left(2\delta_{\mathbf{d}\boldsymbol{s}_k,\mathbf{d}\boldsymbol{s}'_k}\delta_{\mathbf{d}\boldsymbol{r}_k,\mathbf{d}\boldsymbol{r}'_k}+\mu\cos\left[\frac{\pi}{2}(\mathbf{d}\boldsymbol{s}_k+\mathbf{d}\boldsymbol{s}'_k-\mathbf{d}\boldsymbol{r}_k-\mathbf{d}\boldsymbol{r}'_k)\right]\right)^{1-\delta_{k,1}} \\
&\quad\times\left(2\delta_{\mathbb{P}(a)\boldsymbol{s}_1,\mathbb{P}(a')\boldsymbol{s}'_1}\delta_{\mathbb{P}(b)\boldsymbol{r}_1,\mathbb{P}(b')\boldsymbol{r}'_1}+(1+(\mu-1)\delta_{k,1})\cos\left[\frac{\pi}{4}(\mathbb{P}(a)\boldsymbol{s}_1+\mathbb{P}(a')\boldsymbol{s}'_1-\mathbb{P}(b)\boldsymbol{r}_1-\mathbb{P}(b')\boldsymbol{r}'_1)\right]\right).
\end{aligned}
\tag{D15}
$$

■

由公式（D12）直接得出以下推论

推论 11 如果三个 o f 符号向量 $\boldsymbol{s},\boldsymbol{s}',\boldsymbol{r},\boldsymbol{r}'$ 相等，那么公式（D11）中的系综平均为零。

证明。如果三个符号向量相同，那么可以验证 $2\delta_{\mathbf{d}\boldsymbol{s}_j,\ \mathbf{d}\boldsymbol{s}'_j}\delta_{\mathbf{d}\boldsymbol{r}_j,\ \mathbf{d}\boldsymbol{r}'_j}=0$ 且 $\cos\left[\frac{\pi}{2}(\mathbf{d}\boldsymbol{s}_j+\mathbf{d}\boldsymbol{s}'_j-\mathbf{d}\boldsymbol{r}_j-\mathbf{d}\boldsymbol{r}'_j)\right]=\cos(\pm\pi/2)=0$，这使得公式（D12）中的平均值为零。 ■

附录 E：单模连续变量态制备中的梯度方差

本节将详细证明制备目标能量为 $\langle m^\dagger m\rangle=E_t$ 的单模连续变量态 $|\psi\rangle$ 时梯度方差的界（不等式 (6) 和 (7)，其中 $M=1$），而辅助量子比特的目标态则简单地选为 $|0\rangle_q$。

对于量子比特目标态 $|0\rangle_q$，梯度方差中的各项展开式（见公式 (D4)）可简化为

$$
\mathbb{E}\left[\langle O\rangle_{k^{(+1)}}\langle O\rangle_{k^{(\mu)}}\right]=\sum_{\boldsymbol{s},\boldsymbol{s}',\boldsymbol{r},\boldsymbol{r}'}\mathbb{E}_{\boldsymbol{\theta},\boldsymbol{\phi}}\left[w_{\boldsymbol{s},k^{(+1)}}w^*_{\boldsymbol{s}',k^{(+1)}}w_{\boldsymbol{r},k^{(\mu)}}w^*_{\boldsymbol{r}',k^{(\mu)}}\right]\mathbb{E}_{\boldsymbol{\beta}}\left[\langle\psi|B_{\boldsymbol{s}}\rangle\langle B_{\boldsymbol{s}'}|\psi\rangle\langle\psi|B_{\boldsymbol{r}}\rangle\langle B_{\boldsymbol{r}'}|\psi\rangle\right],
\tag{E1}
$$

其中，为简化表达式，我们省略了与量子比特态 a,a',b,b' 相关的下标，因为 $a=a'=b=b'=0$。对 $\boldsymbol{s},\boldsymbol{s}',\boldsymbol{r},\boldsymbol{r}'$ 的求和在以下四种情况下可能非零：(i) $\boldsymbol{s}=\boldsymbol{s}'=\boldsymbol{r}=\boldsymbol{r}'$；(ii) $\boldsymbol{s}-\boldsymbol{s}'=\boldsymbol{r}-\boldsymbol{r}'=\boldsymbol{0}$

但是 $\boldsymbol{s}\neq\boldsymbol{r}$；（iii） $\boldsymbol{s}-\boldsymbol{r}'=\boldsymbol{r}-\boldsymbol{s}'=\boldsymbol{0}$ 但是 $\boldsymbol{s}\neq\boldsymbol{r}$；（iv） $\boldsymbol{s}$，$\boldsymbol{s}'$，$\boldsymbol{r}$，$\boldsymbol{r}'$不相等。然而，$\boldsymbol{s}-\boldsymbol{r}=\boldsymbol{s}'-\boldsymbol{r}'=\boldsymbol{0}$ 与 $\boldsymbol{s}\neq\boldsymbol{s}'$ 并不 产生贡献，因为对 $\boldsymbol{\phi}$ 取平均后变为 $\mathbb{E}[e^{2i(\Phi_{\boldsymbol{s},a}-\Phi_{\boldsymbol{s}',a})}]=0$，其中 $\boldsymbol{s}\neq\boldsymbol{s}'$。

根据上述分析，公式（D3）中所示的梯度方差变为

$$
\begin{aligned}
&\mathrm{Var}\left[\partial_{\theta_k}\mathcal{C}\right]=\frac{1}{2}\left(\mathbb{E}\left[\langle O\rangle^2_{k^{(+1)}}\right]-\mathbb{E}\left[\langle O\rangle_{k^{(+1)}}\langle O\rangle_{k^{(-1)}}\right]\right)\\
&=\frac{1}{2}\Bigg(\sum_{\boldsymbol{s}}\Delta_\mu\left\{\mathbb{E}_{\boldsymbol{\theta},\boldsymbol{\phi}}\left[|w_{\boldsymbol{s},k^{(+1)}}|^2|w_{\boldsymbol{s},k^{(\mu)}}|^2\right]\right\}\mathbb{E}_{\boldsymbol{\beta}}\left[|\langle\psi|B_{\boldsymbol{s}}\rangle|^4\right]\\
&\quad+\sum_{\boldsymbol{s}\neq\boldsymbol{r}}\Delta_\mu\left\{\mathbb{E}_{\boldsymbol{\theta},\boldsymbol{\phi}}\left[|w_{\boldsymbol{s},k^{(+1)}}|^2|w_{\boldsymbol{r},k^{(\mu)}}|^2\right]+\mathbb{E}_{\boldsymbol{\theta},\boldsymbol{\phi}}\left[w_{\boldsymbol{s},k^{(+1)}}w^*_{\boldsymbol{r},k^{(+1)}}w_{\boldsymbol{r},k^{(\mu)}}w^*_{\boldsymbol{s},k^{(\mu)}}\right]\right\}\mathbb{E}_{\boldsymbol{\beta}}\left[|\langle\psi|B_{\boldsymbol{s}}\rangle|^2|\langle\psi|B_{\boldsymbol{r}}\rangle|^2\right]\\
&\quad+\sum_{\substack{\boldsymbol{s},\boldsymbol{s}',\boldsymbol{r},\boldsymbol{r}'\\ \text{unequal}}}\Delta_\mu\left\{\mathbb{E}_{\boldsymbol{\theta},\boldsymbol{\phi}}\left[w_{\boldsymbol{s},k^{(+1)}}w^*_{\boldsymbol{s}',k^{(+1)}}w_{\boldsymbol{r},k^{(\mu)}}w^*_{\boldsymbol{r}',k^{(\mu)}}\right]\right\}\mathbb{E}_{\boldsymbol{\beta}}\left[\langle\psi|B_{\boldsymbol{s}}\rangle\langle B_{\boldsymbol{s}'}|\psi\rangle\langle\psi|B_{\boldsymbol{r}}\rangle\langle B_{\boldsymbol{r}'}|\psi\rangle\right]\Bigg)
\end{aligned}
\tag{E2}
$$

$$
\equiv\frac{1}{2}\left(S_1+S_2+S_3\right). \tag{E3}
$$

符号 $\Delta_\mu\{X\}\equiv X|_{\mu=1}-X|_{\mu=-1}$ 表示 X 与 $\mu=1$ 和 $\mu=-1$之间的量的差，且该差仅在对 θ_ℓ求平均时计算。为方便起见，我们在最后一行引入 $\{S_1,S_2,S_3\}$，用来表示上方大括号中的三个求和项。接下来，我们将逐项计算它们。

对于 S_1，公式（D12）中的权重系综平均为

$$
\mathbb{E}\left[|w_{\boldsymbol{s},k^{(+1)}}|^2|w_{\boldsymbol{s},k^{(\mu)}}|^2\right]=\mathbb{E}\left[T^2_{\boldsymbol{s},k^{(+1)}}T^2_{\boldsymbol{s},k^{(\mu)}}\right]=\frac{1}{8^L}3^{L-2+\delta_{k,1}}(2+\mu)^{1-\delta_{k,1}}\left(3+(\mu-1)\delta_{k,1}\right)=\frac{3^{L-1}}{8^L}(2+\mu). \tag{E4}
$$

与 $\mu=\pm1$ 的区别是

$$
\Delta_\mu\left\{\mathbb{E}\left[|w_{\boldsymbol{s},k^{(+1)}}|^2|w_{\boldsymbol{s},k^{(\mu)}}|^2\right]\right\}=\frac{2\cdot3^{L-1}}{8^L}. \tag{E5}
$$

对于位移部分，回顾一下，$|B_{\boldsymbol{s}}\rangle=e^{i\chi_{\boldsymbol{s}}}|\boldsymbol{s}\cdot\boldsymbol{\beta}\rangle$ 其中 $\chi_{\boldsymbol{s}}$ 是纯相位，$\boldsymbol{s}\cdot\boldsymbol{\beta}$ 是均值为零且方差为 E 的复高斯变量，如 $(\boldsymbol{s}\cdot\boldsymbol{\beta})\sim\mathcal{N}_E^{\mathrm{C}}$，而对 $\boldsymbol{\beta}$ 的平均变为

$$
\mathbb{E}\left[|\langle\psi|B_{\boldsymbol{s}}\rangle|^4\right]=\mathbb{E}\left[|\langle\psi|\boldsymbol{s}\cdot\boldsymbol{\beta}\rangle|^4\right]=\mathbb{E}_{\alpha\sim\mathcal{N}_E^{\mathrm{C}}}\left[|\langle\psi|\alpha\rangle|^4\right]\equiv C_1, \tag{E6}
$$

其中，在第二个等式中，$|\alpha\rangle$ 是位移为 $\alpha\sim\mathcal{N}_E^{\mathrm{C}}$ 的相干态，我们将该平均值记为关联函数 C_1。由于式 (E5) 和 (E6) 均与 $\boldsymbol{s}$ 无关，求和 S_1 仅为

$$
S_1=\sum_{\boldsymbol{s}}\Delta_\mu\left\{\mathbb{E}\left[|w_{\boldsymbol{s},k^{(+1)}}|^2|w_{\boldsymbol{s},k^{(\mu)}}|^2\right]\right\}\mathbb{E}\left[|\langle\psi|B_{\boldsymbol{s}}\rangle|^4\right]=2^{L-1}\frac{2\cdot3^{L-1}}{8^L}C_1=\frac{3^{L-1}}{4^L}C_1. \tag{E7}
$$

量子比特旋转角 $\boldsymbol{\phi}$，$\boldsymbol{\theta}$ 在 S_2 中利用公式 D12 的平均值为

$$
\begin{aligned}
&\mathbb{E}\left[|w_{\boldsymbol{s},k^{(+1)}}|^2|w_{\boldsymbol{r},k^{(\mu)}}|^2\right]+\mathbb{E}\left[w_{\boldsymbol{s},k^{(+1)}}w^*_{\boldsymbol{r},k^{(+1)}}w_{\boldsymbol{r},k^{(\mu)}}w^*_{\boldsymbol{s},k^{(\mu)}}\right]=\mathbb{E}\left[T^2_{\boldsymbol{s},k^{(+1)}}T^2_{\boldsymbol{r},k^{(\mu)}}\right]+\mathbb{E}\left[T_{\boldsymbol{s},k^{(+1)}}T_{\boldsymbol{r},k^{(+1)}}T_{\boldsymbol{r},k^{(\mu)}}T_{\boldsymbol{s},k^{(\mu)}}\right]\\
&=\frac{1}{8^L}\prod_{\substack{\ell=2\\ \ell\neq k}}^{L}\left(1+2\delta_{\mathbf{d}\boldsymbol{s}_\ell,\mathbf{d}\boldsymbol{r}_\ell}\right)\left(2-\mu+2\mu\delta_{\mathbf{d}\boldsymbol{s}_k,\mathbf{d}\boldsymbol{r}_k}\right)^{1-\delta_{k,1}}\left[2+\left(1+(\mu-1)\delta_{k,1}\right)\left(2\delta_{\mathbf{d}\boldsymbol{s}_1,\mathbf{d}\boldsymbol{r}_1}-1\right)\right]\\
&\quad+\frac{1}{8^L}\prod_{\substack{\ell=2\\ \ell\neq k}}^{L}\left(1+2\delta_{\mathbf{d}\boldsymbol{s}_\ell,\mathbf{d}\boldsymbol{r}_\ell}\right)\left(2\delta_{\mathbf{d}\boldsymbol{s}_k,\mathbf{d}\boldsymbol{r}_k}+\mu\right)^{1-\delta_{k,1}}\left[2\delta_{\mathbf{d}\boldsymbol{s}_1,\mathbf{d}\boldsymbol{r}_1}+\left(1+(\mu-1)\delta_{k,1}\right)\right]\\
&=\frac{1}{8^L}\prod_{\substack{\ell=1\\ \ell\neq k}}^{L}\left(1+2\delta_{\mathbf{d}\boldsymbol{s}_\ell,\mathbf{d}\boldsymbol{r}_\ell}\right)\left(2-\mu+2\mu\delta_{\mathbf{d}\boldsymbol{s}_k,\mathbf{d}\boldsymbol{r}_k}\right)+\frac{1}{8^L}\prod_{\substack{\ell=1\\ \ell\neq k}}^{L}\left(1+2\delta_{\mathbf{d}\boldsymbol{s}_\ell,\mathbf{d}\boldsymbol{r}_\ell}\right)\left(2\delta_{\mathbf{d}\boldsymbol{s}_k,\mathbf{d}\boldsymbol{r}_k}+\mu\right)\\
&=\frac{2}{8^L}\prod_{\substack{\ell=1\\ \ell\neq k}}^{L}\left(1+2\delta_{\mathbf{d}\boldsymbol{s}_\ell,\mathbf{d}\boldsymbol{r}_\ell}\right)\left(1+(1+\mu)\delta_{\mathbf{d}\boldsymbol{s}_k,\mathbf{d}\boldsymbol{r}_k}\right).
\end{aligned}
\tag{E8}
$$

