扩散模型的数学导论

JIANFENG LU

摘要. 这些笔记从采样的观点出发，给出一个以证明为导向的扩散模型导论，沿着一条主线从经典采样动力学走向现代扩散采样器、它们的误差分析以及推理时控制。全篇分为五个部分。第一部分建立采样语言和朗之万工具箱——目标分布与差异、马尔可夫核、Fokker–Planck 演化和熵耗散——并由此得到朗之万扩散、未调整朗之万算法及其 Metropolis 调整修正的收敛保证。第二部分通过高斯加噪、Tweedie 恒等式、反向时间 SDE 和概率流 ODE，构造连续时间的基于得分的扩散模型，并通过随机定位和 Polchinski 流重新考察同一个高斯通道。第三部分把这些连续动力学离散化为可实现的 DDPM 采样器——精确反向核、高斯反向核和去噪-得分等价性——并进行采样误差分析，把早停、KL 望远镜求和和得分误差分开处理，覆盖 Euler–Maruyama、Hessian 控制和高精度一阶拒绝采样。第四部分研究有限状态空间上的离散扩散，其中连续时间马尔可夫链取代加噪 SDE，反向核及其误差分析也改写为有限状态形式。第五部分转向对训练好模型的推理时操控：引导、奖励倾斜、路径空间控制和推理时强化学习。贯穿全篇，材料被分成三个层次：完整证明的核心定义与恒等式，在简化假设下证明的代表性估计，以及附有证明路线图的研究级定理。预期读者是具备概率背景的低年级研究生，不要求此前接触过随机微分方程、随机数值方法或扩散模型。

目录

1. 采样与朗之万动力学简介 4
2. 朗之万扩散与 ULA 的收敛 9
3. 基于得分的扩散模型 14
4. 随机定位与 Polchinski 流 20
5. 连续扩散模型的离散化 26
6. 扩散模型的误差分析 29
7. 离散扩散模型 40
8. 引导、奖励倾斜与推理时强化学习 50
附录 A. Itô 微积分与 Girsanov 定理 58
附录 B. 高斯工具箱 60
参考文献 61

如何阅读这些笔记. 预期读者是已经学过概率、线性代数和多元微积分的低年级研究生，但他们可能还没有接触过随机微分方程或扩散模型。因此，这些笔记把材料分成三个层次。

日期: 2026 年 7 月 2 日.

这些讲义用于 SLMath 的 John Tukey Summer Graduate School on Mathematics of Generative Models（2026 年 6 月 22 日–2026 年 7 月 2 日），该暑期学校由作者与 Eric Vanden-Eijnden 共同组织并共同授课。作者感谢 SLMath 的接待，以及暑期学校学生的有益反馈。本工作部分得到美国国家科学基金会 DMS-2309378 和 IIS-2403276 项目的支持。

- *核心定义和等式*的证明详尽地进行了。这些内容贯穿于每一节，并包括 Fokker–Planck 方程、熵耗散、Tweedie 的等式、反向 SDE、概率流 ODE、精确和高斯 DDPM 反向核、去噪得分等价、有限状态反向核的离散扩散模型、奖励倾斜等式。
- *代表性估计*证明了这些结果在简化假设下成立，以展示其最简洁的形式。例如，我们证明在对数 Sobolev 不等式下存在 KL 收缩，精确计算了高斯目标的 ULA 偏移，并对扩散模型的 KL 进行了一步离散化和得分误差的上界。
- *研究级定理*这些结果以简化形式陈述，并附有证明路线图。这些包括 ULA 和 MALA 收敛保证、高精度的扩散采样（通过第一类拒绝采样）以及离散扩散误差分析。目标不是复制原始论文中的每一个技术引理，而是清楚地说明每个定理的含义、为什么假设出现以及证明是如何组织的。

两个附录收集了反复使用的分析工具：Itô 微积分和附录 A 中的 Girsanov 变换公式，以及附录 B 中的高斯恒等式。一个对这些工具熟悉的读者可以跳过它们，并在需要时随时查阅。

这些笔记是自洽的，但故意选择性地介绍，它们与多个读者希望并行查阅的优秀处理相重叠。Yuansi Chen 为 ETH 课程 *扩散模型的计算与统计方面* [15] 撰写的讲义覆盖了许多相同内容，并更强调收敛证明、引导和离散扩散，是后面几节的有用伴读材料。对于朗之万和对数凹采样方面，Chewi 的书 [18] 是标准参考。对于 SDE 和 CTMC 的随机微积分与数值分析，我们参考 E、Li 和 Vanden-Eijnden 的书 [21]。全文中，当某个结果最早出现在研究论文中时，我们引用原始论文；上面这些讲义只在这里统一标出，作为后续若干主题的入口。

这些笔记并不试图综述完整的现代生成建模图景。我们聚焦于扩散和朗之万机制，其中反向时间 SDE、概率流 ODE、得分估计恒等式和推理时控制是核心角色。这些笔记把学习得分误差作为采样器保证的输入，但不讨论从有限数据估计得分的统计学习理论。我们也不详细处理基于似然的正规化流、对抗模型和变分模型、自回归架构、大规模潜扩散工程，或若干更新的连续时间替代方法。关于后者的入口，可参见 flow matching [34]、rectified flows [35]、stochastic interpolants [1]、consistency models [43]，以及 mean flow [23] 等近期一步生成形式。这些方法共享许多相同语言：概率路径、输运方程、去噪器以及 ODE/SDE 采样器。

最小概率背景. 我们重复使用以下约定。X 是一个随机变量，$\mathrm{Law}(X)$ 表示其分布。X 的密度为 p，期望通常写作 $\mathbb{E}[f(X)]$ 或 $\int f(x)p(x)\,\mathrm{d}x$。对于条件期望 $\mathbb{E}[Y \mid X = x]$，它最好被看作是 Y 作为观测值 x 的函数所作的最佳预测。对于马尔可夫链，$P(x, \mathrm{d}y)$ 表示给定当前状态 x 的下一个状态的分布。对于 SDE，所有形式微分都可以在光滑性和衰减假设下严格化；这些笔记使用这种形式化计算来让主要思想保持清晰可见。

我们经常使用同一个符号来表示概率分布和其密度，只要上下文足够清楚。例如，p_{data} 可以表示数据分布、该分布的密度，或者测度 $p_{\mathsf{data}}(\mathrm{d}x)$。同样，$p_t$ 可以表示 $\mathrm{Law}(X_t)$ 或其密度。这种常见的记号滥用能让公式更易读；但当确实需要密度值时，我们会写出诸如 $\nabla \log p_t(x)$ 的表达式。

循环符号. 以下符号在多个部分中被使用。在需要时，引入了更具体的符号表示。

符号	意义
p_{data}	数据分布，生成模型的目标。
X_0	干净数据样本，$X_0 \sim p_{\mathsf{data}}$.
X_t, p_t	连续时间加噪变量及其概率密度。
B_t	布朗运动驱动的向前加噪 SDE。
a_t, σ_t	连续高斯边缘参数：$X_t \mid X_0 \sim \mathcal{N}(a_t X_0, \sigma_t^2 I)$.
$\mathsf{s}_t^\star, \mathsf{s}_t$	真实的和学习得到的连续时间得分。
$Y_s^{\leftarrow}, B_s^{\leftarrow}$	反向时间 SDE 的样本路径及其反向布朗运动，反向时间为 $s = T - t$，且 $Y_0^{\leftarrow} \sim p_T$.
$Z_s^{\leftarrow}$	反向时间的概率流 ODE 样本路径，$Z_0^{\leftarrow} \sim p_T$.
X_k	离散时间的前向加噪变量 k.
p_k	X_k 的密度/概率分布。
α_k, η_k	一阶加噪参数 $X_{k+1} \mid X_k \sim \mathcal{N}(\alpha_k X_k, \alpha_k^2 \eta_k I)$；$\eta_k$ 是一步方差参数。
h_k	离散化分析中的反向步长增量 $t_{k+1} - t_k$；在方差爆炸情形下等于 η_k。
a_k, σ_k	边际参数：$X_k \mid X_0 \sim \mathcal{N}(a_k X_0, \sigma_k^2 I)$.
$P_k^{\rightarrow}(x \to x')$	前向转移密度 $\mathrm{Law}(X_{k+1} \mid X_k = x)$.
$P_k^{\leftarrow}(x' \to \cdot)$	精确的反向转移密度 $\mathrm{Law}(X_k \mid X_{k+1} = x')$.
$\mathsf{s}_k^\star$	真值得分 $\nabla \log p_k$.
s_k	学习得到的得分估计。
$\varepsilon_{k,\mathsf{score}}^2$	均方得分误差 $\mathbb{E}_{p_k} \lVert \mathsf{s}_k - \mathsf{s}_k^\star \rVert^2$.

算法缩写. 以下的缩写表示在几部分中使用的算法或算法离散化。

缩写	含义
MCMC	马尔可夫链蒙特卡洛法，通过运行一个马尔可夫链，该链的长期分布是目标分布。
ULA	未调整朗之万算法，即过阻尼朗之万扩散的 Euler–Maruyama 离散化。
MALA	修正朗之万算法，是使用密度比的接受-拒绝修正的未调整朗之万算法。
DDPM	去噪扩散概率模型，是离散高斯加噪和学习反向采样的框架。
DDIM	去噪扩散隐式模型，是一个确定性或部分随机的 DDPM 兼容采样器，与概率流离散化密切相关。
EM	Euler–Maruyama，SDE 的基本时间离散化规则。
FORS	一阶拒绝采样，一种仅依赖得分的高斯倾斜修正机制。

人工智能使用披露. 这些笔记的编写使用了大型语言模型工具，以帮助草拟、编辑、重新组织和一致性检查。数学内容、阐述选择、参考文献以及任何剩余的错误，均是作者的责任。

1. 采样与朗之万动力学简介

采样是生成看起来像从给定分布中抽样出来的随机点的问题。这很容易表述，但通常很难实现：分布可能只在未知归一化常数的意义下给出，也可能只通过数据给出，因而无法直接从它抽样。这些笔记中的策略是间接的。我们不会一次性从这个难抽样的目标中抽样，而是构造一个容易模拟的随机过程，使其分布逐渐逼近目标，并在过程运行足够长后读取样本。要精确表达这个想法，最重要的习惯是：不要只跟踪一条轨道，而要跟踪整个分布在算法中的运动。本节刻意放慢速度，先固定这种习惯，并固定我们用来衡量两个分布相距多远的语言。

1.1. **采样问题.** 设 π 是一个概率分布于 $\mathbb{R}^d$ 上。在大多数应用中，π 无法通过精确采样获得。相反，我们可能只能访问：

- 未归一化密度 $\pi(x) \propto e^{-U(x)}$
- 梯度 $\nabla \log \pi(x) = -\nabla U(x)$
- 从与 π 相关的噪声分布中采样。

目标是构造一个随机变量 $\widehat{X}$，其分布为 $\widehat{\pi}$，并满足 $D(\widehat{\pi}, \pi) \le \varepsilon$，其中 D 是适当的差异度量。D 没有唯一最优的选择。令 μ 和 ν 是 $\mathbb{R}^d$ 上的两个概率律。我们回顾三种基本差异。

(1) 总变差问的是事件的概率是否接近正确：

$$\mathrm{D}_{\mathsf{TV}}(\mu, \nu) = \sup_A |\mu(A) - \nu(A)| = \frac{1}{2} \int \left| \frac{\mathrm{d}\mu}{\mathrm{d}\lambda} - \frac{\mathrm{d}\nu}{\mathrm{d}\lambda} \right| \mathrm{d}\lambda.$$

这里 λ 是任何使 μ 和 ν 都对其绝对连续的参考测度，即 $\mu, \nu \ll \lambda$；例如可取 $\lambda = \mu + \nu$。$\mathrm{d}\mu / \mathrm{d}\lambda$ 是 μ 相对于 λ 的 Radon–Nikodym 导数，由

$$\mu(A) = \int_A \frac{\mathrm{d}\mu}{\mathrm{d}\lambda} \, \mathrm{d}\lambda$$

对所有可测集 A 成立。同样的解释适用于 $\mathrm{d}\nu / \mathrm{d}\lambda$；总变差距离的值不依赖于参考测度 λ 的选择。

(2) Wasserstein-2 距离询问概率质量是否可以在较短的几何距离下进行运输：

$$W_2^2(\mu, \nu) = \inf_{\gamma \in \Pi(\mu, \nu)} \int \|x - y\|^2 \, \gamma(\mathrm{d}x, \mathrm{d}y).$$

这里 $\Pi(\mu, \nu)$ 是 μ 和 ν 的耦合集合，表示 γ 和 $\mathbb{R}^d \times \mathbb{R}^d$ 上的概率测度，其第一边缘是 μ，其第二边缘是 ν。

(3) KL 散度询问 μ 是否可以通过 ν 作为参考进行高效编码：

$$\mathrm{D}_{\mathsf{KL}}(\mu \| \nu) = \int \log\left(\frac{\mathrm{d}\mu}{\mathrm{d}\nu} \right) \mathrm{d}\mu,$$

当 μ 绝对连续于 ν 时，$+\infty$ 否则。请注意，KL 散度是不对称的：$\mathrm{D}_{\mathsf{KL}}(\mu \| \nu) \ne \mathrm{D}_{\mathsf{KL}}(\nu \| \mu)$。特别是，KL 不是度量。

KL 是我们课程中经常使用的主记账分歧。其他度量将此信息转化为关于事件、几何位移或弱测试函数的陈述。基本比较始于 Csiszár–Kullback–Pinsker 不等式；例如参见 Bakry、Gentil 和 Ledoux [6]：

$$\mathrm{D}_{\mathsf{TV}}(\mu,\nu) \le \sqrt{\frac{1}{2}\,\mathrm{D}_{\mathsf{KL}}(\mu\|\nu)}.$$

因此，一个 KL 保证立即给出一个 TV 保证，从而也给出一个弱测试函数保证。*数据处理*：把同一个观测映射施加到两个随机对象上不会增加 KL。如果 T 可测，且 $T_{\#}\mu$ 表示 $X \sim \mu$ 时 $T(X)$ 的分布，则

$$\mathrm{D}_{\mathsf{KL}}(T_{\#}\mu\|T_{\#}\nu) \le \mathrm{D}_{\mathsf{KL}}(\mu\|\nu).$$

同样的单调性也适用于总变差。我们会反复使用这一原则，从路径空间上的比较过渡到端点分布；证明见引理 A.1。Wasserstein 距离更具几何性。如果 μ 和 ν 都支撑在直径为 R 的集合上，则最大耦合给出

$$W_2^2(\mu,\nu) \le R^2\,\mathrm{D}_{\mathsf{TV}}(\mu,\nu) \le R^2\sqrt{\frac{1}{2}\,\mathrm{D}_{\mathsf{KL}}(\mu\|\nu)}.$$

如果没有有界直径假设，或矩/泛函不等式假设，反方向并不存在普适比较：小 W_2 不必推出小 TV 或小 KL；而如果一小部分质量可以移动到很远处，小 TV 或小 KL 也不必控制 W_2。

与固定参考定律相比，这是一个更有力的比较。一个概率律 ν 若满足

$$W_2^2(\mu,\nu) \le 2C_T\,\mathrm{D}_{\mathsf{KL}}(\mu\|\nu) \qquad \text{for all } \mu \ll \nu.$$

则称其满足常数为 C_T 的 Talagrand T_2 输运-熵不等式。通向这个不等式的一条标准路径是 log-Sobolev。不妨采用这些笔记中的约定：如果 ν 满足

$$\mathrm{D}_{\mathsf{KL}}(\mu\|\nu) \le \frac{C_{\mathsf{LSI}}}{2}\,\mathrm{FI}(\mu\|\nu),$$

其中

$$\mathrm{FI}(\mu\|\nu) = \int \left\|\nabla \log \frac{\mathrm{d}\mu}{\mathrm{d}\nu}\right\|^2 \mathrm{d}\mu$$

其中导数良定义时，

$$W_2^2(\mu,\nu) \le 2C_{\mathsf{LSI}}\,\mathrm{D}_{\mathsf{KL}}(\mu\|\nu).$$

则 Otto–Villani 定理给出上述输运界。这也是 log-Sobolev 不等式和输运不等式自然出现在采样理论中的原因之一；见 Villani [47] 以及 Bakry、Gentil 和 Ledoux [6]。

例 (三种距离看到不同的东西). 令 $\mu = \delta_0$ 且 $\nu = \delta_\epsilon$，定义在 $\mathbb{R}$ 上。于是 $\mathrm{D}_{\mathsf{TV}}(\mu,\nu) = 1$ 对每个 $\epsilon \neq 0$ 都成立，因为两个点质量落在不相交的集合上。另一方面，$W_2(\mu,\nu) = \epsilon$ 且 $\mathrm{D}_{\mathsf{KL}}(\mu\|\nu) = \mathrm{D}_{\mathsf{KL}}(\nu\|\mu) = +\infty$。因此，TV 和 KL 对比较点质量与自身微小位移过于严格，而 Wasserstein 观察到这两个概率律在几何上很接近。

1.2. **马尔可夫链蒙特卡洛.** 马尔可夫链蒙特卡洛（MCMC）把采样转化为运行马尔可夫链的问题。给定 π，我们选择一条转移规则，使得反复应用这条规则后应以 π 为平衡律。马尔可夫核 P 把当前状态 x 映到一个分布 $P(x,\cdot)$。如果 $\pi P = \pi$，则称 π 是不变的；理想的 MCMC 链运行方式为

$$X_{n+1} \sim P(X_n, \cdot), \qquad n = 0, 1, \ldots,$$

并希望当 $n \to \infty$ 时 $\mathrm{Law}(X_n)$ 趋近于 π。这个观点立即分离出两个问题。第一，所设计的转移核是否真的保持目标分布不变？第二，链需要运行多久才接近平衡？对于诸如 W_2 这样的度量 D，典型的算法分解为

$$D(\mathrm{Law}(X_n), \pi) \leq \underbrace{D(\mathrm{Law}(X_n), \pi_h)}_{\text{mixing to the algorithm's invariant law}} + \underbrace{D(\pi_h, \pi)}_{\text{bias from approximation}},$$

其中 π_h 是可实现算法的不变分布。这是一个三角不等式论证，因此适用于真正的度量，但不适用于 KL 散度；一般而言并不存在形如 $\mathrm{D}_{\mathsf{KL}}(\mu\|\rho) \leq \mathrm{D}_{\mathsf{KL}}(\mu\|\nu) + \mathrm{D}_{\mathsf{KL}}(\nu\|\rho)$ 的不等式。这个分解把算法未运行足够久造成的误差，与使用近似转移规则造成的误差分开。对于朗之万采样，第二类误差出现在连续扩散被时间步长规则替代时；这种离散化可能产生一个满足 $\pi_h \neq \pi$ 的可实现链。在精确 MCMC 方法中，如 Metropolis 调整朗之万算法，转移被构造成满足 $\pi_h = \pi$.

定义 1.1 (不变且可逆的核)**.** 若对 P 有

$$\int \pi(\mathrm{d}x) P(x, A) = \pi(A) \qquad \text{for every measurable set } A.$$

则概率分布 π 对 P 是*不变的*。若详细平衡恒等式

$$\pi(\mathrm{d}x) P(x, \mathrm{d}y) = \pi(\mathrm{d}y) P(y, \mathrm{d}x)$$

作为 (x, y) 对上的测度成立，则 π 对 P 是*可逆的*。对两边关于 x 积分即可看出，可逆性推出不变性。

警告：不变性不是收敛。不变性仅是一个固定点陈述：如果 $X_0 \sim \pi$，那么链的一步仍然遵循 π 的概率分布。它本身并不直接说明，从另一个概率分布开始的链会趋向于 π。例如，$P(x,\cdot) = \delta_x$ 的身份核对每个分布都保持不变，但它从不混合。即使不完全的不变性法则是不够的，除非排除周期行为：在两点空间 $\{0,1\}$ 上，确定性的翻转 $0 \mapsto 1$，$1 \mapsto 0$ 使均匀分布成为可逆的不变概率分布，但从 0 开始的链会永无止境地振荡。因此，详细平衡是验证不变性的方便方法，而不是收敛定理。为了证明 MCMC，还需要一个马尔可夫链的对称性机制，例如有限状态空间中的不可约性和周期性，或者一般状态空间中的适当的哈里斯重叠和小数条件。然后，定量的采样界面对结构又添加了更多：谱间隙、导电性、收缩或功能不等式。请参阅 Meyn 和 Tweedie [37] 的经典处理这些马尔可夫链条件。因此，MCMC 的一种常见设计原则是先设计一个连续时间的马尔可夫过程，其 π 的不变分布，并且可以分析其收敛机制。然后，将该过程离散化以获得一个可计算的马尔可夫链，如果需要的话，可以添加蒙特卡洛修正以去除离散化偏差。Langevin 动力学是这一原则的典型例子：连续扩散的不变分布为 π，而 ULA 和 MALA 则是将这种扩散转化为算法的两种不同方法。

1.3. **过阻尼朗之万扩散.** 假设 $\pi(x) \propto e^{-U(x)}$，其中 $U : \mathbb{R}^d \to \mathbb{R}$ 光滑。过阻尼朗之万扩散是连续时间马尔可夫过程：

$$\mathrm{d}X_t = -\nabla U(X_t)\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}B_t = \nabla \log \pi(X_t)\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}B_t. \tag{1.1}$$

漂移 $-\nabla U = \nabla \log \pi$ 指向 π 更大的区域，就像梯度下降会走向更低的势能。布朗项中的扩散系数经过调节，使样本整体恰好保持 π 所规定的分散程度。这正是优化和采样的基本区别：优化要寻找一个目标值较小的点 $x_\star$，而采样要得到一个概率律。如果 π 是双峰的，并且一半质量集中在两个相距较远的模态附近，那么优化器可以正确地返回其中一个模态，但采样器必须以正确频率访问两个模态。

有两种互补的方式理解 (1.1)。从路径看，每个粒子被 $-\nabla U$ 沿势能景观向下拉动，而布朗运动保持整体分散。从分布看，X_t 的密度 q_t 由一个确定性 PDE 演化。相应的 Fokker–Planck 方程为

$$\partial_t q_t = \nabla \cdot (q_t \nabla U) + \Delta q_t = \nabla \cdot \left(q_t \nabla \log \frac{q_t}{\pi} \right). \tag{1.2}$$

在这里，(1.2) 的要点只是：随机粒子系统有一个确定性的分布层面描述。SDE 图像直观且易于模拟；PDE 图像适合证明不变性和熵耗散。

为了推导 Fokker–Planck 方程，我们回顾计算所需形式的 Itô 公式；更多细节见附录 A。如果

$$\mathrm{d}X_t = b(X_t)\,\mathrm{d}t + \sigma\,\mathrm{d}B_t$$

其中扩散矩阵 σ 为常数，则对光滑测试函数 φ，

$$\mathrm{d}\varphi(X_t) = \langle \nabla \varphi(X_t), b(X_t) \rangle\,\mathrm{d}t + \frac{1}{2} \operatorname{Tr}\left(\sigma\sigma^\top \nabla^2 \varphi(X_t) \right) \mathrm{d}t + \langle \nabla \varphi(X_t), \sigma\,\mathrm{d}B_t \rangle .$$

最后一项是鞅增量，因此取期望后消失。对 (1.1)，有 $b = -\nabla U$ 且 $\sigma = \sqrt{2}\, I$。将上述公式应用到 (1.1)，得到

$$\frac{\mathrm{d}}{\mathrm{d}t} \mathbb{E}[\varphi(X_t)] = \mathbb{E}[-\langle \nabla U(X_t), \nabla \varphi(X_t) \rangle + \Delta \varphi(X_t)].$$

如果 X_t 具有密度 q_t，则

$$\frac{\mathrm{d}}{\mathrm{d}t} \int \varphi q_t\,\mathrm{d}x = \int \left[-\langle \nabla U, \nabla \varphi \rangle + \Delta \varphi \right] q_t\,\mathrm{d}x.$$

通过积分分部，假设边界项为零。

$$\int -q_t \langle \nabla U, \nabla \varphi \rangle\,\mathrm{d}x = \int \varphi\, \nabla \cdot (q_t \nabla U)\,\mathrm{d}x, \qquad \int q_t \Delta \varphi\,\mathrm{d}x = \int \varphi\, \Delta q_t\,\mathrm{d}x.$$

由于这对所有测试函数 φ 都成立，我们得到

$$\partial_t q_t = \nabla \cdot (q_t \nabla U) + \Delta q_t.$$

最后，$\pi \propto e^{-U}$ 蕴含 $\nabla U = -\nabla \log \pi$，并且

$$\nabla \cdot (q_t \nabla U) + \Delta q_t = \nabla \cdot \left(q_t \nabla \log \frac{q_t}{\pi} \right).$$

在测试函数计算中出现的微分算子，

$$\mathcal{L}\varphi = -\langle \nabla U, \nabla \varphi \rangle + \Delta \varphi,$$

被称为朗之万扩散的微小生成器。Fokker–Planck 方程是密度的伴随方程。在弱形式下，上述计算说

$$\frac{\mathrm{d}}{\mathrm{d}t}\int \varphi q_t \,\mathrm{d}x = \int (\mathcal{L}\varphi) q_t \,\mathrm{d}x = \int \varphi (\mathcal{L}^* q_t)\,\mathrm{d}x,$$

所以 $\partial_t q_t = \mathcal{L}^* q_t$，其中

$$\mathcal{L}^* q = \nabla \cdot (q \nabla U) + \Delta q.$$

命题 1.2 (平稳性)**.** *目标密度 $\pi \propto e^{-U}$ 满足*

$$\mathcal{L}^* \pi = 0.$$

因此，如果 $q_0 = \pi$，那么 Fokker–Planck 方程的解对于所有 t 都保持 $q_t = \pi$。因此 π 是 (1.1) 的不变分布。

证明. 因为 $\nabla \log \pi = -\nabla U$，所以 $\nabla \pi = -\pi \nabla U$。因此

$$\mathcal{L}^* \pi = \nabla \cdot (\pi \nabla U) + \Delta \pi = \nabla \cdot (\pi \nabla U + \nabla \pi) = 0.$$

因此，伴随方程 $\partial_t q_t = \mathcal{L}^* q_t$ 的右端在 $q_t = \pi$ 处为零，所以密度 π 是平稳的。 □

例 (Ornstein–Uhlenbeck 过程)*.* 取 $\pi = \mathcal{N}(0, I)$，于是 $U(x) = \|x\|^2/2$。朗之万动力学变为

$$\mathrm{d}X_t = -X_t \,\mathrm{d}t + \sqrt{2}\,\mathrm{d}B_t.$$

显式解为

$$X_t = e^{-t} X_0 + \sqrt{2}\int_0^t e^{-(t-s)}\,\mathrm{d}B_s.$$

若 X_0 是确定的，则

$$X_t \sim \mathcal{N}\left(e^{-t}X_0, (1 - e^{-2t})I\right).$$

此例值得记住：均值以指数方式收敛，方差填充到目标方差。

1.4. **熵耗散.** 沿着流对 KL 求导，得到

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\mathrm{D}_{\mathsf{KL}}(q_t \| \pi) = -\int \left\| \nabla \log \frac{q_t}{\pi} \right\|^2 q_t \,\mathrm{d}x = -\,\mathrm{FI}(q_t \| \pi), \tag{1.3}$$

其中 $\mathrm{FI}(q\|\pi)$ 是相对 Fisher 信息。

(1.3) 的证明如下。. 我们有

$$\mathrm{D}_{\mathsf{KL}}(q_t \| \pi) = \int q_t \log \frac{q_t}{\pi}\,\mathrm{d}x.$$

因为 $\int \partial_t q_t \,\mathrm{d}x = 0$，求导得到

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\mathrm{D}_{\mathsf{KL}}(q_t\|\pi) = \int \partial_t q_t \,\log\frac{q_t}{\pi}\,\mathrm{d}x.$$

使用 Fokker–Planck 方程的等价形式，

$$\partial_t q_t = \nabla\cdot\Big(q_t\nabla\log\frac{q_t}{\pi}\Big),$$

并分部积分，

$$\int \nabla\cdot\Big(q_t\nabla\log\frac{q_t}{\pi}\Big)\log\frac{q_t}{\pi}\,\mathrm{d}x = -\int q_t\left\|\nabla\log\frac{q_t}{\pi}\right\|^2\mathrm{d}x.$$

这正是 $-\,\mathrm{FI}(q_t\|\pi)$。 □

一旦与泛函不等式结合，这个恒等式就变成定量收敛估计。若 π 满足 log-Sobolev 不等式

$$\mathrm{D}_{\mathsf{KL}}(q\|\pi) \le \frac{C_{\mathsf{LSI}}}{2}\,\mathrm{FI}(q\|\pi),$$

则 (1.3) 蕴含

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\mathrm{D}_{\mathsf{KL}}(q_t\|\pi) \le -\frac{2}{C_{\mathsf{LSI}}}\,\mathrm{D}_{\mathsf{KL}}(q_t\|\pi).$$

由 Grönwall 不等式，

$$\mathrm{D}_{\mathsf{KL}}(q_t\|\pi) \le e^{-2t/C_{\mathsf{LSI}}}\,\mathrm{D}_{\mathsf{KL}}(q_0\|\pi).$$

因此，连续时间过程的 KL 距离误差呈指数级减小。

特别地，如果势 U 是 m-强凸的，即 $\nabla^2 U \succeq mI$ 且 $m>0$，那么由 Bakry–Émery 判据 [6]，目标满足 $C_{\mathsf{LSI}} \le 1/m$ 的 log-Sobolev 不等式，上述估计变成干净的指数速率

$$\mathrm{D}_{\mathsf{KL}}(q_t\|\pi) \le e^{-2mt}\,\mathrm{D}_{\mathsf{KL}}(q_0\|\pi).$$

主要教训是机制：KL 散度是一个 Lyapunov 函数，相对的 Fisher 信息是它的耗散率，而目标结构不等式将耗散转化为收敛。采样方法通常遵循这种模式：选择一个差异，计算其沿着动力学的方向导数，并使用目标概率分布的结构不等式将导数转化为率。

同样，这个恒等式也有几何读法。在最优输运语言中，过阻尼朗之万是 $\mathrm{D}_{\mathsf{KL}}(\cdot\|\pi)$ 的 W_2 梯度流，而 (1.3) 是它的能量-耗散恒等式；见 Ambrosio、Gigli 和 Savaré [3]。关于扩散半群、熵耗散和 log-Sobolev 不等式的背景，见 Bakry、Gentil 和 Ledoux [6]；关于 Wasserstein 几何背后的输运观点，见 Villani [47]。

2. 朗之万扩散与 ULA 的收敛

上一节通过熵恒等式和 log-Sobolev 不等式建立了连续时间收敛。现在我们要问：当扩散被一个可实现的马尔可夫链替代后，这个故事还剩下什么？新的主要问题是数值误差：时间步进规则本身可以作为马尔可夫链收敛，但它的不变律不一定是目标律。关于对数凹采样算法的现代处理，包括朗之万算法及其 Metropolis 调整变体，可参见 Chewi 的优秀著作 [18]。

2.1. **未调整朗之万算法.** 步长为 $h > 0$ 的 Euler–Maruyama 离散化给出未调整朗之万算法（ULA）

$$X_{n+1} = X_n - h\nabla U(X_n) + \sqrt{2h}\,\xi_n, \qquad \xi_n \sim \mathcal{N}(0, I)\,. \tag{2.1}$$

ULA 常常是人们首先考虑的采样器，因为它简单、基于梯度且成本低。但它的局限同样重要：(2.1) 并不是 (1.1) 的精确转移，一般也不会保持 π 不变。

一个诱人的想法是把 ULA 看成“几乎是朗之万”，并因此假设它必然具有几乎正确的平稳律。即使对高斯目标，我们也能精确看出这里的“几乎”是什么意思。当步长较小时，偏差较小；但只要固定 h，无论链运行多久，这个偏差都不会消失。这就是混合误差和离散化偏差之间的区别：前者会随着迭代次数增加而减小，后者则内置在转移规则本身之中。

例 (标准高斯情形下的 ULA 偏差). 设一维情形下 $U(x) = x^2/2$, 因此目标为 $\pi = \mathcal{N}(0,1)$。ULA 变为马尔可夫链

$$X_{n+1} = (1-h)X_n + \sqrt{2h}\xi_n.$$

回顾高斯工具箱中的引理 B.1：高斯变量的仿射变换仍是高斯，且独立高斯变量的方差相加。在平稳状态下，X_n 和 X_{n+1} 具有相同的分布。因此，如果不变律是均值为零、方差为 v_h 的高斯分布，则

$$X_n \sim \mathcal{N}(0, v_h) \quad \Longrightarrow \quad X_{n+1} \sim \mathcal{N}\big(0, (1-h)^2 v_h + 2h\big)\,,$$

因为 $\xi_n \sim \mathcal{N}(0,1)$ 且与 X_n 独立。平稳输入和输出方差相等给出恒等式

$$v_h = (1-h)^2 v_h + 2h.$$

解得

$$v_h = \frac{2h}{1-(1-h)^2} = \frac{2}{2-h}.$$

因此，ULA 的不变律为 $\pi_h = \mathcal{N}(0, 2/(2-h))$，除非 $h = 0$，否则它不是 $\pi = \mathcal{N}(0,1)$。当 h 较小时，方差偏差为 $2/(2-h) - 1 = h/(2-h) = O(h)$。同一个偏差也可以表示为两个不变律之间的散度：

$$\mathrm{D}_{\mathsf{KL}}(\pi_h \| \pi) = \frac{1}{2}\left(v_h - 1 - \log v_h\right), \qquad \mathrm{D}_{\mathsf{KL}}(\pi \| \pi_h) = \frac{1}{2}\left(\frac{1}{v_h} - 1 + \log v_h\right),$$

以及

$$W_2^2(\pi_h, \pi) = (\sqrt{v_h} - 1)^2.$$

当 $h \downarrow 0$ 时，这三者都是 $h^2/16 + O(h^3)$。这个例子是看出未调整离散化为何产生偏差的最干净方式。

2.2. **一步 KL 离散化误差.** 令 $P_h(x, \cdot)$ 为精确朗之万过程运行时间 h 的转移，令 $\widehat{P}_h(x, \cdot)$ 为 ULA 转移。为了在 KL 中比较它们，把一个 ULA 步在区间 $[0, h]$ 上插值为

$$\mathrm{d}\widehat{X}_s = -\nabla U(x)\,\mathrm{d}s + \sqrt{2}\,\mathrm{d}B_s, \qquad \widehat{X}_0 = x.$$

从同一点出发的精确朗之万律具有漂移场 $y \mapsto -\nabla U(y)$。由于下面的 KL 是 $\mathrm{Law}(\widehat{X}_{[0,h]})$ 相对于 $\mathrm{Law}(X_{[0,h]})$ 的 KL，Girsanov 公式会沿第一条路径，即在 $\widehat{X}_s$ 处，评价两个漂移场。因此，附录 A 回顾的公式给出路径空间比较

$$\mathrm{D}_{\mathsf{KL}}\Big(\mathrm{Law}(\widehat{X}_{[0,h]})\Big\|\mathrm{Law}(X_{[0,h]})\Big) = \frac{1}{4}\int_0^h \mathbb{E}\left\|\nabla U(\widehat{X}_s) - \nabla U(x)\right\|^2 \mathrm{d}s.$$

由引理 A.1 中的数据处理不等式，端点律之间的 KL 不会更大。如果 ∇U 是 L-Lipschitz 的，则

$$\mathrm{D}_{\mathsf{KL}}\Big(\widehat{P}_h(x,\cdot)\Big\|P_h(x,\cdot)\Big) \le \frac{L^2}{4}\int_0^h \mathbb{E}\left\|\widehat{X}_s - x\right\|^2 \mathrm{d}s.$$

由于 $\widehat{X}_s = x - s\nabla U(x) + \sqrt{2}\,B_s$，

$$\mathbb{E}\left\|\widehat{X}_s - x\right\|^2 = s^2\left\|\nabla U(x)\right\|^2 + 2ds,$$

因此

$$\mathrm{D}_{\mathsf{KL}}\Big(\widehat{P}_h(x,\cdot)\Big\|P_h(x,\cdot)\Big) \le \frac{L^2}{4}\left(dh^2 + \frac{h^3}{3}\left\|\nabla U(x)\right\|^2\right). \tag{2.2}$$

这就是局部截断误差的 KL 类比：在一个短时间区间上冻结漂移，会付出 dh^2 阶的代价，外加一个依赖局部漂移大小的项。

2.3. **ULA 的直接 KL 递推.** 一步估计解释了离散化误差的尺度，但它本身并不是收敛到 π 的定理。事实上，诸如 $\mathrm{D}_{\mathsf{KL}}(\widehat{q}_k\|q_{kh})$ 这样的比较不能简单地加到 $\mathrm{D}_{\mathsf{KL}}(q_{kh}\|\pi)$ 上，因为 KL 没有三角不等式。现代基于 KL 的分析转而直接跟踪相对于目标的量 $\mathrm{D}_{\mathsf{KL}}(\widehat{q}_k\|\pi)$。

下面的结果是 Vempala 和 Wibisono [46] 的 KL 定理在我们记号下的表述。

定理 2.1 (ULA 在 KL 中的收敛)**.** *假设 $\pi(\mathrm{d}x) \propto e^{-U(x)}\,\mathrm{d}x$ 满足 log-Sobolev 不等式，其常数在我们的约定下为 C_{LSI}：*

$$\mathrm{D}_{\mathsf{KL}}(q\|\pi) \le \frac{C_{\mathsf{LSI}}}{2}\,\mathrm{FI}(q\|\pi),$$

并且 ∇U 是 L-Lipschitz 的。令 $\widehat{q}_k$ 为步长

$$0 < h \le \frac{1}{4C_{\mathsf{LSI}}L^2}$$

的 ULA (2.1) 的分布。如果 $\mathrm{D}_{\mathsf{KL}}(\widehat{q}_0\|\pi) < \infty$，则

$$\mathrm{D}_{\mathsf{KL}}(\widehat{q}_k\|\pi) \le e^{-kh/C_{\mathsf{LSI}}}\,\mathrm{D}_{\mathsf{KL}}(\widehat{q}_0\|\pi) + 8C_{\mathsf{LSI}}L^2dh. \tag{2.3}$$

这个定理不假设凸性。log-Sobolev 不等式提供全局混合机制，而 Hessian 有界假设提供离散化扩散所需的局部控制。

因此，为了达到 KL 目标 $\mathrm{D}_{\mathsf{KL}}(\widehat{q}_k\|\pi) \le \varepsilon^2$，取

$$h \lesssim \min\left\{\frac{1}{C_{\mathsf{LSI}}L^2}, \frac{\varepsilon^2}{C_{\mathsf{LSI}}L^2d}\right\}$$

并运行

$$k \gtrsim \frac{C_{\mathsf{LSI}}}{h} \log \frac{\mathrm{D}_{\mathsf{KL}}(\widehat{q}_0 \| \pi)}{\varepsilon^2}$$

次迭代就足够。等价地，迭代复杂度为

$$k = \widetilde{O}\left(\frac{C_{\mathsf{LSI}}^2 L^2 d}{\varepsilon^2} \right).$$

在这个 KL 分析中，对环境维数的依赖是线性的。由 log-Sobolev 不等式的 Bakry–Émery 判据（见 Bakry、Gentil 和 Ledoux [6]），在 m-强凸下有 $C_{\mathsf{LSI}} \le 1/m$，所以复杂度界变为 $\widetilde{O}(\kappa^2 d/\varepsilon^2)$，其中 $\kappa = L/m$ 是条件数。

证明思路. 固定一步，并从分布 $\widehat{q}_k$ 开始。通过如下冻结漂移扩散来插值 ULA 更新：在经过时间 $t \in [0,h]$ 时，

$$\mathrm{d}\bar{X}_t = -\nabla U(\bar{X}_0)\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}B_t, \qquad \bar{X}_0 \sim \widehat{q}_k.$$

记 $\nu_t = \mathrm{Law}(\bar{X}_t)$，所以 $\nu_h = \widehat{q}_{k+1}$。沿着这个插值过程对 $\mathrm{D}_{\mathsf{KL}}(\nu_t\|\pi)$ 求导，会得到和连续朗之万计算中相同的负 Fisher 信息项，以及由冻结漂移造成的误差：

$$\frac{\mathrm{d}}{\mathrm{d}t} \mathrm{D}_{\mathsf{KL}}(\nu_t\|\pi) = -\,\mathrm{FI}(\nu_t\|\pi) + \mathbb{E}\left\langle \nabla U(\bar{X}_t) - \nabla U(\bar{X}_0), \nabla \log \frac{\nu_t}{\pi}(\bar{X}_t) \right\rangle.$$

Young 不等式直接给出上述内积项的界：

$$\begin{aligned}\mathbb{E}\left\langle \nabla U(\bar{X}_t) - \nabla U(\bar{X}_0), \nabla \log \frac{\nu_t}{\pi}(\bar{X}_t) \right\rangle &\le \frac{1}{2}\mathbb{E}\left\| \nabla \log \frac{\nu_t}{\pi}(\bar{X}_t) \right\|^2 + \frac{1}{2}\mathbb{E}\left\| \nabla U(\bar{X}_t) - \nabla U(\bar{X}_0) \right\|^2 \\ &= \frac{1}{2}\,\mathrm{FI}(\nu_t\|\pi) + \frac{1}{2}\mathbb{E}\left\| \nabla U(\bar{X}_t) - \nabla U(\bar{X}_0) \right\|^2.\end{aligned}$$

由 L-光滑性，$\left\|\nabla U(\bar{X}_t) - \nabla U(\bar{X}_0)\right\| \le L\left\|\bar{X}_t - \bar{X}_0\right\|$，所以

$$\frac{\mathrm{d}}{\mathrm{d}t} \mathrm{D}_{\mathsf{KL}}(\nu_t\|\pi) \le -\frac{1}{2}\,\mathrm{FI}(\nu_t\|\pi) + \frac{L^2}{2}\mathbb{E}\left\|\bar{X}_t - \bar{X}_0\right\|^2.$$

log-Sobolev 不等式把第一项转化为收缩：$\mathrm{FI}(\nu_t\|\pi) \ge 2C_{\mathsf{LSI}}^{-1}\,\mathrm{D}_{\mathsf{KL}}(\nu_t\|\pi)$。于是

$$\frac{\mathrm{d}}{\mathrm{d}t} \mathrm{D}_{\mathsf{KL}}(\nu_t\|\pi) \le -\frac{1}{C_{\mathsf{LSI}}}\,\mathrm{D}_{\mathsf{KL}}(\nu_t\|\pi) + \frac{L^2}{2}\mathbb{E}\left\|\bar{X}_t - \bar{X}_0\right\|^2.$$

剩下的是控制冻结漂移插值的位移。由于

$$\mathbb{E}\left\|\bar{X}_t - \bar{X}_0\right\|^2 = t^2 \mathbb{E}_{\widehat{q}_k}\left\|\nabla U\right\|^2 + 2dt,$$

在 $t \in [0,h]$ 上积分后，布朗部分贡献局部 dh^2 项。对漂移部分，使用 $\widehat{q}_k$ 与 π 之间的最优 W_2 耦合。沿着这样的耦合，光滑性给出 $\left\|\nabla U(x)\right\|^2 \le 2\left\|\nabla U(y)\right\|^2 + 2L^2\left\|x-y\right\|^2$，平均后得到

$$\mathbb{E}_{\widehat{q}_k}\left\|\nabla U\right\|^2 \le 2\mathbb{E}_\pi\left\|\nabla U\right\|^2 + 2L^2 W_2^2(\widehat{q}_k, \pi).$$

第一项至多为 $2Ld$，因为在 $\pi \propto e^{-U}$ 下分部积分给出 $\mathbb{E}_\pi\left\|\nabla U\right\|^2 = \mathbb{E}_\pi \Delta U \le Ld$。第二项由第 1.1 小节中的 Talagrand 比较 $W_2^2(\widehat{q}_k, \pi) \le 2C_{\mathsf{LSI}}\,\mathrm{D}_{\mathsf{KL}}(\widehat{q}_k\|\pi)$ 控制。因此，漂移贡献是一个低

阶的 dh^2 项，加上 $C_{\mathsf{LSI}}L^4h^3\,\mathrm{D}_{\mathsf{KL}}(\widehat{q}_k\|\pi)$ 的倍数；上述步长条件使最后这部分足够小，可以被收缩吸收。

Vempala 和 Wibisono [46] 的一步不等式，在我们的记号下为

$$\mathrm{D}_{\mathsf{KL}}(\widehat{q}_{k+1}\|\pi) \le e^{-h/C_{\mathsf{LSI}}}\,\mathrm{D}_{\mathsf{KL}}(\widehat{q}_k\|\pi) + 6L^2dh^2.$$

迭代这个递推得到

$$\mathrm{D}_{\mathsf{KL}}(\widehat{q}_k\|\pi) \le e^{-kh/C_{\mathsf{LSI}}}\,\mathrm{D}_{\mathsf{KL}}(\widehat{q}_0\|\pi) + 6L^2dh^2\sum_{j=0}^{k-1} e^{-jh/C_{\mathsf{LSI}}}.$$

最后，在同样的小步长假设下，$\sum_{j=0}^{k-1} e^{-jh/C_{\mathsf{LSI}}} \le (1-e^{-h/C_{\mathsf{LSI}}})^{-1} \le 4C_{\mathsf{LSI}}/(3h)$。因此最后一项至多为 $8C_{\mathsf{LSI}}L^2dh$，这就得到 (2.3)。这正是熵耗散的离散类似物：log-Sobolev 提供收缩，而光滑性控制 Euler–Maruyama 误差。 □

我们看一下这个定理的含义。对维数的依赖是线性的，但对精度的依赖是多项式的：为了让 KL 误差达到 ε^2 阶，ULA 需要 $\widetilde{O}(C_{\mathsf{LSI}}^2L^2d/\varepsilon^2)$ 步。这是因为 Euler 离散化有 h 阶的平稳偏差，所以高精度迫使步长很小。因此自然要问：能否保留朗之万提议，但通过精确修正去掉离散化偏差？当密度比可用时，经典的 Metropolis 调整正是这样做的。

2.4. **Metropolis 调整朗之万算法.** Metropolis 调整朗之万算法（MALA）用一个接受-拒绝步骤修正 ULA 提议。它与只使用梯度的设定之间有一个核心区别：密度值让经典的高精度拒绝思想成为可能，而仅有梯度信息并不能提供这些思想所需的密度比。

令 $\pi(\mathrm{d}x) \propto e^{-U(x)}\,\mathrm{d}x$。从当前状态 x 出发，MALA 先抽取 ULA 提议

$$y \sim q_h(x,\cdot) = \mathcal{N}(x - h\nabla U(x), 2hI)\,.$$

然后以概率

$$a_h(x,y) = 1 \wedge \frac{\pi(y)q_h(y,x)}{\pi(x)q_h(x,y)} \tag{2.4}$$

接受 y。如果提议被拒绝，链就停留在 x。Metropolis 比率 (2.4) 强制详细平衡：

$$\pi(\mathrm{d}x)P_h^{\mathsf{MALA}}(x,\mathrm{d}y) = \pi(\mathrm{d}y)P_h^{\mathsf{MALA}}(y,\mathrm{d}x).$$

因此，π 对 MALA 是精确平稳的。这是它相对于 ULA 的基本优势：离散化偏差被移除了，代价是需要密度比 $\pi(y)/\pi(x)$。

MALA 是修正机制中最干净的例子。提议和 ULA 一样只使用梯度信息，但接受-拒绝步骤会询问所提议的移动在目标密度下是否具有正确概率。这个单一的密度比检查，把不变分布从近似的 π_h 改回精确目标 π。

不过，精确平稳性只是故事中“不变性”的部分。要把 MALA 变成一个定量采样算法，仍需界定需要多少个被接受或被拒绝的提议之后，链才接近 π。现代高精度理论把这个分析分成

两个任务：在已有良好初始化时高效采样，以及首先产生这个初始化。两者都相对于暖启动来表述。标准的暖启动条件如下：若对每个可测集 A 都有

$$\mu_0(A) \le M\pi(A),$$

或等价地

$$\left\| \frac{\mathrm{d}\mu_0}{\mathrm{d}\pi} \right\|_{L^\infty(\pi)} \le M, \tag{2.5}$$

则称初始律 μ_0 相对于 π 是 M-暖启动的。有些暖启动准备结果使用有限阶 Rényi 版本：

$$D_q(\mu_0\|\pi) = \frac{1}{q-1}\log\int\left(\frac{\mathrm{d}\mu_0}{\mathrm{d}\pi}\right)^q \mathrm{d}\pi = O(1)$$

其中 $q>1$。注意，在 Rényi 记号下，(2.5) 可写为 $D_\infty(\mu_0\|\pi)\le \log M$。

对于第一个任务，固定总变差目标精度 $\varepsilon\in(0,1)$。假设有一个 M-暖启动，Wu、Schmidler 和 Chen [51] 证明，对于 $\mathbb{R}^d$ 中 m-强对数凹且 L-光滑的目标，MALA 的混合迭代次数为

$$O\left(\kappa\sqrt{d}\,\log^3\max\left\{\kappa, d, \frac{M}{\varepsilon}\right\}\right), \qquad \kappa = L/m,$$

其中隐含常数是普适常数。Chen 和 Gatmiry [17] 在光滑性和等周性假设下扩展了这个暖启动 MALA 图像；在强对数凹情形下，他们恢复了相同的主导 $\kappa\sqrt{d}$ 行为，并且同样只对 M/ε 有对数依赖。

对于第二个任务，Altschuler 和 Chewi [2] 表明，动能朗之万（kinetic Langevin），也称欠阻尼朗之万，可以以同样的 $\widetilde{O}(\sqrt{d})$ 维数依赖产生所需的有限阶 Rényi 暖启动。从维数依赖的层面看，高精度图像是：用动能朗之万准备暖启动，然后用 MALA 作为精确 Metropolis 修正采样器。

上面的讨论也显示了经典调整型 MCMC 的局限。MALA 在密度值可用时去除了离散化偏差，但它的尖锐收敛保证需要结构性假设，例如强对数凹性、log-Sobolev 或等周不等式、光滑性以及暖启动。一般的数据分布可能是多模态的、奇异的，或者只能通过样本获得，因此这些假设不是自然的出发点。这引出了一个不同的问题：如果给定的是数据而不是显式目标密度，我们应当如何设计一个算法来生成新样本?

3. 基于得分的扩散模型

上一节处理的是从显式密度中采样：算法可以使用 ∇U，MALA 甚至可以使用密度比来去除离散化偏差。扩散模型从一个不同前提出发。目标律由数据表示，而不是由一个易处理的公式表示，因此我们并不试图直接在数据分布上运行朗之万动力学。相反，我们给数据加噪，形成一条由更光滑的概率律组成的路径。在每个正噪声水平上，加噪律的得分就是一个局部向量场，它引导反向动力学回到数据。

支撑这种方法的“去噪-得分”联系，在一定程度上可以追溯到 Ho、Jain 和 Abbeel 的 DDPM 表述 [26]。这里使用的连续时间 SDE 和概率流表述遵循 Yang Song、Sohl-Dickstein、Kingma、

Kumar、Ermon 和 Poole [44]。基本图像很简单：在时间 t，加噪分布 p_t 是数据的模糊版本，而 $\nabla \log p_t(x)$ 指向附近使这个模糊密度更大的区域。扩散模型学习的正是这个以时间为索引的局部去噪方向场。

3.1. **连续时间前向加噪.** 概念上最干净的起点是连续时间。令 $X_0 \sim p_{\mathsf{data}}$，并令 $(X_t)_{t\in[0,T]}$ 解一个前向加噪 SDE

$$\mathrm{d}X_t = f_t(X_t)\,\mathrm{d}t + g_t\,\mathrm{d}B_t, \qquad X_0 \sim p_{\mathsf{data}}. \tag{3.1}$$

这里 B_t 是布朗运动，f_t 是漂移场，g_t 是标量扩散系数。我们用 p_t 表示 X_t 的密度。大多数扩散模型使用线性高斯前向过程；对某些确定性函数 a_t 和 σ_t，它们满足

$$X_t \mid X_0 \sim \mathcal{N}\big(a_t X_0, \sigma_t^2 I\big)\,. \tag{3.2}$$

例 (两种常见的连续调度). 对于方差爆炸热流，

$$X_t = X_0 + \sqrt{t}\,Z, \qquad p_t = p_{\mathsf{data}} * \mathcal{N}(0, tI)\,,$$

其高斯转移核为

$$X_t \mid X_0 \sim \mathcal{N}(X_0, tI)\,.$$

对于常速率的 Ornstein–Uhlenbeck 或方差保持流，

$$\mathrm{d}X_t = -\frac{1}{2}X_t\,\mathrm{d}t + \mathrm{d}B_t,$$

有

$$X_t \mid X_0 \sim \mathcal{N}\big(e^{-t/2}X_0, (1-e^{-t})I\big)\,.$$

两个例子都符合 (3.2)；区别只在于 a_t 和 σ_t。方差保持约定满足 $a_t^2 + \sigma_t^2 = 1$，因此当数据具有单位尺度时，总边际方差被保持。

两个标量函数 a_t 和 σ_t 概括了信噪比。系数 a_t 表示原始样本在条件均值中还剩下多少可见信号，而 σ_t 表示加入了多少独立高斯不确定性。在前向过程的早期，σ_t 很小，得分可能很复杂，因为它反映数据的精细结构。在高噪声处，p_t 更光滑，并且更接近一个简单参考律，因此反向采样器有一个更容易的起点。

3.2. **连续时间 Tweedie 恒等式.** 撤销前向加噪最自然的方式是去噪：给定一个带噪观测 $X_t = x$，估计产生它的干净样本，也就是后验均值 $\mathbb{E}[X_0 \mid X_t = x]$。编码这种去噪信息的对象，是加噪律的得分

$$\mathsf{s}_t^\star(x) = \nabla \log p_t(x), \tag{3.3}$$

也即时间 t 的真实得分；实践中它由学习模型 $\mathsf{s}_t(x) \approx \mathsf{s}_t^\star(x)$ 近似。高斯边际公式 (3.2) 通过连续时间形式的 Tweedie 恒等式，精确连接了得分与去噪：

$$\mathsf{s}_t^\star(x) = \frac{1}{\sigma_t^2}\mathbb{E}[a_t X_0 - X_t \mid X_t = x]. \tag{3.4}$$

等价地，只要 $a_t \neq 0$，定义最优去噪器

$$\mathsf{D}_t^\star(x) := \mathbb{E}[X_0 \mid X_t = x] = a_t^{-1}\big(x + \sigma_t^2 \mathsf{s}_t^\star(x)\big).$$

学习得分 s_t 也类似地确定一个学习去噪器

$$\mathsf{D}_t(x) := a_t^{-1}\big(x + \sigma_t^2 \mathsf{s}_t(x)\big).$$

因此，去噪和得分估计是描述同一个后验信息的两种方式。

证明. X_t 的密度是高斯混合

$$p_t(x) = \int \frac{1}{(2\pi\sigma_t^2)^{d/2}} \exp\left(-\frac{\|x - a_t x_0\|^2}{2\sigma_t^2}\right) p_{\mathsf{data}}(\,\mathrm{d}x_0).$$

把这个积分中的高斯核写作

$$\varphi_t(x \mid x_0) = \frac{1}{(2\pi\sigma_t^2)^{d/2}} \exp\left(-\frac{\|x - a_t x_0\|^2}{2\sigma_t^2}\right).$$

对固定的 x_0，

$$\nabla_x \varphi_t(x \mid x_0) = \frac{a_t x_0 - x}{\sigma_t^2}\, \varphi_t(x \mid x_0).$$

因此，在积分号下求导，

$$\nabla p_t(x) = \frac{1}{\sigma_t^2} \int (a_t x_0 - x)\, \varphi_t(x \mid x_0)\, p_{\mathsf{data}}(\,\mathrm{d}x_0).$$

另一方面，Bayes 公式给出带噪观测下干净样本的条件律：

$$\mathbb{P}(X_0 \in \,\mathrm{d}x_0 \mid X_t = x) = \frac{\varphi_t(x \mid x_0)\, p_{\mathsf{data}}(\,\mathrm{d}x_0)}{p_t(x)}.$$

因此，把上一个展示式除以 $p_t(x)$，得到

$$\nabla \log p_t(x) = \frac{1}{\sigma_t^2} \mathbb{E}[a_t X_0 - x \mid X_t = x],$$

这与 (3.4) 相同，因为条件事件设置了 $X_t = x$。 □

这个 Tweedie 恒等式解释了为什么得分可以从经验数据中学习。固定 $t > 0$。在高斯腐蚀 $X_t = a_t X_0 + \sigma_t Z$ 下，其中 $Z \sim \mathcal{N}(0, I)$ 且与 X_0 独立，这个条件期望内部的随机向量变为

$$\frac{a_t X_0 - X_t}{\sigma_t^2} = -\frac{Z}{\sigma_t}.$$

因此，在给定 $a_t X_0 + \sigma_t Z = x$ 时，$-Z/\sigma_t$ 的条件均值恰好是 $\mathsf{s}_t^\star(x)$。这就引出了如下回归总体损失：

$$\mathbb{E}\left\|\mathsf{s}_t(a_t X_0 + \sigma_t Z) + \frac{Z}{\sigma_t}\right\|^2.$$

把对 $X_0 \sim p_{\mathsf{data}}$ 的期望替换为对数据点的经验平均，并重新采样高斯噪声，就得到基本的去噪得分匹配目标。

这也解释了为什么学习到的场具有去噪解释，而不是任意向量场。对每一次具体腐蚀，回归目标是有噪声的；但它的条件平均会从观测指向干净样本的后验均值 $\mathsf{D}_t^\star(x)$。等价地，得分记录了由加噪数据分布编码的 Bayes 修正。

3.3. **连续时间反向 SDE.** 前向 SDE (3.1) 被设计成把数据推向一个简单律。对于采样，我们需要的对象并不是布朗运动路径层面的逆。只要能构造一个马尔可夫过程，使它的一时边际按相反顺序经过同一组密度，就足够了：如果 $(p_t)_{0\le t\le T}$ 是前向加噪边际，那么反向采样器 $(Y_s^{\leftarrow})_{0\le s\le T}$ 应满足

$$\mathrm{Law}(Y_s^{\leftarrow}) = p_{T-s} \qquad \text{for every } s\in[0,T].$$

这个要求刻意是分布层面的。它说反向采样器具有正确快照，而不是说它逐条重走前向布朗路径。因此我们首先匹配密度的演化。令 $q_s = p_{T-s}$，并写 $t = T-s$。前向密度满足

$$\partial_t p_t = -\nabla\cdot(f_t p_t) + \frac{1}{2}g_t^2\Delta p_t.$$

因此，当 $t = T-s$ 时，

$$\partial_s q_s = \nabla\cdot(f_t p_t) - \frac{1}{2}g_t^2\Delta p_t.$$

关键步骤使用恒等式 $\Delta p_t = \nabla\cdot(p_t\nabla\log p_t)$，它恰好引入了 (3.3) 中的得分 $\mathsf{s}_t^\star = \nabla\log p_t$。于是密度演化可以改写为

$$\partial_s q_s = -\nabla\cdot\left[\left(-f_t + g_t^2\mathsf{s}_t^\star\right)p_t\right] + \frac{1}{2}g_t^2\Delta p_t.$$