因此，相对于 μ 的差为

$$\Delta_\mu \left\{ \mathbb{E}\left[|w_{\boldsymbol{s},k^{(+1)}}|^2 |w_{\boldsymbol{r},k^{(\mu)}}|^2 \right] + \mathbb{E}\left[w_{\boldsymbol{s},k^{(+1)}} w^*_{\boldsymbol{r},k^{(+1)}} w_{\boldsymbol{r},k^{(\mu)}} w^*_{\boldsymbol{s},k^{(\mu)}} \right] \right\} = \frac{4}{8^L} \prod_{\ell=1,\ell\neq k}^{L} (1 + 2\delta_{\mathbf{d}\boldsymbol{s}_\ell,\mathbf{d}\boldsymbol{r}_\ell}) \delta_{\mathbf{d}\boldsymbol{s}_k,\mathbf{d}\boldsymbol{r}_k}. \tag{E9}$$

注意，上述最后一个δ函数仅当 $\mathbf{d}\boldsymbol{s}_k = \mathbf{d}\boldsymbol{r}_k$ 时非零；由于 N_T 是根据命题 8 满足 $\mathbf{d}\boldsymbol{s}_\ell = \mathbf{d}\boldsymbol{r}_\ell$ 的元素总数，因此有 $N_T - 1$ 个元素满足 $\mathbf{d}\boldsymbol{s}_\ell = \mathbf{d}\boldsymbol{r}_\ell$，且 $\ell \neq k$，这使得公式 (E9) 等于 $4 \cdot 3^{N_T-1}/8^L$，其概率为 $\binom{L-1}{N_T-1}/(2^{L-1}-1)$（根据命题 8）。对所有 $\boldsymbol{s} \neq \boldsymbol{r}$ 的差值求和为

$$\begin{aligned} &\sum_{\boldsymbol{s}\neq\boldsymbol{r}} \Delta_\mu \left\{ \mathbb{E}\left[|w_{\boldsymbol{s},k^{(+1)}}|^2 |w_{\boldsymbol{r},k^{(\mu)}}|^2 \right] + \mathbb{E}\left[w_{\boldsymbol{s},k^{(+1)}} w^*_{\boldsymbol{r},k^{(+1)}} w_{\boldsymbol{r},k^{(\mu)}} w^*_{\boldsymbol{s},k^{(\mu)}} \right] \right\} \\ &= 2^{L-1}\left(2^{L-1}-1\right) \sum_{\substack{N_T=0 \\ \mathbb{P}(N_T)=\mathbb{P}(L)}}^{L-2} \frac{4 \cdot 3^{N_T-1}}{8^L} \frac{\binom{L-1}{N_T-1}}{2^{L-1}-1} = \frac{1}{4} - \frac{2 \cdot 3^{L-1} - 2^{L-1}}{4^L}. \end{aligned} \tag{E10}$$

对于 S_2 中位移 $\boldsymbol{\beta}$ 的平均，注意，对于任意两个 $\boldsymbol{s}\cdot\boldsymbol{\beta}$，$\boldsymbol{r}\cdot\boldsymbol{\beta}$，若满足 $\boldsymbol{s} \neq \boldsymbol{r}$，总可以将它们写成 $\boldsymbol{s}\cdot\boldsymbol{\beta} = \alpha_z + \alpha_{1-z}$ 和 $\boldsymbol{r}\cdot\boldsymbol{\beta} = \alpha_z - \alpha_{1-z}$，其中 α_z，α_{1-z} 是服从分布 $\mathcal{N}^{\mathrm{C}}_{zE}$，$\mathcal{N}_{\mathrm{C}(1-z)E}$ 的复高斯变量。注意，$z = \ell/L$，其中 ℓ 是范围为 $[1, L-1]$ 内的整数。因此，位移平均可以简化为

$$\mathbb{E}_{\boldsymbol{\beta}}\left[|\langle\psi|B_{\boldsymbol{s}}\rangle|^2 |\langle\psi|B_{\boldsymbol{r}}\rangle|^2 \right] = \mathbb{E}_{\boldsymbol{\beta}}\left[|\langle\psi|\boldsymbol{s}\cdot\boldsymbol{\beta}\rangle|^2 |\langle\psi|\boldsymbol{r}\cdot\boldsymbol{\beta}\rangle|^2 \right] = \mathbb{E}_{\alpha_y\sim\mathcal{N}^{\mathrm{C}}_{yE}} \left[\prod_{h=0}^{1} |\langle\psi|\alpha_z + (-1)^h \alpha_{1-z}\rangle|^2 \right] \equiv C_2(z), \tag{E11}$$

其中，在第二个等式中，我们对满足 $\alpha_z \sim \mathcal{N}^{\mathrm{C}}_{zE}$ 和 $\alpha_{1-z} \sim \mathcal{N}_{\mathrm{C}(1-z)E}$ 的独立变量 α_z 和 α_{1-z} 取平均，为简便起见记作 $\alpha_y \sim \mathcal{N}_{\mathrm{C}yE}$。由式 (E10) 和 (E11)，我们可以得到 S_2 的界：

$$\begin{aligned} S_2 &\geq \sum_{\boldsymbol{s}\neq\boldsymbol{r}} \Delta_\mu \left\{ \mathbb{E}\left[|w_{\boldsymbol{s},k^{(+1)}}|^2 |w_{\boldsymbol{r},k^{(\mu)}}|^2 \right] + \mathbb{E}\left[w_{\boldsymbol{s},k^{(+1)}} w^*_{\boldsymbol{r},k^{(+1)}} w_{\boldsymbol{r},k^{(\mu)}} w^*_{\boldsymbol{s},k^{(\mu)}} \right] \right\} \min \mathbb{E}\left[|\langle\psi|B_{\boldsymbol{s}}\rangle|^2 |\langle\psi|B_{\boldsymbol{r}}\rangle|^2 \right] \\ &= \left(\frac{1}{4} - \frac{2 \cdot 3^{L-1} - 2^{L-1}}{4^L} \right) \min_\ell C_2\left(\frac{\ell}{L}\right), \end{aligned} \tag{E12}$$

$$\begin{aligned} S_2 &\leq \sum_{\boldsymbol{s}\neq\boldsymbol{r}} \Delta_\mu \left\{ \mathbb{E}\left[|w_{\boldsymbol{s},k^{(+1)}}|^2 |w_{\boldsymbol{r},k^{(\mu)}}|^2 \right] + \mathbb{E}\left[w_{\boldsymbol{s},k^{(+1)}} w^*_{\boldsymbol{r},k^{(+1)}} w_{\boldsymbol{r},k^{(\mu)}} w^*_{\boldsymbol{s},k^{(\mu)}} \right] \right\} \max \mathbb{E}\left[|\langle\psi|B_{\boldsymbol{s}}\rangle|^2 |\langle\psi|B_{\boldsymbol{r}}\rangle|^2 \right] \\ &= \left(\frac{1}{4} - \frac{2 \cdot 3^{L-1} - 2^{L-1}}{4^L} \right) \max_\ell C_2\left(\frac{\ell}{L}\right), \end{aligned} \tag{E13}$$

其中，最小化和最大化是在所有整数 $\ell \in [1, L-1]$ 上进行的。

求和 S_3 涉及四个不同的符号向量，尽管求和项总数约为 16^{L-1}，其中很大一部分可以通过对相位取平均而排除。要使 $\mathbb{E}\left[w_{\boldsymbol{s},k^{(+1)}} w^*_{\boldsymbol{s}',k^{(+1)}} w_{\boldsymbol{r},k^{(\mu)}} w^*_{\boldsymbol{r}',k^{(\mu)}} \right]$ 非零，$\Phi_{\boldsymbol{s}} - \Phi_{\boldsymbol{s}'} + \Phi_{\boldsymbol{r}} - \Phi_{\boldsymbol{r}'}$ 只能是一个与任何 ϕ_ℓ 无关的常数。根据式（C5），这要求每个 ϕ_ℓ 对应的系数都为零，

$$(\mathbf{d}\boldsymbol{s}_\ell - 1)\,\boldsymbol{s}_\ell - (\mathbf{d}\boldsymbol{s}'_\ell - 1)\,\boldsymbol{s}'_\ell + (\mathbf{d}\boldsymbol{r}_\ell - 1)\,\boldsymbol{r}_\ell - (\mathbf{d}\boldsymbol{r}'_\ell - 1)\,\boldsymbol{r}'_\ell = 0. \tag{E14}$$

N注意，$\mathbf{d}\boldsymbol{s}_\ell = |\boldsymbol{s}_\ell - \boldsymbol{s}_{\ell-1}|/2 = (1 - \boldsymbol{s}_\ell \boldsymbol{s}_{\ell-1})/2$，那么上述约束可以化简为 o

$$\begin{aligned} &\boldsymbol{s}_\ell(1 - \boldsymbol{s}_\ell\boldsymbol{s}_{\ell-1}) - \boldsymbol{s}'_\ell(1 - \boldsymbol{s}'_\ell\boldsymbol{s}'_{\ell-1}) + \boldsymbol{r}_\ell(1 - \boldsymbol{r}_\ell\boldsymbol{r}_{\ell-1}) - \boldsymbol{r}'_\ell(1 - \boldsymbol{r}'_\ell\boldsymbol{r}'_{\ell-1}) = 2(\boldsymbol{s}_\ell - \boldsymbol{s}'_\ell + \boldsymbol{r}_\ell - \boldsymbol{r}'_\ell) \\ &\Rightarrow \boldsymbol{s}_\ell - \boldsymbol{s}'_\ell + \boldsymbol{r}_\ell - \boldsymbol{r}'_\ell = -(\boldsymbol{s}_{\ell-1} - \boldsymbol{s}'_{\ell-1} + \boldsymbol{r}_{\ell-1} - \boldsymbol{r}'_{\ell-1}), \end{aligned} \tag{E15}$$

其中，我们使用了 $\boldsymbol{s}_\ell^2 = \boldsymbol{s}_\ell'^2 = \boldsymbol{r}_\ell^2 = \boldsymbol{r}_\ell'^2 = 1$ 得到了最后一行。由于 $\boldsymbol{s}_L - \boldsymbol{s}'_L + \boldsymbol{r}_L - \boldsymbol{r}'_L = 0$，上述约束变为

$$\boldsymbol{s}_\ell + \boldsymbol{r}_\ell - \boldsymbol{s}'_\ell - \boldsymbol{r}'_\ell = 0, \quad \forall \ell \in [1, L] \cap \mathbb{N}. \tag{E16}$$

在表 II 中，我们列出了 $\boldsymbol{s}_\ell$，$\boldsymbol{s}'_\ell$，$\boldsymbol{r}_\ell$，$\boldsymbol{r}'_\ell$ 的所有组合（符号的整体反转等价），并检验是否满足约束方程（E16）。例如，将 $\boldsymbol{s}_\ell$，$\boldsymbol{s}'_\ell$，$\boldsymbol{r}_\ell$，$\boldsymbol{r}'_\ell$ 的所有符号整体反转后，仅有 $\boldsymbol{s}_\ell = -1$ 且仅有

$\boldsymbol{s}_\ell$	$\boldsymbol{s}'_\ell$	$\boldsymbol{r}_\ell$	$\boldsymbol{r}'_\ell$	是否满足公式（E16）
-1	-1	-1	-1	Yes
-1	-1	-1	$+1$	No
-1	-1	$+1$	-1	No
-1	$+1$	-1	-1	No
$+1$	-1	-1	-1	No
-1	-1	$+1$	$+1$	Yes
-1	$+1$	-1	$+1$	No
$+1$	-1	-1	$+1$	Yes

表 II。对 $\boldsymbol{s}_\ell$、$\boldsymbol{s}'_\ell$、$\boldsymbol{r}_\ell$、$\boldsymbol{r}'_\ell$ 的所有可能组合进行式 (E16) 的可满足性检验，其中 $1 \le \ell \le L-1$ 最多进行符号的整体反转。

$\boldsymbol{s}_\ell = +1$ 得出相同的可满足性结果。根据表 II 的总结，$\boldsymbol{s}_\ell, \boldsymbol{s}'_\ell, \boldsymbol{r}_\ell, \boldsymbol{r}'_\ell$仅存在三种允许的组合，因此 $\boldsymbol{s}\cdot\boldsymbol{\beta}$, $\boldsymbol{s}'\cdot\boldsymbol{\beta}$, $\boldsymbol{r}\cdot\boldsymbol{\beta}$, $\boldsymbol{r}'\cdot\boldsymbol{\beta}$的划分为

$$\begin{aligned}
\boldsymbol{s}\cdot\boldsymbol{\beta} &= \alpha_z + \alpha_{\tilde{z}} + \alpha_{1-z-\tilde{z}} \\
\boldsymbol{s}'\cdot\boldsymbol{\beta} &= \alpha_z + \alpha_{\tilde{z}} - \alpha_{1-z-\tilde{z}} \\
\boldsymbol{r}\cdot\boldsymbol{\beta} &= \alpha_z - \alpha_{\tilde{z}} - \alpha_{1-z-\tilde{z}} \\
\boldsymbol{r}'\cdot\boldsymbol{\beta} &= \alpha_z - \alpha_{\tilde{z}} + \alpha_{1-z-\tilde{z}},
\end{aligned} \tag{E17}$$

其中 $\alpha_z \sim \mathcal{N}_{zE}^{\mathrm{C}}$ 是高斯变量， $\alpha_{\tilde{z}}$ 和 $\alpha_{1-z-\tilde{z}}$ 同理。由于 $\boldsymbol{s}$， $\boldsymbol{r}$， $\boldsymbol{s}'$， $\boldsymbol{r}'$ 不同， z， $\tilde{z}$ 仅限于 $z=\ell_1/L$， $\tilde{z}=\ell_2/L$，其中 ℓ_1, $\ell_2 \in [1, L-2]\cap\mathbb{N}$ 和 $\ell_1+\ell_2 \le L-1$。

S_3 中对位移取平均为

$$\begin{aligned}
&\mathbb{E}\left[\langle\psi|B_{\boldsymbol{s}}\rangle\langle B_{\boldsymbol{s}'}|\psi\rangle\langle\psi|B_{\boldsymbol{r}}\rangle\langle B_{\boldsymbol{r}'}|\psi\rangle\right] \\
&= \mathbb{E}_{\alpha_y\sim\mathcal{N}_{yE}^{\mathrm{C}}}\left[e^{i(\chi_{\boldsymbol{s}}-\chi_{\boldsymbol{s}'}+\chi_{\boldsymbol{r}}-\chi_{\boldsymbol{r}'})}\prod_{a=0}^{1}\langle\psi|\alpha_z+(-1)^a\alpha_{\tilde{z}}+(-1)^a\alpha_{1-z-\tilde{z}}\rangle\langle\alpha_z+(-1)^a\alpha_{\tilde{z}}-(-1)^a\alpha_{1-z-\tilde{z}}|\psi\rangle\right] \qquad \text{(E18)} \\
&\le \mathbb{E}_{\alpha_y\sim\mathcal{N}_{yE}^{\mathrm{C}}}\left[\prod_{a=0}^{1}|\langle\psi|\alpha_z+(-1)^a\alpha_{\tilde{z}}+(-1)^a\alpha_{1-z-\tilde{z}}\rangle||\langle\alpha_z+(-1)^a\alpha_{\tilde{z}}-(-1)^a\alpha_{1-z-\tilde{z}}|\psi\rangle|\right] \equiv C_3(z,\tilde{z}), \qquad \text{(E19)}
\end{aligned}$$