这正是反向时间扩散的 Fokker–Planck 方程：

$$\mathrm{d}Y_s^{\leftarrow} = \left[-f_{T-s}(Y_s^{\leftarrow}) + g_{T-s}^2\,\mathsf{s}_{T-s}^\star(Y_s^{\leftarrow})\right]\mathrm{d}s + g_{T-s}\,\mathrm{d}B_s^{\leftarrow}, \qquad Y_0^{\leftarrow}\sim p_T. \tag{3.5}$$

这里 $B_s^{\leftarrow}$ 是反向采样时间中的布朗运动。由构造可知，如果 $Y_0^{\leftarrow}\sim p_T$，那么对每个 $s\in[0,T]$ 都有 $Y_s^{\leftarrow}\sim p_{T-s}$。

反向 SDE 的漂移有两部分。项 $-f_{T-s}$ 反转前向动力学中的确定性输运部分。额外项 $g_{T-s}^2\,\mathsf{s}_{T-s}^\star$ 正是第 3.2 小节中作为去噪方向引入的得分：它通过把每个样本推回到加噪密度更高的区域，也就是推向干净样本的后验均值，来修正由扩散导致的密度演化。这就是上一小节的去噪器正是反向动力学引擎的精确含义。把真实得分 $\mathsf{s}_t^\star$ 换成学习得分 s_t，就得到基于得分的反向 SDE 采样器；它是 DDPM 型算法所离散化的连续对象。

把 (3.5) 特化到贯穿示例的两种调度是有启发的；下一小节我们还会从概率流观点重新考察这两个过程。

例 (方差爆炸反向 SDE). 对 $X_t = X_0 + \sqrt{t}Z$，有 $f_t = 0$ 和 $g_t = 1$，所以反向 SDE 为

$$\mathrm{d}Y_s^{\leftarrow} = \mathsf{s}_{T-s}^\star(Y_s^{\leftarrow})\,\mathrm{d}s + \mathrm{d}B_s^{\leftarrow}, \qquad Y_0^{\leftarrow}\sim p_T.$$

整个漂移就是得分：从高噪声样本出发，采样器沿去噪方向 $\mathsf{s}_{T-s}^\star$ 反复前进，同时注入新的布朗噪声。

例 (方差保持 Ornstein–Uhlenbeck 反向 SDE). 对于

$$\mathrm{d}X_t = -\frac{1}{2}X_t\,\mathrm{d}t + \mathrm{d}B_t,$$

有 $f_t = -\frac{1}{2}x$ 和 $g_t = 1$，因此反向 SDE 为

$$\mathrm{d}Y_s^{\leftarrow} = \left[\frac{1}{2}Y_s^{\leftarrow} + \mathsf{s}^\star_{T-s}(Y_s^{\leftarrow})\right]\mathrm{d}s + \mathrm{d}B_s^{\leftarrow}, \qquad Y_0^{\leftarrow} \sim p_T.$$

项 $\frac{1}{2}Y_s^{\leftarrow}$ 抵消 OU 向原点的收缩，而得分项 $\mathsf{s}^\star_{T-s}$ 提供去噪修正。

在实践中，反向 SDE 以离散时间运行，而这种离散化正是 DDPM。人们在采样时间 s 上固定一个网格，用学习模型 s_{T-s} 替换真实得分 $\mathsf{s}^\star_{T-s}$，并对 (3.5) 采取 Euler–Maruyama 步：每一步沿去噪漂移推动当前样本，然后加入具有适当方差的新高斯噪声。这就是 Ho、Jain 和 Abbeel [26] 的 DDPM 采样器，是下一小节将从概率流 ODE 中读出的确定性 DDIM 更新的随机对应物。详细的高斯转移核以及由此产生的离散化误差，将在第 5 节讨论。

3.4. **概率流 ODE.** 存在一个确定性 ODE，它的一时边际与前向 SDE (3.1) 的一时边际相同。这个 ODE 称为*概率流 ODE*。概率流 ODE 用一个输运场取代布朗随机性，并在边际密度上产生相同效果。定义速度场

$$v_t(x) = f_t(x) - \frac{1}{2}g_t^2\mathsf{s}_t^\star(x) = f_t(x) - \frac{1}{2}g_t^2\nabla\log p_t(x). \tag{3.6}$$

概率流 ODE 为

$$\frac{\mathrm{d}X_t}{\mathrm{d}t} = v_t(X_t). \tag{3.7}$$

如果 $X_0 \sim p_0$，那么只要 (3.7) 适定，它的解在每个时刻 t 都具有边际密度 p_t。

从 *Fokker–Planck 方程推导*. 前向 SDE (3.1) 的 Fokker–Planck 方程为

$$\partial_t p_t = -\nabla\cdot(f_t p_t) + \frac{1}{2}g_t^2\Delta p_t. \tag{3.8}$$

使用恒等式

$$\Delta p_t = \nabla\cdot(\nabla p_t) = \nabla\cdot(p_t\nabla\log p_t) = \nabla\cdot(p_t\mathsf{s}_t^\star).$$

于是 (3.8) 变为

$$\partial_t p_t = -\nabla\cdot(f_t p_t) + \frac{1}{2}g_t^2\nabla\cdot(p_t\mathsf{s}_t^\star) = -\nabla\cdot\left[\left(f_t - \frac{1}{2}g_t^2\mathsf{s}_t^\star\right)p_t\right].$$

这正是由确定性流 $\dot{X}_t = v_t(X_t)$ 输运的密度所满足的连续性方程

$$\partial_t p_t + \nabla\cdot(p_t v_t) = 0. \tag{3.9}$$

由于 p_t 以初始条件 p_0 满足这个方程，ODE 流把 p_0 推前到 p_t。 □

例 (方差爆炸概率流). 对 $X_t = X_0 + \sqrt{t}Z$，有 $f_t = 0$ 和 $g_t = 1$。概率流速度为

$$v_t(x) = -\frac{1}{2}\nabla\log p_t(x).$$

得分指向更高密度，因此 $-\frac{1}{2}\nabla\log p_t$ 把质量向外推，并复现方差爆炸加噪过程的平滑效果。为了生成样本，人们沿时间反向积分同一个 ODE，从而反转这个速度。

例 (方差保持 Ornstein–Uhlenbeck 流). 对于

$$\mathrm{d}X_t = -\frac{1}{2}X_t\,\mathrm{d}t + \mathrm{d}B_t,$$

概率流速度为

$$v_t(x) = -\frac{1}{2}x - \frac{1}{2}\nabla \log p_t(x).$$

第一项是向原点收缩的确定性 OU 项。第二项是布朗平滑的输运表示。

反向 SDE 和概率流 ODE 是两台不同机器；当得分精确时，它们产生相同快照。SDE 机器持续注入随机性；ODE 机器则确定性地输运粒子（从随机初始数据出发）。由于许多数值问题和控制问题依赖路径，而不只依赖快照，因此在没有检查所分析量的情况下，不应随意把一台机器替换为另一台机器。

为了采样，我们使用与 (3.5) 相同的反向时间约定：令 $s = T - t$，从分布为 p_T 的高噪声样本出发，从 $s = 0$ 积分到 $s = T$。反向时间概率流采样器是 ODE

$$\frac{\mathrm{d}Z_s^{\leftarrow}}{\mathrm{d}s} = -v_{T-s}(Z_s^{\leftarrow}) = -f_{T-s}(Z_s^{\leftarrow}) + \frac{1}{2}g_{T-s}^2 \mathsf{s}_{T-s}^\star(Z_s^{\leftarrow}), \qquad Z_0^{\leftarrow} \sim p_T.$$

由同一个连续性方程计算，当得分精确时，$Z_s^{\leftarrow}$ 的分布是 p_{T-s}。把它和 (3.5) 中的反向 SDE 漂移 $-f_{T-s} + g_{T-s}^2 \mathsf{s}_{T-s}^\star$ 相比较。Song、Meng 和 Ermon [42] 的 DDIM 采样器是概率流这一侧的标准采样器：它使用与 DDPM 相同的训练好去噪模型，但遵循 ODE 观点所提示的确定性更新。

ODE 的确定性对似然有一个重要后果。沿着 $\dot{X}_t = v_t(X_t)$ 的解，连续性方程 (3.9) 蕴含

$$\frac{\mathrm{d}}{\mathrm{d}t}\log p_t(X_t) = -\nabla \cdot v_t(X_t). \tag{3.10}$$

确实，

$$\frac{\mathrm{d}}{\mathrm{d}t}\log p_t(X_t) = \partial_t \log p_t(X_t) + \langle v_t(X_t), \nabla \log p_t(X_t)\rangle\,,$$

而将 (3.9) 除以 p_t 给出

$$\partial_t \log p_t + \langle v_t, \nabla \log p_t\rangle = -\nabla \cdot v_t.$$

因此，如果映射 $X_0 \mapsto X_T$ 是通过向前积分概率流 ODE 得到的，那么

$$\log p_0(X_0) = \log p_T(X_T) + \int_0^T \nabla \cdot v_t(X_t)\,\mathrm{d}t.$$

这就是用于从概率流 ODE 计算似然的连续正规化流恒等式。

评注. 在实践中，真实得分 $\mathsf{s}_t^\star$ 会被学习得分 s_t 替代。于是，实际实现的反向 SDE 和概率流 ODE 都不再精确具有目标边际。它们的最终采样误差来自得分误差和数值离散化误差：对 SDE 来说是时间步进误差，对概率流来说是 ODE 积分误差。对于使用概率流 ODE 进行似然计算，还有一个额外的数值问题：还必须近似散度 $\nabla \cdot v_t$，通常通过自动微分或随机迹估计器完成。

4. 随机定位与 Polchinski 流

上一节通过高斯加噪介绍了扩散模型。我们从数据出发，跟随加噪边际 p_t，学习得分 $\nabla \log p_t$，并用这个得分运行反向 SDE 或概率流 ODE。本节从两个互补角度考察同一个高斯通道：随机定位和 Polchinski 流。三种观点都从一个干净随机变量和带噪观测出发；这里以及下文我们取方差爆炸形式

$$X_t = X_\star + \sqrt{t} Z,$$

但它们以不同方式组织信息。扩散模型强调反向时间采样。随机定位提出一个 Bayes 问题：当带噪观测越来越精确时，隐藏信号 $X_\star$ 的后验律如何演化？Polchinski 流提出一个密度层面的问题：当噪声水平变化时，平滑密度 p_t 和有效势 $U_t = -\log p_t$ 如何演化？

这两种视角转变很有用，因为它们暴露了仅从反向 SDE 不那么容易看见的结构。在随机定位中，得分表现为后验去噪修正。在 Polchinski 流中，得分是一个粗粒化能量景观的负梯度。这两个观点都会自然导向定量工具，尤其是协方差恒等式；这些工具在后面会很有用。

我们先把随机定位看成同一个高斯通道的观测模型。这样可以比较扩散时间与定位精度，推出后验律，并把后验均值与得分联系起来。然后记录使定位方法有用的鞅结构。最后，我们回到密度 p_t 本身，并写出有效势 U_t 的相应 Polchinski 流。

4.1. 随机定位.

令 $X_\star \sim \mu$ 为信号，或干净数据点。这里把它的分布写作 μ 而不是 p_{data}，是为了强调这个构造适用于一般分布，而不只适用于数据律。从高斯平滑通道出发：

$$X_t = X_\star + \sqrt{t} Z, \qquad Z \sim \mathcal{N}(0, I)\,.$$

这里 t 是噪声方差，$p_t = \mathrm{Law}(X_t)$ 是数据律与协方差 tI 的高斯卷积。当 t 增大时，密度变得更光滑；当 t 减小时，通道接近原始数据分布。

随机定位使用同一个高斯通道，但用精度而不是方差来参数化它。通过连续时间高斯观测过程观察同一个未知量 $X_\star$：

$$\mathrm{d}Y_u = X_\star \,\mathrm{d}u + \,\mathrm{d}W_u, \qquad Y_0 = 0. \tag{4.1}$$

这里 W_u 是与 $X_\star$ 独立的布朗运动。在给定 $X_\star = x$ 时，有 $Y_u \sim \mathcal{N}(ux, uI)$，因此重标度后

$$\bar{Y}_u := u^{-1} Y_u \sim \mathcal{N}\big(x, u^{-1} I\big)\,.$$

所以定位时间 u 对应于扩散噪声方差 $t = u^{-1}$：

$$\bar{Y}_u = X_\star + \sqrt{t} Z, \qquad t = \frac{1}{u}.$$

同一个带噪观测既可以用方差 t 索引，也可以用精度 u 索引。扩散记号强调 t 增大时的平滑。定位记号强调 u 增大、观测更有信息时的 Bayes 推断。这种精度参数化很有用，因为后验均值和协方差会满足简单的鞅恒等式，下面会讨论。

现在显式写出这个后验律。在精度 u 处，观测为 Y_u，基本对象是隐藏信号 $X_\star$ 的条件律。令 $\mu_u(\cdot \mid y) = \mathrm{Law}(X_\star \mid Y_u = y)$ 表示观测到 $Y_u = y$ 后的后验律。Bayes 公式给出

$$\mu_u(\,\mathrm{d}x \mid Y_u = y) = \frac{1}{Z_u(y)} \exp\Big\{\langle y, x\rangle - \frac{u}{2}\|x\|^2\Big\}\,\mu(\,\mathrm{d}x). \tag{4.2}$$

归一化常数为

$$Z_u(y) = \int \exp\left\{\langle y, x\rangle - \frac{u}{2}\|x\|^2\right\} \mu(\mathrm{d}x).$$

这解释了为什么这个方法被称为定位：后验是原始测度经过随机线性场倾斜后，再乘上一个逐渐增强的二次惩罚得到的。随着 u 增大，后验越来越集中在隐藏信号附近。

(4.2) 的推导. 给定 $X_\star = x$ 时，观测 Y_u 的密度正比于

$$\exp\left(-\frac{\|y - ux\|^2}{2u}\right).$$

展开平方项，

$$-\frac{\|y - ux\|^2}{2u} = -\frac{\|y\|^2}{2u} + \langle y, x\rangle - \frac{u}{2}\|x\|^2.$$

第一项与 x 无关，因此被吸收到归一化常数中。再乘以前验测度 $\mu(\mathrm{d}x)$，就得到 (4.2)。 □

现在把随机观测放回去，并写

$$\mu_u = \mu_u(\cdot \mid Y_u), \qquad M_u = \int x\, \mu_u(\mathrm{d}x).$$

随机定位的原始观点是：随机测度 $(\mu_u)_{u\geq 0}$ 本身由一个测度值 SDE 演化；下面推导这一点。令 $\mathcal{F}_u^Y = \sigma(Y_r : 0 \leq r \leq u)$，并定义创新过程

$$I_u = Y_u - \int_0^u M_s\, \mathrm{d}s, \qquad \mathrm{d}I_u = \mathrm{d}Y_u - M_u\, \mathrm{d}u.$$

这减去了下一段观测中已经能由当前后验预测的部分。直观上，I_u 只记录观测流中的“惊讶”：在条件化于 $\mathcal{F}_u^Y$ 后，$\mathrm{d}Y_u$ 的漂移为 $M_u\,\mathrm{d}u$，因此减去它后，剩下的增量没有可预测成分，并且累积协方差与原始布朗噪声相同。更形式化地，可以检查 I_u 是一个连续的 $\mathcal{F}_u^Y$-鞅，其在 $[0, u]$ 上的累积协方差为 uI；于是 Levy 表征把它识别为相对于 $\mathcal{F}_u^Y$ 的布朗运动。后验律满足

$$\mathrm{d}\mu_u(\mathrm{d}x) = \langle x - M_u,\, \mathrm{d}I_u\rangle\ \mu_u(\mathrm{d}x). \tag{4.3}$$

等价地，对每个有界测试函数 f，

$$\mathrm{d}\int f(x)\, \mu_u(\mathrm{d}x) = \left\langle \int f(x)(x - M_u)\, \mu_u(\mathrm{d}x),\, \mathrm{d}I_u \right\rangle.$$

当相关矩有限时，同一个恒等式可逐分量用于向量值测试函数。特别地，对每个 Borel 集 A，过程 $u \mapsto \mu_u(A) = \Pr(X_\star \in A \mid \mathcal{F}_u^Y)$ 是一个鞅：

$$\mathbb{E}[\mu_v(A) \mid \mathcal{F}_u^Y] = \mu_u(A), \qquad v \geq u.$$

因此，在这种表述中，随机定位是一个测度值鞅；它的随机倾斜使先验集中，同时保持条件期望。

(4.3) 的证明. 只需计算归一化后验权重的微分。由 (4.2)，沿随机观测路径有

$$\frac{\mu_u(\,\mathrm{d}x)}{\mu(\,\mathrm{d}x)} = \frac{1}{Z_u}\exp\Big\{\langle Y_u, x\rangle - \frac{u}{2}\|x\|^2\Big\}.$$

首先，Itô 公式给出

$$\begin{aligned}\mathrm{d}\exp\Big\{\langle Y_u, x\rangle - \frac{u}{2}\|x\|^2\Big\} &= \exp\Big\{\langle Y_u, x\rangle - \frac{u}{2}\|x\|^2\Big\}\Big(\langle x,\,\mathrm{d}Y_u\rangle - \frac{1}{2}\|x\|^2\,\mathrm{d}u\Big)\\ &\quad + \frac{1}{2}\exp\Big\{\langle Y_u, x\rangle - \frac{u}{2}\|x\|^2\Big\}\|x\|^2\,\mathrm{d}u\\ &= \exp\Big\{\langle Y_u, x\rangle - \frac{u}{2}\|x\|^2\Big\}\langle x,\,\mathrm{d}Y_u\rangle.\end{aligned}$$

中间一行包含 Itô 修正；它抵消了指数中 $-\frac{u}{2}\|x\|^2$ 项带来的漂移。对这个恒等式关于 x 积分，得到

$$\frac{\mathrm{d}Z_u}{Z_u} = \langle M_u,\,\mathrm{d}Y_u\rangle, \qquad \mathrm{d}\log Z_u = \langle M_u,\,\mathrm{d}Y_u\rangle - \frac{1}{2}\|M_u\|^2\,\mathrm{d}u.$$

因此，μ_u 相对于 μ 的对数密度满足

$$\mathrm{d}\log\frac{\mu_u(\,\mathrm{d}x)}{\mu(\,\mathrm{d}x)} = \langle x - M_u,\,\mathrm{d}Y_u\rangle - \frac{1}{2}\big(\|x\|^2 - \|M_u\|^2\big)\,\mathrm{d}u.$$

再次应用 Itô 公式，这次作用在该对数密度的指数上，得到

$$\frac{\mathrm{d}\mu_u(\,\mathrm{d}x)}{\mu_u(\,\mathrm{d}x)} = \langle x - M_u,\,\mathrm{d}Y_u\rangle - \langle x - M_u, M_u\rangle\,\mathrm{d}u = \langle x - M_u,\,\mathrm{d}I_u\rangle.$$

这正是 (4.3)；对有界测试函数 f 积分即得弱形式。 □

4.2. **后验均值与鞅结构.** 上一小节构造了后验律及其测度值鞅方程。现在转向它的一阶和二阶矩。我们先冻结观测，并把后验均值与扩散模型的得分联系起来；随后把随机观测放回去，并记录鞅动力学。

对确定性观测 y，定义后验均值和协方差

$$m_u(y) = \mathbb{E}[X_\star \mid Y_u = y], \qquad \Sigma_u(y) = \mathrm{Cov}(X_\star \mid Y_u = y).$$

后验均值正是一个去噪器。确实，由于 $\bar{Y}_u = Y_u/u = X_\star + u^{-1/2}Z$，确定性观测 y 对应于带噪样本 y/u；Tweedie 恒等式 (3.4) 给出

$$m_u(y) = \frac{y}{u} + \frac{1}{u}\nabla\log p_{1/u}\Big(\frac{y}{u}\Big), \tag{4.4}$$

其中 p_t 是 $X_\star + \sqrt{t}Z$ 的密度。等价地，

$$\nabla\log p_{1/u}(y/u) = u\big(m_u(y) - y/u\big).$$

因此，学习得分就是学习定位后验均值。

把 Tweedie 特化到定位通道. 重标度观测 $\bar{Y}_u = X_\star + u^{-1/2}Z$ 是方差爆炸约定下的前向通道 $X_t = X_\star + \sqrt{t}Z$ 在时间 $t = 1/u$ 处的形式，此时 $a_t = 1$，$\sigma_t^2 = t = 1/u$。因此，连续时间 Tweedie 恒等式 (3.4) 在当前记号下读作

$$\nabla \log p_{1/u}(y/u) = u\,\mathbb{E}[X_\star - y/u \mid Y_u = y] = u\big(m_u(y) - y/u\big),$$

其中使用了 $\mathbb{E}[X_\star \mid Y_u = y] = m_u(y)$。解出 $m_u(y)$ 就得到 (4.4)。 □

协方差则是这个去噪器的线性响应：把后验倾斜 (4.2) 对观测求导，得到

$$\nabla_y m_u(y) = \Sigma_u(y). \tag{4.5}$$

把它与 Tweedie 恒等式 (4.4) 结合，并写 $t = 1/u$，得到得分-Jacobian 恒等式

$$\nabla \mathsf{s}_t^\star(x) = t^{-2}\Sigma_{1/t}(x/t) - t^{-1}I. \tag{4.6}$$

因此，在这个时间变换下，加噪对数密度的 Hessian 正是定位过程的后验协方差。

现在回到随机定位路径。下面把 M_u 和 Σ_u 分别写作随机取值 $m_u(Y_u)$ 和 $\Sigma_u(Y_u)$。测度值方程 (4.3) 把固定观测下的后验矩变成随机动力学。

从均值开始。把 (4.3) 的弱形式逐分量应用于测试函数 $f(x) = x$，在标准可积性假设下得到

$$\mathrm{d}M_u = \left(\int x(x - M_u)^\top \mu_u(\,\mathrm{d}x)\right)\,\mathrm{d}I_u = \Sigma_u\,\mathrm{d}I_u. \tag{4.7}$$

因此，均值只响应创新布朗运动而移动，而这种移动的大小和方向由当前后验协方差控制。特别地，后验均值是一个鞅：在看到下一小段数据之前，当前后验均值就是下一时刻后验均值的最佳预测。

为什么后验均值是鞅. 首先注意，虽然 M_u 是通过条件化于单个观测 Y_u 定义的，但如果条件化于整个历史 $\mathcal{F}_u^Y = \sigma(Y_r : 0 \le r \le u)$，它并不会改变。原因是 Y_u 是关于 $X_\star$ 的充分统计量：由 Girsanov 定理（附录 A），在信号值 $X_\star = x$ 下路径 $(Y_r)_{0\le r\le u}$ 相对于布朗运动的似然为

$$\exp\!\left(\int_0^u \langle x,\,\mathrm{d}Y_r\rangle - \frac{u}{2}\,\|x\|^2\right) = \exp\!\left(\langle x, Y_u\rangle - \frac{u}{2}\,\|x\|^2\right),$$

它对路径的依赖只通过端点 Y_u。这正是 (4.2) 中出现的倾斜，因此给定整个历史的后验等于给定 Y_u 的后验，特别地

$$\mathbb{E}[X_\star \mid \mathcal{F}_u^Y] = \mathbb{E}[X_\star \mid Y_u] = M_u.$$

现在鞅性质由塔式性质推出。对 $v \ge u$，由于 $\mathcal{F}_u^Y \subseteq \mathcal{F}_v^Y$，

$$\mathbb{E}[M_v \mid \mathcal{F}_u^Y] = \mathbb{E}[\mathbb{E}[X_\star \mid \mathcal{F}_v^Y] \mid \mathcal{F}_u^Y] = \mathbb{E}[X_\star \mid \mathcal{F}_u^Y] = M_u.$$ □

协方差决定这些均值波动的大小，它自身也有一个演化。不同于均值，它带有漂移，并且这个漂移是严格耗散的：在一维中，

$$\frac{\mathrm{d}}{\mathrm{d}u}\mathbb{E}[\Sigma_u] = -\mathbb{E}[\Sigma_u^2] \le 0,$$

而在多维中，迹满足

$$\frac{\mathrm{d}}{\mathrm{d}u}\mathbb{E}[\operatorname{Tr}\Sigma_u] = -\mathbb{E}[\operatorname{Tr}(\Sigma_u^2)] \le 0. \tag{4.8}$$

把 (4.8) 在 $u \in [0,\infty)$ 上积分，微积分基本定理给出

$$\int_0^\infty \mathbb{E}\operatorname{Tr}(\Sigma_u^2)\,\mathrm{d}u = \mathbb{E}\operatorname{Tr}\Sigma_0 - \lim_{u\to\infty}\mathbb{E}\operatorname{Tr}\Sigma_u \le \mathbb{E}\operatorname{Tr}\Sigma_0,$$

不等式来自 $\operatorname{Tr}\Sigma_u \ge 0$。在 $u = 0$ 时，还没有做任何观测，所以 $\Sigma_0 = \operatorname{Cov}(X_\star)$；因此如果 $\operatorname{Cov}(X_\star) \preceq I$，

$$\int_0^\infty \mathbb{E}\operatorname{Tr}(\Sigma_u^2)\,\mathrm{d}u \le \mathbb{E}\operatorname{Tr}\Sigma_0 \le d, \tag{4.9}$$

所以沿整条路径累积的总平方协方差由维数控制。这就是定位的几何内容：随机后验测度稳步集中，而它的均值按鞅增量漂移，增量大小由剩余不确定性设定。

一维协方差计算. 在一维中，后验方差分解为二阶矩和均值平方之差：

$$\Sigma_u = \mathbb{E}[X_\star^2 \mid Y_u] - M_u^2,$$

我们分别跟踪这两部分。对二阶矩，把 (4.3) 的弱形式用于测试函数 $f(x) = x^2$，得到

$$\mathrm{d}\,\mathbb{E}[X_\star^2 \mid Y_u] = \Big(\int x^2(x - M_u)\,\mu_u(\mathrm{d}x)\Big)\,\mathrm{d}I_u = \Big(\mathbb{E}[X_\star^3 \mid Y_u] - M_u\mathbb{E}[X_\star^2 \mid Y_u]\Big)\,\mathrm{d}I_u,$$

其中使用了 $\mathbb{E}[X_\star \mid Y_u] = M_u$。对均值平方，$\mathrm{d}M_u = \Sigma_u\,\mathrm{d}I_u$ 和 Itô 公式给出

$$\mathrm{d}(M_u^2) = 2M_u\Sigma_u\,\mathrm{d}I_u + \Sigma_u^2\,\mathrm{d}u.$$

两式相减，

$$\mathrm{d}\Sigma_u = \Big(\mathbb{E}[X_\star^3 \mid Y_u] - M_u\mathbb{E}[X_\star^2 \mid Y_u] - 2M_u\Sigma_u\Big)\,\mathrm{d}I_u - \Sigma_u^2\,\mathrm{d}u.$$

$\mathrm{d}I_u$ 项是鞅增量，均值为零，所以取期望后只剩漂移：

$$\frac{\mathrm{d}}{\mathrm{d}u}\mathbb{E}[\Sigma_u] = -\mathbb{E}[\Sigma_u^2].$$

在 d 维中，同样的计算现在用于矩阵值测试函数 $f(x) = xx^\top$，得到协方差 SDE，其分量为

$$\mathrm{d}(\Sigma_u)_{ij} = \sum_k \mathbb{E}\big[(X_\star - M_u)_i(X_\star - M_u)_j(X_\star - M_u)_k \mid Y_u\big]\,(\mathrm{d}I_u)_k - (\Sigma_u^2)_{ij}\,\mathrm{d}u. \tag{4.10}$$

第一项是均值为零的鞅增量，所以取迹并取期望，就得到 (4.8)：$\frac{\mathrm{d}}{\mathrm{d}u}\mathbb{E}[\operatorname{Tr}\Sigma_u] = -\mathbb{E}[\operatorname{Tr}(\Sigma_u^2)]$。 □

这种“鞅加协方差”的结构使随机定位成为有用的分析工具。它不是直接界定原始高维测度，而是把它嵌入到一条随机路径中；路径上的后验被逐步倾斜，均值是鞅，协方差耗散。许多静态问题由此转化为估计定位路径上的平均量，而协方差恒等式 (4.8)–(4.9) 提供了记账工具。

随机定位的早期主要应用 [22] 出现在高维几何中。对各向同性对数凹测度，Kannan–Lovász–Simonovits（KLS）猜想的一种表述，是要求 Cheeger 等周常数具有与维数无关的下界。这个问题与谱隙、Poincaré 不等式、薄壳估计和测度集中密切相关。Eldan 引入随机定位，把薄壳界和谱隙/KLS 界联系起来，误差至多为对数因子 [22]。Lee 和 Vempala 进一步发展该方法，用于等周性、集中和混合，并改进了各向同性对数凹测度的已知 KLS 型界 [32]。Chen 后来的工作使用相关定位技术，得到了 KLS 等周系数的近似常数下界 [14]。Chen 和 Eldan 还明确建立了采样联系：他们的定位方案框架把马尔可夫链与测度定位鞅联系起来，并通过分析定位过程证明混合界 [16]。对我们来说，教训是随机定位把全局几何或采样问题转化为对随机后验协方差过程的估计。同样的协方差预算观点，后来也重新出现在扩散模型分析中 [38, 9]。