其中，在最后一行中，我们将其上界限定为 $\mathbb{E}[x] \le \mathbb{E}[|x|]$ 。对于所考虑的每一类态，我们将证明， $C_3(z,\ \tilde{z})$ 相较于公式（E6）和（E11）会产生更高阶项，因此在后续讨论的 E 渐近区域中可以忽略。

总之，公式（E3）中梯度方差的上下界为

$$\begin{aligned}
\mathrm{Var}\left[\partial_{\theta_k}\mathcal{C}(\boldsymbol{x})\right] &= \frac{1}{2}\left[\frac{3^{L-1}}{4^L}C_1 + \left(\frac{1}{4} - \frac{2\cdot 3^{L-1}-2^{L-1}}{4^L}\right)\min_\ell C_2\left(\frac{\ell}{L}\right)\right] + \mathcal{O}\left(C_2\right) \\
&\ge \frac{1}{2}\left[\frac{3^{L-1}}{4^L}C_1 + \left(\frac{1}{4} - \frac{3^L}{4^L}\right)\min_\ell C_2\left(\frac{\ell}{L}\right)\right] + \mathcal{O}\left(C_2\right), \qquad \text{(E20)}
\end{aligned}$$

and

$$\begin{aligned}
\mathrm{Var}\left[\partial_{\theta_k}\mathcal{C}(\boldsymbol{x})\right] &= \frac{1}{2}\left[\frac{3^{L-1}}{4^L}C_1 + \left(\frac{1}{4} - \frac{2\cdot 3^{L-1}-2^{L-1}}{4^L}\right)\max_\ell C_2\left(\frac{\ell}{L}\right)\right] + \mathcal{O}\left(C_2\right) \\
&\le \frac{1}{2}\left[\frac{3^{L-1}}{4^L}C_1 + \left(\frac{1}{4} + \frac{2^{L-1}}{4^L}\right)\max_\ell C_2\left(\frac{\ell}{L}\right)\right] + \mathcal{O}\left(C_2\right), \qquad \text{(E21)}
\end{aligned}$$

其中，最小化和最大化均针对整数 $1 \le \ell \le L-1$ 进行。上述界中的关联函数 C_1， C_2的具体表达式取决于特定的目标态。下文将针对不同的目标态 $|\psi\rangle_m$ 计算这些关联函数，以深入了解其渐近行为，从而了解梯度的行为。在小节 E 1 中，我们考虑单模高斯态；在小节 E 2 中，我们考虑福克数态。

单模高斯态

假设目标量子模式态是任意高斯（纯）态，由于关联函数 C_1 和 C_2 只取决于目标态 $|\psi\rangle_m$ 与相干态之间的保真度，借助参考文献 [87–89]，即可进行解析计算。

关于高斯态的简要介绍，请参阅附录 A。

由公式 (E6) 和 (E11) 可见，我们需要计算的关联函数仅取决于目标态 $|\psi\rangle_m$ 与相干态之间的保真度。我们首先计算公式 (A3) 中的高斯态与公式 (A8) 中的相干态 $|\alpha\rangle$ 之间的保真度。

引理 12 任意单模态高斯态 $|\psi\rangle_m = D(\gamma)R(\tau)S(\zeta)\,|0\rangle_m$ 与相干态 $|\alpha\rangle_m = D(\alpha)\,|0\rangle_m$ 之间的保真度为

$$F(\psi,|\alpha\rangle) = \operatorname{sech}(\zeta)e^{-(1+\kappa_1)(\mathrm{Re}\{\gamma\}-\mathrm{Re}\{\alpha\})^2-(1-\kappa_1)(\mathrm{Im}\{\gamma\}-\mathrm{Im}\{\alpha\})^2+2\kappa_2(\mathrm{Re}\{\gamma\}-\mathrm{Re}\{\alpha\})(\mathrm{Im}\{\gamma\}-\mathrm{Im}\{\alpha\})} \tag{E22}$$

其中 $\kappa_1 \equiv \cos(2\tau)\tanh(\zeta)$ 且 $\kappa_2 \equiv \sin(2\tau)\tanh(\zeta)$。

关联函数 C_1 就是保真度的平方，即

$$C_1^{\mathrm{Gauss}} \equiv \mathbb{E}_{\alpha\sim\mathcal{N}_E^{\mathbb{C}}}\left[|\langle\psi|\alpha\rangle|^4\right] = \mathbb{E}_{\alpha\sim\mathcal{N}_E^{\mathbb{C}}}\left[F(\psi,|\alpha\rangle)^2\right] = \frac{\operatorname{sech}^2(\zeta)e^{-R(E)/G_1(E)}}{\sqrt{G_1(E)}}, \tag{E23}$$

其中最后一行是 分别对 α 的实部和虚部取平均值获得。这里我们定义了 ne

$$G_1(x) = 1 + 4x + 4\operatorname{sech}^2(\zeta)x^2 \tag{E24}$$

$$R(x) = 2|\gamma|^2 + 4\operatorname{sech}^2(\zeta)|\gamma|^2 x + 2\tanh(\zeta)|\gamma|^2\cos(2(\varphi+\tau)), \tag{E25}$$

其中，$\varphi = \arctan(\mathrm{Im}\{\gamma\}/\mathrm{Re}\{\gamma\})$ 是复数 γ 的辐角。在 E 的渐近区域中，可以看到 $C_1^{\mathrm{Gauss}} \sim 1/2E$。对于 $\zeta = 0$, $\tau = 0$ 的相干态和 $\gamma = 0, \tau = 0$ 的单模压缩真空态（SMSV），上述 C_1^{Gauss} 可简化为

$$C_1^{\mathrm{Coh}} = \frac{e^{-2|\gamma|^2/(1+2E)}}{1+2E}, \tag{E26}$$

$$C_1^{\mathrm{SMSV}} = \frac{\operatorname{sech}^2(\zeta)}{\sqrt{1+4E+4\operatorname{sech}^2(\zeta)E^2}}. \tag{E27}$$

The 关联函数 C_2 是 ψ 与相干态 $|\alpha_z \pm \alpha_{1-z}\rangle$ a 之间保真度的乘积 s

$$\begin{aligned} C_2^{\mathrm{Gauss}}(z) &\equiv \mathbb{E}_{\alpha_y\sim\mathcal{N}_{yE}^{\mathbb{C}}}\left[\prod_{h=0}^{1}|\langle\psi|\alpha_z+(-1)^h\alpha_{1-z}\rangle|^2\right] \\ &= \mathbb{E}_{\alpha_y\sim\mathcal{N}_{yE}^{\mathbb{C}}}\left[\prod_{h=0}^{1}F(\psi,|\alpha_z+(-1)^h\alpha_{1-z}\rangle)\right] = \frac{\operatorname{sech}^2(\zeta)e^{-R(zE)/G_1(zE)}}{\sqrt{G_1(E-zE)G_1(zE)}}. \end{aligned} \tag{E28}$$

在 E的渐近区域中，我们还看到$C_2^{\mathrm{Gauss}}(z) \sim 1/4z(1-z)E^2$。对于相干态和单模压缩真空态，我们还具有

$$C_2^{\mathrm{Coh}}(z) = \frac{-2|\gamma|^2/(1+2zE)}{[1+2(1-z)E](1+2zE)}, \tag{E29}$$

$$C_2^{\mathrm{SMSV}}(z) = \frac{\operatorname{sech}^2(\zeta)}{\sqrt{G_1(E-zE)G_1(zE)}}. \tag{E30}$$

对于 C_3，我们有

$$\begin{aligned} C_3^{\mathrm{Gauss}}(z,\tilde{z}) &= \mathbb{E}_{\alpha_y\sim\mathcal{N}_{yE}^{\mathbb{C}}}\left[\prod_{h=0}^{1}|\langle\psi|\alpha_z+(-1)^h\alpha_{\tilde{z}}+(-1)^h\alpha_{1-z-\tilde{z}}\rangle||\langle\alpha_z+(-1)^h\alpha_{\tilde{z}}-(-1)^h\alpha_{1-z-\tilde{z}}|\psi\rangle|\right] \\ &= \mathbb{E}_{\alpha_y\sim\mathcal{N}_{yE}^{\mathbb{C}}}\left[\prod_{h=0}^{1}\sqrt{F(\psi,|\alpha_z+(-1)^h\alpha_{\tilde{z}}+(-1)^h\alpha_{1-z-\tilde{z}}\rangle)}\sqrt{F(\psi,|\alpha_z+(-1)^h\alpha_{\tilde{z}}-(-1)^h\alpha_{1-z-\tilde{z}}\rangle)}\right] \\ &= \frac{\operatorname{sech}^2(\zeta)e^{-R(zE)/G_1(zE)}}{\sqrt{G_1(zE)G_1(\tilde{z}E)G_1[(1-z-\tilde{z})E]}}, \end{aligned} \tag{E31}$$

这显然趋近于 $1/E^3$ 在 E的渐近区域中的标度，因此可以省略。我们预计 t

其标度也可以推广到其他非高斯态，不过求解可能并不容易。制备任意高斯态时，梯度方差的界为

$$\mathrm{Var}\left[\partial_{\theta_k}\mathcal{C}(\boldsymbol{x})\right] \geq \frac{1}{2}\left[\frac{3^{L-1}}{4^L}C_1^{\mathrm{Gauss}} + \left(\frac{1}{4}-\frac{3^L}{4^L}\right)\min_{\ell} C_2^{\mathrm{Gauss}}\left(\frac{\ell}{L}\right)\right] + \mathcal{O}\left(\frac{1}{E^3}\right), \tag{E32}$$

$$\mathrm{Var}\left[\partial_{\theta_k}\mathcal{C}(\boldsymbol{x})\right] \leq \frac{1}{2}\left[\frac{3^{L-1}}{4^L}C_1^{\mathrm{Gauss}} + \left(\frac{1}{4}+\frac{2^{L-1}}{4^L}\right)\max_{\ell} C_2^{\mathrm{Gauss}}\left(\frac{\ell}{L}\right)\right] + \mathcal{O}\left(\frac{1}{E^3}\right), \tag{E33}$$

其中，最小化和最大化是在所有整数 $1 \leq \ell \leq L-1$ 上进行的。在 E 的渐近区域内，由于 C_1^{Gauss} 和 C_2^{Gauss} 呈现不同的标度，我们可以发现，当.....时，梯度方差由 $1/E$ 主导

$$\frac{1/4E^2}{(3/4)^L/6E} \in \mathcal{O}(1) \Rightarrow E \in \Omega(1)\frac{3}{2}\left(\frac{4}{3}\right)^L \in \Omega(\exp L), \tag{E34}$$

或者等价地,

$$L \in \frac{1}{\log(4/3)}\log\left(\mathcal{O}(1)\frac{2E}{3}\right) \in \mathcal{O}(\log E). \tag{E35}$$

根据能　　量，我们可以将渐近 E 区域内的连续变量变分量子电路分为浅层电路和深层电路　　s.

当线路浅至 $L \in \mathcal{O}(\log E)$ 时，梯度方差的界由关联函数 C_1^{Gauss} 的首个 $\sim 1/E$ 项主导；二者相同，因此梯度方差可表示为

$$\mathrm{Var}\left[\partial_{\theta_k}\mathcal{C}\right] = \frac{1}{6}\left(\frac{3}{4}\right)^L C_1^{\mathrm{Gauss}} + \mathcal{O}\left(\frac{1}{E^2}\right). \tag{E36}$$

在制备非零均值高斯态（即相干态）时，上述主导阶会在约为 $E \sim |\gamma|^2 = E_t$ （即目标态能量）处带来一个方差峰值，并且当$E \gtrsim E_t$时，方差呈现出 $1/E$ 的标度。而对于零均值高斯态（即单模压缩真空态），式 (E27) 随着 E 单调递减；可以估计，当 $E \geq \cosh(\zeta) = \sqrt{1+E_t}$ 时，梯度方差趋近于 $1/E$ 的标度。

另一方面，当线路深度达到 $L \in \Omega(\log E)$ 时，方差界由关联函数 C_2^{Gauss} 中的第二个 $\sim 1/E^2$ 项决定，其界为

$$\frac{1}{8}\min_{\ell} C_2^{\mathrm{Gauss}}\left(\frac{\ell}{L}\right) \leq \mathrm{Var}\left[\partial_{\theta_k}\mathcal{C}\right] \leq \frac{1}{8}\max_{\ell} C_2^{\mathrm{Gauss}}\left(\frac{\ell}{L}\right). \tag{E37}$$

在 E的渐近区域中，两侧都遵循 $1/E^2$ 的标度，梯度方差本身也如此。对于相干态这类非零均值高斯态，通过求解式 (E29) 的极值点，也会得到一个方差峰值，其位于 $[E_t/2, E_t]$ 范围内。超过该范围后，方差开始趋近于 $1/E^2$。然而，对于单模压缩真空态这类零均值高斯态，式 (E30) 中的 $C_2^{\mathrm{SMSV}}(z)$ 只会随着 E 递减；通过比较分母中涉及 $E^{3/2}$ 和 E^2 的项，可以估计 $1/E^2$ 的标度从 $E \sim L\cosh\zeta = L\sqrt{1+E_t}$ 开始。

2. 福克数态

对于非高斯态，保真度的评估通常较为困难。本部分中，我们考虑制备一种保真度具有闭式表达的福克数态，以便从物理角度理解制备非高斯态时方差的标度。福克数态构成希尔伯特空间的一组完备正交基。在本节中，我们将数态记为$|E_t\rangle^{\mathrm{F}}$。我们先介绍关于高斯复变量 $\alpha \sim \mathcal{N}_{\sigma^2}^{\mathrm{C}}$ 的模长和复辐角分布的引理。由于这些结果广为人知，我们仅陈述结果。

引理 13 给定一个复高斯分布随机变量 $\alpha \sim \mathcal{N}_{\sigma^2}^{\mathrm{C}}$，其模的平方服从伽马分布 $|\alpha|^2 \sim \mathrm{Gamma}(1,\sigma^2)$，其概率密度函数为 $p(|\alpha|^2) = e^{-|\alpha|^2/\sigma^2}/\sigma^2$。辐角$\arg\{\alpha\} \equiv \tan^{-1}(\mathrm{Im}\{\alpha\}/\mathrm{Re}\{\alpha\})$ 在 $[-\pi,\pi]$ 上均匀分布。

还可以进一步发现，来自同一系综的两个复高斯变量 α_1、α_2 的辐角之差满足如下三角分布。

推论14 对于复高斯变量 $\alpha_i \in \mathcal{N}_{\mathrm{C}}{}^{\sigma_i^2}$，其中 $i=1$、2，其辐角之差 $\delta \equiv \arg\{\alpha_1\} - \arg\{\alpha_2\}$ 服从三角分布 $\delta \sim \mathrm{Tri}(-2\pi, 2\pi, 0)$，分布为 $p(\delta) = (2\pi - |\delta|)/4\pi^2$。