4.3. **有效势的 Polchinski 流.** 定位图像通过后验量描述高斯通道。Polchinski 流则在加噪边际层面描述同一个通道。用扩散模型语言说，这就是上一节中的方差爆炸前向过程：

$$X_t = X_0 + \sqrt{t}Z,$$

其边际密度是高斯平滑后的密度

$$p_t = p_0 * \mathcal{N}(0, tI)\,.$$

它满足热方程

$$\partial_t p_t = \frac{1}{2}\Delta p_t. \tag{4.11}$$

定义有效势

$$U_t(x) = -\log p_t(x).$$

噪声水平 t 处的扩散模型得分因此为 $\mathsf{s}_t^\star = \nabla \log p_t = -\nabla U_t$。所以 Polchinski 流并不是一个新的采样过程；它是跟踪同一个得分场如何随着前向扩散平滑数据分布而演化的一种方式。热方程变成非线性 PDE

$$\partial_t U_t = \frac{1}{2}\Delta U_t - \frac{1}{2}\left\|\nabla U_t\right\|^2. \tag{4.12}$$

这就是*有限维 Polchinski 方程*。它是 Polchinski 重整化流 [41, 7] 的各向同性高斯卷积版本。密度表述是线性的，但有效势表述是非线性的，因为取对数会把平滑变成带黏性的 Hamilton–Jacobi 方程。

证明. 因为 $p_t = e^{-U_t}$，

$$\Delta p_t = \Delta(e^{-U_t}) = e^{-U_t}\big(\|\nabla U_t\|^2 - \Delta U_t\big).$$

使用 (4.11)，

$$\partial_t U_t = -\frac{\partial_t p_t}{p_t} = -\frac{1}{2}\frac{\Delta p_t}{p_t} = \frac{1}{2}\Delta U_t - \frac{1}{2}\left\|\nabla U_t\right\|^2. \qquad \square$$

由于得分是有效势的负梯度，

$$\mathsf{s}_t^\star(x) = \nabla \log p_t(x) = -\nabla U_t(x),$$

得分满足带黏性的 Burgers 方程

$$\partial_t \mathsf{s}_t^\star = \frac{1}{2}\Delta \mathsf{s}_t^\star + \nabla\left(\frac{1}{2}\left\|\mathsf{s}_t^\star\right\|^2\right). \tag{4.13}$$

扩散模型通常通过得分匹配训练，而不是求解 Burgers 方程。尽管如此，这个方程解释了为什么得分在正噪声水平处正则性会改善，也解释了为什么小噪声估计很微妙：当 $t \downarrow 0$ 时，控制导数的平滑作用消失。

Polchinski 方程 (4.12) 很有用，因为它说明前向加噪对能量景观做了什么。$U_0 = -\log p_0$ 的尖锐势阱、脊和非光滑特征会被卷积平均掉。在正噪声处，U_t 更光滑，其梯度表现更好。反向模型试图沿着这一族平滑景观向后走，而几何会逐渐变得更尖锐。

这就是扩散模型的重整化解释。历史上，重整化进入统计物理，是通过这样一个想法：应当在选定分辨率上描述一个系统，并跟踪当微观自由度被平均掉时，有效描述如何变化。Kadanoff 的块自旋缩放图像 [29] 给出了临界附近粗粒化的物理图景。Wilson 的重整化群把这个图景变成有效理论的系统流，并解释了临界点处的普适性 [48, 49]；Wilson 和 Kogut 的综述 [50] 是经典参考。随后，Polchinski 的精确重整化群方程 [41] 直接在有效 Lagrangian 层面表达了这个流。有限维方程 (4.12) 是同一个思想在扩散模型中的影子。

重整化群语言说，前向加噪积分掉了微观信息。在图像模型中，这种说法是隐喻性的，但很有用：数据分布被逐渐模糊，高频细节变得更难区分，而有效势 U_t 描述粗粒化律的对数密度。在场论应用中，同一个想法可以是字面意义的：人们选择一个按尺度去除自由度的协方差调度，相应有效作用量遵循 Polchinski 型方程。在 Bauerschmidt 及其合作者的工作中，这个流正是以这种严格意义使用的：它是尺度依赖有效势的演化方程，因此沿流的估计把重整化转化为对随机动力学和 Gibbs 测度的解析控制 [7]。

4.4. **一张词典.** 现在，可以用三种语言概括同一个高斯平滑操作：扩散动力学、后验定位，以及通过高斯粗粒化进行的重整化。

扩散模型视角	随机定位视角	Polchinski/RG 视角
前向加噪 $X_t = X_0 + \sqrt{t}Z$	观测 $\mathrm{d}Y_u = X_\star\,\mathrm{d}u + \mathrm{d}W_u$	通过高斯卷积进行粗粒化
加噪律 p_t	后验律 $\mu_u(\cdot \mid Y_u)$，其中 $t = 1/u$	有效势 $U_t = -\log p_t$
得分 $\mathsf{s}_t^\star = \nabla \log p_t$	后验均值 $M_u = \mathbb{E}[X_\star \mid Y_u]$	作用量梯度 $-\nabla U_t$
Tweedie 去噪器 $x + t\mathsf{s}_t^\star(x)$	定位均值 $m_u(y)$	用于采样的逆 RG 漂移
反向 SDE 或概率流 ODE	后验均值鞅过程	Hamilton–Jacobi/Polchinski PDE

5. 连续扩散模型的离散化

现在我们从理想的连续动力学转向算法。采样器必须选择有限时间网格，并把每一个精确反向转移替换为可计算的转移核。本节建立数值分析所需的公式。

保持表述清晰的一种有用方式，是把三个实现问题分开。第一，反向链必须从一个简单的高噪声律初始化，而不是从精确终端律初始化。第二，反向动力学使用学习得到的得分或去噪器，而不是真实对象。第三，连续反向动力学，或者等价地，网格上的精确 Bayes 反向核，必须被离散化。后续分析正是围绕这个三分法展开：初始化误差、统计得分误差和离散化误差。

5.1. **DDPM 网格、得分与精确反向核.** 令 $X_0 \sim p_{\mathsf{data}}$，定义在 $\mathbb{R}^d$ 上。选择时间网格

$$0 = t_0 < t_1 < \cdots < t_K = T.$$

目标是用一个有限马尔可夫链替代连续加噪过程，并保持规定的一时边际。回顾 (3.2)，这些边际满足

$$X_t \mid X_0 \sim \mathcal{N}\big(a_t X_0, \sigma_t^2 I\big)\,.$$

在网格上，我们写

$$X_k = X_{t_k}, \qquad a_k = a_{t_k}, \qquad \sigma_k = \sigma_{t_k}.$$

网格调度决定一步加噪参数。我们选择高斯前向转移

$$X_{k+1} \mid X_k \sim \mathcal{N}\big(\alpha_k X_k, \alpha_k^2 \eta_k I\big)\,. \tag{5.1}$$

如果

$$X_k = a_k X_0 + \sigma_k Z_k, \qquad Z_k \sim \mathcal{N}(0, I)\,,$$

则 (5.1) 等价于

$$X_{k+1} = \alpha_k X_k + \alpha_k \sqrt{\eta_k}\xi_k, \qquad \xi_k \sim \mathcal{N}(0, I)\,,$$

从而

$$X_{k+1} \mid X_0 \sim \mathcal{N}\big(\alpha_k a_k X_0, \alpha_k^2(\sigma_k^2 + \eta_k) I\big)\,.$$

将它与目标边际 $X_{k+1} \mid X_0 \sim \mathcal{N}\big(a_{k+1} X_0, \sigma_{k+1}^2 I\big)$ 匹配，得到

$$\alpha_k = \frac{a_{k+1}}{a_k}, \qquad \eta_k = \frac{\sigma_{k+1}^2}{\alpha_k^2} - \sigma_k^2.$$

方差保持情形满足 $a_k^2 + \sigma_k^2 = 1$。方差爆炸情形满足 $a_k \equiv 1$，并且噪声尺度 $\sigma_k^2 = t_k$ 单调增大。

反向采样器必须反转一个网格步：给定第 $k+1$ 层噪声处的点，它需要上一层 k 的条件律。这正是去噪和得分信息进入的地方。令 p_k 表示 X_k 的密度。时间 k 的真实得分为

$$\mathsf{s}_k^\star(x) = \nabla \log p_k(x).$$

得分决定干净样本的后验均值。这是连续时间 Tweedie 恒等式 (3.4) 的网格时间特化，其中 a_t, σ_t 在网格上读取：

$$\mathsf{s}_k^\star(x) = \frac{1}{\sigma_k^2}\mathbb{E}[a_k X_0 - X_k \mid X_k = x]. \tag{5.2}$$

等价地，第 3.2 小节中最优去噪器的网格时间版本为

$$\mathsf{D}_k^\star(x) := \mathbb{E}[X_0 \mid X_k = x] = a_k^{-1}\big(x + \sigma_k^2 \mathsf{s}_k^\star(x)\big).$$

这也是与随机定位中的后验均值 $m_u(y)$ 对应的网格时间记号：在方差爆炸归一化 $\sigma_t^2 = t$ 下，有 $\mathsf{D}_t^\star(x) = x + t\mathsf{s}_t^\star(x) = m_{1/t}(x/t)$。得分模型 s_k 等价地给出一个去噪器模型

$$\mathsf{D}_k(x) := a_k^{-1}\big(x + \sigma_k^2 \mathsf{s}_k(x)\big), \qquad \mathsf{D}_k(x) - \mathsf{D}_k^\star(x) = a_k^{-1}\sigma_k^2\big(\mathsf{s}_k(x) - \mathsf{s}_k^\star(x)\big).$$

在方差爆炸归一化 $a_t \equiv 1$ 且 $\sigma_t^2 = t$ 下，反向漂移可写为

$$\mathsf{s}_t^\star(x) = \frac{\mathsf{D}_t^\star(x) - x}{t}.$$

因此，可以把采样器理解为使用得分场，也可以把它理解为使用一个最优去噪器，并在每个离散反向步中冻结该去噪器的值。

在用得分模型近似反向步之前，先用边际密度 p_k 写出精确 Bayes 核。令 $P_k^{\leftarrow}(x' \to \cdot)$ 为反向转移 $\mathrm{Law}(X_k \mid X_{k+1} = x')$ 的密度，并令 $P_k^{\rightarrow}(x \to x')$ 表示 (5.1) 中的前向转移密度。对固定的 x，

$$P_k^{\rightarrow}(x \to x') = (2\pi\alpha_k^2\eta_k)^{-d/2} \exp\left(-\frac{\|x' - \alpha_k x\|^2}{2\alpha_k^2\eta_k}\right).$$

Bayes 公式给出

$$P_k^{\leftarrow}(x' \to x) = \frac{p_k(x) P_k^{\rightarrow}(x \to x')}{p_{k+1}(x')}, \qquad p_{k+1}(x') = \int p_k(y) P_k^{\rightarrow}(y \to x')\, \mathrm{d}y.$$

因此精确反向核为

$$P_k^{\leftarrow}(x' \to x) \propto p_k(x) \exp\left(-\frac{\left\|x - \alpha_k^{-1}x'\right\|^2}{2\eta_k}\right). \tag{5.3}$$

于是每一个反向步都是一个 Bayes 更新：p_k 是前一噪声层的先验，高斯因子是重标度观测 $\alpha_k^{-1}x'$ 在通道 $\alpha_k^{-1}X_{k+1} = X_k + \sqrt{\eta_k}\xi_k$ 下的似然。如果 p_k 作为密度已知，就可以直接使用 (5.3)。在扩散模型中，学习到的对象只是得分 $\nabla \log p_k$，所以下一步是局部近似这个 Bayes 更新。

5.2. **高斯反向更新.** 常见采样器把 (5.3) 替换为一个高斯转移。在本节记号下，

$$\widehat{P}_k^{\leftarrow}(x' \to \cdot) = \mathcal{N}\big(\alpha_k^{-1}x' + \eta_k\alpha_k\mathsf{s}_{k+1}(x'), \eta_k I\big), \tag{5.4}$$

其中同一个步长参数 η_k 用作高斯方差。应把它与精确反向核 (5.3) 对照：由于先验因子 $p_k(x)$ 的存在，精确反向核并不一定是高斯的。

为了看出 (5.4) 中均值为何具有这种形式，对单个前向步本身应用 Tweedie 恒等式。重标度下一步观测后，

$$\alpha_k^{-1}X_{k+1} = X_k + \sqrt{\eta_k}\xi_k.$$

这个通道的 Tweedie 恒等式给出

$$\begin{aligned}\mathbb{E}[X_k \mid X_{k+1} = x'] &= \alpha_k^{-1}x' + \alpha_k\eta_k\nabla\log\big(p_{k+1}(x')\big) \\ &= \alpha_k^{-1}x' + \alpha_k\eta_k\mathsf{s}_{k+1}^\star(x').\end{aligned}$$

因此，(5.4) 的中心就是把 $\mathsf{s}_{k+1}^\star$ 替换为学习得分 s_{k+1} 后的精确后验均值。剩下的近似，是把精确后验协方差替换为局部高斯尺度 $\eta_k I$：当先验密度 p_k 光滑时，(5.3) 中的似然把 X_k 限制在观测附近半径约为 $\sqrt{\eta_k}$ 的球内，这是主导阶行为。

我们还给出一个来自数值 SDE 观点的互补推导。选择 $[0,T]$ 上的前向加噪 SDE，使其边际与通道 $X_t \mid X_0 \sim \mathcal{N}(a_t X_0, \sigma_t^2 I)$ 一致。假设调度光滑，$a_t > 0$，并具有通常归一化 $a_0 = 1$、$\sigma_0 = 0$；定义

$$\lambda_t = \frac{\dot{a}_t}{a_t}, \qquad f_t(x) = \lambda_t x, \qquad g_t^2 = \frac{\mathrm{d}}{\mathrm{d}t}\sigma_t^2 - 2\lambda_t \sigma_t^2.$$

假设最后这个量非负，从而 g_t 良定义，则线性 SDE

$$\mathrm{d}X_t = \lambda_t X_t \,\mathrm{d}t + g_t \,\mathrm{d}B_t$$

具有条件均值 $a_t X_0$ 和条件方差 $\sigma_t^2 I$：方差满足 $\dot{v}_t = 2\lambda_t v_t + g_t^2$ 且 $v_0 = 0$，因此 $v_t = \sigma_t^2$。把 (3.5) 应用于这个特定的 f_t 和 g_t，得到反向 SDE

$$\mathrm{d}Y_s^{\leftarrow} = \left\{ -\lambda_{T-s} Y_s^{\leftarrow} + g_{T-s}^2 \nabla \log p_{T-s}(Y_s^{\leftarrow}) \right\} \mathrm{d}s + g_{T-s} \,\mathrm{d}B_s^{\leftarrow}.$$

现在在网格上对这个反向 SDE 应用 Euler–Maruyama。对从时间 t_{k+1} 到时间 t_k 的反向步，记 $h_k = t_{k+1} - t_k$，并从噪声层 t_{k+1} 的点 x' 出发。把系数和得分冻结在这个反向步的起点，即前向时间 t_{k+1} 处，并使用网格简记 $\lambda_{k+1} := \lambda_{t_{k+1}}$ 与 $g_{k+1} := g_{t_{k+1}}$，得到

$$x' - h_k \lambda_{k+1} x' + h_k g_{k+1}^2 \mathsf{s}_{k+1}(x') + g_{k+1}\sqrt{h_k}\,\xi.$$

离散参数正是同一调度在网格上的版本：

$$\alpha_k = \frac{a_{k+1}}{a_k}, \qquad \eta_k = \frac{\sigma_{k+1}^2}{\alpha_k^2} - \sigma_k^2.$$

在 t_{k+1} 处作 Taylor 展开，得到

$$\alpha_k^{-1} = 1 - h_k \lambda_{k+1} + O(h_k^2), \qquad \alpha_k \eta_k = h_k g_{k+1}^2 + O(h_k^2),$$

这与上面讨论的高斯反向更新一致。

这里值得明确区分两个步长参数，因为它们都会在误差分析中反复出现：$h_k = t_{k+1} - t_k$ 是反向步的时间增量，而 η_k 是一步高斯核 (5.1) 的方差。在一般调度中，它们是不同量，通过 $\alpha_k \eta_k = h_k g_{k+1}^2 + O(h_k^2)$ 相联系；但在第 6 节所用的方差爆炸归一化中（$\alpha_k = 1$、$g_t \equiv 1$），二者相同，即 $\eta_k = h_k$。因此，更新 (5.4) 既可看作对精确 Bayes 核的局部近似，也可看作对同一高斯加噪通道所对应反向 SDE 的 Euler–Maruyama 离散化。离散化分析要问的是：在每一个反向步都作这种局部替换，会引入多少误差。

6. 扩散模型的误差分析

上一节分离出了局部的数值问题：每一步都有一个精确的 Bayes 核 (5.3)，而实际采样器使用诸如 (5.4) 这样的可实现近似。现在我们把这个局部比较转化为全局保证。除了平滑目标与原始数据律之间的早停间隙之外，反向链误差分解为上一节已经分离出的三个来源：链如何初始化，得分学习得有多准确，以及每个反向核离散化得有多准确。

在进入细节之前，有必要先给出整个分析的路线图；后续小节会把每一部分精确化。贯穿本节，$P_k^{\leftarrow}(x_{k+1}\to\cdot)$ 表示精确反向核 (5.3)，而 $\widehat{P}_k^{\leftarrow}(x_{k+1}\to\cdot)$ 表示某个采样器实际实现的反向核。对每个时间 k，$L^2(p_k)$ 得分误差为

$$\varepsilon_{k,\mathsf{score}}^2 = \mathbb{E}_{X_k\sim p_k}\left\|\mathsf{s}_k(X_k)-\mathsf{s}_k^\star(X_k)\right\|^2. \tag{6.1}$$

实际反向链在高噪声端点初始化，其误差为

$$\mathrm{D}_{\mathsf{KL}}(p_K\|\widehat{p}_K)\le\varepsilon_{\mathsf{init}}^2,$$

累积的一步核误差分解为纯离散化部分和得分误差部分，

$$\sum_{k=1}^{K-1}\mathbb{E}_{X_{k+1}\sim p_{k+1}}\mathrm{D}_{\mathsf{KL}}\left(P_k^{\leftarrow}(X_{k+1}\to\cdot)\middle\|\widehat{P}_k^{\leftarrow}(X_{k+1}\to\cdot)\right)\le\varepsilon_{\mathsf{disc}}^2+c\sum_{k=1}^{K-1}\eta_k\varepsilon_{k+1,\mathsf{score}}^2.$$

把这些项合在一起，我们将建立的误差分解为

$$\mathrm{D}_{\mathsf{KL}}(p_1\|\widehat{p}_1)\le\varepsilon_{\mathsf{init}}^2+\varepsilon_{\mathsf{disc}}^2+c\sum_{k=1}^{K-1}\eta_k\varepsilon_{k+1,\mathsf{score}}^2, \tag{6.2}$$

而 p_1 与 p_{data} 之间的早停比较则作为几何平滑估计单独处理，并自然地由 W_2 控制。

这个分解主要是一个记账陈述。我们不只比较最终样本，而是把精确反向链和实际反向链作为完整路径来比较。随后，数据处理不等式说明终端边际之间的 KL 不大于路径空间 KL。路径空间的链式法则又把后者拆成两部分：高噪声端点处两个初始律的不匹配，以及沿反向链累积的一步 KL 误差之和。

乘在得分误差前面的因子有一个简单来源。在高斯 DDPM 或 Euler–Maruyama 更新中，学习得分会把一步高斯核的均值改变约 $\eta_k(\mathsf{s}_{k+1}-\mathsf{s}_{k+1}^\star)$。具有相同协方差 $\eta_k I$ 的高斯分布之间的 KL，会把均值误差平方除以方差，因此留下量级为 $\eta_k\varepsilon_{k+1,\mathsf{score}}^2$ 的贡献。因此，在反向步长较大的时间处，得分误差更重要；而很小步长上的误差则自然被降权。这里 η_k 是第 5.1 节中的一步方差参数；在本节其余部分分析的方差爆炸链中，它正好是反向时间步长 $h_k=t_{k+1}-t_k$，这也解释了为什么这些权重表现为步长。

项 $\varepsilon_{\mathsf{disc}}^2$ 收集即便使用精确得分仍然存在的误差。对 Euler–Maruyama 来说，这是在一步中冻结得分或去噪器所付出的代价。随机定位论证通过后验协方差预算来控制这一代价，而后面的高精度修正通过更忠实地采样局部高斯倾斜来降低它。下面各小节按如下顺序展开这些部分：早停、KL 望远镜分解、得分误差贡献、Euler–Maruyama 离散化及其定位改进，最后是高精度修正。

6.1. **早停.** 第一个问题是反向过程的终点。我们停在一个小的正噪声水平，并以平滑后的律 p_1 为目标。这个早停律是 KL 分析的正确对象：路径空间链式法则和 Girsanov 型估计可以直接控制采样器输出 $\widehat{p}_1$ 与它之间的 $\mathrm{D}_{\mathsf{KL}}(p_1\|\widehat{p}_1)$。原始数据分布可能是奇异的、支撑在低维集合上，或者与任何平滑采样器输出互相奇异，因此一般不能在 KL 中与 $\widehat{p}_1$ 比较。

因此我们把两种误差度量分开。早停误差是几何平滑误差，并由 W_2 控制。p_1 与 $\widehat{p}_1$ 之间的算法误差则在后面用 KL 控制。

W_2 估计直接来自前向加噪耦合。在上一节的网格记号中，$X_1 = a_1 X_0 + \sigma_1 Z$，因此把 X_1 与同一个干净样本 X_0 耦合，得到

$$W_2^2(p_{\mathsf{data}}, p_1) \le (1-a_1)^2 \mathbb{E}\,\|X_0\|^2 + \sigma_1^2 d. \tag{6.3}$$

因此，当 σ_1 和 $1-a_1$ 相对于数据尺度很小时，早停是无害的。

早停估计的证明. 用同一个干净数据点 X_0 以及形成

$$X_1 = a_1 X_0 + \sigma_1 Z, \qquad Z \sim \mathcal{N}(0, I)$$

所用的同一份高斯噪声来耦合 p_1 和 p_{data}。于是

$$X_1 - X_0 = (a_1 - 1) X_0 + \sigma_1 Z.$$

利用 $\mathbb{E}Z = 0$ 以及 Z 与 X_0 的独立性，

$$\mathbb{E}\,\|X_1 - X_0\|^2 = (1-a_1)^2 \mathbb{E}\,\|X_0\|^2 + \sigma_1^2 \mathbb{E}\,\|Z\|^2 = (1-a_1)^2 \mathbb{E}\,\|X_0\|^2 + \sigma_1^2 d.$$

由于 W_2^2 是所有耦合代价的下确界，它不超过这个特定耦合的代价。 □

6.2. **KL 望远镜转移误差.** 本节的算法比较是

$$\mathrm{D}_{\mathsf{KL}}(p_1 \| \widehat{p}_1),$$

其中 p_1 是精确反向链产生的律，$\widehat{p}_1$ 是实际实现链产生的律，二者都停在第一个正噪声水平。我们先把两条反向链作为完整路径进行比较，再投影到它们的端点，由此控制最终 KL。

这种做法有用的原因很简单：马尔可夫链路径密度是初始密度与转移密度的乘积。取对数把乘积变成和，再取期望得到加性的 KL 恒等式。这就是误差分解背后的基本望远镜机制：一个长的反向采样过程被化约为许多一步反向核的比较。

引理 6.1 (反向马尔可夫核的 KL 链式法则)**.** *令 $\mu_K P_{K-1} \cdots P_1$ 和 $\nu_K Q_{K-1} \cdots Q_1$ 是 $(X_K, \dots, X_1)$ 上的两个路径律，其中 $P_k(x_{k+1} \to \cdot)$ 和 $Q_k(x_{k+1} \to \cdot)$ 从层 $k+1$ 转移到层 k。令 μ_{k+1} 为第一条路径律下 X_{k+1} 的边际律。则*

$$\mathrm{D}_{\mathsf{KL}}(\mu_K P_{K-1} \cdots P_1 \| \nu_K Q_{K-1} \cdots Q_1) = \mathrm{D}_{\mathsf{KL}}(\mu_K \| \nu_K) + \sum_{k=1}^{K-1} \mathbb{E}_{X_{k+1} \sim \mu_{k+1}} \mathrm{D}_{\mathsf{KL}}(P_k(X_{k+1} \to \cdot) \| Q_k(X_{k+1} \to \cdot)).$$

证明. 把两个路径密度写成

$$\mu_K(x_K) \prod_{k=1}^{K-1} P_k(x_{k+1} \to x_k), \qquad \nu_K(x_K) \prod_{k=1}^{K-1} Q_k(x_{k+1} \to x_k).$$

因此对数比为

$$\log \frac{\mu_K(x_K)}{\nu_K(x_K)} + \sum_{k=1}^{K-1} \log \frac{P_k(x_{k+1} \to x_k)}{Q_k(x_{k+1} \to x_k)}.$$

在第一条路径律下取期望，得到初始项 $\mathrm{D}_{\mathsf{KL}}(\mu_K\|\nu_K)$。对第 k 个转移项，以 X_{k+1} 为条件。在第一条路径律下，给定 X_{k+1} 时 X_k 的条件律为 $P_k(X_{k+1}\to\cdot)$，而 X_{k+1} 本身的律为 μ_{k+1}。因此

$$\mathbb{E}\log\frac{P_k(X_{k+1}\to X_k)}{Q_k(X_{k+1}\to X_k)} = \mathbb{E}_{X_{k+1}\sim\mu_{k+1}}\mathrm{D}_{\mathsf{KL}}(P_k(X_{k+1}\to\cdot)\|Q_k(X_{k+1}\to\cdot)).$$

对 k 求和即得恒等式。一般的测度论陈述可通过把密度替换为 Radon–Nikodym 导数得到。 □

评注. 由数据处理不等式，最终时刻的边际 KL 至多为路径空间 KL；见附录 A 的引理 A.1。因此，路径空间计算给出了最终采样误差的一个有效且常常方便的上界。这是 KL 所缺少的三角不等式的路径空间替代品：可加性来自对两个完整路径律的比较，而这两个路径律具有分解的初始分布和转移核。

命题 6.2 (从局部反向核误差到最终误差)**.** *令精确反向链从 p_K 出发，并使用来自 (5.3) 的精确反向核 $P_k^{\leftarrow}(x_{k+1}\to\cdot)$，$k=1,\dots,K-1$，产生最终律 p_1。令近似反向链从 $\widehat{p}_K$ 出发，并使用核 $\widehat{P}_k^{\leftarrow}(x_{k+1}\to\cdot)$，产生最终律 $\widehat{p}_1$。则*

$$\mathrm{D}_{\mathsf{KL}}(p_1\|\widehat{p}_1)\le\mathrm{D}_{\mathsf{KL}}(p_K\|\widehat{p}_K)+\sum_{k=1}^{K-1}\mathbb{E}_{X_{k+1}\sim p_{k+1}}\mathrm{D}_{\mathsf{KL}}\Big(P_k^{\leftarrow}(X_{k+1}\to\cdot)\Big\|\widehat{P}_k^{\leftarrow}(X_{k+1}\to\cdot)\Big).$$

证明. 考虑 $(X_K,\dots,X_1)$ 的精确路径律和 $(\widehat{X}_K,\dots,\widehat{X}_1)$ 的近似路径律。对这些路径律应用引理 6.1。初始贡献为 $\mathrm{D}_{\mathsf{KL}}(p_K\|\widehat{p}_K)$，反向步 k 处的转移贡献为

$$\mathbb{E}_{X_{k+1}\sim p_{k+1}}\mathrm{D}_{\mathsf{KL}}\Big(P_k^{\leftarrow}(X_{k+1}\to\cdot)\Big\|\widehat{P}_k^{\leftarrow}(X_{k+1}\to\cdot)\Big).$$

于是两条完整路径律之间的 KL 散度至多为上式右端。把一条路径投影到它的最终坐标是一个可测映射，所以 KL 的数据处理不等式给出关于 $\mathrm{D}_{\mathsf{KL}}(p_1\|\widehat{p}_1)$ 的所需界。 □

这个命题是误差分解 (6.2) 的路径空间部分。剩下的工作是估计求和中的一步项。

评注 (把度量合在一起). 上面的主要估计存在于各自自然的度量中：W_2 用于早停，KL 用于反向链误差。如果想得到一个单一的弱可观测量保证，可以引入有界 Lipschitz 距离

$$\mathrm{D}_{\mathsf{BL}}(\mu,\nu)=\sup_{\|f\|_\infty\le1,\,\|f\|_{\mathrm{Lip}}\le1}|\mathbb{E}_\mu f-\mathbb{E}_\nu f|.$$

那么由 W_2 定义中的耦合论证，有 $\mathrm{D}_{\mathsf{BL}}(p_{\mathsf{data}},p_1)\le W_2(p_{\mathsf{data}},p_1)$：对任意 1-Lipschitz 的 f 以及 μ,ν 的任意耦合 (X,Y)，

$$|\mathbb{E}f(X)-\mathbb{E}f(Y)|\le\mathbb{E}\|X-Y\|\le\sqrt{\mathbb{E}\|X-Y\|^2},$$