有了这些引理，我们可以得到关联函数 C_1 as

$$C_1^{\mathrm{Fock}} \equiv \mathbb{E}_{\alpha\sim\mathcal{N}_E^{\mathrm{C}}}\left[|^{\mathrm{F}}\langle E_t\,|\alpha\rangle\,|^4\right] = \mathbb{E}_{\alpha\sim\mathcal{N}_E^{\mathrm{C}}}\left[\frac{e^{-2|\alpha|^2}}{(E_t!)^2}|\alpha|^{4E_t}\right] = \mathbb{E}_{|\alpha|^2\sim\mathrm{Gamma}(1,E)}\left[\frac{e^{-2|\alpha|^2}}{(E_t!)^2}|\alpha|^{4E_t}\right] = \frac{(2E_t)!}{(2^{E_t}E_t!)^2}\frac{(1+1/2E)^{-2E_t}}{1+2E}. \tag{E38}$$

同样，关联函数 C_2 变为

$$\begin{aligned}
C_2^{\mathrm{Fock}}(x) &\equiv \mathbb{E}_{\alpha_y\sim\mathcal{N}_{yE}^{\mathrm{C}}}\left[\prod_{h=0}^{1}|^{\mathrm{F}}\langle E_t\,|\alpha_z + (-1)^h\alpha_{1-z}\rangle\,|^2\right] = \mathbb{E}_{\alpha_y\sim\mathcal{N}_{yE}^{\mathrm{C}}}\left[\frac{e^{-|\alpha_z+\alpha_{1-z}|^2}}{E_t!}|\alpha_z+\alpha_{1-z}|^{2E_t}\frac{e^{-|\alpha_z-\alpha_{1-z}|^2}}{E_t!}|\alpha_z-\alpha_{1-z}|^{2E_t}\right] \\
&= \frac{1}{(E_t!)^2}\mathbb{E}_{|\alpha_y|^2\sim\mathrm{Gamma}(1,yE),\delta\sim\mathrm{Tri}(-2\pi,2\pi,0)}\left[e^{-2|\alpha_z|^2-2|\alpha_{1-z}|^2}\left(|\alpha_z|^4+|\alpha_{1-z}|^4+2|\alpha_z|^2|\alpha_{1-z}|^2-4|\alpha_z|^2|\alpha_{1-z}|^2\cos^2\delta\right)^{E_t}\right] \\
&= \frac{1}{(E_t!)^2}\mathbb{E}_{|\alpha_y|^2\sim\mathrm{Gamma}(1,yE)}\left[e^{-2(|\alpha_z|^2+|\alpha_{1-z}|^2)}\mathbb{E}_{\delta\sim\mathrm{Tri}(-2\pi,2\pi,0)}\left[\left(|\alpha_z|^4+|\alpha_{1-z}|^4-2|\alpha_z|^2|\alpha_{1-z}|^2\cos(2\delta)\right)^{E_t}\right]\right] \\
&= \frac{1}{(E_t!)^2}\mathbb{E}_{|\alpha_y|^2\sim\mathrm{Gamma}(1,yE)}\left[e^{-2(|\alpha_z|^2+|\alpha_{1-z}|^2)}\left(|\alpha_z|^2+|\alpha_{1-z}|^2\right)^{2E_t}{}_2F_1\left(\frac{1}{2},-E_t,1,\frac{4|\alpha_z|^2|\alpha_{1-z}|^2}{(|\alpha_z|^2+|\alpha_{1-z}|^2)^2}\right)\right],
\end{aligned} \tag{E39}$$

其中，${}_2F_1(a,\ b,\ c,\ z)$ 是超几何函数。上面的 $|\alpha_z|^2,\ |\alpha_{1-z}|^2$ 积分难以求解，但注意到 $0 \le 4|\alpha_z|^2|\alpha_{1-z}|^2/(|\alpha_z|^2+|\alpha_{1-z}|^2)^2 \le 1$,

$${}_2F_1\left(\frac{1}{2},-E_t,1,1\right) \le {}_2F_1\left(\frac{1}{2},-E_t,1,\frac{4|\alpha_z|^2|\alpha_{1-z}|^2}{(|\alpha_z|^2+|\alpha_{1-z}|^2)^2}\right) \le 1, \tag{E40}$$

其中，左端（L.H.S.）是一个仅依赖于 E_t 的常数。式 (E39) 中不含超几何函数的系综平均为

$$\begin{aligned}
&\frac{1}{(E_t!)^2}\mathbb{E}_{|\alpha_y|^2\sim\mathrm{Gamma}(1,yE)}\left[e^{-2(|\alpha_z|^2+|\alpha_{1-z}|^2)}\left(|\alpha_z|^2+|\alpha_{1-z}|^2\right)^{2E_t}\right] \\
&= \frac{(2E_t)!}{(2^{E_t}E_t!)^2}\frac{\left[(1-z)(1+2zE)\left(1+\frac{1-2z}{z+2(1-z)zE}\right)^{2E_t}-2(1-z)zE-z\right]}{1-2z}\frac{\left(1+\frac{1}{2zE}\right)^{-2E_t}}{[1+2(1-z)E]\,(1+2zE)}.
\end{aligned} \tag{E41}$$

因此，我们得到关联函数 C_2^{Fock}as

$$C_2^{\mathrm{Fock}}(z) = \eta\frac{(2E_t)!}{(2^{E_t}E_t!)^2}\frac{\left[(1-z)(1+2zE)\left(1+\frac{1-2z}{z+2(1-z)zE}\right)^{2E_t}-2(1-z)zE-z\right]}{1-2z}\frac{\left(1+\frac{1}{2zE}\right)^{-2E_t}}{[1+2(1-z)E]\,(1+2zE)}. \tag{E42}$$

其中，η 在下界中等于 ${}_2F_1(1/2,\ -E_t,\ 1,\ 1)$，在上界中等于 1。在 E 的渐近区域中，中间的长分式可以化简为 $1+2E_t$，因此关联函数变为

$$C_2^{\mathrm{Fock}}(x) = \eta\frac{(1+2E_t)(2E_t)!}{(2^{E_t}E_t!)^2}\frac{(1+1/2zE)^{-2E_t}}{[1+2(1-z)E]\,(1+2zE)}. \tag{E43}$$

福克态的关联函数 C_3 为

$$\begin{aligned}
C_3^{\mathrm{Fock}}(z,\tilde{z}) &= \mathbb{E}_{\alpha_y\sim\mathcal{N}_{yE}^{\mathrm{C}}}\left[\prod_{h=0}^{1}|^{\mathrm{F}}\langle E_t\,|\alpha_z+(-1)^h\alpha_{\tilde{z}}+(-1)^h\alpha_{1-z-\tilde{z}}\rangle\,||\,\langle\alpha_z+(-1)^h\alpha_{\tilde{z}}-(-1)^h\alpha_{1-z-\tilde{z}}|\,E_t\rangle^{\mathrm{F}}|\right] \\
&= \mathbb{E}_{\alpha_y\sim\mathcal{N}_{yE}^{\mathrm{C}}}\left[\prod_{h=0}^{1}\left(\frac{e^{-|\alpha_z+(-1)^h\alpha_{\tilde{z}}+(-1)^h\alpha_{1-z-\tilde{z}}|^2/2}}{\sqrt{E_t!}}|\alpha_z+(-1)^h\alpha_{\tilde{z}}+(-1)^h\alpha_{1-z-\tilde{z}}|^{E_t}\right.\right. \\
&\qquad\left.\left.\times\frac{e^{-|\alpha_z+(-1)^h\alpha_{\tilde{z}}-(-1)^h\alpha_{1-z-\tilde{z}}|^2/2}}{\sqrt{E_t!}}|\alpha_z+(-1)^h\alpha_{\tilde{z}}-(-1)^h\alpha_{1-z-\tilde{z}}|^2\right)\right]
\end{aligned} \tag{E44}$$

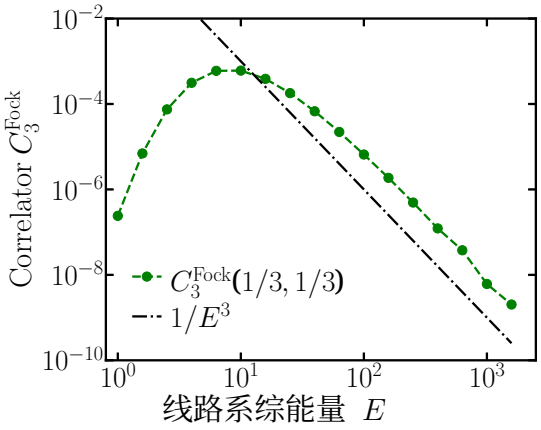


图 8. 式 (E44) 中福克态的关联函数 $C_3^{\mathrm{Fock}}(z,\ \tilde{z})$ 的数值结果。这里我们选择 $z=\tilde{z}=1/3$。点划线为 $1/E^3$，以供参考。

解析求解该问题很困难，因此我们通过蒙特卡洛计算展示其渐近标度，如图 8 所示。这里我们取 $z=\tilde{z}=1/3$。显然，在 E 的渐近区域，$C_3^{\mathrm{Fock}}(z,\ \tilde{z})$ 的标度为 $\sim 1/E^3$ （点划线），这是相对于 C_1^{Fock} 和 $C_2^{\mathrm{Fock}}(z)$ 的高阶项，也可以忽略。

结合式（E38）和（E43），我们得到制备福克态时梯度方差的界为

$$\mathrm{Var}\left[\partial_{\theta_k}\mathcal{C}(\boldsymbol{x})\right] \geq \frac{1}{2}\left[\frac{3^{L-1}}{4^L}C_1^{\mathrm{Fock}} + \left(\frac{1}{4}-\frac{3^L}{4^L}\right)\min_{\ell} C_2^{\mathrm{Fock}}\left(\frac{\ell}{L}\right)\right] + \mathcal{O}\left(\frac{1}{E^3}\right), \tag{E45}$$

$$\mathrm{Var}\left[\partial_{\theta_k}\mathcal{C}(\boldsymbol{x})\right] \leq \frac{1}{2}\left[\frac{3^{L-1}}{4^L}C_1^{\mathrm{Fock}} + \left(\frac{1}{4}+\frac{2^{L-1}}{4^L}\right)C_2^{\mathrm{Fock}}\left(1-\frac{1}{L}\right)\right] + \mathcal{O}\left(\frac{1}{E^3}\right). \tag{E46}$$

与高斯态制备的讨论类似，我们也可以确定在固定 L 时，从 $1/E^2$ 到 $1/E$ 发生标度转变的临界 E。

$$\frac{\eta(1+2E_t)/4E^2}{(3/4)^L/6E} \in \mathcal{O}(1) \Rightarrow E \in \Omega(1)\frac{3\eta(1+2E_t)}{2}\left(\frac{4}{3}\right)^L \in \Omega\left(\exp L\right). \tag{E47}$$

或者等价地，我们有

$$L \in \frac{1}{\log(4/3)}\log\left(\mathcal{O}(1)\frac{2E}{3\eta(1+2E_t)}\right) \in \mathcal{O}(\log E). \tag{E48}$$

当电路深度浅至 $L\in\mathcal{O}(\log E)$ 时，梯度方差的界由关联函数 C_1^{Fock} 的第一$\sim 1/E$ 项主导；两者相同，因此梯度方差可描述为

$$\mathrm{Var}\left[\partial_{\theta_k}\mathcal{C}\right] = \frac{1}{6}\left(\frac{3}{4}\right)^L C_1^{\mathrm{Fock}} + \mathcal{O}\left(\frac{1}{E^2}\right). \tag{E49}$$

在 $E\sim E_t$ 处也可以找到方差峰值。另一方面，当电路深度深至 $L\in\Omega(\log E)$ 时，方差的界由关联函数 C_2^{Fock} 的第二 $\sim 1/E^2$ 项主导，可表示为

$$\frac{1}{8}\min_{\ell} C_2^{\mathrm{Fock}}\left(\frac{\ell}{L}\right) \leq \mathrm{Var}\left[\partial_{\theta_k}\mathcal{C}\right] \leq \frac{1}{8}C_2^{\mathrm{Fock}}\left(1-\frac{1}{L}\right), \tag{E50}$$

其中，不等式左侧和右侧的 C_2^{Fock} 中的系数 η 分别取为 ${}_2F_1(1/2,\ -E_t,\ 1,\ 1)$ 和 1。在渐近区域，方差也遵循标度关系 $1/E^2$，峰值位于区域 $[E_t/2, E_t]$。

附录 F：多模量子模式连续变量态制备中的梯度方差

在本节中，我们将说明附录 E 中单模情形的结果可以推广到任意多模连续变量态 $|\psi\rangle_{\boldsymbol{m}}$ 制备过程中的梯度方差。引理 1（见附录 B）指出，可以

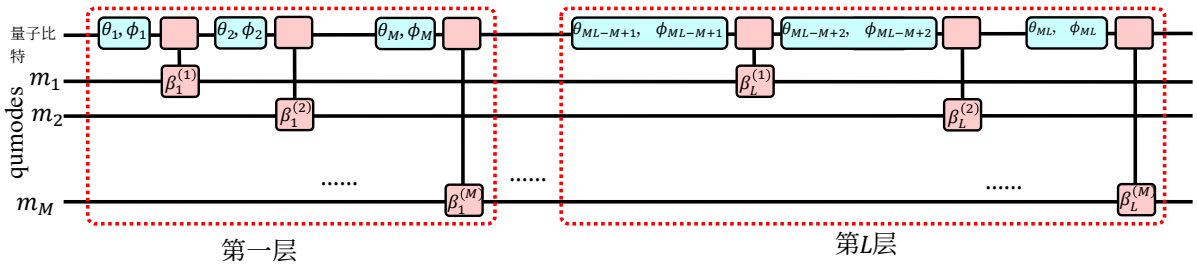


图 9. M 模式、 L 层连续变量变分量子电路示意图。青色方框（ θ_ℓ, ϕ_ℓ 范围为 $1 \le \ell \le ML$ ）表示量子比特旋转U_{R}（θ_ℓ, ϕ_ℓ）；粉色方框（ $\beta_\ell^{(j)}$ ）表示作用于量子比特和第 j模式的 ECD 门 $U^{(j)}{}_{\mathrm{ECD}}$（$\beta_\ell^{(j)}$）。

通过应用 ECD门和单量子比特旋转这一组操作，实现对多个模式和一个量子比特的通用控制。因此，我们考虑量子电路中的梯形结构，如图 9 所示。下文中，我们用上标 in$A^{(j)}$ 表示算符 A；该算符除作用于第 j个量子模式外，对其他量子模式均平凡作用。