然后对耦合取下确界。在算法侧，

$$\mathrm{D}_{\mathsf{BL}}(p_1,\widehat{p}_1)\le2\,\mathrm{D}_{\mathsf{TV}}(p_1,\widehat{p}_1)\le2\sqrt{\mathrm{D}_{\mathsf{KL}}(p_1\|\widehat{p}_1)/2}.$$

因此，在原生的 W_2 和 KL 估计被证明之后，可以得到最终的 D_{BL} 陈述；但它不是分析中任一部分的驱动度量。

6.3. **得分误差如何贡献到 KL.** 命题 6.2 把全局 KL 误差化约为精确反向核 $P_k^{\leftarrow}$ 与实际实现核 $\widehat{P}_k^{\leftarrow}$ 之间的局部 KL 比较。每个局部比较可能包含若干误差来源。一个来源是数值性的：即便使用精确得分，实际转移也可能只是近似 Bayes 反向核。另一个来源是统计性的：采样器使用学习得分 s 代替真实得分 $\mathsf{s}^\star$。

在本节其余部分，我们把第 5.1 节的 DDPM 网格专门化到方差爆炸情形 $(\alpha_k = 1, \sigma_k^2 = t_k)$，现在从大噪声水平 T 反向运行到早停水平 $\delta > 0$：

$$\delta = t_1 < t_2 < \cdots < t_K = T, \qquad h_k = t_{k+1} - t_k.$$

指标为 k 的反向步从噪声水平 t_{k+1} 移动到 t_k，其中 $k = 1, \dots, K-1$，所以链结束于 $t_1 = \delta$；前面小节中写作 p_1 的最终律因此就是早停律 p_δ。一步方差参数于是化为时间增量，$\eta_k = h_k$；从现在起我们写作 h_k，以强调反向步是 SDE 离散化中的真实时间步。

本小节分离统计部分。实际反向步是高斯更新 (5.4)，等价地说，是反向 SDE 的一个 Euler–Maruyama 离散化；得分只通过核的均值进入其中。因此得分误差会变成均值误差，我们现在将其量化。

为了把记号分开，先固定高斯离散化。令

$$P_k^{\mathsf{EM}}(x' \to \cdot) = \mathcal{N}\big(x' + h_k \mathsf{s}_{k+1}^\star(x'), h_k I\big) \tag{6.4}$$

为精确得分 Euler 核，并令

$$\widehat{P}_k^{\mathsf{EM}}(x' \to \cdot) = \mathcal{N}(x' + h_k \mathsf{s}_{k+1}(x'), h_k I) \tag{6.5}$$

为学习得分 Euler 核。它们不是精确 Bayes 反向核 $P_k^{\leftarrow}$；而是同一个高斯近似 (5.4) 的真实得分版本和学习得分版本。因此，下面的 KL 在离散化已经固定之后分离出统计得分误差。对固定的 x'，应用引理 B.2 得

$$\mathrm{D}_{\mathsf{KL}}\Big(P_k^{\mathsf{EM}}(x' \to \cdot)\Big\|\widehat{P}_k^{\mathsf{EM}}(x' \to \cdot)\Big) = \frac{1}{2h_k}\left\|h_k\big(\mathsf{s}_{k+1}(x') - \mathsf{s}_{k+1}^\star(x')\big)\right\|^2.$$

对 $X_{k+1} \sim p_{k+1}$ 取平均，并使用 $\varepsilon_{k+1,\mathsf{score}}^2$ 的定义，得到

$$\mathbb{E}_{X_{k+1}\sim p_{k+1}} \mathrm{D}_{\mathsf{KL}}\Big(P_k^{\mathsf{EM}}(X_{k+1} \to \cdot)\Big\|\widehat{P}_k^{\mathsf{EM}}(X_{k+1} \to \cdot)\Big) = \frac{h_k}{2}\varepsilon_{k+1,\mathsf{score}}^2, \tag{6.6}$$

这正是在分离离散化选择之后，误差分解 (6.2) 中的得分估计贡献；此时 $\eta_k = h_k$ 且 $c = \frac{1}{2}$。

沿反向链把这些一步项相加，总得分贡献具有形式

$$\sum_{k=1}^{K-1} h_k \varepsilon_{k+1,\mathsf{score}}^2, \tag{6.7}$$

这是路径空间估计

$$\int_\delta^T \mathbb{E}_{X_t\sim p_t}\left\|\mathsf{s}_t(X_t) - \mathsf{s}_t^\star(X_t)\right\|^2 \,\mathrm{d}t.$$

的离散类似物。权重 h_k 使这个估计具有信息量：得分模型不需要在所有时间上同样准确，因为采样器只走很小一步时误差影响较小，而大反向步中的误差会产生更大的影响。

6.4. **Euler–Maruyama.** 现在讨论最简单的反向时间积分器，也就是 Euler–Maruyama 格式的数值误差，并保持本节开头固定的方差爆炸设定。此时误差分解 (6.2) 中的初始化项是显式的。如果实际反向链从 $\mathcal{N}(0, TI)$ 初始化，则

$$\varepsilon_{\mathsf{init}}^2 := \mathrm{D}_{\mathsf{KL}}(p_T \| \mathcal{N}(0, TI)) \le \int \mathrm{D}_{\mathsf{KL}}(\mathcal{N}(x, TI) \| \mathcal{N}(0, TI))\, p_{\mathsf{data}}(\,\mathrm{d}x) = \frac{\mathbb{E}\,\|X_0\|^2}{2T}. \tag{6.8}$$

这里我们使用表示 $p_T = \int \mathcal{N}(x, TI)\; p_{\mathsf{data}}(\,\mathrm{d}x)$，以及 KL 对其第一个自变量的凸性。

另一个更实质性的贡献是沿反向步累积的离散化误差，它的大小由步长 h_k 控制。因此，如何选择网格是核心设计问题，而衡量 h_k 的相关局部尺度是当前噪声水平。确实，反向漂移包含 $\mathsf{s}_t^\star = (\mathsf{D}_t^\star - x)/t$，所以同一个绝对步长在早停水平附近比在大噪声处激进得多。均匀网格因而会迫使所有步长都像最低噪声尺度 δ 一样小，导致对 T/δ 的依赖很差。标准的非均匀网格补救是使用相对控制，如 Chen、Lee 和 Lu [12]：

$$h_k \le \bar{h}\, t_{k+1}, \qquad \bar{h} \le \frac{1}{2}. \tag{6.9}$$

典型例子是几何网格

$$t_k = (1 - \bar{h})^{K-k} T, \qquad \delta = t_1 = (1 - \bar{h})^{K-1} T. \tag{6.10}$$

对这个网格，

$$K \asymp \bar{h}^{-1} \log \frac{T}{\delta}, \qquad \bar{h} \asymp \frac{\log(T/\delta)}{K}.$$

在这个约定下，精确 Bayes 反向核为

$$P_k^{\leftarrow}(x' \to \,\mathrm{d}x) \propto_x p_k(x) \exp\left(-\frac{\|x - x'\|^2}{2h_k} \right) \mathrm{d}x, \tag{6.11}$$

其中 x' 是噪声水平 t_{k+1} 处的状态。一阶高斯近似把得分冻结在反向步开始处。这就是 (6.4) 中的精确得分 Euler 核 P_k^{EM}；学习采样器使用 (6.5) 中的 $\widehat{P}_k^{\mathsf{EM}}$，即用 s_{k+1} 代替 $\mathsf{s}_{k+1}^\star$。上标 EM 只是提醒我们得分在步内被冻结。在 (5.4) 的一般 DDPM 记号中，这些公式由令 $\alpha_k = 1$ 且 $\eta_k = h_k$ 得到。

局部误差有两个来源。得分估计部分已经在上一小节分离出来：在当前方差爆炸归一化下，(6.6) 在常数意义下贡献 $h_k \varepsilon_{k+1,\mathsf{score}}^2$。本小节的新问题是 Euler–Maruyama 本身造成的数值误差。因此我们先比较精确反向核 $P_k^{\leftarrow}$ 与精确得分 Euler 核 P_k^{EM}。这是在一个时间步内冻结真实反向漂移所产生的误差。

命题 6.3 (Euler–Maruyama 离散化误差)**.** *假设在反向步 k 中，精确得分场在过程典型所在区域内以尺度 L_k 为 Lipschitz。则精确得分 Euler 核满足如下形式的一步估计：*

$$\mathbb{E}_{X_{k+1} \sim p_{k+1}} \mathrm{D}_{\mathsf{KL}}\big(P_k^{\leftarrow}(X_{k+1} \to \cdot) \big\| P_k^{\mathsf{EM}}(X_{k+1} \to \cdot)\big) \lesssim d\, L_k^2\, h_k^2,$$

其中略去了较低阶的调度因子。

证明. 我们在反向时间中写出单个反向步。给定噪声水平 t_{k+1} 处的起点 $X_{k+1} = x'$，令 $0 \le r \le h_k$。这个区间上的精确反向扩散具有漂移 $\mathsf{s}^\star_{t_{k+1}-r}$：

$$\mathrm{d}Y_r^{\leftarrow} = \mathsf{s}^\star_{t_{k+1}-r}(Y_r^{\leftarrow})\,\mathrm{d}r + \mathrm{d}B_r^{\leftarrow}, \qquad Y_0^{\leftarrow} = x'.$$

它的端点律为 $P_k^{\leftarrow}(x' \to \cdot)$。精确得分 Euler–Maruyama 插值把漂移冻结在步开始处：

$$\mathrm{d}\bar{Y}_r^{\leftarrow} = \mathsf{s}^\star_{k+1}(x')\,\mathrm{d}r + \mathrm{d}B_r^{\leftarrow}, \qquad \bar{Y}_0^{\leftarrow} = x',$$

它的端点律正是 $P_k^{\mathsf{EM}}(x' \to \cdot)$。由附录 A 中的 Girsanov KL 公式，再结合数据处理不等式（引理 A.1），端点比较可由路径比较控制：

$$\mathrm{D}_{\mathsf{KL}}\big(P_k^{\leftarrow}(x' \to \cdot)\big\|P_k^{\mathsf{EM}}(x' \to \cdot)\big) \le \frac{1}{2}\mathbb{E}\int_0^{h_k} \left\|\mathsf{s}^\star_{t_{k+1}-r}(Y_r^{\leftarrow}) - \mathsf{s}^\star_{k+1}(x')\right\|^2 \mathrm{d}r.$$

这个式子就是基本的 Euler 离散化估计：全部代价就是用左端点处冻结的得分替代理想反向路径上移动得分的代价。

在局部 Lipschitz 界下，得分差由反向路径在该步中的运动控制，另有来自噪声水平确定性变化的较低阶项。由于时间 r 内的 Brownian 位移平方量级为 dr，局部估计为

$$\mathbb{E}\left\|\mathsf{s}^\star_{t_{k+1}-r}(Y_r^{\leftarrow}) - \mathsf{s}^\star_{k+1}(x')\right\|^2 \lesssim L_k^2\, d\, r,$$

这里同样压低了依赖调度的较低阶项。对 $r \in [0, h_k]$ 积分得到

$$\frac{1}{2}L_k^2 d\int_0^{h_k} r\ \mathrm{d}r \lesssim d\, L_k^2 h_k^2.$$

最后再对 $X_{k+1} \sim p_{k+1}$ 平均。 □

命题 6.3 只是精确得分离散化估计。当采样器使用 s_{k+1} 而非 $\mathsf{s}^\star_{k+1}$ 时，额外贡献就是 (6.6) 中计算的统计高斯均值误差项；在方差爆炸归一化下它变为 $h_k\varepsilon^2_{k+1,\mathsf{score}}$。

结合命题 6.3、得分误差贡献 (6.6) 和命题 6.2，得到标准的 Euler–Maruyama 扩散采样界：

$$\mathrm{D}_{\mathsf{KL}}(p_\delta\|\widehat{p}_\delta) \lesssim \varepsilon^2_{\mathsf{init}} + \sum_{k=1}^{K-1} d\, L_k^2 h_k^2 + \sum_{k=1}^{K-1} h_k\varepsilon^2_{k+1,\mathsf{score}}. \tag{6.12}$$

三个项分别是高噪声端点初始化、Euler 离散化和得分估计。这是命题 6.2 的 KL 望远镜估计加上局部高斯计算，因此也是误差分解 (6.2) 的 Euler–Maruyama 特化。

评注 (与 ULA 的比较). (6.12) 中的离散化项与 (2.3) 中的 ULA 误差有相同来源。在两种情况下，数值格式都在一个短时间区间内冻结漂移。该区间内的 Brownian 位移产生量级为 dL^2h^2 的局部 KL 代价；在非均匀扩散网格上则为 $dL_k^2h_k^2$。对网格求和给出与网格尺度成比例的累积离散化偏差：ULA 中为 dL^2h，这里为 $\sum_k dL_k^2h_k^2$。差别在于外围论证。对 ULA 来说，(2.3) 是一个以固定密度为目标的马尔可夫链混合定理，所以它需要诸如对数 Sobolev 不等式这样的全局收缩机制，同时需要光滑性来控制 Euler 误差。对扩散模型来说，参考路径 p_t 已经由前向加噪过程提供，反向采样器通过 KL 望远镜分解和 Girsanov 与这条路径比较。因此，不需要数据分布上的全局结构假设，如凸性、对数 Sobolev 或等周不等式、暖启动等；它们被初始化项和得分估计项所替代。

事实上，估计 (6.12) 中的离散化误差可以进一步锐化，如下一小节所示。一旦初始化和得分估计被分离出来，精确得分 Euler 采样器只为在每个反向步中冻结真实得分付费。令 t^+ 表示不低于当前噪声水平的下一个网格点，所以对 $t \in [t_k, t_{k+1}]$ 有 $t^+ = t_{k+1}$。离散化代价由路径量

$$\int_\delta^T \mathbb{E} \left\| \mathsf{s}_t^\star(X_t) - \mathsf{s}_{t^+}^\star(X_{t^+}) \right\|^2 \, \mathrm{d}t, \tag{6.13}$$

来度量，在更一般调度中还需乘以确定性调度因子。这是有限网格项 $\sum_k dL_k^2 h_k^2$ 的路径版本：Euler–Maruyama 在每个反向步的高噪声端点 t^+ 处冻结得分，而离散化误差由理想反向路径中真实得分在噪声时间从 t^+ 降到 t 时的变化控制。

在上面的误差分析中，从这个得分差到界 $dL_k^2 h_k^2$ 的过渡使用了最坏情形控制：得分由 Lipschitz 常数 L_k 控制，或等价地由 $\log p_t$ 的点态 Hessian 界控制。因子 d 来自一步中的 Brownian 位移；潜在损失来自把它乘以 Hessian 的一致算子范数界。在弱假设下，这种最坏情形曲率可能比反向过程看到的平均曲率大得多，尤其是在早停时间附近。下一小节用平均的后验协方差量替代这种点态控制，从而导出近乎线性依赖于 d 的离散化界。

6.5. **Hessian 控制.** 上一小节留下的困难是 (6.13) 中的路径得分差。本小节的目标是把最坏情形 Hessian 估计替换为更平均的界

$$\text{得分冻结代价} \ \lesssim \ \bar{h} \times \text{后验协方差预算}.$$

我们先推出局部得分冻结估计，其中会出现因子 $\bar{h}$；随后把剩下的加权二次变差改写成后验协方差。

对 $t \in [t_k, t_{k+1}]$，写 $t^+ = t_{k+1}$。(6.13) 中的 Girsanov 得分冻结代价，忽略常数 $\frac{1}{2}$ 和调度因子后，是所有网格区间上的和

$$\int_{t_k}^{t_{k+1}} \mathbb{E} \left\| \mathsf{s}_t^\star(X_t) - \mathsf{s}_{t^+}^\star(X_{t^+}) \right\|^2 \, \mathrm{d}t.$$

有用的变量不是得分本身，而是最优去噪器。在方差爆炸归一化下，

$$\mathsf{D}_t^\star(x) := \mathbb{E}[X_0 \mid X_t = x] = x + t\mathsf{s}_t^\star(x), \qquad \mathsf{s}_t^\star(x) = \frac{\mathsf{D}_t^\star(x) - x}{t}.$$

随机定位鞅恒等式 (4.7) 改写到噪声时间坐标后说明，对 $t \le t^+$，

$$\mathbb{E} \left\| \mathsf{D}_t^\star(X_t) - \mathsf{D}_{t^+}^\star(X_{t^+}) \right\|^2 = \int_t^{t^+} \mathbb{E} \left\| \nabla \mathsf{D}_r^\star(X_r) \right\|_{\mathrm{F}}^2 \, \mathrm{d}r.$$

这取代了点态 Hessian 界：它度量沿加噪路径的去噪器真实二次变差。

现在估计一个网格区间。使用 $\mathsf{s}_t^\star = (\mathsf{D}_t^\star - x)/t$，并只显式保留曲率项，得分增量由去噪器增量和自然的 t^{-2} 权重控制：

$$\mathbb{E} \left\| \mathsf{s}_t^\star(X_t) - \mathsf{s}_{t^+}^\star(X_{t^+}) \right\|^2 \lesssim \frac{1}{t^2} \mathbb{E} \left\| \mathsf{D}_t^\star(X_t) - \mathsf{D}_{t^+}^\star(X_{t^+}) \right\|^2.$$

这里省略了来自显式因子 x/t 的一些误差项，这些项可以容易处理。代入鞅恒等式并应用 Fubini，得到

$$\begin{aligned}
&\int_{t_k}^{t_{k+1}} \mathbb{E}\left\|\mathsf{s}_t^\star(X_t) - \mathsf{s}_{t^+}^\star(X_{t^+})\right\|^2 \,\mathrm{d}t \\
&\qquad \lesssim \int_{t_k}^{t_{k+1}} \frac{1}{t^2} \int_t^{t_{k+1}} \mathbb{E}\left\|\nabla \mathsf{D}_r^\star(X_r)\right\|_{\mathrm{F}}^2 \,\mathrm{d}r \,\mathrm{d}t \\
&\qquad = \int_{t_k}^{t_{k+1}} \mathbb{E}\left\|\nabla \mathsf{D}_r^\star(X_r)\right\|_{\mathrm{F}}^2 \left(\int_{t_k}^{r} \frac{\mathrm{d}t}{t^2}\right) \mathrm{d}r \\
&\qquad = \int_{t_k}^{t_{k+1}} \mathbb{E}\left\|\nabla \mathsf{D}_r^\star(X_r)\right\|_{\mathrm{F}}^2 \frac{r - t_k}{r\, t_k} \,\mathrm{d}r.
\end{aligned}$$

由于 $r - t_k \le h_k$，且网格条件 (6.9) 与 $\bar{h} \le 1/2$ 给出 $t_k \ge t_{k+1}/2$，我们得到

$$\int_{t_k}^{t_{k+1}} \mathbb{E}\left\|\mathsf{s}_t^\star(X_t) - \mathsf{s}_{t^+}^\star(X_{t^+})\right\|^2 \,\mathrm{d}t \lesssim \frac{h_k}{t_{k+1}} \int_{t_k}^{t_{k+1}} \frac{\mathbb{E}\left\|\nabla \mathsf{D}_r^\star(X_r)\right\|_{\mathrm{F}}^2}{r} \,\mathrm{d}r.$$

这正是 $\bar{h}$ 进入的地方：比值 h_k/t_{k+1} 是该区间上的相对步长，而 (6.9) 说明它至多为 $\bar{h}$。因此对 k 求和给出

$$\sum_k \int_{t_k}^{t_{k+1}} \mathbb{E}\left\|\mathsf{s}_t^\star(X_t) - \mathsf{s}_{t^+}^\star(X_{t^+})\right\|^2 \,\mathrm{d}t \lesssim \bar{h} \int_\delta^T \frac{\mathbb{E}\left\|\nabla \mathsf{D}_t^\star(X_t)\right\|_{\mathrm{F}}^2}{t} \,\mathrm{d}t.$$

还需把加权二次变差表达为更内在的形式。令

$$\Sigma_t(x) = \mathrm{Cov}(X_0 \mid X_t = x)$$

为方差爆炸加噪通道中的后验协方差。线性响应恒等式 (4.5) 和得分 Jacobian 恒等式 (4.6) 翻译到这个记号中，给出

$$\nabla \mathsf{D}_t^\star(x) = I + t \nabla \mathsf{s}_t^\star(x) = t^{-1} \Sigma_t(x).$$

协方差耗散恒等式 (4.8)，$\frac{\mathrm{d}}{\mathrm{d}u}\mathbb{E}[\mathrm{Tr}\,\Sigma_u] = -\mathbb{E}[\mathrm{Tr}(\Sigma_u^2)]$，在变量替换 $t = 1/u$ 后给出

$$\mathbb{E}\left\|\nabla \mathsf{D}_t^\star(X_t)\right\|_{\mathrm{F}}^2 = \partial_t\, \mathbb{E}\,\mathrm{Tr}\,\Sigma_t(X_t).$$

因此

$$\int_\delta^T \frac{\mathbb{E}\left\|\nabla \mathsf{D}_t^\star(X_t)\right\|_{\mathrm{F}}^2}{t} \,\mathrm{d}t = \int_\delta^T \frac{\partial_t\, \mathbb{E}\,\mathrm{Tr}\,\Sigma_t(X_t)}{t} \,\mathrm{d}t.$$

分部积分得

$$\int_\delta^T \frac{\partial_t\, \mathbb{E}\,\mathrm{Tr}\,\Sigma_t(X_t)}{t} \,\mathrm{d}t = \frac{\mathbb{E}\,\mathrm{Tr}\,\Sigma_T(X_T)}{T} - \frac{\mathbb{E}\,\mathrm{Tr}\,\Sigma_\delta(X_\delta)}{\delta} + \int_\delta^T \frac{\mathbb{E}\,\mathrm{Tr}\,\Sigma_t(X_t)}{t^2} \,\mathrm{d}t.$$

中间项非正，因此第一项和第三项给出上界。沿用 Chewi [18, Chapter 12] 的记号，定义协方差预算

$$\mathfrak{D}_{\delta,T}(p_{\mathsf{data}}) := \frac{\mathbb{E}_{X_T \sim p_T}\,\mathrm{Tr}\,\Sigma_T(X_T)}{T} + \int_\delta^T \frac{\mathbb{E}_{X_t \sim p_t}\,\mathrm{Tr}\,\Sigma_t(X_t)}{t^2} \,\mathrm{d}t. \tag{6.14}$$

结合前面的式子，得到简洁总结

$$\sum_k \int_{t_k}^{t_{k+1}} \mathbb{E}\left\|\mathsf{s}_t^\star(X_t) - \mathsf{s}_{t^+}^\star(X_{t^+})\right\|^2 \, \mathrm{d}t \lesssim \bar{h}\, \mathfrak{D}_{\delta,T}(p_{\mathsf{data}}).$$

这就是关键识别：离散化误差由最优去噪器的加权二次变差控制，而 $\mathfrak{D}_{\delta,T}$ 在计入 DDPM 时间权重之后封装了这个变差。该估计的严格版本，加上得分误差项，给出

$$\mathrm{D}_{\mathsf{KL}}(p_\delta \| \widehat{p}_\delta) \le \varepsilon_{\mathsf{init}}^2 + 4\sum_{k=1}^{K-1} h_k\, \mathbb{E}_{p_{k+1}} \left\|\mathsf{s}_{k+1} - \mathsf{s}_{k+1}^\star\right\|^2 + 2\bar{h}\, \mathfrak{D}_{\delta,T}(p_{\mathsf{data}}), \tag{6.15}$$

其中中间项是时间加权的得分误差。重要教训是，$\mathfrak{D}_{\delta,T}(p_{\mathsf{data}})$ 取代得分的全局 Lipschitz 常数，成为离散化复杂度中的数据依赖因子。

剩下的问题是这个协方差预算可能有多大。一个有用的环境维数估计直接来自它的后验方差形式：对任意具有有限二阶矩的分布，

$$\mathbb{E}\,\mathrm{Tr}\,\Sigma_t(X_t) = \mathbb{E}\left\|X_0 - \mathbb{E}[X_0 \mid X_t]\right\|^2 \le \mathbb{E}\left\|X_0 - X_t\right\|^2 = dt,$$

因为 X_t 本身是 X_0 的一个可行估计量。把它代入 (6.14) 得到

$$\mathfrak{D}_{\delta,T}(p_{\mathsf{data}}) \le d\left(1 + \log\frac{T}{\delta}\right).$$

将这个估计与 (6.15) 和几何网格关系 (6.10) 结合，精确得分离散化项由

$$O\left(\frac{d\log^2(T/\delta)}{K}\right)$$

控制。因此，当初始化误差和得分估计误差也都是 $O(\varepsilon^2)$ 时，取

$$K = \widetilde{O}\left(\frac{d\log^2(T/\delta)}{\varepsilon^2}\right)$$

个反向步就足够。这是 Benton、De Bortoli、Doucet 和 Deligiannidis [9] 在有限二阶矩假设下证明的近乎 d 线性依赖。在有限 Fisher 信息假设下，Conforti、Durmus 和 Gentiloni Silveri [19] 得到了相关的 KL 保证。当数据本质上是低维时，协方差预算可以更锐利地有界：平滑律 p_δ 的熵或覆盖数估计可用尺度 $\sqrt{\delta}$ 上的内在维数替换环境维数。

6.6. **一阶拒绝采样与高精度.** Hessian 控制界解释了为什么 Euler–Maruyama 可以在弱假设下被分析，但它也暴露出采样器本身的一个限制。Euler–Maruyama 并不试图采样精确的一步 Bayes 核；它线性化对数密度，并依赖步长足够小，使非线性余项可以忽略。因此，就像 Metropolis 修正之前的 ULA 一样，高精度是用细网格换来的，从而带来对目标精度的多项式依赖。自然的 MALA 式问题是：这种依赖能否降低为对 $1/\varepsilon$ 的对数依赖，或至少是多重对数依赖？

这里，与 MALA 的类比也揭示了主要障碍。MALA 通过对数密度比得到修正。在扩散模型中，一步目标涉及加噪密度 p_k，但 $p_k(x)$ 和 $\log p_k(x)$ 不是可用的数值量；学习到的对象只有得分 $\mathsf{s}_k^\star = \nabla \log p_k$。因此问题比简单应用一个 Metropolis 步更尖锐：只用得分的修正能否模拟缺失的对数密度比检验，并把局部高斯提议校正到精确 Bayes 核？

对方差爆炸链而言，这个问题具有有用的局部结构。精确反向核 (6.11) 不只是一个均值平移的高斯；它是如下形式的高斯倾斜

$$p_k(x)\exp\left(-\frac{\|x-x'\|^2}{2h_k}\right), \tag{6.16}$$

其中 x' 是噪声水平 t_{k+1} 处的状态。Euler–Maruyama 对应于用 $\log p_k$ 的一阶展开所得的高斯来替代这个倾斜。高精度思想是在仍然只使用一阶信息的情况下，更忠实地采样倾斜律 (6.16)。

算法创新就在这里出现。一阶拒绝采样，记为 FORS [11]，用随机化的一阶估计替代直接的对数密度比计算。虽然 $\log p_k(x)$ 和 $\log p_k(y)$ 分别不可用，但二者之差可以写成只涉及得分的路径积分：对任意光滑路径 γ，其中 $\gamma_0 = y$ 且 $\gamma_1 = x$，

$$\log p_k(x) - \log p_k(y) = \int_0^1 \langle \dot\gamma_r, \mathsf{s}_k^\star(\gamma_r)\rangle \, \mathrm{d}r.$$

沿路径随机抽取一点，就把这个积分变成对数密度差的无偏估计。因此，虽然端点对数密度本身不可用，得分查询仍然可以产生拒绝修正所需的精确对数倾斜的随机估计。

还需要再做一次转换。拒绝采样需要与对数倾斜的指数成比例的接受权重，而不只是对数倾斜本身的一个估计。对一个无偏估计取指数通常不会给出正确平均值，因此 FORS 使用 Poisson 乘积恒等式从一阶随机估计构造指数倾斜。抽象地说，假设提议从律 Q 中抽取，目标律是 Q 的指数倾斜，其 Radon–Nikodym 导数正比于 $e^{w(x)}$。假设给定提议 x，可以采样有界随机变量 $W_1, W_2, \dots$，满足 $\mathbb{E}[W_1 \mid x] = w(x)$。在最简单的有界形式中，假设存在确定性包络 R 使得几乎必然有 $|W_j| \le R$。取 $J \sim \mathsf{Poisson}(2R)$，并以概率

$$\prod_{j=1}^{J} \frac{R + W_j}{2R}.$$

接受提议 x。Poisson 随机变量的概率母函数使平均接受因子正比于 $e^{w(x)}$。因此，被接受的提议服从指数倾斜后的律。如果 $R = \Theta(1)$，则一阶查询次数的期望为常数，并且高概率下为对数级。

Poisson 乘积计算. 以提议点 x 为条件，写

$$A(x) = \mathbb{E}\left[\prod_{j=1}^{J} \frac{R + W_j}{2R} \middle| x\right].$$

给定 J 时，各因子相互独立，并且条件均值为 $(R + w(x))/(2R)$。因此

$$A(x) = \mathbb{E}\left[\left(\frac{R + w(x)}{2R}\right)^{J} \middle| x\right].$$

$J \sim \mathsf{Poisson}(2R)$ 的母函数给出

$$A(x) = \exp\left(2R\left(\frac{R + w(x)}{2R} - 1\right)\right) = e^{w(x) - R}.$$

额外因子 e^{-R} 与 x 无关。因此，被接受提议的联合律正比于 $e^{w(x)}Q(\mathrm{d}x)$，这正是指数倾斜。得分查询次数为 J，所以均值为 $2R$；当 R 有界时，标准 Poisson 尾界给出高概率下的对数控制。□

因此，FORS 是 Metropolis 修正的只用得分的类似物。MALA 使用密度比来修正高斯提议；FORS 使用随机化一阶估计，在不评估密度的情况下修正高斯倾斜。算法不把高斯线性化直接接受为采样器，而是把这个局部提议校正到真实高斯倾斜。

在误差分解 (6.2) 的记号中，FORS 改变了精确得分算法项 $\varepsilon_{\mathsf{disc}}^2$ 的大小。通过足够准确地采样每个局部高斯倾斜，累积的局部算法误差可以用仅多重对数依赖于目标精度 ε 的反向步数做到 $O(\varepsilon^2)$。在一个简化的环境维数形式中，Chen、Chewi、Daskalakis 和 Rakhlin [11] 的结果说明，略去调度常数，并令初始化预算和局部算法预算均为 ε^2 量级，存在一个使用得分评估的反向采样器，使得

$$\mathrm{D}_{\mathsf{KL}}(p_\delta \| \widehat{p}_\delta) \lesssim \varepsilon^2 + \sum_{k=1}^{K-1} h_k \varepsilon_{k+1,\mathsf{score}}^2 \tag{6.17}$$

且

$$K = \widetilde{O}(d \ \mathrm{polylog}(1/\varepsilon))$$

在基本环境维数解读下成立。完整定理包含更精细的维数度量，以及在额外结构下的改进；但这里的核心信息更简单：昂贵的精度参数依赖被替换为多重对数依赖。结合方差爆炸早停估计 $W_2^2(p_{\mathsf{data}}, p_\delta) \le \delta d$，这分离了两种误差度量：初始平滑由 W_2 控制，而实际反向链由 KL 控制。注意，(6.17) 中仍然出现同一个时间积分得分误差项：FORS 改进了局部采样步，但它并没有消除对准确学习得分的需求。

7. 离散扩散模型

到目前为止，我们一直在连续状态空间上构造和分析扩散模型。现在转向另一个建模问题：如果状态空间本身不是连续的，会发生什么？反向时间的 Bayes 规则、去噪目标和路径空间误差分解在有限状态空间上仍然存在；但 Euclidean 梯度和 Brownian 运动不再存在。因此，离散扩散模型迫使我们用一种不依赖微积分的形式来表述采样思想。

7.1. 有限状态空间上的前向核与反向核.