我们首先将附录 C 中的态表示推广到多模情形，以此展开分析。为简化符号，我们定义 $\boldsymbol{\beta}^{(j)} = (\beta_1^{(j)}, \ldots, \beta_L^{(j)})^T$, $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_{ML})^T$, $\boldsymbol{\phi} = (\phi_1, \ldots, \phi_{ML})^T$ 以及整体参数$\boldsymbol{x} = (\{\boldsymbol{\beta}^{(j)}\}_{j=1}^M, \boldsymbol{\theta}, \boldsymbol{\phi})$。我们还将 $\boldsymbol{m} = (m_1, \cdots, m_M)$ 记作所有模式。对于一个 M 模式系统，每一层 L都包含 M个单量子比特旋转，以及在控制量子比特与模式 m_j之间施加的 ECD 门，模式范围从 $j = 1$到M （见图 9 中红色虚线框所圈出的门集合）。深度为 L的线路中，相应的变分参数为 $\{\theta_\ell\}_{\ell=1}^{ML}$、 $\{\phi_\ell\}_{\ell=1}^{ML}$和 $\bigcup_{j=1}^M \{\beta_\ell^{(j)}\}_{\ell=1}^L$，上标（$j$）表示第 j个模式，如上文所述。图 9 中 M 模式、深度为 L的线路的酉算符为

$$
\begin{aligned}
U &= \prod_{\ell=1}^{L} \prod_{j=1}^{M} U_{\mathrm{ECD}}^{(j)}\left(\beta_\ell^{(j)}\right) U_{\mathrm{R}}(\theta_{M(\ell-1)+j}, \phi_{M(\ell-1)+j}) \\
&= \prod_{\ell=1}^{L} \prod_{j=1}^{M} \begin{pmatrix} e^{i\phi_{M(\ell-1)+j}} \sin\frac{\theta_{M(\ell-1)+j}}{2} D^{(j)}(-\beta_\ell^{(j)}) & \cos\frac{\theta_{M(\ell-1)+j}}{2} D^{(j)}(-\beta_\ell^{(j)}) \\ \cos\frac{\theta_{M(\ell-1)+j}}{2} D^{(j)}(\beta_\ell^{(j)}) & e^{i(\pi-\phi_{M(\ell-1)+j})} \sin\frac{\theta_{M(\ell-1)+j}}{2} D^{(j)}(\beta_\ell^{(j)}) \end{pmatrix},
\end{aligned} \tag{F1}
$$

其中，上述分块矩阵表示取自公式 (C1)。注意，位移算符 $D^{(j)}(\cdot)$ 作用于所有 M 模式，其中第 j 个模式具有位移，其余模式均为平凡的恒等算符。酉算符 $U_L^{(M)}$ 作用于初始态 $|0\rangle_q \otimes_{j=1}^M |0\rangle_{m_j}$ 后得到的输出态为

$$
\begin{aligned}
&|\psi(\boldsymbol{\theta}, \boldsymbol{\phi}, \{\boldsymbol{\beta}^{(j)}\}_{j=1}^M)\rangle_{q,\boldsymbol{m}} \equiv U\left(|0\rangle_q \otimes_{j=1}^M |0\rangle_{m_j}\right) \\
&= \sum_{a=0}^{1} \sum_{\boldsymbol{s}} w_{\boldsymbol{s},a}(\boldsymbol{\theta}, \boldsymbol{\phi}) |a\rangle_q e^{i\sum_{j=1}^M \chi_{\boldsymbol{s}^{(j)}}(\boldsymbol{\beta}^{(j)})} \bigotimes_{j=1}^{M} |(-1)^a \boldsymbol{s}^{(j)} \cdot \boldsymbol{\beta}^{(j)}\rangle_{m_j} = \sum_{a=0}^{1} \sum_{\boldsymbol{s}} w_{\boldsymbol{s},a}(\boldsymbol{\theta}, \boldsymbol{\phi}) |a\rangle_q |\boldsymbol{B}_{\boldsymbol{s},a}\rangle_{\boldsymbol{m}},
\end{aligned} \tag{F2}
$$

其中， $\boldsymbol{s}$是一个长度为ML的符号向量，表示为 $\boldsymbol{s} = (\boldsymbol{s}_{1:ML-1}, -1)$，其中 $\boldsymbol{s}_{1:ML-1} \in \{-1, 1\}^{ML-1}$的定义方式与公式 (C2) 中的$\boldsymbol{s}$ 相同。权重$w_{\boldsymbol{s},a}$ 通过替换 $L \to ML$，按照公式 (C2) 中相同的方式，根据 $\boldsymbol{s}$ 和 a 定义。这里， $\boldsymbol{s}^{(j)}$ 是第 j个模式对应的符号向量，可以通过按顺序收集 $\boldsymbol{s}$ 中所有具有 $\ell \in [1, L-1]$ 个元素的第$(M(\ell-1) + j)$个元素，轻松生成，如下所示。

$$
\boldsymbol{s}^{(j)} = \left(\boldsymbol{s}_j, \boldsymbol{s}_{M+j}, \ldots, \boldsymbol{s}_{(L-1)M+j}\right), \tag{F3}
$$

其中 $\boldsymbol{s}_j$ 表示整个长度为ML 的符号向量 $\boldsymbol{s}$中的第 j个元素。注意，所有模式$\{\boldsymbol{s}^{(j)}\}_{j=1}^M$ 的符号向量合在一起构成 $\boldsymbol{s}$ 的一个划分。反过来，生成所有 $\{\boldsymbol{s}^{(j)}\}_{j=1}^M$ 的另一种显式方法是

$$
\boldsymbol{s}^{(j)} \in \{-1, 1\}^L, \qquad \text{if } 1 \le j \le M-1, \tag{F4}
$$

$$
\boldsymbol{s}^{(M)} = (\boldsymbol{s}_{1:L-1}^{(M)}, -1), \quad \text{where } \boldsymbol{s}_{1:L-1}^{(M)} \in \{-1, 1\}^{L-1}. \tag{F5}
$$

并按照划分的逆过程将它们组合起来，以构造完整的符号向量 $\boldsymbol{s}$。每个模式的位移 $B_{\boldsymbol{s}^{(j)},a}$定义与公式 (C2) 中相同，而公式 (F2) 中所有量子模式上的态定义为

$$
|\boldsymbol{B}_{\boldsymbol{s},a}\rangle_{\boldsymbol{m}} \equiv e^{i\sum_{j=1}^M \chi_{\boldsymbol{s}^{(j)}}} \otimes_{j=1}^M |(-1)^a \boldsymbol{s}^{(j)} \cdot \boldsymbol{\beta}\rangle_{m_j} \tag{F6}
$$

为方便起见。如果我们定义 $v_{\boldsymbol{s},a}(\boldsymbol{\theta},\boldsymbol{\phi},\boldsymbol{\beta}) \equiv e^{i\sum_{j=1}^{M}\chi_{\boldsymbol{s}^{(j)}}(\boldsymbol{\beta}^{(j)})}w_{\boldsymbol{s},a}(\boldsymbol{\theta},\boldsymbol{\phi})$，则有

$$|\psi(\boldsymbol{x})\rangle_{q,\boldsymbol{m}} = \sum_{a=0}^{1}\sum_{\boldsymbol{s}} v_{\boldsymbol{s},a}(\boldsymbol{\theta},\boldsymbol{\phi},\boldsymbol{\beta})\,|a\rangle_q \bigotimes_{j=1}^{M} |(-1)^a \boldsymbol{s}^{(j)}\cdot\boldsymbol{\beta}^{(j)}\rangle_{m_j}\,, \tag{F7}$$

这是对公式（3）的推广。

总之，公式 (C2) 与（F2）之间的对应关系表明，可以将由U_L 生成的 M模式态映射为由 U_{ML}生成的单模式态

$$\mathbb{M}: \psi_{L,M}(\boldsymbol{\theta},\boldsymbol{\phi},\{\boldsymbol{\beta}^{(j)}\}_{j=1}^{M}) \to \psi_{ML,1}(\boldsymbol{\theta},\boldsymbol{\phi},\boldsymbol{\beta}). \tag{F8}$$

通过 $L=1$ 和 $M=2$ 的一个具体示例即可很容易地看出该证明，然后利用数学归纳法进行推广，这与附录 C 中的方法相同。 $M=2$ 模式和 $L=1$ 线路的输出态为

$$|\psi\rangle = \begin{pmatrix} e^{i\phi_2}\sin\frac{\theta_2}{2}D^{(2)}(-\beta^{(2)}) & \cos\frac{\theta_2}{2}D^{(2)}(-\beta^{(2)}) \\ \cos\frac{\theta_2}{2}D^{(2)}(\beta^{(2)}) & e^{i(\pi-\phi_2)}\sin\frac{\theta_2}{2}D^{(2)}(\beta^{(2)}) \end{pmatrix}\begin{pmatrix} e^{i\phi_1}\sin\frac{\theta_1}{2}D^{(1)}(-\beta^{(1)}) & \cos\frac{\theta_1}{1}D^{(1)}(-\beta^{(2)}) \\ \cos\frac{\theta_1}{2}D^{(1)}(\beta^{(1)}) & e^{i(\pi-\phi_1)}\sin\frac{\theta_1}{2}D^{(1)}(\beta^{(1)}) \end{pmatrix}\begin{pmatrix} |0\rangle_{m_1}|0\rangle_{m_2} \\ 0 \end{pmatrix} \tag{F9}$$

$$\begin{aligned} &= |0\rangle_q \otimes \left(e^{i(\phi_1+\phi_2)}\sin\frac{\theta_1}{2}\sin\frac{\theta_2}{2}\,|-\beta^{(1)}\rangle_{m_1}|-\beta^{(2)}\rangle_{m_2} + \cos\frac{\theta_1}{2}\cos\frac{\theta_2}{2}\,|\beta^{(1)}\rangle_{m_1}|-\beta^{(2)}\rangle_{m_2}\right) \\ &\quad + |1\rangle_q \otimes \left(e^{i\phi_1}\sin\frac{\theta_1}{2}\cos\frac{\theta_2}{2}\,|-\beta^{(1)}\rangle_{m_1}|\beta^{(2)}\rangle_{m_2} + e^{i(\pi-\phi_2)}\cos\frac{\theta_1}{2}\sin\frac{\theta_2}{2}\,|\beta^{(1)}\rangle_{m_1}|\beta^{(2)}\rangle_{m_2}\right), \end{aligned} \tag{F10}$$

这表明其与公式（C10）中所示的态$|\psi_{L=2,\ M=1}(\boldsymbol{\theta},\ \boldsymbol{\phi},\ \boldsymbol{\beta})\rangle$ 之间存在明确的映射 .

对于能量正则化，每个 ECD 门中的位移仍服从高斯分布， $\beta_\ell^{(j)} \sim \mathcal{N}_{E/L}^{\mathrm{C}}$，因此每模能量的系综平均值也为 $\mathbb{E}\langle m_j^\dagger m_j\rangle = E$。

我们仍然考虑关于第 k 个量子比特旋转角 θ_k 的梯度。对于一般的 M 模算符，很容易验证梯度的系综平均值仍为零，其方差可以写成与公式（D3）相同的形式。与附录 E 中的研究一致，控制量子比特的目标态也为 $|0\rangle_q$ ，公式（E1）变为

$$\mathbb{E}\left[\langle O\rangle_{k^{(+1)}}\langle O\rangle_{k^{(\mu)}}\right] = \sum_{\boldsymbol{s},\boldsymbol{s}',\boldsymbol{r},\boldsymbol{r}'} \mathbb{E}\left[w_{\boldsymbol{s},k^{(+1)}}w^*_{\boldsymbol{s}',k^{(+1)}}w_{\boldsymbol{r},k^{(\mu)}}w^*_{\boldsymbol{r}',k^{(\mu)}}\right]\mathbb{E}\left[\langle\psi|\boldsymbol{B}_{\boldsymbol{s}}\rangle\langle\boldsymbol{B}_{\boldsymbol{s}'}|\psi\rangle\langle\psi|\boldsymbol{B}_{\boldsymbol{r}}\rangle\langle\boldsymbol{B}_{\boldsymbol{r}'}|\psi\rangle\right]. \tag{F11}$$

通过从 $\psi_{E,L,M}$ 到 $\psi_{E,ML,1}$的映射，多模情形下的方差与式 (E3) 形式相同，只需将单模关联函数替换为多模关联函数即可。我们将在下文讨论它们。对于C_1，每个 $\boldsymbol{s}^{(j)}\cdot\boldsymbol{\beta}^{(j)} \sim \mathcal{N}_E^{\mathrm{C}}$都服从高斯分布，并且我们有

$$C_1 = \mathbb{E}\left[|\langle\psi|\boldsymbol{B}_{\boldsymbol{s}}\rangle|^4\right] = \mathbb{E}_{\alpha\sim\mathcal{N}_E^{\mathrm{C}}}\left[|\langle\psi|\left(\otimes_{j=1}^{M}|\alpha_j\rangle\right)|^4\right]. \tag{F12}$$

注意，系综平均是对每个 $\{\alpha_j\}_{j=1}^{M}$ 独立进行的。

对于关联函数 C_2，对于每个 $\boldsymbol{s}^{(j)}\cdot\boldsymbol{\beta} = \alpha_{z_j} + \alpha_{1-z_j}$和 $\boldsymbol{r}^{(j)}\cdot\boldsymbol{\beta} = \alpha_{z_j} - \alpha_{1-z_j}$，我们仍然有，因此 C_2可以写为

$$C_2(\boldsymbol{z}) = \mathbb{E}\left[|\langle\psi|\boldsymbol{B}_{\boldsymbol{s}}\rangle|^2|\langle\psi|\boldsymbol{B}_{\boldsymbol{r}}\rangle|^2\right] = \mathbb{E}_{\alpha_y\sim\mathcal{N}_{yE}^{\mathrm{C}}}\left[\prod_{a=0}^{1}|\langle\psi|\left(\otimes_{j=1}^{M}|\alpha_{z_j}+(-1)^a\alpha_{1-z_j}\rangle\right)|^2\right], \tag{F13}$$

其中我们定义 $\boldsymbol{z}=(z_1,\ldots,z_M)$。然而，与单模情形不同，一般而言， $\boldsymbol{z}$中的某些元素可能为零，只要 $\mathbf{1}-\boldsymbol{z}$中至少有一个元素非零即可（$\mathbf{1}=(1,\ldots,1)$是一个长度为-$M$的向量），因此 $\boldsymbol{s}\neq\boldsymbol{r}$。假设 $\boldsymbol{z}$ 中属于 (0,1) 的元素个数为 $N_{\boldsymbol{z}}$，那么 $N_{\boldsymbol{z}}=M$ 的概率为

$$p\left(N_{\boldsymbol{z}}=M\right) = \frac{(2^{ML-1})(2^L-2)^{M-1}(2^{L-1}-1)}{(2^{ML-1})(2^{ML-1}-1)} = \frac{(2^L-2)^M}{2^{ML}-2}. \tag{F14}$$

在固定的 M 值下，随着 L 的增加，概率呈指数级趋近于 1。在展示一些典型态的关联函数标度之后，我们将讨论这种指数缩放所带来的后果。

最后一个关联函数 C_3 是

$$\begin{aligned}
&\mathbb{E}\left[\left|\left\langle\psi|\boldsymbol{B}_{\boldsymbol{s}}\right\rangle\right|\left|\left\langle\boldsymbol{B}_{\boldsymbol{s}'}|\psi\right\rangle\right|\left|\left\langle\psi|\boldsymbol{B}_{\boldsymbol{r}}\right\rangle\right|\left|\left\langle\boldsymbol{B}_{\boldsymbol{r}'}|\psi\right\rangle\right|\right] \\
&=\mathbb{E}\left[\left|\left\langle\psi\right|\left(\otimes_{j=1}^{M}\left|\boldsymbol{s}^{(j)}\cdot\boldsymbol{\beta}^{(j)}\right\rangle\right)\right|\left|\left(\otimes_{j=1}^{M}\left\langle\boldsymbol{s}'^{(j)}\cdot\boldsymbol{\beta}^{(j)}\right|\right)\left|\psi\right\rangle\right|\left|\left(\otimes_{j=1}^{M}\left|\boldsymbol{r}^{(j)}\cdot\boldsymbol{\beta}^{(j)}\right\rangle\right)\right|\left|\left(\otimes_{j=1}^{M}\left\langle\boldsymbol{r}'^{(j)}\cdot\boldsymbol{\beta}^{(j)}\right|\right)\left|\psi\right\rangle\right|\right] \\
&=\mathbb{E}_{\alpha_y\sim\mathcal{N}_{yE}^{\mathrm{C}}}\left[\prod_{a=0}^{1}\left|\left\langle\psi\right|\left(\otimes_{j=1}^{M}\left|\alpha_{z_j}+(-1)^a\alpha_{\tilde{z}_j}+(-1)^a\alpha_{1-z_j-\tilde{z}_j}\right\rangle\right)\right|\left|\left(\otimes_{j=1}^{M}\left\langle\alpha_{z_j}+(-1)^a\alpha_{\tilde{z}_j}-(-1)^a\alpha_{1-z_j-\tilde{z}_j}\right|\right)\left|\psi\right\rangle\right|\right] \\
&\equiv C_3(\boldsymbol{z},\tilde{\boldsymbol{z}}),
\end{aligned}\tag{F15}$$

其中我们采用与式（E19）相同的方法。与上文关于 C_2 的讨论类似， $\boldsymbol{z}$、 $\tilde{\boldsymbol{z}}$ 中的某些元素也可能为零，只要 $\tilde{\boldsymbol{z}}$、 $\mathbf{1}-\boldsymbol{z}-\tilde{\boldsymbol{z}}$ 中至少有一个非零元素，从而使 $\boldsymbol{s}$、 $\boldsymbol{s}'$、 $\boldsymbol{r}$、 $\boldsymbol{r}'$ 彼此不同。其标度也留待后续讨论。

因此，我们得到梯度方差的界为

$$\begin{aligned}
\mathrm{Var}\left[\partial_{\theta_k}\mathcal{C}(\boldsymbol{x})\right]&=\frac{1}{2}\left[\frac{3^{ML-1}}{4^{ML}}C_1+\left(\frac{1}{4}-\frac{2\cdot3^{ML-1}-2^{ML-1}}{4^{ML}}\right)\min_{\boldsymbol{z}}C_2(\boldsymbol{z})\right]+\mathcal{O}\left(C_2\right)\\
&\geq\frac{1}{2}\left[\frac{3^{ML-1}}{4^{ML}}C_1+\left(\frac{1}{4}-\frac{3^{ML}}{4^{ML}}\right)\min_{\boldsymbol{z}}C_2(\boldsymbol{z})\right]+\mathcal{O}\left(C_2\right),
\end{aligned}\tag{F16}$$

and

$$\begin{aligned}
\mathrm{Var}\left[\partial_{\theta_k}\mathcal{C}(\boldsymbol{x})\right]&=\frac{1}{2}\left[\frac{3^{ML-1}}{4^{ML}}C_1+\left(\frac{1}{4}-\frac{2\cdot3^{ML-1}-2^{ML-1}}{4^{ML}}\right)\max_{\{x^{(j)}\}}C_2(\boldsymbol{z})\right]+\mathcal{O}\left(C_2\right)\\
&\leq\frac{1}{2}\left[\frac{3^{ML-1}}{4^{ML}}C_1+\left(\frac{1}{4}+\frac{2^{ML-1}}{4^{ML}}\right)\max_{\boldsymbol{z}}C_2(\boldsymbol{z})\right]+\mathcal{O}\left(C_2\right),
\end{aligned}\tag{F17}$$

其中，我们省略了 C_3 在 $E\gg1$渐近区域中的贡献。注意，根据式（F8）中的映射替换 $L\to ML$ 后，每个关联函数前的系数与式（E20）和式（E21）中的系数完全相同。我们考虑这样的渐近区域：线路系综的每模能量大于任意模式下的目标态能量， $E\geq\max_j\left\langle\psi\right|\left(m_j^{\dagger}m_j\right)\left|\psi\right\rangle$。一般而言，上述关联函数难以求值；我们通过两个例子来了解它们的性质。

1. 乘积态

一个简单的例子是各模式之间没有任何关联的乘积态 $\left|\psi\right\rangle_{\boldsymbol{m}}=\otimes_{j=1}^{M}\left|\psi_j\right\rangle_{m_j}$，其中 $\left|\psi_j\right\rangle_{m_j}$ 是第 j 个模式的状态。我们证明，这与单模关联函数直接相关。在这种情况下， C_1 简化为乘积形式，

$$C_1^{\mathrm{Prod}}=\mathbb{E}_{\alpha_j\sim\mathcal{N}_E^{\mathrm{C}}}\left[\left|\left(\otimes_{j=1m_j}^{M}\left\langle\psi_j\right|\right)\left(\otimes_{j=1}^{M}\left|\alpha_j\right\rangle\right)\right|^4\right]=\prod_{j=1}^{M}\left(\mathbb{E}_{\alpha_j\sim\mathcal{N}_E^{\mathrm{C}}}\left[\left|_{m_j}\left\langle\psi_j\left|\alpha_j\right\rangle\right.\right|^4\right]\right).\tag{F18}$$

注意，括号内项的系综平均就是式（E6）中单模式连续变量态的关联函数 C_1，其在渐近区域的标度已被证明为 $1/E$。因此，对于M模式乘积态，我们有 $C_1\sim1/E^M$。

$C_2(\boldsymbol{z})$ 可化简为

$$\begin{aligned}
C_2^{\mathrm{Prod}}(\boldsymbol{z})&=\mathbb{E}_{\alpha_y\sim\mathcal{N}_{yE}^{\mathrm{C}}}\left[\prod_{a=0}^{1}\left|\left(\otimes_{j=1}^{M}\left\langle\psi_j\right|\right)\left(\otimes_{j=1}^{M}\left|\alpha_{z_j}+(-1)^a\alpha_{1-z_j}\right\rangle\right)\right|^2\right]\\
&=\prod_{j=1}^{M}\left(\mathbb{E}_{\alpha_y\sim\mathcal{N}_{yE}^{\mathrm{C}}}\left[\prod_{a=0}^{1}\left|\left\langle\psi_j|\alpha_{z_j}+(-1)^a\alpha_{1-z_j}\right\rangle\right|^2\right]\right).
\end{aligned}\tag{F19}$$

注意，对于特定的 z_j，如果 $z_j\in(0,1)$，则式 (F19) 括号内对应的项是单模关联函数 C_2；而如果 $z_j=0,1$，对应的项则变为 C_1。假设 $\boldsymbol{z}$ 中

区间 $(0,1)$ 内的元素个数为 $N_{\boldsymbol{z}}$，那么 C_2^{Prod} 的标度为 $\sim 1/E^{M+N_{\boldsymbol{z}}}$。根据式 (F14)，$N_{\boldsymbol{z}}=M$ 为 $p=(2^L-2)^M/(2^{ML}-2)\sim 1-1/2^L$ 的概率为。

同样，C_3 变为

$$
\begin{aligned}
&C_3^{\mathrm{Prod}}(\boldsymbol{z},\tilde{\boldsymbol{z}})\\
&=\mathbb{E}_{\alpha_y\sim\mathcal{N}_{yE}^{\mathbb{C}}}\left[\prod_{a=0}^{1}\left|\langle\psi|\left(\otimes_{j=1}^{M}|\alpha_{z_j}+(-1)^a\alpha_{\tilde{z}_j}+(-1)^a\alpha_{1-z_j-\tilde{z}_j}\rangle\right)\right|\left|\left(\otimes_{j=1}^{M}\langle\alpha_{z_j}+(-1)^a\alpha_{\tilde{z}_j}-(-1)^a\alpha_{1-z_j-\tilde{z}_j}|\right)|\psi\rangle\right|\right]\\
&=\prod_{j=1}^{M}\left(\mathbb{E}_{\alpha_y\sim\mathcal{N}_{yE}^{\mathbb{C}}}\left[\prod_{a=0}^{1}\left|\langle\psi_j|\alpha_{z_j}+(-1)^a\alpha_{\tilde{z}_j}+(-1)^a\alpha_{1-z_j-\tilde{z}_j}\rangle\right|\left|\langle\alpha_{z_j}+(-1)^a\alpha_{\tilde{z}_j}-(-1)^a\alpha_{1-z_j-\tilde{z}_j}|\psi_j\rangle\right|\right]\right),
\end{aligned}
\tag{F20}
$$

为方便起见，我们将式 (F20) 括号内的项记为 $C^{(j)}$。对于 z_j、$\tilde{z}_j$、$1-z_j-\tilde{z}_j$ 的某种特定组合，如果它们全都属于 $(0,1)$，那么 $C^{(j)}$ 就是式 (E19) 中的单模关联函数 $C_3\sim 1/E^3$；另一方面，如果其中只有两个属于 $(0,1)$，而另一个为零，那么 $C^{(j)}$ 变为 $C_2\sim 1/E^2$；此外，如果三者中只有一个非零，那么 $C^{(j)}$ 就是 $C_1\sim 1/E$。因此，对于不同的 $\boldsymbol{z}$、$\tilde{\boldsymbol{z}}$，式 (F20) 的标度可以是$1/E^{\nu}$，其中 $\nu\in[M,3M]$ 为整数。如表 II 所示，$\boldsymbol{s}$、$\boldsymbol{s}'$、$\boldsymbol{r}$、$\boldsymbol{r}'$ 中的每个元素只有三种允许的赋值，因此，对于第 j 个模式，$C^{(j)}\sim 1/E^{\nu_j}$ 的概率为 $\nu_j\in\{1,2,3\}$

$$
p(C^{(j)}\sim 1/E^{\nu_j})=\begin{cases}1/3^{L-1}, & if\ \nu_j=1\\ 2^L/3^{L-1}, & if\ \nu_j=2\\ 1-(1+2^L)/3^{L-1}, & otherwise.\end{cases}
\tag{F21}
$$

假设 ν_j 中 $\{\nu_j\}_{j=1}^M$ 为 1,2 的数量为 n_1,n_2，那么$\sum_{j=1}^M\nu_j\le\nu_c$的概率为

$$
p\left(\sum_{j=1}^{M}\nu_j\le\nu_c\right)=\sum_{\substack{n_1,n_2\ge 0,\\ n_1+n_2\le M,\\ n_1+2n_2+3(M-n_1-n_2)\le\nu_c}}\frac{M!}{n_1!n_2!(M-n_1-n_2)!}\left(\frac{1}{3^{L-1}}\right)^{n_1}\left(\frac{2^L}{3^{L-1}}\right)^{n_2}\left(\frac{3^{L-1}-2^L-1}{3^{L-1}}\right)^{M-n_1-n_2}
\tag{F22}
$$

设置 $\nu_c=2M$ 使我们能够确定C_3^{Prod}的阶数是否不 高于C_2^{Prod}。解析计算较为困难，因此我们在图10中展示了数值计算结果。非高阶项的概率呈指数抑制，这表明我们可以像在单模式情形中所做的那样，将C_3^{Prod}视为相较于C_1^{Prod}和C_2^{Prod}的高阶项。

在乘积态部分的末尾，我们讨论浅层电路和深层电路的判据。回顾一下，$C_2^{\mathrm{Prod}}\sim 1/E^{2M}$为 $p\sim 1-1/2^L$的概率。对于浅层电路，在约束条件 $^{1/4E^{2M}}{}_{(3/4)^{ML}/3E^M}\in\mathcal{O}(1)$ 和 $_{(1-p)/4E^{M+1}}{}^{(3/4)^{ML}/3E^M}\in\mathcal{O}(1)$ 下，主导阶为~~$1/E^M$~~，因此得到条件~~$L\in\mathcal{O}(\log E)$~~。另一方面，对于深层电路，在条件 $^{1/4E^{2M}}{}_{(3/4)^{ML}/3E^M}\in\Omega(1)$ 和 $_{(1-p)}{}_{1/4E^{2M}}^{/4E^{M+1}}\in\mathcal{O}(1)$ 下，梯度方差为 $1/E^{2M}$，从而得到$L\in\Omega(\log E)$。

2. 多模高斯态

本节给出制备任意多模高斯态（无论是否纠缠）的结果。为简单起见，我们将目标态设为一个由正交分量均值 $\mathcal{X}_{\boldsymbol{m}}$ and 协方差矩阵$\boldsymbol{V}_{\boldsymbol{m}}$ 描述的 M模高斯态。由于所有关联函数都是目标态与相干态乘积之间保真度的函数，我们首先给出相干态乘积 $\otimes_{j=1}^M|\alpha_j\rangle_{m_j}$的相空间表示，其正交分量均值和协方差矩阵为

$$
\overline{\mathcal{X}}=2(\mathrm{Re}\{\alpha_1\},\mathrm{Im}\{\alpha_1\},\ldots,\mathrm{Re}\{\alpha_M\},\mathrm{Im}\{\alpha_M\})^T\equiv 2\left(\xi_1,\xi_2,\ldots,\xi_{2M-1},\xi_{2M}\right)^T=2\boldsymbol{\xi},
\tag{F23}
$$

$$
\boldsymbol{V}_{\mathrm{coh}}=\boldsymbol{I},
\tag{F24}
$$

其中，我们将所有 α_j 的实部和虚部统一表示为向量 $\boldsymbol{\xi}$，其每个元素均服从分布 $\mathcal{N}_{E/2}$。CM $\boldsymbol{V}_{\mathrm{coh}}$ 是一个 $2M\times 2M$ 单位矩阵。类似地，我们可以定义 $\mathcal{X}_{\boldsymbol{m}}=2\boldsymbol{\xi}_{\boldsymbol{m}}$。通过应用一般保真度