像素空间中的图像常常可以当作连续对象处理，但语言、符号序列、分子、图以及许多科学构型本质上是离散的。如果 x 是一个 token 序列，那么像 $x + \sqrt{t}Z$ 这样的表达式就没有意义。在词表内部，“cat”这个词不存在小的高斯扰动。离散扩散模型不添加 Brownian 噪声，而是在有限状态空间上施加一列马尔可夫核。

概率骨架与我们一直使用的完全相同：一条把数据退化到参考律的前向链，以及一条由 Bayes 规则恢复的反向链。因此我们保留连续情形中的同一组符号。第 5.1 小节中的前向转移 $P_k^{\rightarrow}$ 和反向转移 $P_k^{\leftarrow}$，现在应读作有限集合上的随机矩阵，而不是 $\mathbb{R}^d$ 上的转移密度。唯一真正失去的是微积分：没有梯度 $\nabla \log p_k$，所以得分必须由某种有限差分对象替代。

令 $\mathcal{X}$ 为有限状态空间。数据律 p_{data} 可以支撑在某个子集 $\mathcal{X}_0 \subseteq \mathcal{X}$ 上；我们把它视为 $\mathcal{X}$ 上的律，即在 $\mathcal{X}_0$ 外赋零质量。对普通类别序列，可以取 $\mathcal{X}_0 = \mathcal{X} = \mathcal{V}^L$；而对下面讨论的掩码扩

散，噪声状态空间 $\mathcal{X}$ 还包含掩码符号。离散前向扩散是一条马尔可夫链

$$X_0 \sim p_{\mathsf{data}}, \qquad X_{k+1} \sim P_k^{\rightarrow}(X_k \to \cdot), \qquad k = 0, \dots, K-1, \tag{7.1}$$

其中每个 $P_k^{\rightarrow}$ 都是 $\mathcal{X}$ 上的随机矩阵：

$$P_k^{\rightarrow}(x \to y) \geq 0, \qquad \sum_{y \in \mathcal{X}} P_k^{\rightarrow}(x \to y) = 1.$$

令 p_k 为 X_k 的律。按行向量记号，

$$p_{k+1} = p_k P_k^{\rightarrow}, \qquad p_k = p_{\mathsf{data}} P_0^{\rightarrow} P_1^{\rightarrow} \cdots P_{k-1}^{\rightarrow}.$$

前向核被选择成使 p_K 接近一个简单参考分布 π_{ref}，例如全掩码 token、独立均匀 token 或一个乘积分布。

设计原则与连续扩散相同。前向过程应当足够简单，使我们能够采样并知道它的转移概率；同时它应当足够破坏性，使得许多步之后分布接近某个参考律。反向过程随后成为需要学习的部分。改变的是破坏的性质：token 被替换、掩码或重新采样，而不是由小的高斯增量移动。

7.2. **精确反向核与比值得分.** 精确反向核同样由 Bayes 规则给出。对固定的 $x' \in \mathcal{X}$，

$$P_k^{\leftarrow}(x' \to x) = \mathbb{P}(X_k = x \mid X_{k+1} = x') = \frac{p_k(x) P_k^{\rightarrow}(x \to x')}{p_{k+1}(x')}. \tag{7.2}$$

这是 (5.3) 的离散类似物，使用同一个 $P_k^{\leftarrow}$ 记号：在连续情形中，反向核等于 p_k 乘以高斯似然再归一化；这里它等于 p_k 乘以前向转移似然 $P_k^{\rightarrow}(x \to x')$ 再归一化。

如果精确知道 p_k，(7.2) 就会给出精确反向采样器。在学习中，我们近似反向核 $P_k^{\leftarrow}(x' \to \cdot)$ 本身，或者近似一个可以用来计算它的对象，例如给定 X_k 时 X_0 的后验分布。

反向核恒等式 (7.2) 准确显示了必须学习什么。前向转移 $P_k^{\rightarrow}$ 由设计给定并已知，但中间律 p_k 包含关于数据分布的信息，因而未知。学习一个离散扩散模型，就是学习足够多关于这些中间律的信息，以模拟反向 Bayes 核。表示这种缺失信息的一种方式，是使用未知律的比值。事实上，对同一当前状态 x' 的两个候选前驱 x 和 y，(7.2) 中的归一化因子 $p_{k+1}(x')$ 会抵消：

$$\frac{P_k^{\leftarrow}(x' \to y)}{P_k^{\leftarrow}(x' \to x)} = \frac{p_k(y) P_k^{\rightarrow}(y \to x')}{p_k(x) P_k^{\rightarrow}(x \to x')}.$$

由于 $P_k^{\rightarrow}$ 已知，未知部分就是同一时刻的密度比

$$\mathsf{s}_k^{\star}(x, y) = \frac{p_k(y)}{p_k(x)}, \qquad x, y \in \mathcal{X}. \tag{7.3}$$

在连续空间中，得分告诉我们从 x 移动到 $x + \mathrm{d}x$ 时的无穷小对数密度变化。在图或有限集合上，比值 $p_k(y)/p_k(x)$ 告诉我们从 x 移动到相邻状态 y 时的有限对数密度变化。如果前向核只允许局部变化，这些比值就足以指定局部反向转移概率。

然而，对许多现代离散扩散模型来说，去噪后验是更方便的学习目标：

$$\mathsf{D}_k^{\star}(x_0 \mid x) = \mathbb{P}(X_0 = x_0 \mid X_k = x), \qquad x_0 \in \mathcal{X}_0,\ x \in \mathcal{X}.$$

这是连续 Tweedie 恒等式中后验均值 $\mathbb{E}[X_0 \mid X_k = x]$ 的离散对应物。

比值得分视角和去噪器视角通过 Bayes 规则紧密相关：

$$\mathsf{D}_k^\star(x_0 \mid x) = \frac{p_{\mathsf{data}}(x_0)\mathbb{P}(X_k = x \mid X_0 = x_0)}{p_k(x)}.$$

因此，比较两个噪声状态处的后验概率，可以在已知前向似然比的意义下恢复密度比 $p_k(y)/p_k(x)$。去噪器因此包含比值得分信息，但其形式通常更容易通过监督式干净 token 预测来学习。

7.3. **掩码扩散.** 最干净的例子是 token 序列的掩码扩散。令词表为 $\mathcal{V}$，并加入特殊掩码符号 $[\mathsf{M}]$。这里 $\mathcal{X}_0 = \mathcal{V}^L$，而噪声状态空间是扩充后的序列空间

$$\mathcal{X}_{\mathsf{M}} := (\mathcal{V} \cup \{[\mathsf{M}]\})^L.$$

对 $x, y \in \mathcal{X}_{\mathsf{M}}$，用 x_i 和 y_i 表示它们的第 i 个坐标。定义前向步为

$$P_k^{\rightarrow}(x \to y) = \prod_{i=1}^{L} \begin{cases} 1-\beta_k, & y_i = x_i,\ x_i \in \mathcal{V}, \\ \beta_k, & y_i = [\mathsf{M}],\ x_i \in \mathcal{V}, \\ 1, & x_i = y_i = [\mathsf{M}], \\ 0, & \text{otherwise.} \end{cases} \tag{7.4}$$

因此，未被掩码的 token 要么保持不变，要么变成掩码；一旦 token 被掩码，它就保持掩码。可以使用坐标相关的掩码率作为变体，但同一个序列层面的形式仍然适用。

令

$$\bar{\alpha}_k = \prod_{s=0}^{k-1}(1-\beta_s).$$

对每个坐标 i，在给定 $X_{0,i} = a \in \mathcal{V}$ 时，

$$\mathbb{P}(X_{k,i} = a \mid X_{0,i} = a) = \bar{\alpha}_k, \qquad \mathbb{P}(X_{k,i} = [\mathsf{M}] \mid X_{0,i} = a) = 1 - \bar{\alpha}_k.$$

这类似于高斯公式 $X_k \mid X_0 \sim \mathcal{N}(a_k X_0, \sigma_k^2 I)$：标量衰减衡量关于 X_0 的信息还剩多少。

掩码扩散特别透明，因为前向方向上的破坏是不可逆的。一旦 token 被掩码，前向过程就忘记了它的身份。因此，反向模型必须使用周围上下文和学习到的数据分布来推断可能的干净 token。这使得去噪的 Bayes 本质在不使用微积分的情况下变得清楚。在坐标层面，反向步只有两种情形。

- 如果 $X_{k+1,i} = a \in \mathcal{V}$ 未被掩码，则必有 $X_{k,i} = a$。该坐标上没有需要采样的内容。
- 如果 $X_{k+1,i} = [\mathsf{M}]$，则 $X_{k,i}$ 可能已经是掩码，也可能是原始干净 token。反向采样器必须决定是否去掩码；若去掩码，还要决定放置哪个 token。

把反向核 (7.2) 专门化到 (7.4)，得到序列层面的公式

$$P_k^{\leftarrow}(x' \to x) = \frac{p_k(x)}{p_{k+1}(x')} P_k^{\rightarrow}(x \to x'). \tag{7.5}$$

通过已知因子 $P_k^{\rightarrow}(x \to x')$，这仍然是一条坐标式掩码规则；但未知的数据依赖权重是完整序列律 $p_k(x)$。因此，未掩码 token 的分布通常依赖于整个部分掩码序列。

7.4. **掩码扩散的训练目标.** 模型通常被训练成从破坏后的序列中预测被掩码的干净 token。令 $M_k \subseteq \{1,\dots,L\}$ 为时间 k 的被掩码坐标集合，令 x_k 为掩码序列。去噪模型 $\mathsf{D}_k(i,\cdot \mid x_k)$ 对坐标 i 输出一个词表 token 上的分布。总体损失为

$$\mathsf{L}_{\mathsf{mask}}(\mathsf{D}) = \mathbb{E}\left[\sum_{i\in M_k} -\log \mathsf{D}_k(i, X_{0,i} \mid X_k)\right]. \tag{7.6}$$

这是以干净 token 为标签的交叉熵。总体极小化子是后验的坐标边际：

$$\mathsf{D}_k^\star(i,a \mid x_k) = \mathbb{P}(X_{0,i} = a \mid X_k = x_k).$$

因此，交叉熵目标不是任意的语言建模损失。它是去噪得分匹配的离散类似物。在高斯情形中，最优预测是一个后验均值，并可由 Tweedie 恒等式转换为得分。在掩码 token 情形中，最优预测是隐藏干净 token 的完整后验分布。

证明. 对固定的 (i,x_k,k)，损失中的条件贡献为

$$\sum_{a\in\mathcal{V}} \mathbb{P}(X_{0,i} = a \mid X_k = x_k)\,[-\log \mathsf{D}_k(i,a \mid x_k)].$$

这是真实后验分布和模型分布 $\mathsf{D}_k(i,\cdot \mid x_k)$ 之间的交叉熵。它唯一地在模型分布等于真实后验时取最小。 □

7.5. **从掩码模型采样.** 训练好的去噪器把上述后验预测变成反向时间更新规则。在吸收式情形中，终端状态是全掩码序列。从这个状态出发，采样器反复选择当前被掩码坐标中的一个集合 $A \subseteq M_k$，并通过抽取

$$\widehat{x}_i \sim \mathsf{D}_k(i,\cdot \mid x_k), \qquad i \in A.$$

来填充这些坐标。除非算法有意重新掩码或重采样，A 外的坐标保持不变。因此，调度就是每个反向阶段选择集合 A 和时间标签 k 的规则。

这个视角使掩码扩散的任意顺序特征变得显式。模型不绑定到从左到右的分解：它被训练为从部分掩码上下文中去噪，所以推理时采样器可以以任意顺序揭示坐标。这种不依赖顺序的视角可追溯到 NADE 的顺序无关训练 [45]，并由 Hoogeboom 等人的自回归扩散模型 [28] 直接联系到吸收式/掩码扩散。对一个排序或块序列 $A_1, A_2, \dots, A_m \subseteq \{1,\dots,L\}$，采样器使用以当前部分序列为条件的后验预测来更新 A_j 中仍被掩码的坐标。单元素块给出带有某个选定顺序的类自回归采样器；较大的块给出并行解码；根据模型置信度自适应选择的块给出从粗到细或从易到难的揭示调度。

这种自由之所以合理，是因为训练一开始就没有固定顺序。目标 (7.6) 掩码一个随机坐标子集，并要求模型从未掩码坐标预测被掩码位置上的干净 token。随着训练样本、噪声水平和随机掩码变化，被掩码集合 M_k 基本遍历 $\{1,\dots,L\}$ 的所有子集，所以模型隐式地被训练来近似整族条件分布

$$\mathbb{P}\big(X_{0,i} = a \mid X_{k,M_k^c} = x_{M_k^c}\big), \qquad M_k \subseteq \{1,\dots,L\},\ i \in M_k,$$

即在给定部分掩码序列中可见条目时，干净坐标 i 的条件律。掩码扩散的两个特征使这件事精确成立。第一，前向过程只掩码 token，从不改变它们，所以 $X_{k,M_k^c} = X_{0,M_k^c}$：以 X_k 的可见部分为条件，等同于以相应干净值为条件。第二，某个坐标是否被掩码与 token 值独立决定，所以掩码模式除了说明哪些坐标被观测到之外不携带额外信息。因此，(7.6) 的总体最优解 $\mathsf{D}_k^\star(i, \cdot \mid x_k)$ 正是上述条件分布。

推理顺序因此只是一个选择下一步查询哪个条件分布的规则。从当前观测集合 M_k^c 中揭示坐标 i，使用的是 $\mathbb{P}(X_{0,i} \mid X_{k,M_k^c} = x_{M_k^c})$，这是模型已经学过的那族条件分布中的一个成员。每一种顺序都把联合分布分解成来自同一族的条件分布，这就是同一个训练好模型支持任意顺序的原因。不过，块大小的选择仍然是采样选择。如果一个块包含若干坐标，而采样器从同一上下文中独立抽取它们，那么它就是在使用该块联合条件分布的因子化近似。接下来两个小节把这个区别精确化：先写出近似反向核的一般 KL 记账，再分离并行块更新带来的额外误差。

7.6. **与连续情形平行的误差分析.** 一个调度加上一个去噪器定义近似反向核 $\widehat{P}_k^{\leftarrow}(x' \to \cdot)$，而精确核是来自 (7.2) 的 $P_k^{\leftarrow}(x' \to \cdot)$。令 $\widehat{p}_0$ 为近似反向链生成的最终分布，令 $\widehat{p}_K$ 为该链在时间 K 初始化时使用的律。路径空间 KL 论证给出

$$\mathrm{D}_{\mathsf{KL}}(p_0 \| \widehat{p}_0) \le \mathrm{D}_{\mathsf{KL}}(p_K \| \widehat{p}_K) + \sum_{k=0}^{K-1} \mathbb{E}_{X_{k+1} \sim p_{k+1}} \mathrm{D}_{\mathsf{KL}}\Big(P_k^{\leftarrow}(X_{k+1} \to \cdot) \Big\| \widehat{P}_k^{\leftarrow}(X_{k+1} \to \cdot) \Big). \tag{7.7}$$

这是连续扩散分解的精确类似：

连续扩散	离散扩散	含义
$p_K \approx \mathcal{N}(0, I)$	$p_K \approx \pi_{\mathsf{ref}}$	在噪声端初始化
$P_k^{\leftarrow} \propto p_k \times$ 高斯似然	$P_k^{\leftarrow} \propto p_k \times P_k^{\rightarrow}$	精确反向核
$\mathsf{s}_k \approx \mathsf{s}_k^\star$	$\widehat{P}_k^{\leftarrow} \approx P_k^{\leftarrow}$ 或 $\mathsf{D}_k \approx \mathsf{D}_k^\star$	学习到的去噪信息
$\sum_k \eta_k \varepsilon_{k,\mathsf{score}}^2$	$\sum_k \mathbb{E}\ \mathrm{D}_{\mathsf{KL}}(P_k^{\leftarrow} \| \widehat{P}_k^{\leftarrow})$	统计/模型误差

证明与第 6 节的 KL 望远镜论证逐字相同：对精确和近似反向路径律应用引理 6.1，然后用数据处理把路径投影到最终状态。

分解 (7.7) 分离了与连续情形相同的两个误差来源：初始化间隙 $\mathrm{D}_{\mathsf{KL}}(p_K \| \widehat{p}_K)$ 和沿反向步求和的一步反向核 KL。在吸收式掩码扩散中，这些 KL 项是否有限本身就有信息量。如果采样器从确定性的全掩码状态出发，则 $\widehat{p}_K = \delta_{[\mathsf{M}]^L}$，所以只有当前向终端律 p_K 也支撑在全掩码状态上时，初始化项才有限。对独立掩码链 (7.4)，这要求 $\bar{\alpha}_K = 0$，例如通过一个 $\beta_k = 1$ 的终端步实现。

一旦支撑兼容从而 KL 项有限，(7.7) 中的第二项就与模型实际训练的量相关。对掩码扩散而言，模型不直接学习反向核；它通过交叉熵目标 (7.6) 学习后验干净 token 分布。示意性地，我们有界

$$\mathbb{E}\ \mathrm{D}_{\mathsf{KL}}(P_k^{\leftarrow}(X_{k+1} \to \cdot) \| \widehat{P}_k^{\leftarrow}(X_{k+1} \to \cdot)) \lesssim \mathbb{E} \sum_{i \in M_{k+1}} \mathrm{D}_{\mathsf{KL}}\big(\mathsf{D}_{k+1}^\star(i, \cdot \mid X_{k+1}) \big\| \mathsf{D}_{k+1}(i, \cdot \mid X_{k+1}) \big).$$

这个界是后验预测的交叉熵后悔项，也就是训练所控制的部分：当学习后验 D_{k+1} 接近真实后验 $\mathsf{D}_{k+1}^\star$ 时它消失。当采样器一次只去掩码一个坐标时，它就是全部一步误差，因为每个单坐

标反向步查询一个真实条件分布（如第 7.5 小节所解释），所以近似反向链会复制精确反向链。因此，对完美去噪器而言，任意顺序的单坐标解码都是精确的。它的缺点是成本：每一步只揭示一个坐标需要与坐标数一样多的网络评估。这引出下一小节的问题：能否一次揭示多个坐标？这样会牺牲多少精度？

7.7. **并行掩码推理.** 为了避免每个坐标花一次网络评估，实际掩码采样器常常在一个反向步中揭示整块坐标。这可能表示两件不同的事。最简单的实现是*因子化块解码*：如果 A 是要从当前部分观测序列 x_k 中揭示的坐标集合，就采样

$$a_A \sim \prod_{i\in A} \mathsf{D}_k(i, a_i \mid x_k).$$

然而，精确对象是联合条件律

$$\mathbb{P}(X_{0,A} = a_A \mid X_k = x_k).$$

等价地，对任意排序 $A = \{i_1, \dots, i_m\}$，链式法则把这个联合条件分布写成

$$\prod_{\ell=1}^{m} \mathbb{P}\big(X_{0,i_\ell} = a_{i_\ell} \mid X_k = x_k,\, X_{0,i_1} = a_{i_1}, \dots, X_{0,i_{\ell-1}} = a_{i_{\ell-1}}\big)\,.$$

因此，一个块可以通过顺序查询更新后的条件分布来精确采样，但从同一上下文得到的一坐标后验的乘积，只有在条件独立时才是精确的。这两个律之间的间隙就是因子化误差。因此，推理调度设置了速度–精度权衡：更大的块缩短反向链，但引入条件独立偏差；单元素块消除偏差，代价是反向评估次数与坐标数相同。Lavenant 和 Zanella [31] 直接分析了使用因子化近似的掩码扩散中的这一误差，把相对熵误差分解为学习项和因子化项，并把最优块大小调度与数据分布的信息剖面联系起来。

这并不是并行推理的唯一可能做法。上面的块采样器是对联合条件分布的因子化近似；另一个问题是，能否并行化精确顺序条件采样过程本身。在一个可访问目标律在 $\mathcal{V}^L$ 上的精确条件边际的 oracle 模型中，Anari、Gao 和 Rubinstein [5] 展示了如何组织查询，从而在 $\widetilde{O}(L^{2/3})$ 并行时间内采样任意乘积空间律。Anari 等人 [4] 使用自推测拒绝采样把并行时间改进到 $\widetilde{O}(L^{1/2})$。这些结果最好被理解为简单块更新的一种修正并行版本：它们不是用独立边际替代联合条件分布，而是使用并行查询并配合修正来模拟顺序条件分布。

7.8. **CTMC 的设置与学习.** 到目前为止，我们使用离散时间掩码链，因为它让 Bayes 规则和去噪后验容易看清。不过，对实际采样器的分析而言，更锐利的语言是连续时间：它让我们把精确反向动力学描述为一次一个坐标跳变，然后把并行更新看作这些动力学的数值近似。

这也是现代离散扩散模型大多不使用随机矩阵 $P_k^{\rightarrow}$，而是表述为*连续时间马尔可夫链*（CTMC）的原因之一；CTMC 是前向 SDE 在有限状态上的类似物。前向破坏作为一个 CTMC 运行，其速率矩阵按坐标分解，所以每个坐标独立加噪。这对反向过程有一个决定性后果：任一瞬间只有一个坐标改变，因为两个独立坐标在同一个无穷小区间中同时跳变的概率是 $o(\mathrm{d}t)$。因此，精确反向动力学没有歧义，并且一次更新一个坐标；上面描述的并行更新偏差以受控形式再次出现，即作为在有限时间网格上近似这些动力学的误差（*τ-leaping*）。连续时间因而清

晰分离了后验/得分误差与离散化误差，产生最干净的得分类似物，并允许一族灵活的反向采样器 [10, 36]。

有限集合 $\mathcal{X}$ 上的 CTMC 由一个时间依赖的前向速率矩阵 $R_t^{\rightarrow}$ 生成，其中

$$R_t^{\rightarrow}(x,y) \geq 0 \ (y \neq x), \qquad R_t^{\rightarrow}(x,x) = -\sum_{y \neq x} R_t^{\rightarrow}(x,y).$$

因此，非对角元是跳变速率，对角元使每一行和为零。在无穷小步上，

$$\mathbb{P}(X_{t+\,\mathrm{d}t} = y \mid X_t = x) = \begin{cases} R_t^{\rightarrow}(x,y)\,\mathrm{d}t + o(\,\mathrm{d}t), & y \neq x, \\ 1 + R_t^{\rightarrow}(x,x)\,\mathrm{d}t + o(\,\mathrm{d}t), & y = x. \end{cases}$$

等价地，

$$\mathbb{P}(X_{t+\,\mathrm{d}t} = y \mid X_t = x) = \delta_{xy} + R_t^{\rightarrow}(x,y)\,\mathrm{d}t + o(\,\mathrm{d}t).$$

如果把 $p_t(x) = \mathbb{P}(X_t = x)$ 写成行向量，则其演化为

$$\partial_t p_t = p_t R_t^{\rightarrow}, \qquad \partial_t p_t(y) = \sum_{x \in \mathcal{X}} p_t(x) R_t^{\rightarrow}(x,y). \tag{7.8}$$

这是 Fokker–Planck 方程 (1.2) 的有限状态对应物：速率矩阵替代了作用于连续密度上的漂移–扩散算子。

速率矩阵也是指定破坏族的位置。吸收式(掩码)生成元以一个依赖调度的速率把每个 token 发送到 [M]，它是前面小节掩码扩散的连续时间版本；均匀生成元把每个 token 推向 $\mathcal{V}$ 上的均匀分布；其他选择可以编码结构化 token 图。因此，掩码扩散是这个 CTMC 图景的一个特例。

反向过程同样是一个 CTMC，它的速率由一个 Bayes 比值公式固定；这是反向核 (7.2) 的时间连续版本。对 $y \neq x$，

$$R_t^{\leftarrow}(y,x) = R_t^{\rightarrow}(x,y)\,\frac{p_t(x)}{p_t(y)}. \tag{7.9}$$

对角元被选择为使行和为零时，这些速率正好在反向时间中追踪同一个边际。事实上，对每个满足 $p_t(y) > 0$ 的 y，

$$\begin{aligned} (p_t R_t^{\leftarrow})(y) &= \sum_{x \neq y} p_t(x) R_t^{\leftarrow}(x,y) - p_t(y) \sum_{x \neq y} R_t^{\leftarrow}(y,x) \\ &= \sum_{x \neq y} p_t(x) R_t^{\rightarrow}(y,x) \frac{p_t(y)}{p_t(x)} - p_t(y) \sum_{x \neq y} R_t^{\rightarrow}(x,y) \frac{p_t(x)}{p_t(y)} \\ &= p_t(y) \sum_{x \neq y} R_t^{\rightarrow}(y,x) - \sum_{x \neq y} p_t(x) R_t^{\rightarrow}(x,y) \\ &= -\partial_t p_t(y), \end{aligned}$$

其中最后一个等号是前向方程 (7.8)。因此 $-\partial_t p_t = p_t R_t^{\leftarrow}$，这就是链从时间 T 向时间 0 模拟时的前向方程。从 p_T 开始反向 CTMC，就会在每个中间时间得到边际 p_t。与连续情形一样，前

向速率 $R_t^{\rightarrow}$ 由设计给定并已知，唯一未知的成分是相邻状态上同一时刻的密度比集合。我们把精确得分写作

$$\mathsf{s}_t^\star(x,y) := \frac{p_t(y)}{p_t(x)}, \qquad y \neq x,\ R_t^{\rightarrow}(y,x) > 0.$$

这是比值得分 (7.3) 的连续时间形式，也是 $\nabla \log p_t$ 的离散类似物，记录加噪律沿每个可行移动的有限相对变化。令 $\mathsf{s}_t(x,y) > 0$ 为学习预测，$\mathsf{s}_t(x,y) \approx \mathsf{s}_t^\star(x,y)$。方向按反向采样来选择：当反向链位于 x 时，候选前驱 y 可行，正好意味着前向 CTMC 能从 y 跳到 x。在这个约定下，学习反向速率为

$$\widehat{R}_t^{\leftarrow}(x,y) = R_t^{\rightarrow}(y,x)\,\mathsf{s}_t(x,y), \qquad y \neq x.$$

一开始训练目标看起来不可达，因为比值含有未知边际律 p_t。关键的去噪恒等式是：这个未知比值可以写成已知前向似然比的后验平均。

$$\mathsf{s}_t^\star(x,y) = \mathbb{E}\left[\frac{\mathbb{P}(X_t = y \mid X_0)}{\mathbb{P}(X_t = x \mid X_0)} \,\middle|\, X_t = x\right]. \tag{7.10}$$

因此，训练可以采样干净数据 X_0，把它破坏到 $X_t = x$，并在损失内部使用已知的前向似然比。

得分熵训练 [36] 先为一条反向边选择一个标量差异度，从而把这个恒等式变成正比值的监督损失。令 $\mathsf{s}^\star > 0$ 表示精确比值，令 $\mathsf{s} > 0$ 为学习预测，并在 $u > 0$ 上设 $F(u) = -\log u$。得分熵使用缩放后的 Bregman 散度

$$\mathsf{s}^\star D_F(\mathsf{s},\mathsf{s}^\star) = \mathsf{s}^\star\big(F(\mathsf{s}) - F(\mathsf{s}^\star) - F'(\mathsf{s}^\star)(\mathsf{s}-\mathsf{s}^\star)\big) = \mathsf{s} - \mathsf{s}^\star\log\mathsf{s} + \mathsf{s}^\star\log\mathsf{s}^\star - \mathsf{s}^\star.$$