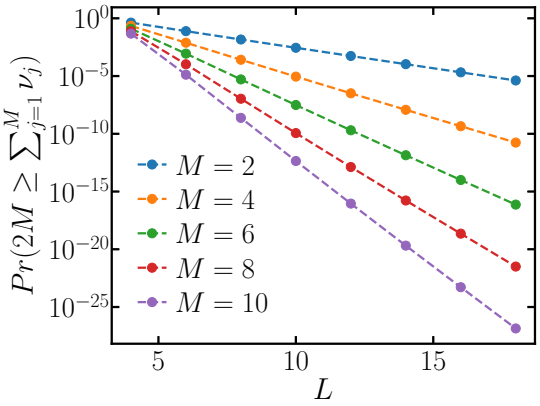


图 10。{v1}的概率 在不同电路深度 L 和模式数量 M下，式（F22）中的 $\sum_{j=1}^{M} \nu_j \leq 2M$ 。

公式 (A8)，我们可以得到 C_1

$$
\begin{aligned}
&C_1^{\text{Gauss}} = \mathbb{E}_{\alpha\sim\mathcal{N}_E^{\mathbb{C}}}\left[|\langle\psi|\left(\otimes_{j=1}^{M}|\alpha_j\rangle\right)|^4\right] = \mathbb{E}_{\alpha\sim\mathcal{N}_E^{\mathbb{C}}}\left[F(|\psi\rangle_{\boldsymbol{m}}, \otimes_{j=1}^{M}|\alpha_j\rangle)^2\right] \\
&= \int d\boldsymbol{\xi}\frac{4^M}{\det(\boldsymbol{V_m}+\boldsymbol{I})}e^{-4(\boldsymbol{\xi}-\boldsymbol{\xi_m})^T(\boldsymbol{V_m}+\boldsymbol{I})^{-1}(\boldsymbol{\xi}-\boldsymbol{\xi_m})}\frac{1}{(\pi E)^M}e^{-\boldsymbol{\xi}^T\boldsymbol{\xi}/E} \\
&= \frac{4^M}{(\pi E)^M\det(\boldsymbol{V_m}+\boldsymbol{I})}\int d\boldsymbol{\xi}\exp\left[-\boldsymbol{\xi}^T\left(4(\boldsymbol{V_m}+\boldsymbol{I})^{-1}+\frac{\boldsymbol{I}}{E}\right)\boldsymbol{\xi}+2\boldsymbol{\xi_m}^T4(\boldsymbol{V_m}+\boldsymbol{I})^{-1}\boldsymbol{\xi}-4\boldsymbol{\xi_m}^T(\boldsymbol{V_m}+\boldsymbol{I})^{-1}\boldsymbol{\xi_m}\right] \\
&= \frac{4^M\det(\boldsymbol{K})e^{-4\boldsymbol{\xi_m}^T\boldsymbol{K}\boldsymbol{\xi_m}}}{(\pi E)^M}\int d\boldsymbol{\xi}\exp\left[-\boldsymbol{\xi}^T(4\boldsymbol{K}+\boldsymbol{I}/E)\boldsymbol{\xi}+2\boldsymbol{\xi}_{\boldsymbol{m}}'^T\boldsymbol{\xi}\right] \qquad \text{(F25)} \\
&= \frac{4^Me^{-4\boldsymbol{\xi_m}^T\boldsymbol{K}\boldsymbol{\xi_m}}}{(\pi E)^M\det(\boldsymbol{K})^{-1}}\int d\boldsymbol{\xi}\exp\left[-(\boldsymbol{\xi}-(4\boldsymbol{K}+\boldsymbol{I}/E)^{-1}\boldsymbol{\xi}_{\boldsymbol{m}}')^T(4\boldsymbol{K}+\boldsymbol{I}/E)(\boldsymbol{\xi}-(4\boldsymbol{K}+\boldsymbol{I}/E)^{-1}\boldsymbol{\xi}_{\boldsymbol{m}}')+\boldsymbol{\xi}_{\boldsymbol{m}}'^T(4\boldsymbol{K}+\boldsymbol{I}/E)^{-1}\boldsymbol{\xi}_{\boldsymbol{m}}'\right], \qquad \text{(F26)}
\end{aligned}
$$

其中，为方便起见，在倒数第二行中我们将 $\boldsymbol{K} = (\boldsymbol{V_m}+\boldsymbol{I})^{-1}$ 和 $\boldsymbol{\xi}_{\boldsymbol{m}}' = 4\boldsymbol{K}\boldsymbol{\xi_m}$ 记作。在最后一行中，我们将指数写成配方形式。由于 $4\boldsymbol{K}+\boldsymbol{I}/E$ 是一个对称实矩阵，我们可以通过正交矩阵 Q 将其对角化为 $4\boldsymbol{K}+\boldsymbol{I}/E = Q^T\Lambda Q$，其中 $\Lambda = \text{diag}(\lambda_1,\ldots,\lambda_{2M})$为对角矩阵。我们可以进行变量变换 $\tilde{\boldsymbol{\xi}} = Q(\boldsymbol{\xi}-\boldsymbol{K}^{-1}\boldsymbol{\xi}_{\boldsymbol{m}}')$，于是被积函数化为

$$
\int d\tilde{\boldsymbol{\xi}}\left|\frac{\partial\boldsymbol{\xi}}{\partial\tilde{\boldsymbol{\xi}}}\right|\exp\left[-\tilde{\boldsymbol{\xi}}^T\Lambda\tilde{\boldsymbol{\xi}}\right] = \int d\tilde{\boldsymbol{\xi}}\exp\left[-\sum_{i=1}^{2M}\lambda_i\tilde{\xi}_i^2\right] = \frac{\pi^M}{\sqrt{\det(4\boldsymbol{K}+\boldsymbol{I}/E)}}, \qquad \text{(F27)}
$$

因此，关联函数 C_1^{Gauss}变为

$$
C_1^{\text{Gauss}} = \frac{4^M\det(\boldsymbol{K})\exp\left\{-4\boldsymbol{\xi}_{\boldsymbol{m}}^T\left[\boldsymbol{K}-4\boldsymbol{K}(4\boldsymbol{K}+\boldsymbol{I}/E)^{-1}\boldsymbol{K}\right]\boldsymbol{\xi_m}\right\}}{\sqrt{\det(4\boldsymbol{K}+\boldsymbol{I}/E)}E^M}. \qquad \text{(F28)}
$$

在对 C_2 进行评估时，既有 $\otimes_{j=1}^{M}|\alpha_{z_j}\pm\alpha_{1-z_j}\rangle$，也可以用 $\mathcal{X}_\pm = 2(\boldsymbol{\xi_z}\pm\boldsymbol{\xi_{1-z}})$ 和 $V_\pm = \boldsymbol{I}$ 来表征。 $\boldsymbol{\xi_z}$ 的分布为 $p(\boldsymbol{\xi_z}) = \exp\left[-\boldsymbol{\xi}_{\boldsymbol{z}}^T\boldsymbol{S_z}\boldsymbol{\xi_z}\right]\sqrt{\det\boldsymbol{S_z}}/\pi^M$，其中 $\boldsymbol{S_z} = \oplus_{j=1}^{M}\boldsymbol{I}_2/(\overline{z_jE})$，而 $\boldsymbol{I}_2$ 是一个

2×2 单位矩阵。关联函数 C_2 变为

$$C_2^{\mathrm{Gauss}}(\boldsymbol{z})=\mathbb{E}_{\alpha_y\sim\mathcal{N}_{yE}^{\mathrm{C}}}\left[\prod_{h=0}^{1}|\left\langle\psi\right|\left(\otimes_{j=1}^{M}\left|\alpha_{z_j}+(-1)^h\alpha_{1-z_j}\right\rangle\right)|^2\right]=\mathbb{E}_{\alpha_y\sim\mathcal{N}_{yE}^{\mathrm{C}}}\left[\prod_{h=0}^{1}F(|\psi\rangle_{\boldsymbol{m}}\left(\otimes_{j=1}^{M}\left|\alpha_{z_j}+(-1)^h\alpha_{1-z_j}\right\rangle\right)\right]$$

$$\begin{aligned}
&=\int d\boldsymbol{\xi_z}d\boldsymbol{\xi_{1-z}}\left(\frac{2^M}{\sqrt{\det(\boldsymbol{V_m}+\boldsymbol{I})}}e^{-2(\boldsymbol{\xi_z}+\boldsymbol{\xi_{1-z}}-\boldsymbol{\xi_m})^T(\boldsymbol{V_m}+\boldsymbol{I})^{-1}(\boldsymbol{\xi_z}+\boldsymbol{\xi_{1-z}}-\boldsymbol{\xi_m})}\right.\\
&\qquad\times\frac{2^M}{\sqrt{\det(\boldsymbol{V_m}+\boldsymbol{I})}}e^{-2(\boldsymbol{\xi_z}-\boldsymbol{\xi_{1-z}}-\boldsymbol{\xi_m})^T(\boldsymbol{V_m}+\boldsymbol{I})^{-1}(\boldsymbol{\xi_z}-\boldsymbol{\xi_{1-z}}-\boldsymbol{\xi_m})}\\
&\qquad\left.\times\frac{\sqrt{\det\boldsymbol{S_z}}}{\pi^M}e^{-\boldsymbol{\xi_z}^T\boldsymbol{S_z}\boldsymbol{\xi_z}}\frac{\sqrt{\det\boldsymbol{S_{1-z}}}}{\pi^M}e^{-\boldsymbol{\xi_{1-z}}^T\boldsymbol{S_{1-z}}\boldsymbol{\xi_{1-z}}}\right)
\end{aligned}\tag{F29}$$

$$=\frac{4^M\det(\boldsymbol{K})\sqrt{\det\boldsymbol{S_z}\det\boldsymbol{S_{1-z}}}e^{-4\boldsymbol{\xi}_{\boldsymbol{m}}^T\boldsymbol{K}\boldsymbol{\xi_m}}}{\pi^{2M}}\int d\boldsymbol{\xi_z}d\boldsymbol{\xi_{1-z}}e^{-\boldsymbol{\xi_z}^T(4\boldsymbol{K}+\boldsymbol{S_z})\boldsymbol{\xi_z}+2\boldsymbol{\xi}_{\boldsymbol{m}}'^T\boldsymbol{\xi_z}-\boldsymbol{\xi}_{\boldsymbol{1-z}}^T(4\boldsymbol{K}+\boldsymbol{S_{1-z}})\boldsymbol{\xi_{1-z}}}\tag{F30}$$

$$=\frac{4^M\det(\boldsymbol{K})\sqrt{\det\boldsymbol{S_z}\det\boldsymbol{S_{1-z}}}e^{-4\boldsymbol{\xi}_{\boldsymbol{m}}^T\boldsymbol{K}\boldsymbol{\xi_m}+\boldsymbol{\xi}_{\boldsymbol{m}}'^T(4\boldsymbol{K}+\boldsymbol{S_z})^{-1}\boldsymbol{\xi}_{\boldsymbol{m}}'^T}}{\pi^{2M}\sqrt{\det\boldsymbol{S_z}\det\boldsymbol{S_{1-z}}}}\frac{\pi^M}{\sqrt{\det(4\boldsymbol{K}+\boldsymbol{S_z})}}\frac{\pi^M}{\sqrt{\det(4\boldsymbol{K}+\boldsymbol{S_{1-z}})}}$$

$$=\frac{4^M\det(\boldsymbol{K})\exp\left\{-4\boldsymbol{\xi}_{\boldsymbol{m}}^T\left[\boldsymbol{K}-4\boldsymbol{K}(4\boldsymbol{K}+\boldsymbol{S_z})^{-1}\boldsymbol{K}\right]\boldsymbol{\xi_m}\right\}}{\sqrt{\det(4\boldsymbol{K}+\boldsymbol{S_z})\det(4\boldsymbol{K}+\boldsymbol{S_{1-z}})}}\frac{1}{\left[\prod_{j=1}^{M}z_j(1-z_j)\right]E^{2M}}\tag{F32}$$

其中，在公式 (F30) 中，我们将 $\boldsymbol{K}$ 和 $\boldsymbol{\xi}_{\boldsymbol{m}}'$ 记作与公式 (F26) 中相同的形式。在公式 (F31) 中，我们应用了推导公式 (F28) 时使用的相同对角化方法，其中 $\Lambda_{\boldsymbol{z}}=Q_{\boldsymbol{z}}$（$4\boldsymbol{K}+\boldsymbol{S_z}$）$Q_{\boldsymbol{z}}^T$，依此类推，$\Lambda_{\boldsymbol{1-z}}$、$\tilde{\boldsymbol{\xi}}_{\boldsymbol{z}}=Q_{\boldsymbol{z}}$（$\boldsymbol{\xi_z}-$（$4\boldsymbol{K}+\boldsymbol{S_z}$）$^{-1}\boldsymbol{\xi}_{\boldsymbol{m}}'$）以及 $\tilde{\boldsymbol{\xi}}_{\boldsymbol{1-z}}=Q_{\boldsymbol{1-z}}\boldsymbol{\xi_{1-z}}$。与对 C_2^{Prod} 的讨论类似，如果 $\boldsymbol{z}$ 中有 $M-N_{\boldsymbol{z}}$ 个元素为零或一，那么 $C_2^{\mathrm{Gauss}}(\boldsymbol{z})\sim 1/E^{M+N_{\boldsymbol{z}}}$，这是因为公式 (F29) 中方差为零的高斯分布会变为狄拉克δ函数。