这个量非负，导数为 $1 - \mathsf{s}^\star/\mathsf{s}$，并在 $\mathsf{s} = \mathsf{s}^\star$ 处最小；注意它只对正比值有定义，所以学习到的 $\mathsf{s}_t(x,y)$ 必须被约束为保持正值。在学习目标中，项 $\mathsf{s}^\star\log\mathsf{s}^\star - \mathsf{s}^\star$ 与学习得分无关。因此，一条边上依赖模型的部分是 $\mathsf{s}_t(x,y) - \mathsf{s}_t^\star(x,y)\log\mathsf{s}_t(x,y)$。训练时，未知比值 $\mathsf{s}_t^\star(x,y)$ 被加噪过程给出的可计算前向似然比替代。一个示意性的去噪得分熵损失为

$$\mathbb{E}_{t,X_0}\sum_{x\in\mathcal{X}}\mathbb{P}(X_t = x \mid X_0)\sum_{y:\,R_t^{\rightarrow}(y,x)>0} w_t(x,y)\left[\mathsf{s}_t(x,y) - \frac{\mathbb{P}(X_t = y \mid X_0)}{\mathbb{P}(X_t = x \mid X_0)}\log\mathsf{s}_t(x,y)\right], \tag{7.11}$$

其中省略了与学习得分无关的项。这里 $w_t(x,y) \geq 0$ 是选择的权重，常常取为进入的前向速率 $R_t^{\rightarrow}(y,x)$ 或其倍数。

这个目标具有正确总体目标的原因，正是上面的去噪恒等式 (7.10)。由于展示的损失对 (7.10) 中出现的似然比是仿射的，对未知干净数据平均只是把它替换为条件均值。由此得到的逐边风险，在忽略常数后为 $\mathsf{s}_t(x,y) - \mathsf{s}_t^\star(x,y)\log\mathsf{s}_t(x,y)$，并在每条反向可行边上的精确得分 $\mathsf{s}_t(x,y) = \mathsf{s}_t^\star(x,y)$ 处最小。

特殊选择 $w_t(x,y) = R_t^{\rightarrow}(y,x)$ 赋予这个去噪损失采样解释。在从当前状态 x 回到可能前驱 y 的边上，真实反向速率为 $R_t^{\rightarrow}(y,x)\mathsf{s}_t^\star(x,y)$，而学习反向速率为 $R_t^{\rightarrow}(y,x)\mathsf{s}_t(x,y)$。相应的 CTMC 路径空间相对熵的逐边贡献为（将在 (7.12) 中显示）

$$R_t^{\rightarrow}(y,x)\left[\mathsf{s}_t(x,y) - \mathsf{s}_t^\star(x,y) + \mathsf{s}_t^\star(x,y)\log\frac{\mathsf{s}_t^\star(x,y)}{\mathsf{s}_t(x,y)}\right].$$

去掉与学习得分无关的项后，这正是 (7.11) 中 $w_t(x,y) = R_t^{\rightarrow}(y,x)$ 时的目标。因此，速率加权选择与下面 CTMC 采样分析中使用的路径空间误差量相匹配。

7.9. **CTMC 采样.** 一旦学习比值指定了反向速率，剩下的问题就是对所得跳过程进行数值模拟。τ-leaping 是化学反应网络中随机模拟算法的一种标准加速，由 Gillespie [24] 引入。其思想是在一个短区间内聚合许多小跳变，同时假装跳变速率在该区间内基本保持常数。

对学习到的反向 CTMC，仍然保留前面的前向时间网格约定：

$$0 = t_0 < t_1 < \cdots < t_K = T, \qquad h_k = t_{k+1} - t_k.$$

反向采样器从 t_{k+1} 移动到 t_k。如果当前状态在时间 t_{k+1} 为 x，最简单的 tau-leap 把学习反向速率冻结在 (t_{k+1}, x)，并使用一阶核

$$\mathbb{P}(\widehat{X}_{t_k} = z \mid \widehat{X}_{t_{k+1}} = x) = h_k \, \widehat{R}^{\leftarrow}_{t_{k+1}}(x, z) + O(h_k^2), \qquad z \neq x,$$

其余概率分配给停留在 x。等价地，对乘积空间，可以为可行的局部移动抽取独立 Poisson 时钟 $N_z \sim \mathsf{Poisson}(h_k \widehat{R}^{\leftarrow}_{t_{k+1}}(x, z))$，并并行施加所提议的坐标改变；若若干不相容时钟同时响起，则使用固定的打破平局规则。这是 Euler–Maruyama 步的离散类似物：它用冻结的反向生成元替代随时间变化的反向生成元，并在并行实现中由于允许多个坐标在一步内更新而引入局部独立性误差。关于 τ-leaping 及相关算法的更详细讨论，参见 [21, Chapter 13]。

Tau-leaping 方便，但它仍然是近似，因为速率在整个区间上被冻结。一个紧密相关的精确构造是均匀化，最初由 Grassmann [25] 引入，随后由 Beentjes 和 Baker [8] 应用于化学反应模拟，并由 Chen 和 Ying [13] 应用于离散扩散模型。

我们说明均匀化如何工作。在区间 $[t_k, t_{k+1}]$ 上，选择一个时钟速率

$$\Lambda_k \geq \sup_{\substack{s \in [t_k, t_{k+1}] \\ x \in \mathcal{X}}} \sum_{z \neq x} \widehat{R}^{\leftarrow}_s(x, z).$$

从时间 t_{k+1} 的状态出发，以速率 Λ_k 在 $[t_k, t_{k+1}]$ 中抽取 Poisson 事件时刻，并按时间递减顺序处理它们。在事件时刻 s，如果当前状态为 x，则以概率 $\widehat{R}^{\leftarrow}_s(x, z)/\Lambda_k$ 跳到 $z \neq x$，否则停在 x。停留事件是虚跳。由于该时钟支配所有流出速率，这个 thinning 构造精确具有学习反向 CTMC 在该区间上的跳变律，没有冻结速率近似。

7.10. **CTMC 采样器的误差分析.** 描述完两种 CTMC 模拟格式后，我们现在问它们的输出律与期望数据律相差多少。数值分析可以按照第 6 节相同的方式组织。有三个贡献。第一，精确反向过程应从 p_T 出发，而采样器从某个可实现律 $\widehat{p}_T$ 初始化。这给出共同的初始化间隙。定义

$$\varepsilon^2_{\mathsf{init}} := \mathrm{D}_{\mathsf{KL}}(p_T \| \widehat{p}_T).$$

对 $\widehat{p}_T$ 的两种常见选择应作不同解释。如果 $\widehat{p}_T$ 是点质量，则 $\varepsilon^2_{\mathsf{init}}$ 主要是一个支撑条件。例如，在吸收式掩码扩散中，常取 $\widehat{p}_T = \delta_{[\mathsf{M}]^L}$；只有当前向终端律 p_T 也支撑在全掩码状态上时，KL 才有限。在有限积分掩码速率下，这可能失败，所以应当要么强制精确终端吸收，要么用更弱的度量来衡量端点误差。

如果 $\widehat{p}_T$ 具有满支撑，这一项就是普通初始化误差。例如，如果 $\widehat{p}_T$ 为均匀分布，并且前向 CTMC 以速率 ρ 把 KL 收缩到 $\widehat{p}_T$，那么

$$\varepsilon^2_{\mathsf{init}} \leq e^{-\rho T} \, \mathrm{D}_{\mathsf{KL}}(p_0 \| \widehat{p}_T) \leq e^{-\rho T} \log |\mathcal{X}|.$$

在乘积 token 空间 $\mathcal{X} = \mathcal{V}^L$ 上，对通常的独立坐标破坏而言，这一项的量级为 $L(\log|\mathcal{V}|)e^{-T}$。

初始化之后，固定一个模拟器，并先想象用真实反向速率运行该模拟器，再把真实比值得分替换为学习比值得分：

$$p_0 \quad\longrightarrow\quad p_0^{\mathsf{num},*} \quad\longrightarrow\quad \widehat{p}_0^{\mathsf{num}}.$$

这里 $p_0^{\mathsf{num},*}$ 表示所选数值模拟器在使用真实反向速率时的输出律。第一个比较分离数值误差，第二个比较分离得分估计误差。严格说来，KL 没有三角不等式，所以严格证明通常在路径空间上进行这个分解，或直接在对数强度的被积函数中进行，而不是把两个边际 KL 散度相加。不过，这种记账有助于组织误差分析。

得分估计项具有一个与扩散 Girsanov 定理类似的路径空间形式。我们按照连续得分误差相同的 epsilon 平方约定来写它：

(7.12)

$$\varepsilon_{\mathsf{SE}}^2 := \int_0^T \mathbb{E}_{X_t\sim p_t} \sum_{y:\, R_t^{\rightarrow}(y,X_t)>0} R_t^{\rightarrow}(y,X_t)\left[\mathsf{s}_t(X_t,y) - \mathsf{s}_t^\star(X_t,y) + \mathsf{s}_t^\star(X_t,y)\log\frac{\mathsf{s}_t^\star(X_t,y)}{\mathsf{s}_t(X_t,y)}\right]\mathrm{d}t.$$

这正是学习目标 (7.11) 中出现的同一个得分熵 Bregman 损失，这解释了它为什么是自然的训练损失。

(7.12) 的证明. 比较精确反向 CTMC 与学习反向 CTMC，前者速率为 $R_t^{\leftarrow}(x,y) = R_t^{\rightarrow}(y,x)\mathsf{s}_t^\star(x,y)$，后者速率为 $\widehat{R}_t^{\leftarrow}(x,y) = R_t^{\rightarrow}(y,x)\mathsf{s}_t(x,y)$，并假设凡精确速率为正处学习速率也为正。在一个长度为 $\mathrm{d}t$ 的小区间上，以当前状态 x 为条件。使用对角元表示停留概率，一步 KL 为

$$\begin{aligned}
&\sum_{y\neq x} R_t^{\leftarrow}(x,y)\,\mathrm{d}t\log\frac{R_t^{\leftarrow}(x,y)\,\mathrm{d}t}{\widehat{R}_t^{\leftarrow}(x,y)\,\mathrm{d}t} + \big(1+R_t^{\leftarrow}(x,x)\,\mathrm{d}t\big)\log\frac{1+R_t^{\leftarrow}(x,x)\,\mathrm{d}t}{1+\widehat{R}_t^{\leftarrow}(x,x)\,\mathrm{d}t} + o(\mathrm{d}t)\\
&\quad= \mathrm{d}t\sum_{y\neq x}\left[\widehat{R}_t^{\leftarrow}(x,y) - R_t^{\leftarrow}(x,y) + R_t^{\leftarrow}(x,y)\log\frac{R_t^{\leftarrow}(x,y)}{\widehat{R}_t^{\leftarrow}(x,y)}\right] + o(\mathrm{d}t).
\end{aligned}$$

代入两个反向速率，会在对数内部抵消公共因子 $R_t^{\rightarrow}(y,x)$，并给出

$$\sum_{y:\, R_t^{\rightarrow}(y,x)>0} R_t^{\rightarrow}(y,x)\left[\mathsf{s}_t(x,y) - \mathsf{s}_t^\star(x,y) + \mathsf{s}_t^\star(x,y)\log\frac{\mathsf{s}_t^\star(x,y)}{\mathsf{s}_t(x,y)}\right]\mathrm{d}t + o(\mathrm{d}t).$$

精确反向过程在噪声时间 t 的边际为 p_t。对 $X_t\sim p_t$ 平均并在 $t\in[0,T]$ 上积分，得到 (7.12)；如果两条反向链从不同的律出发，还会有初始化 KL 误差作为额外贡献。 □

对 τ-leaping 而言，精确得分数值律并不是真实反向 CTMC 律。在每个区间 $[t_k,t_{k+1}]$ 上，它用冻结速率 $R_{t_{k+1}}^{\leftarrow}$ 替代随时间变化的真实反向速率。随后用 $\widehat{R}_{t_{k+1}}^{\leftarrow}$ 替代这些冻结的真实速率，添加得分估计部分。令 $\bar{h}$ 表示控制反向步长的网格参数，令 $\bar{R}$ 为总流出反向速率的上界，即在 $[0,T]$ 上 $\sum_{z\neq x} R_s^{\leftarrow}(x,z)\le\bar{R}$。在有界速率和正则性假设下，对组合比较的随机积分分析给出如下示意形式的界：

$$\mathrm{D}_{\mathsf{KL}}(p_0\|\widehat{p}_0^{\tau}) \lesssim \varepsilon_{\mathsf{init}}^2 + \varepsilon_{\mathsf{SE}}^2 + \bar{R}^2\bar{h}T. \tag{7.13}$$

三个项分别是初始化误差、得分估计误差和精确得分 tau-leaping 离散化误差。

对乘积 token 空间 $\mathcal{X} = \mathcal{V}^L$，较新的工作给出了同一信息的更锐利版本，并显式展示了对词表大小的依赖。如果学习比值被裁剪到 $[B^{-1}, B]$，且反向网格由网格参数 $\bar{h}$ 控制，那么标准 tau-leaping 采样器满足如下界，略去对数因子和端点正则性常数：

$$\mathrm{D}_{\mathsf{KL}}(p_0 \| \widehat{p}_0^{\tau}) \lesssim \varepsilon_{\mathsf{init}}^2 + \varepsilon_{\mathsf{SE}}^2 + \bar{h}\, L^2 |\mathcal{V}|\, T, \tag{7.14}$$

如 Liang 等人 [33] 的分析所示。对通常的独立坐标破坏，可以代入 $\varepsilon_{\mathsf{init}}^2 \lesssim L(\log|\mathcal{V}|)e^{-T}$。因此，要让 tau-leaping 离散化误差为 ε 量级，需要 $\bar{h}$ 为 $\widetilde{O}(\varepsilon/(L^2|\mathcal{V}|T))$ 量级，进而在该分析中需要大约 $\widetilde{O}(L^2|\mathcal{V}|/\varepsilon)$ 个反向步的确定性网格。

使用均匀化可以实现高精度采样。若使用真实比值得分，它会精确模拟真实反向 CTMC，所以如果反向过程从 p_T 初始化，则 $p_0^{\mathsf{uni},*} = p_0$。因此，对学习得分，均匀化的误差估计化为

$$\mathrm{D}_{\mathsf{KL}}\big(p_0 \big\| \widehat{p}_0^{\mathrm{uni}}\big) \le \varepsilon_{\mathsf{init}}^2 + \varepsilon_{\mathsf{SE}}^2. \tag{7.15}$$

在使用通常独立坐标破坏的乘积 token 空间 $\mathcal{X} = \mathcal{V}^L$ 上，可以把 $\varepsilon_{\mathsf{init}}^2 \lesssim L(\log|\mathcal{V}|)e^{-T}$ 代入 (7.15)，得到 $\mathrm{D}_{\mathsf{KL}}(p_0\|\widehat{p}_0^{\mathrm{uni}}) \lesssim L(\log|\mathcal{V}|)e^{-T} + \varepsilon_{\mathsf{SE}}^2$。选择 $T \asymp \log(L\log|\mathcal{V}|/\varepsilon^2)$，并学习比值得分使 $\varepsilon_{\mathsf{SE}}^2 \lesssim \varepsilon^2$，得到 $\mathrm{D}_{\mathsf{KL}}(p_0\|\widehat{p}_0^{\mathrm{uni}}) \lesssim \varepsilon^2$。Chen 和 Ying [13] 进一步证明，使用自适应支配速率时，Poisson 事件的期望次数在超立方体维数上近乎线性。他们的定理针对 $\{0,1\}^d$ 陈述；对大小为 $|\mathcal{V}|$ 的字母表，一个固定二进制编码每个 token 使用 $q = \lceil \log_2 |\mathcal{V}| \rceil$ 个 bit，因此 $d = Lq$。用 token 记号来说，这给出了 ε^2 KL 误差下

$$\widetilde{O}(L \log |\mathcal{V}|)$$

的期望事件数。

8. 引导、奖励倾斜与推理时强化学习

到目前为止，我们已经构造并分析了连续和离散的扩散采样器，其目标是复现数据分布，或复现它的一个轻微平滑版本。在许多应用中，人们想要的是对这个基础采样器进行受控修改：生成满足某个条件的样本，偏好高奖励输出，或在推理时适配模型。引导、奖励倾斜和推理时 RL 都是在保持接近预训练分布的同时，把基础模型偏向偏好输出的机制；它们共同的数学语言，是本节发展的 KL 正则化测度变换。

8.1. 奖励倾斜目标与 KL 正则化优化.

令 p_0 表示预训练采样器产生的干净输出分布；我们保留扩散约定，把干净样本写作 x_0。此时我们只指定干净输出上的期望律；相应的反向时间动力学会在下面推出。给定奖励 $r : \mathbb{R}^d \to \mathbb{R}$ 和逆温度 $\beta \ge 0$，定义指数倾斜目标

$$p_0^{\beta}(x_0) = \frac{1}{Z_\beta} p_0(x_0) e^{\beta r(x_0)}. \tag{8.1}$$

这里

$$Z_\beta = \int p_0(x_0) e^{\beta r(x_0)}\, \mathrm{d}x_0 = \mathbb{E}_{X_0 \sim p_0} e^{\beta r(X_0)}$$

是归一化常数，假设它有限。等价地，对任意测试函数 f，

$$\mathbb{E}_{p_0^\beta} f(X_0) = \frac{\mathbb{E}_{p_0}\left[f(X_0)e^{\beta r(X_0)}\right]}{\mathbb{E}_{p_0} e^{\beta r(X_0)}}.$$

因此，目标由基础模型按干净输出奖励重加权样本得到。情形 $\beta = 0$ 恢复 p_0，而更大的 β 把更多质量放在高奖励区域，从而用基础模型下的多样性换取奖励提升。

条件生成也适合相同形式。如果 y 是观测、类别标签或提示词，而 $p(y \mid x_0)$ 是相应的似然或兼容性模型，则 Bayes 规则给出

$$p_0(x_0 \mid y) \propto p_0(x_0)p(y \mid x_0).$$

这是一个奖励函数为 $x_0 \mapsto \log p(y \mid x_0)$ 且 $\beta = 1$ 的指数倾斜；当显式似然不可用时，分类器得分或学习得到的偏好模型扮演同样角色。推理时问题是修改反向采样器，使其干净输出律近似这样的倾斜或条件目标，同时仍然复用预训练基础动力学。

同一个倾斜律有一个有用的变分含义。在最终样本上的所有分布 q 中，它是

$$\sup_q \left\{\beta \mathbb{E}_q r - \mathrm{D}_{\mathsf{KL}}(q\|p_0)\right\} \tag{8.2}$$

的优化解。等价地，如果 $\beta > 0$，

$$p_0^\beta = \arg\max_q \left\{\mathbb{E}_q r - \frac{1}{\beta}\,\mathrm{D}_{\mathsf{KL}}(q\|p_0)\right\}.$$

因此，β 控制奖励提升与保持接近基础模型之间的权衡。大的 β 会强烈推向高奖励，并有模式坍缩或奖励黑客化的风险；小的 β 保持基础分布，但带来较弱的对齐。

证明. 令 $p_0^\beta(x) = Z_\beta^{-1} p_0(x)e^{\beta r(x)}$。对任意 q，

$$\mathrm{D}_{\mathsf{KL}}(q\|p_0^\beta) = \int q(x) \log \frac{q(x)}{p_0(x)e^{\beta r(x)}/Z_\beta}\,\mathrm{d}x = \mathrm{D}_{\mathsf{KL}}(q\|p_0) - \beta\mathbb{E}_q r + \log Z_\beta.$$

整理得

$$\beta\mathbb{E}_q r - \mathrm{D}_{\mathsf{KL}}(q\|p_0) = \log Z_\beta - \mathrm{D}_{\mathsf{KL}}(q\|p_0^\beta) \le \log Z_\beta,$$

等号当且仅当 $q = p_0^\beta$ 时成立。 □

8.2. **作为得分倾斜的引导.** 为了对倾斜目标 p_0^β 运行扩散采样器，我们需要它的加噪边际得分，因此问题是这个得分与已经学习到的基础得分有何不同。令 p_t 和 p_t^β 分别为 p_0 和 p_0^β 的加噪边际。把干净到噪声边际前向核 $\mathrm{Law}(X_t \mid X_0 = x_0)$ 的密度写作 $P_{0,t}^{\rightarrow}(x_0 \to x)$。由于奖励只作用于干净样本 x_0，对倾斜律加噪给出

$$p_t^\beta(x) = \frac{1}{Z_\beta}\int P_{0,t}^{\rightarrow}(x_0 \to x)\, e^{\beta r(x_0)} p_0(x_0)\,\mathrm{d}x_0 = \frac{p_t(x)}{Z_\beta}\, h_t^\beta(x), \qquad h_t^\beta(x) := \mathbb{E}\left[e^{\beta r(X_0)} \mid X_t = x\right].$$

因此，加噪后的倾斜律是基础加噪律乘以后验倾斜因子 h_t^β；这个因子从当前噪声状态读取其会去噪到的干净样本的期望指数奖励。由于这个边际是一个乘积，其对数梯度分解为基础得分加上一个修正项：

$$\nabla \log p_t^\beta(x) = \nabla \log p_t(x) + \nabla \log h_t^\beta(x),$$

所以精确引导只是把倾斜梯度 $\nabla \log h_t^\beta$ 加到基础得分上；实际方法之间的差别只在于如何近似它。

对条件生成，奖励函数是 $x_0 \mapsto \log p(y \mid x_0)$。同样的后验倾斜记号给出噪声似然

$$h_t(x \mid y) := \mathbb{E}[p(y \mid X_0) \mid X_t = x].$$

这是噪声分类器估计的量。它本身不是加噪条件密度；相反，它倾斜基础噪声密度。如果条件固定，加噪条件密度为

$$p_t(x \mid y) \propto p_t(x) h_t(x \mid y),$$

其中缺失的归一化常数与 x 无关。在噪声水平 t 处取对数梯度得到

$$\nabla \log p_t(x \mid y) = \nabla \log p_t(x) + \nabla \log h_t(x \mid y). \tag{8.3}$$

这就是 classifier guidance [20]：训练或使用一个作用在噪声输入上的分类器，然后把它的梯度加到无条件得分上。Classifier-free guidance [27] 不用单独的分类器来估计同一个增量。我们使用已经建立的得分记号：

$$\mathsf{s}_t(x) \approx \nabla \log p_t(x), \qquad \mathsf{s}_t(x \mid y) \approx \nabla \log p_t(x \mid y).$$

于是 (8.3) 提示

$$\mathsf{s}_t(x \mid y) - \mathsf{s}_t(x) \approx \nabla \log h_t(x \mid y).$$

给定引导强度 $\beta \geq 0$，classifier-free guidance 把无条件得分替换为

$$\mathsf{s}_t(x) \quad \longmapsto \quad \mathsf{s}_t(x) + \beta\big(\mathsf{s}_t(x \mid y) - \mathsf{s}_t(x)\big). \tag{8.4}$$

选择 $\beta = 1$ 得到条件得分估计 $\mathsf{s}_t(x \mid y)$，而 $\beta > 1$ 对条件得分增量进行外推；经验上这往往改善条件满足度，代价是离基础分布更远。如果两个学习得分都是精确的并且 $\beta = 1$，近似 (8.4) 会恢复 (8.3) 中的条件得分。对 $\beta \neq 1$，它给出的是噪声水平幂倾斜 $p_t(x) h_t(x \mid y)^\beta$ 的得分；而干净奖励倾斜 (8.1) 会涉及后验因子 $h_t^\beta(x) = \mathbb{E}[p(y \mid X_0)^\beta \mid X_t = x]$。对噪声似然取幂，并不同于用取幂后的干净似然的噪声条件期望来倾斜，因此二者只在 $\beta = 1$ 时重合。

对小倾斜，后验倾斜因子有线性近似。由于 $h_t^\beta(x) = 1 + \beta V_t(x) + O(\beta^2)$，其中 $V_t(x) = \mathbb{E}[r(X_0) \mid X_t = x]$，我们有

$$\nabla \log p_t^\beta(x) \approx \nabla \log p_t(x) + \beta \nabla V_t(x), \qquad V_t(x) = \mathbb{E}[r(X_0) \mid X_t = x]. \tag{8.5}$$

因此，在小倾斜区域中，引导用后验期望奖励 V_t 的梯度扰动得分，所以奖励模型只需要捕捉这个一阶景观。简化之处在于 V_t 平均的是奖励 r 本身，而精确因子 h_t^β 平均的是 $e^{\beta r}$；二者只在 β 的一阶意义下相同。精确引导使用完整的 h_t^β，而小倾斜形式把它替换为更便宜的奖励梯度修正。

8.3. **作为 Polchinski 流的奖励倾斜.** 从实际方法后退一步，倾斜因子 h_t^β 把引导与第 4 节中的重整化视角联系起来。回顾那里使用的方差爆炸归一化 $X_t = X_0 + \sqrt{t}\, Z$，其中 $p_t = p_0 * \mathcal{N}(0, tI)$，而有效势 $U_t = -\log p_t$ 求解有限维 Polchinski 方程 (4.12)。反过来阅读定义 h_t^β 的积分，未归一化倾斜边际 $h_t^\beta\, p_t = (e^{\beta r} p_0) * \mathcal{N}(0, tI)$ 是同一前向通道作用于奖励重加权数据测度 $e^{\beta r} p_0$ 的结果，也就是说，对倾斜数据加噪与用 h_t^β 倾斜加噪数据是相同的；因此它是热流，并像 p_t 一样求解 (4.11)。除以 p_t 后，留下倾斜因子的前向 Kolmogorov 方程

$$\partial_t h_t^\beta = \tfrac{1}{2}\Delta h_t^\beta + \mathsf{s}_t^\star \cdot \nabla h_t^\beta, \tag{8.6}$$

所以 h_t^β 被得分驱动的生成元 $\frac{1}{2}\Delta + \mathsf{s}_t^\star \cdot \nabla$ 输运。等价地，倾斜有效势 $U_t^\beta := -\log(h_t^\beta\, p_t)$ 满足同一个 Polchinski 方程

$$\partial_t U_t^\beta = \tfrac{1}{2}\Delta U_t^\beta - \tfrac{1}{2}\left\|\nabla U_t^\beta\right\|^2, \tag{8.7}$$

只不过裸势从 $U_0 = -\log p_0$ 移动到 $U_0^\beta = -\log p_0 - \beta r$。因此，奖励倾斜就是从一个奖励平移后的初始势出发运行重整化流，而引导得分 $-\nabla U_t^\beta = \mathsf{s}_t^\star + \nabla \log h_t^\beta$ 正是上一小节的加性引导修正。

需要强调的是，这是一种重述，而不是一个配方。求解倾斜方程 (8.6) 或 Polchinski 方程 (8.7) 并不比评估定义它的后验期望 $h_t^\beta(x) = \mathbb{E}[e^{\beta r(X_0)} \mid X_t = x]$ 更容易；这个 PDE 携带着与该条件期望相同的计算困难。

8.4. **路径空间控制与 Doob 变换.** 上一小节中得分倾斜、分类器引导和 Polchinski 流对引导的描述都是精确的，但它们都依赖同一个难处理对象，即后验倾斜 h_t^β。为了理解实践中如何近似它，从边际转向整条反向轨迹会有帮助：在路径空间上，倾斜变成一个随机控制问题，而它的精确解，也就是基础反向链的 Doob 变换，准确暴露了必须估计哪些条件期望。

把一个连续反向采样器在生成时间中写作

$$\mathrm{d}Y_s^{\leftarrow} = b_s(Y_s^{\leftarrow})\,\mathrm{d}s + \sigma_s\,\mathrm{d}B_s^{\leftarrow}, \qquad 0 \le s \le T,$$

并令 $\mathbb{P}^0$ 为完整反向轨迹 $(Y_s^{\leftarrow})_{0\le s\le T}$ 上的路径律，其中干净输出为 $Y_T^{\leftarrow}$。

奖励倾斜采样在终端奖励与偏离该基础律的代价之间折中。在路径律层面，它是 KL 正则化优化

$$\sup_{\mathbb{Q}} \left\{\beta \mathbb{E}_{\mathbb{Q}} r(Y_T^{\leftarrow}) - \mathrm{D}_{\mathsf{KL}}(\mathbb{Q}\|\mathbb{P}^0)\right\} = \log \mathbb{E}_{\mathbb{P}^0} e^{\beta r(Y_T^{\leftarrow})}, \tag{8.8}$$

这是 Gibbs 变分原理 (8.2) 的路径空间版本。其优化解是 Gibbs 路径律

$$\frac{\mathrm{d}\mathbb{Q}^\star}{\mathrm{d}\mathbb{P}^0}((Y_s^{\leftarrow})_{0\le s\le T}) = \frac{e^{\beta r(Y_T^{\leftarrow})}}{\mathbb{E}_{\mathbb{P}^0} e^{\beta r(Y_T^{\leftarrow})}}. \tag{8.9}$$