剩下的最后一个关联函数是 $C_3^{\mathrm{Gauss}}(\boldsymbol{z}_1,\ \boldsymbol{z}_2)$

$$C_3^{\mathrm{Gauss}}(\boldsymbol{z},\tilde{\boldsymbol{z}})=\mathbb{E}_{\alpha_y\sim\mathcal{N}_{yE}^{\mathrm{C}}}\left[\prod_{h=0}^{1}\left|\left\langle\psi\right|\left(\otimes_{j=1}^{M}\left|\alpha_{z_j}+(-1)^h\alpha_{\tilde{z}_j}+(-1)^h\alpha_{1-z_j-\tilde{z}_j}\right\rangle\right)\right|\left|\left(\otimes_{j=1}^{M}\left\langle\alpha_{z_j}+(-1)^h\alpha_{\tilde{z}_j}-(-1)^h\alpha_{1-z_j-\tilde{z}_j}\right|\right)\left|\psi\right\rangle\right|\right]$$

$$\begin{aligned}
&=4^M\det(\boldsymbol{K})\int d\boldsymbol{\xi_z}d\boldsymbol{\xi_{\tilde{z}}}d\boldsymbol{\xi_{1-z-\tilde{z}}}\left(e^{-(\boldsymbol{\xi_z}+\boldsymbol{\xi_{\tilde{z}}}+\boldsymbol{\xi_{1-z-\tilde{z}}}-\boldsymbol{\xi_m})^T\boldsymbol{K}(\boldsymbol{\xi_z}+\boldsymbol{\xi_{\tilde{z}}}+\boldsymbol{\xi_{1-z-\tilde{z}}}-\boldsymbol{\xi_m})}e^{-(\boldsymbol{\xi_z}+\boldsymbol{\xi_{\tilde{z}}}-\boldsymbol{\xi_{1-z-\tilde{z}}}-\boldsymbol{\xi_m})^T\boldsymbol{K}(\boldsymbol{\xi_z}+\boldsymbol{\xi_{\tilde{z}}}-\boldsymbol{\xi_{1-z-\tilde{z}}}-\boldsymbol{\xi_m})}\right.\\
&\qquad\times e^{-(\boldsymbol{\xi_z}-\boldsymbol{\xi_{\tilde{z}}}-\boldsymbol{\xi_{1-z-\tilde{z}}}-\boldsymbol{\xi_m})^T\boldsymbol{K}(\boldsymbol{\xi_z}-\boldsymbol{\xi_{\tilde{z}}}-\boldsymbol{\xi_{1-z-\tilde{z}}})}e^{-(\boldsymbol{\xi_z}-\boldsymbol{\xi_{\tilde{z}}}+\boldsymbol{\xi_{1-z-\tilde{z}}}-\boldsymbol{\xi_m})^T\boldsymbol{K}(\boldsymbol{\xi_z}-\boldsymbol{\xi_{\tilde{z}}}+\boldsymbol{\xi_{1-z-\tilde{z}}}-\boldsymbol{\xi_m})}\\
&\qquad\left.\times\frac{\sqrt{\det\boldsymbol{S_z}}}{\pi^M}e^{-\boldsymbol{\xi_z}^T\boldsymbol{S_z}\boldsymbol{\xi_z}}\frac{\sqrt{\det\boldsymbol{S_{\tilde{z}}}}}{\pi^M}e^{-\boldsymbol{\xi_{\tilde{z}}}^T\boldsymbol{S_{\tilde{z}}}\boldsymbol{\xi_{\tilde{z}}}}\frac{\sqrt{\det\boldsymbol{S_{1-z-\tilde{z}}}}}{\pi^M}e^{-\boldsymbol{\xi_{1-z-\tilde{z}}}^T\boldsymbol{S_{1-z-\tilde{z}}}\boldsymbol{\xi_{1-z-\tilde{z}}}}\right)
\end{aligned}$$

$$=\frac{4^M\sqrt{\det\boldsymbol{S_z}\det\boldsymbol{S_{\tilde{z}}}\det\boldsymbol{S_{1-z-\tilde{z}}}}e^{-4\boldsymbol{\xi}_{\boldsymbol{m}}^T\boldsymbol{K}\boldsymbol{\xi_m}}}{\pi^{3M}\det(\boldsymbol{K})^{-1}}\int d\boldsymbol{\xi_z}d\boldsymbol{\xi_{\tilde{z}}}d\boldsymbol{\xi_{1-z-\tilde{z}}}e^{-\boldsymbol{\xi_z}^T(4\boldsymbol{K}+\boldsymbol{S_z})\boldsymbol{\xi_z}+2\boldsymbol{\xi}_{\boldsymbol{m}}'^T\boldsymbol{\xi_z}-\boldsymbol{\xi_{\tilde{z}}}^T(4\boldsymbol{K}+\boldsymbol{S_{\tilde{z}}})\boldsymbol{\xi_{\tilde{z}}}-\boldsymbol{\xi}_{\boldsymbol{1-z-\tilde{z}}}^T(4\boldsymbol{K}+\boldsymbol{S_{1-z-\tilde{z}}})\boldsymbol{\xi_{1-z-\tilde{z}}}}$$

$$=\frac{4^M\sqrt{\det\boldsymbol{S_z}\det\boldsymbol{S_{\tilde{z}}}\det\boldsymbol{S_{1-z-\tilde{z}}}}e^{-4\boldsymbol{\xi}_{\boldsymbol{m}}^T\boldsymbol{K}\boldsymbol{\xi_m}+\boldsymbol{\xi}_{\boldsymbol{m}}'^T(4\boldsymbol{K}+\boldsymbol{S_z})^{-1}\boldsymbol{\xi}_{\boldsymbol{m}}'^T}}{\pi^{3M}\det(\boldsymbol{K})^{-1}}\int d\tilde{\boldsymbol{\xi}}_{\boldsymbol{z}}d\tilde{\boldsymbol{\xi}}_{\tilde{\boldsymbol{z}}}d\tilde{\boldsymbol{\xi}}_{\boldsymbol{1-z-\tilde{z}}}e^{-\tilde{\boldsymbol{\xi}}_{\boldsymbol{z}}\Lambda_{\boldsymbol{z}}\tilde{\boldsymbol{\xi}}_{\boldsymbol{z}}}e^{-\tilde{\boldsymbol{\xi}}_{\tilde{\boldsymbol{z}}}\Lambda_{\boldsymbol{z}}\tilde{\boldsymbol{\xi}}_{\tilde{\boldsymbol{z}}}}e^{-\tilde{\boldsymbol{\xi}}_{\boldsymbol{1-z-\tilde{z}}}\Lambda_{\boldsymbol{1-z-\tilde{z}}}\tilde{\boldsymbol{\xi}}_{\boldsymbol{1-z-\tilde{z}}}}\tag{F33}$$

$$=\frac{4^M\sqrt{\det\boldsymbol{S_z}\det\boldsymbol{S_{\tilde{z}}}\det\boldsymbol{S_{1-z-\tilde{z}}}}e^{-4\boldsymbol{\xi}_{\boldsymbol{m}}^T\boldsymbol{K}\boldsymbol{\xi_m}+\boldsymbol{\xi}_{\boldsymbol{m}}'^T(4\boldsymbol{K}+\boldsymbol{S_z})^{-1}\boldsymbol{\xi}_{\boldsymbol{m}}'^T}}{\pi^{3M}\det(\boldsymbol{K})^{-1}}\frac{\pi^M}{\sqrt{\det(4\boldsymbol{K}+\boldsymbol{S_z})}}\frac{\pi^M}{\sqrt{\det(4\boldsymbol{K}+\boldsymbol{S_{\tilde{z}}})}}\frac{\pi^M}{\sqrt{\det(4\boldsymbol{K}+\boldsymbol{S_{1-z-\tilde{z}}})}}$$

$$=\frac{4^M\det(\boldsymbol{K})\exp\left\{-4\boldsymbol{\xi}_{\boldsymbol{m}}^T\left[\boldsymbol{K}-4\boldsymbol{K}(4\boldsymbol{K}+\boldsymbol{S_z})^{-1}\boldsymbol{K}\right]\boldsymbol{\xi_m}\right\}}{\sqrt{\det(4\boldsymbol{K}+\boldsymbol{S_z})}\sqrt{\det(4\boldsymbol{K}+\boldsymbol{S_{\tilde{z}}})}\sqrt{\det(4\boldsymbol{K}+\boldsymbol{S_{1-z-\tilde{z}}})}}\frac{1}{\left[\prod_{j=1}^{M}z_j\tilde{z}_j(1-z_j-\tilde{z}_j)\right]E^{3M}},\tag{F34}$$

其中，在式 (F33) 中，我们使用与式 (F30) 相同的对角化方法，采用 $\Lambda_{\tilde{\boldsymbol{z}}}=Q_{\tilde{\boldsymbol{z}}}(4\boldsymbol{K}+\boldsymbol{S_{\tilde{z}}})Q_{\tilde{\boldsymbol{z}}}^T$和 $\tilde{\boldsymbol{\xi}}_{\tilde{\boldsymbol{z}}}=Q_{\tilde{\boldsymbol{z}}}\boldsymbol{\xi}_{\tilde{\boldsymbol{\xi}}}$，因此也有 $\Lambda_{\boldsymbol{1-z-\tilde{z}}},\tilde{\boldsymbol{\xi}}_{\boldsymbol{1-z-\tilde{z}}}$。根据式 (F22) 的相同分析，$C_3^{\mathrm{Gauss}}(\boldsymbol{z},\tilde{\boldsymbol{z}})$ 的体贡献表现为 $1/E^{\nu}$，其中 $\nu>2M$，因此当 E 较大时，由于它们是 E 的高阶项，可以忽略。

在 E 的渐近极限下，与 $4\boldsymbol{K}$ 相比，我们可以忽略 C_1^{Gauss} 中的 $\boldsymbol{I}/E$ 和 C_2^{Gauss} 中的 $\boldsymbol{S}_{\mathbf{1}-\boldsymbol{z}}$，$\boldsymbol{S}_{\boldsymbol{z}}$，因此，这些关联函数中的行列式化为与 E 无关的常数，从而直接得出 $1/E^M$ 和 $1/E^{2M}$ 各自的标度。因此，制备一般 M 模式高斯态时，浅层电路与深层电路之间的临界深度也由电路系综能量的对数决定。

下面我们给出双模压缩真空态（TMSV）的一个具体示例。协方差矩阵（CM）已在附录 A 中介绍。通过计算式 (F28) 和 (F32)，可以直接得到 TMSV 态的 C_1 和 C_2，如下所示

$$C_1^{\mathrm{TMSV}} = \frac{\mathrm{sech}^4(\zeta)}{G_1(E)}, \tag{F35}$$

$$C_2^{\mathrm{TMSV}}(\boldsymbol{z}) = \frac{\mathrm{sech}^4(\zeta)}{G_2(\boldsymbol{z}E)G_2[(\mathbf{1}-\boldsymbol{z})E]}, \tag{F36}$$

其中我们定义

$$G_2(\boldsymbol{z}) = 1 + 2\|\boldsymbol{z}\|_1 + 4\,\mathrm{sech}^2(\zeta)\prod_j z_j. \tag{F37}$$

这里，$\boldsymbol{x}$ 是一个向量，而上面的$\mathbf{1}$ 是一个与 $\boldsymbol{x}$长度相同的恒等向量。在 E较大的渐近区域，我们还有 $C_2^{\mathrm{TMSV}}(\boldsymbol{z}) \sim 1/64 z_1 z_2(1-z_1)(1-z_2)E^4$。请注意，两个关联函数都随 E单调递减。

总而言之，对于浅层量子电路，$L \in \mathcal{O}(\log E)$，梯度方差为 $\mathrm{Var} \sim 1/E^M$；而对于深层量子电路$L \in \Omega(\log E)$，在 E的渐近极限下，它变为 $\sim 1/E^{2M}$。

附录 G：相干态和单模压缩真空态的结果

在本节中，我们将进一步展示单模高斯态态制备的结果。在图 11(a)、(b) 中，我们展示了制备相干态 $|\gamma\rangle$ 和满足 $E_t = |\gamma|^2 = \sinh^2(r) = 8$ 的单模压缩真空态 $|\zeta\rangle_{\mathrm{SMSV}}$ 时，梯度方差 $\mathrm{Var}\,[\partial_{\theta_k} C]$ 随系综能量 E的变化情况，展示方式与图 2 相同。对于浅层电路，数值结果（橙色圆点）与定理 3 一致，其中 C_1 分别取为式 (E26) 和式 (E27)。对于深层电路，界通过式 (E37) 计算，其中 C_2 见式 (E29) 和式 (E30)。我们观察到的行为与正文中的示例类似，此处也得出相同的结论。

在图 12 中，我们展示了另一个制备单模压缩真空态的训练示例，其中初始系综能量取不同值。随着系综能量的增加，训练性能变得更差，正如我们在式 (E37) 中已经看到的：式 (E30) 中的 $C_2^{\mathrm{SMSV}}(z)$ 单调递减，较大的系综能量将阻碍量子电路的有效训练。

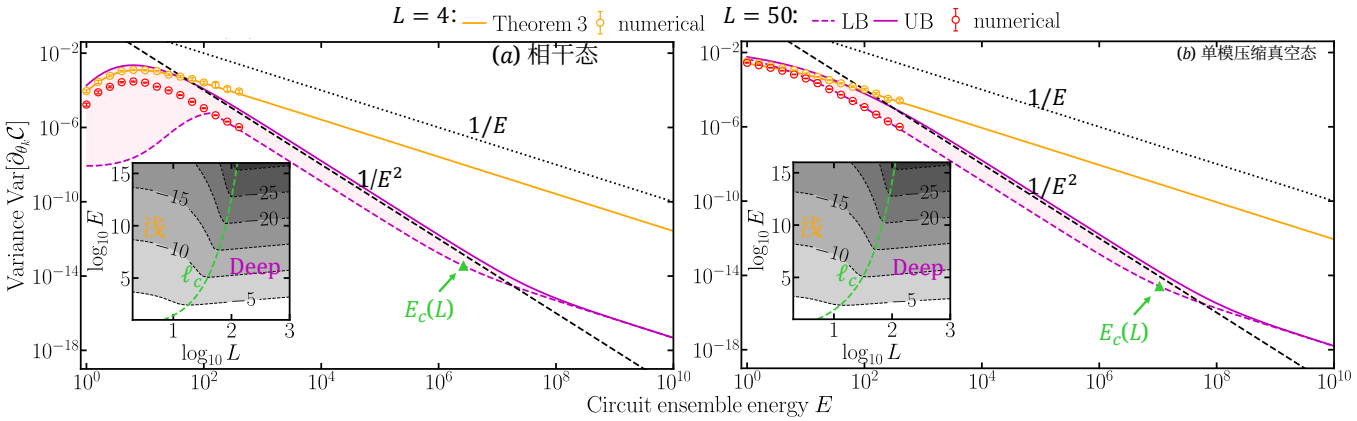


图 11. 制备 (a) 相干态 $|\gamma\rangle$ 和单模压缩真空态 $|\zeta\rangle_{\mathrm{SMSV}}$ 时，在 $k = L/2$ 处的梯度 $[\partial_{\theta_k} \mathcal{C}]$ 的方差 Var。所有图例的含义均与图 2 相同。

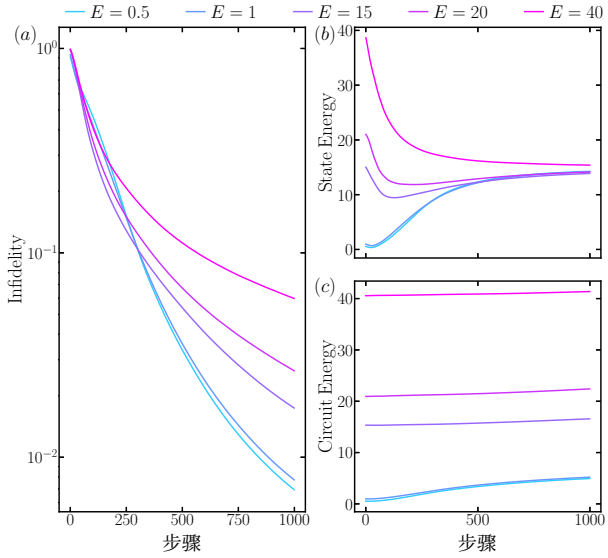


图 12。SMSV 状态的训练 $|\psi\rangle = |\zeta\rangle_{\mathrm{SMSV}}$ 其中 $E_t = \sinh^2 \zeta = 15$，使用以不同系综能量 E 初始化的 $L = 50$ 连续变量变分量子电路。我们展示了 (a) 输出态与目标态之间的平均不保真度，(b) 平均输出态能量，以及 (c) 平均电路能量 $\sum_{j=1}^{L} |\beta_j|^2$ 随训练步数的变化。