证明. 证明与有限维奖励倾斜证明完全相同。定义 $\mathrm{d}\mathbb{Q}^\star \propto e^{\beta r(Y_T^{\leftarrow})}\,\mathrm{d}\mathbb{P}^0$。则

$$\mathrm{D}_{\mathsf{KL}}(\mathbb{Q}\|\mathbb{Q}^\star) = \mathrm{D}_{\mathsf{KL}}(\mathbb{Q}\|\mathbb{P}^0) - \beta \mathbb{E}_{\mathbb{Q}} r(Y_T^{\leftarrow}) + \log \mathbb{E}_{\mathbb{P}^0} e^{\beta r(Y_T^{\leftarrow})}.$$

整理并使用 KL 的非负性，即证明 (8.8)，其上确界在 $\mathbb{Q}^\star$ 处达到。 □

为了把它变成采样器，我们把它表达为漂移上的控制问题。受控采样器只改变漂移，

$$\mathrm{d}Y_s^{\leftarrow} = \big(b_s(Y_s^{\leftarrow}) + \sigma_s u_s(Y_s^{\leftarrow})\big)\,\mathrm{d}s + \sigma_s\,\mathrm{d}B_s^{\leftarrow},$$

路径律为 $\mathbb{P}^u$；使用相同的扩散系数意味着两个路径律只通过漂移不同。在通常的绝对连续性和可积性假设下，附录 A 的 Girsanov KL 公式给出

$$\mathrm{D}_{\mathsf{KL}}(\mathbb{P}^u\|\mathbb{P}^0) = \frac{1}{2}\mathbb{E}_{\mathbb{P}^u}\int_0^T \|u_s(Y_s^{\leftarrow})\|^2\,\mathrm{d}s, \tag{8.10}$$

其中 $u_s(Y_s^{\leftarrow})$ 是渐进可测控制，所以二次控制能量正是把采样器从基础律中引导开来的 KL 代价。把 $\mathbb{Q}$ 限制为受控扩散 $\mathbb{P}^u$，并把这个代价代入 (8.8)，使 Gibbs 问题等价于漂移上的随机控制问题

$$\sup_u \mathbb{E}_{\mathbb{P}^u}\left[\beta r(Y_T^{\leftarrow}) - \frac{1}{2}\int_0^T \|u_s(Y_s^{\leftarrow})\|^2\,\mathrm{d}s\right]. \tag{8.11}$$

二者是同一个优化，只是分别在测度层面和漂移层面陈述。

把优化解 (8.9) 读回漂移层面，就恢复了“引导就是控制”的口号：最优控制的漂移增量正是第 8.2 小节中理想引导反向动力学的价值梯度项，只是现在采用生成时间参数化。KL 代价 (8.10) 在第 6 节中度量离散化误差；在这里它是有意选择的引导预算，被花在最有效提高期望终端奖励的地方。

从这里到本节末尾，我们采用离散时间视角，使用第 5 节中的时间离散反向链 $(X_K,\dots,X_0)$，而不是连续 SDE。原因是我们接下来构造的对象，也就是下面的 Doob 变换、第 8.5 小节的序贯 Monte Carlo 权重，以及第 8.6 小节的策略，都是逐步作用的；有限反向核链能在没有随机微积分记账的情况下陈述它们，连续公式则在小步极限中恢复。在这个记号中，干净输出由 0 标记，所以终端奖励写作 $r(X_0)$。路径空间优化解有显式的马尔可夫核形式：每个基础反向提议按其后续路径的倾斜因子重加权，并用当前状态处的倾斜因子归一化。这个核重加权就是基础反向链的 Doob 变换。

定理 8.1 (最优路径倾斜与 Doob 变换)**.** *令 $\mathbb{P}^0$ 为 $(X_K, X_{K-1},\dots,X_0)$ 上的一条基础反向马尔可夫链，其初始律为 p_K^0，反向核为 $P_k^{0,\leftarrow}(x_{k+1}\to \mathrm{d}x_k)$。对终端奖励 $r(X_0)$，定义*

$$h_k(x_k) = \mathbb{E}_{\mathbb{P}^0}\left[e^{\beta r(X_0)} \mid X_k = x_k\right].$$

则路径空间 Gibbs 问题 (8.8) 的优化解是倾斜路径律

$$\frac{\mathrm{d}\mathbb{Q}^\star}{\mathrm{d}\mathbb{P}^0} = \frac{e^{\beta r(X_0)}}{\mathbb{E}_{\mathbb{P}^0} e^{\beta r(X_0)}}.$$

此外，$\mathbb{Q}^\star$ 是马尔可夫的，其反向核为

$$P_k^{\star,\leftarrow}(x_{k+1}\to \mathrm{d}x_k) = P_k^{0,\leftarrow}(x_{k+1}\to \mathrm{d}x_k)\frac{h_k(x_k)}{h_{k+1}(x_{k+1})}, \qquad h_{k+1}(x_{k+1}) = \int h_k(y)P_k^{0,\leftarrow}(x_{k+1}\to \mathrm{d}y). \tag{8.12}$$

证明. 倾斜路径律的变分最优性来自 KL 的非负性。剩下的是计算它的核。以当前反向时间状态 $X_{k+1} = x_{k+1}$ 为条件。在倾斜律下，下一个状态 X_k 的条件分布正比于基础条件分布乘以未来权重的期望：

$$P_k^{\star,\leftarrow}(x_{k+1} \to \mathrm{d}x_k) \propto P_k^{0,\leftarrow}(x_{k+1} \to \mathrm{d}x_k)\mathbb{E}_{\mathbb{P}^0}\left[e^{\beta r(X_0)} \mid X_k = x_k\right].$$

根据定义，该期望为 $h_k(x_k)$。归一化常数为

$$\int h_k(y)P_k^{0,\leftarrow}(x_{k+1} \to \mathrm{d}y) = \mathbb{E}_{\mathbb{P}^0}\left[e^{\beta r(X_0)} \mid X_{k+1} = x_{k+1}\right] = h_{k+1}(x_{k+1}),$$

其中中间等号使用马尔可夫性。这给出展示的 Doob 变换核，并说明倾斜路径律仍然是马尔可夫的。 □

价值函数 h_k 难以计算，而应对这一点的两条路径组织了本节余下部分。一条路径只把近似价值当作提议，并通过重加权去除其偏差，从而在极限中给出精确采样器（第 8.5 小节）；另一条路径近似价值并接受由此产生的偏差，也就是推理时 RL 视角（第 8.6 小节）。

8.5. **Feynman–Kac 修正与序贯 Monte Carlo.** 定理 8.1 给出精确引导采样器，但它的核需要价值函数 $h_k(x_k) = \mathbb{E}_{\mathbb{P}^0}[e^{\beta r(X_0)} \mid X_k = x_k]$，这与得分图景中的后验 reward-to-go 一样不可得。本小节选择上面岔路的第一条：只把近似价值用作提议，并通过重加权去除其偏差。这就是 Feynman–Kac / 序贯 Monte Carlo 路线，也是引导采样器推理时扩展背后的机制。

这个名称来自价值函数。沿基础反向过程，h 是终端泛函的条件期望，因此是一个鞅：

$$h_{k+1}(x_{k+1}) = \int h_k(y)\, P_k^{0,\leftarrow}(x_{k+1} \to \mathrm{d}y),$$

这正是定理 8.1 中出现的同一个恒等式。在连续时间中，它满足带有终端数据 $h = e^{\beta r}$ 的后向方程 $\partial_s h_s + \mathcal{L}_s h_s = 0$，其中 $\mathcal{L}_s$ 是基础反向扩散的生成元；这是第 8.3 小节中同一个倾斜因子在加噪时间方程 (8.6) 的反向时间伴随。在干净端 $k = 0$，状态就是 X_0 本身，所以终端价值精确为 $h_0(x_0) = e^{\beta r(x_0)}$。

关键事实是，Doob 变换的整个望远镜乘积会坍缩成一个单一终端权重。

命题 8.2 (非扭曲重加权)**.** *令 $\mathbb{P}^0$ 为定理 8.1 中的基础反向链，令 $\mathbb{Q}^\star$ 为满足 $\mathrm{d}\mathbb{Q}^\star/\,\mathrm{d}\mathbb{P}^0 = e^{\beta r(X_0)}/Z_\beta$ 的倾斜路径律。则在轨迹上，*

$$\frac{\mathrm{d}\mathbb{Q}^\star}{\mathrm{d}\mathbb{P}^0}(x_K, \dots, x_0) = \frac{e^{\beta r(x_0)}}{Z_\beta}, \qquad Z_\beta = \mathbb{E}_{\mathbb{P}^0} e^{\beta r(X_0)}.$$

证明. 从初始律和 Doob 核写出倾斜路径律。它的初始律为 $p_K^\star(\,\mathrm{d}x_K) = p_K^0(\,\mathrm{d}x_K)\, h_K(x_K)/Z_\beta$，其核为 $P_k^{\star,\leftarrow} = P_k^{0,\leftarrow}\, h_k(x_k)/h_{k+1}(x_{k+1})$。相乘得到

$$\frac{\mathrm{d}\mathbb{Q}^\star}{\mathrm{d}\mathbb{P}^0} = \frac{h_K(x_K)}{Z_\beta}\prod_{k=0}^{K-1}\frac{h_k(x_k)}{h_{k+1}(x_{k+1})} = \frac{h_K(x_K)}{Z_\beta}\cdot\frac{h_0(x_0)}{h_K(x_K)} = \frac{h_0(x_0)}{Z_\beta},$$

而 $h_0(x_0) = e^{\beta r(x_0)}$。 □

这提示了最简单的精确算法：从基础模型抽取 N 条轨迹，每条以 $e^{\beta r(x_0)}$ 加权，然后重采样。当 $N \to \infty$ 时，加权经验律收敛到 $\mathbb{Q}^\star$，所以干净边际收敛到倾斜目标 p_0^β。问题在于方差：如果基础模型很少产生高奖励输出，几乎全部权重会落到少数轨迹上，有效样本量会坍缩。

序贯 Monte Carlo 通过用近似价值扭曲提议，并沿途重采样来修正这一点。把基础核替换为一个遵循近似 Doob 变换的提议：

(8.13)
$$\widehat{P}_k^{\leftarrow}(x_{k+1} \to \mathrm{d}x_k) = \frac{P_k^{0,\leftarrow}(x_{k+1} \to \mathrm{d}x_k)\,\widehat{h}_k(x_k)}{\widehat{g}_{k+1}(x_{k+1})}, \qquad \widehat{g}_{k+1}(x_{k+1}) = \int \widehat{h}_k(y)\, P_k^{0,\leftarrow}(x_{k+1} \to \mathrm{d}y),$$

其中 $\widehat{h}_k \approx h_k$ 是任意可处理的价值估计（分类器、奖励模型、学习价值，或 (8.5) 的线性化小倾斜价值），并具有精确终端值 $\widehat{h}_0 = e^{\beta r}$。如果粒子从基础噪声律 p_K^0 初始化，它们还携带初始权重 $\widehat{h}_K(X_K)$；如果能改为从正比于 $p_K^0 \widehat{h}_K$ 的扭曲噪声律采样，这个初始权重就是常数。让 N 个粒子通过 $\widehat{P}_k^{\leftarrow}$ 运行，并在每个反向步携带增量重要性权重，再在有效样本量下降时重采样，就得到 $\mathbb{Q}^\star$ 的一致估计量。

在每个反向步，粒子获得一个增量重要性权重：它是“按当前价值 $\widehat{h}_k(x_k)$ 加权的基础转移”与“按下一价值 $\widehat{h}_{k+1}(x_{k+1})$ 加权的扭曲提议”之比。代入提议 (8.13) 后，这个比值坍缩为

$$\frac{P_k^{0,\leftarrow}(x_{k+1} \to \mathrm{d}x_k)\,\widehat{h}_k(x_k)}{\widehat{P}_k^{\leftarrow}(x_{k+1} \to \mathrm{d}x_k)\,\widehat{h}_{k+1}(x_{k+1})} = \frac{\widehat{g}_{k+1}(x_{k+1})}{\widehat{h}_{k+1}(x_{k+1})}.$$

因此，增量权重为

(8.14)
$$w_k(x_{k+1}) = \frac{\widehat{g}_{k+1}(x_{k+1})}{\widehat{h}_{k+1}(x_{k+1})}.$$

它度量在当前状态假定的价值 $\widehat{h}_{k+1}$ 与把 $\widehat{h}_k$ 用一个基础步传播得到的价值 $\widehat{g}_{k+1}$ 之间的一步不匹配。如果扭曲是精确的，即 $\widehat{h} = h$，则由鞅恒等式 $\widehat{g}_{k+1} = h_{k+1}$，所有增量权重都为一；如果采样器从 p_K^0 开始，唯一剩下的非常数权重是初始扭曲权重。因此，所有近似误差都被推到非均匀权重中，而重采样会在大 N 极限中去除这些误差；好的 $\widehat{h}$ 会让权重接近均匀并降低方差。

对条件生成，取终端奖励为 $\log p(y \mid x_0)$。于是 $h_k(x_k) = \mathbb{P}^0(y \mid X_k = x_k)$，而噪声分类器或条件似然模型给出可处理的 $\widehat{h}_k$。当这个扭曲实现为只移动粒子的引导梯度 $\nabla \log \widehat{h}_k$ 时，该提议正是一个分类器引导采样器，而上面的权重则是 SMC 修正，在大粒子极限中恢复条件目标。

评注. 这就是“Feynman–Kac steering”或“twisted diffusion sampler”方法族 [52]。它清晰分离了两个角色：引导提供近似价值 $\widehat{h}$，使提议偏向高奖励区域；而权重 (8.14) 与重采样保证极限律是精确倾斜，而不是有偏近似。花费更多推理时计算，也就是更多粒子、更频繁重采样或更准确的扭曲，会降低修正的方差，而不是改变它的目标。

8.6. **推理时强化学习.** 前几小节描述了奖励倾斜的理想采样器：如果后验价值函数 h_k 已知精确，反向核就会是定理 8.1 中的 Doob 变换。这里的设置是真正的推理时设置：预训练反向采样器作为参考律保持固定，而策略只改变测试时该采样器的运行方式。因此，RL 问题是在生成过程中的采样器选择上进行，而不是重新训练基础得分模型。

在有限时域表述中，反向采样器在采样时采取动作。令 $\pi = (\pi_k)$ 为可能随机的策略：在第 k 步观测到 X_{k+1} 后，它选择动作 $a_k \sim \pi_k(\cdot \mid X_{k+1})$，反向步使用相应的核

$$a_k \sim \pi_k(\cdot \mid X_{k+1}), \qquad X_k \sim P_k^{a_k,\leftarrow}(X_{k+1} \to \cdot), \qquad a_k \in \mathcal{A}.$$

动作可以是引导尺度、时间步选择、注入噪声水平、拒绝或重采样决策，或者加性漂移修正。在掩码语言扩散采样器中，它可以选择要去掩码多少 token、更新哪些位置，或从 token 后验中采样得多尖锐。终端奖励只在最终对象 X_0 产生之后度量。

目标是同一个 KL 正则化测度变换，只是被限制在推理时可用的采样器修改族中：

$$J(\pi) = \mathbb{E}_\pi \left[r(X_0) - \frac{1}{\beta} \sum_k \mathrm{D}_{\mathsf{KL}}\big(P_k^{a_k,\leftarrow}(X_{k+1} \to \cdot) \big\| P_k^{0,\leftarrow}(X_{k+1} \to \cdot)\big) \right]. \tag{8.15}$$

这里 $\beta > 0$ 是变分原理 (8.2) 中同一个逆温度；更大的 β 意味着更弱的 KL 正则化和更强的奖励追求。KL 项使策略诱导的采样器在每个反向步保持接近预训练采样器。这呼应了变分恒等式：奖励提升只有相对于参考分布才有意义。没有这种参考代价，策略优化可能利用奖励模型缺陷，或坍缩到一小组高分样本。

如果动作类足够丰富，可以选择整个下一步核，那么这个优化不会产生一个新采样器：它会恢复定理 8.1 中的 Doob 变换核，只不过现在是通过动态规划而不是通过倾斜路径律得到的。为了看到同一对象出现，从 x_k 定义正则化的价值到达函数

$$v_k(x_k) = \sup_{\pi_0,\dots,\pi_{k-1}} \mathbb{E}_\pi \left[r(X_0) - \frac{1}{\beta} \sum_{j=0}^{k-1} \mathrm{D}_{\mathsf{KL}}\big(P_j^{a_j,\leftarrow}(X_{j+1} \to \cdot) \big\| P_j^{0,\leftarrow}(X_{j+1} \to \cdot)\big) \,\middle|\, X_k = x_k \right],$$

其中空和约定给出 $v_0(x_0) = r(x_0)$。局部 Gibbs 变分原理给出

$$P_k^{\star,\leftarrow}(x_{k+1} \to \mathrm{d}x_k) \propto P_k^{0,\leftarrow}(x_{k+1} \to \mathrm{d}x_k) \exp\left(\beta v_k(x_k)\right).$$

写 $h_k(x_k) = \exp(\beta v_k(x_k))$，这就正是 (8.12) 中的归一化：我们已经得到过的同一个重加权，只是现在价值被呈现为 Bellman 的 value-to-go，而不是后验倾斜因子。两个字母是同一对象的两个坐标：乘性倾斜因子 h_k 进入 Doob 核和 SMC 权重；它在对数域中的软价值为 $v_k = \beta^{-1} \log h_k$，所以在连续图景中，$\beta$ 乘以它的梯度会被加到得分上，并匹配 $\nabla \log h_t^\beta = \beta \nabla v_t$。(8.5) 的后验期望奖励 V_t 是这个软价值在小 β 下的线性化。因此，软动态规划是路径空间倾斜在有限状态中的计算形式；这里真正新的不是最优核，而是价值如何获得，因为引导方法近似这种价值信息，而 RL 方法从采样奖励中估计它。

对带参数策略 π_θ，其反向路径密度为 $p_\theta(x_K, \dots, x_0)$，采样奖励可以通过似然比恒等式转化为更新：

$$\nabla_\theta \mathbb{E}_{(X_K,\dots,X_0)\sim p_\theta}[r(X_0)] = \mathbb{E}_{(X_K,\dots,X_0)\sim p_\theta}\left[r(X_0) \nabla_\theta \log p_\theta(X_K, \dots, X_0)\right],$$

在实现中通常会从奖励中减去一个 baseline 来降低方差，但这不是测度变换恒等式的一部分。沿一条反向扩散链，

$$\log p_\theta(x_K, \dots, x_0) = \log p_{\theta,K}(x_K) + \sum_k \log p_\theta(x_k \mid x_{k+1}),$$

当噪声初始化固定时，初始律项会从梯度中消失。因此，更新沿轨迹中的引导决策分解。对 KL 正则化目标 (8.15)，括号内的回报被替换为正则化回报；当转移族可微时，还要加入 KL 惩罚项的显式导数。

这也是解读当前推理时 RL 方法的视角。人们可以把基础采样器与奖励对齐的提议结合起来，把粒子花在重采样和轨迹修正上，或更直接地估计漂移/Doob 修正。这些都是对同一个理想受控路径律的不同近似。对这些讲义有用的信息不是某个特定近期算法已经取代了其他算法，而是奖励倾斜、KL 正则化控制和 Doob 变换为比较推理时改变采样器的方法提供了共同语言。

8.7. **离散情形中的引导与奖励倾斜.** 本节的一切几乎不加改变地转移到第 7 节的离散扩散模型中，因为目标和目标函数从未使用连续结构。奖励倾斜目标 (8.1) 与 KL 正则化变分原理 (8.2) 可定义在任意空间上；第 8.5 小节的 Feynman–Kac/扭曲 SMC 修正用同样的势函数对粒子重加权；推理时 RL (8.15) 对轨迹似然求导，在离散链中即对 $\sum_k \log P^{\leftarrow}_{k,\theta}(x_{k+1} \to x_k)$ 这样的对数转移概率和相对于 θ 求导，而不是相对于状态求导，因此可以逐字转移。

唯一改变形式的是引导。最优倾斜仍然是定理 8.1 的 Doob h-变换，只是现在作用于跳过程：令软价值 $v_t(x) = \beta^{-1} \log \mathbb{E}[\exp(\beta r(X_0)) \mid X_t = x]$，则引导反向速率由价值差重加权：

$$R_t^{\leftarrow,\mathrm{guided}}(y \to x) = R_t^{\leftarrow}(y \to x)\, \exp\big(\beta(v_t(x) - v_t(y))\big),$$

这正是向反向漂移添加 $\beta\nabla v_t$ 的类似物，其中梯度被跨可行移动的 v_t 有限差替代，也就是 (7.3) 中相同的得分到比值替换。和以前一样，v_t 难以处理，并通过奖励的学习预测器、无预测器插值，或围绕预测干净 token 的 Taylor 展开来近似 [39]。

附录 A. Itô 微积分与 Girsanov 定理

本附录汇集讲义中使用的随机微积分事实。下面的陈述写在光滑、非爆炸的设定中；在这个设定下，正文中的形式计算是有效的。关于严格假设和证明，参见 E、Li 和 Vanden-Eijnden [21]，Karatzas 和 Shreve [30]，或 Øksendal [40] 的教材。

Itô 公式、生成元与伴随. 令 X_t 求解

$$\mathrm{d}X_t = b_t(X_t)\,\mathrm{d}t + \sigma_t(X_t)\,\mathrm{d}B_t,$$

其中 B_t 是 m 维 Brownian 运动，且 $\sigma_t(x) \in \mathbb{R}^{d\times m}$。写 $a_t(x) = \sigma_t(x)\sigma_t(x)^\top$。对光滑测试函数 φ，Itô 公式给出

$$\mathrm{d}\varphi(X_t) = \langle \nabla\varphi(X_t), b_t(X_t)\rangle\,\mathrm{d}t + \frac{1}{2}\,\mathrm{Tr}\big(a_t(X_t)\nabla^2\varphi(X_t)\big)\,\mathrm{d}t + \langle \nabla\varphi(X_t), \sigma_t(X_t)\,\mathrm{d}B_t\rangle\,.$$

最后一项是鞅增量，在通常可积性假设下均值为零。因此

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}[\varphi(X_t)] = \mathbb{E}[(\mathcal{L}_t\varphi)(X_t)],$$

其中无穷小生成元为

$$\mathcal{L}_t\varphi = \langle b_t, \nabla\varphi\rangle + \frac{1}{2}\,\mathrm{Tr}(a_t\nabla^2\varphi).$$

如果 X_t 具有密度 q_t，则 Fokker–Planck 方程就是伴随方程

$$\partial_t q_t = \mathcal{L}_t^* q_t,$$

其中按坐标写为

$$\mathcal{L}_t^* q = -\nabla \cdot (b_t q) + \frac{1}{2} \sum_{i,j=1}^{d} \partial_i \partial_j ((a_t)_{ij} q)\,.$$

对过阻尼 Langevin，$b = -\nabla U$ 且 $a = 2I$，所以它化为

$$\partial_t q_t = \nabla \cdot (q_t \nabla U) + \Delta q_t.$$

Girsanov 定理. Girsanov 定理比较两个具有相同初始律和相同扩散系数、但漂移可能不同的扩散。我们固定系数 $\sqrt{2}\,I$，这是本讲义使用的 Langevin 约定。令 $\mathbb{P}$ 和 $\mathbb{Q}$ 分别为路径空间 $C([0,T];\mathbb{R}^d)$ 上如下过程的律：

$$\mathrm{d}X_t = b_t^{\mathbb{P}}(X_t)\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}B_t \qquad \text{和} \qquad \mathrm{d}X_t = b_t^{\mathbb{Q}}(X_t)\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}B_t.$$

在通常的绝对连续性和可积性假设下，两个律在 $\mathcal{F}_T$，即 Brownian 滤过的终端 σ-代数上等价，其 Radon–Nikodym 密度为

$$\left.\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{Q}}\right|_{\mathcal{F}_T} = \exp\!\left(\frac{1}{2}\int_0^T \left\langle b_t^{\mathbb{P}} - b_t^{\mathbb{Q}},\, \mathrm{d}X_t - b_t^{\mathbb{Q}}\,\mathrm{d}t \right\rangle - \frac{1}{4}\int_0^T \left\| b_t^{\mathbb{P}} - b_t^{\mathbb{Q}} \right\|^2 \mathrm{d}t\right), \tag{A.1}$$

其中所有被积函数都在 X_t 处取值。

为了得到 KL 散度，对 (A.1) 取对数，并在 $\mathbb{P}$ 下平均：

$$\mathrm{D}_{\mathsf{KL}}(\mathbb{P}\|\mathbb{Q}) = \mathbb{E}_{\mathbb{P}}\left[\log\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{Q}}\right] = \frac{1}{2}\mathbb{E}_{\mathbb{P}}\int_0^T \left\langle b_t^{\mathbb{P}} - b_t^{\mathbb{Q}},\, \mathrm{d}X_t - b_t^{\mathbb{Q}}\,\mathrm{d}t\right\rangle - \frac{1}{4}\mathbb{E}_{\mathbb{P}}\int_0^T \left\| b_t^{\mathbb{P}} - b_t^{\mathbb{Q}}\right\|^2 \mathrm{d}t.$$

第二项已经是普通路径积分期望，所以保持原样。对第一项，注意在 $\mathbb{P}$ 下，$\mathrm{d}X_t - b_t^{\mathbb{Q}}\,\mathrm{d}t = (b_t^{\mathbb{P}} - b_t^{\mathbb{Q}})\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}B_t$，并且

$$\frac{1}{2}\mathbb{E}_{\mathbb{P}}\int_0^T \left\langle b_t^{\mathbb{P}} - b_t^{\mathbb{Q}},\, \mathrm{d}X_t - b_t^{\mathbb{Q}}\,\mathrm{d}t\right\rangle = \frac{1}{2}\mathbb{E}_{\mathbb{P}}\int_0^T \left\| b_t^{\mathbb{P}} - b_t^{\mathbb{Q}}\right\|^2 \mathrm{d}t + \frac{1}{\sqrt{2}}\,\mathbb{E}_{\mathbb{P}}\int_0^T \left\langle b_t^{\mathbb{P}} - b_t^{\mathbb{Q}},\, \mathrm{d}B_t\right\rangle.$$

最后一个积分的期望为零，于是得到

$$\mathrm{D}_{\mathsf{KL}}(\mathbb{P}\|\mathbb{Q}) = \frac{1}{4}\mathbb{E}_{\mathbb{P}}\int_0^T \left\| b_t^{\mathbb{P}}(X_t) - b_t^{\mathbb{Q}}(X_t)\right\|^2 \mathrm{d}t. \tag{A.2}$$

到目前为止，两个过程共享同一个初始律。若它们从不同初始律出发，则密度 (A.1) 在 $t=0$ 还会乘上额外因子 $\mathrm{d}\,\mathrm{Law}_{\mathbb{P}}(X_0)/\,\mathrm{d}\,\mathrm{Law}_{\mathbb{Q}}(X_0)$，KL 相应地获得初始项：

$$\mathrm{D}_{\mathsf{KL}}(\mathbb{P}\|\mathbb{Q}) = \mathrm{D}_{\mathsf{KL}}\big(\mathrm{Law}_{\mathbb{P}}(X_0)\,\big\|\,\mathrm{Law}_{\mathbb{Q}}(X_0)\big) + \frac{1}{4}\mathbb{E}_{\mathbb{P}}\int_0^T \left\| b_t^{\mathbb{P}}(X_t) - b_t^{\mathbb{Q}}(X_t)\right\|^2 \mathrm{d}t.$$

引理 A.1 (数据处理)**.** 令 μ 和 ν 为某个可测空间上的概率律，令 T 是到另一个可测空间的可测映射。则

$$\mathrm{D}_{\mathsf{KL}}(T_\#\mu\|T_\#\nu) \le \mathrm{D}_{\mathsf{KL}}(\mu\|\nu) \qquad \text{和} \qquad \mathrm{D}_{\mathsf{TV}}(T_\#\mu, T_\#\nu) \le \mathrm{D}_{\mathsf{TV}}(\mu,\nu).$$

证明. 如果 μ 不关于 ν 绝对连续，KL 界是平凡的。否则令 $Z = \mathrm{d}\mu/\,\mathrm{d}\nu$。在 $T_{\#}\nu$ 下，$T_{\#}\mu$ 关于 $T_{\#}\nu$ 的 Radon–Nikodym 导数为

$$\mathbb{E}_\nu[Z \mid T].$$

因此 Jensen 不等式给出

$$\mathrm{D}_{\mathsf{KL}}(T_{\#}\mu \| T_{\#}\nu) = \mathbb{E}_\nu[\mathbb{E}_\nu[Z \mid T] \log \mathbb{E}_\nu[Z \mid T]] \le \mathbb{E}_\nu[Z \log Z] = \mathrm{D}_{\mathsf{KL}}(\mu \| \nu).$$

对总变差，在目标空间中的事件 B 上取上确界：

$$|T_{\#}\mu(B) - T_{\#}\nu(B)| = \left|\mu(T^{-1}B) - \nu(T^{-1}B)\right| \le \mathrm{D}_{\mathsf{TV}}(\mu, \nu). \qquad \square$$

因此，对路径空间上的任意可测映射 F，

$$\mathrm{D}_{\mathsf{KL}}(F_{\#}\mathbb{P}^u \| F_{\#}\mathbb{P}^0) \le \mathrm{D}_{\mathsf{KL}}(\mathbb{P}^u \| \mathbb{P}^0).$$

取 $F(\omega) = \omega_T$，就得到一步 ULA 计算中使用的端点 KL 界。

附录 B. 高斯工具箱

这个简短工具箱汇集了讲义中反复使用的事实。对最近没有处理过高斯密度的学生来说，这些都是很好的热身练习。

引理 B.1 (仿射高斯映射与和)**.** 令 $\mathbb{R}^d$ 中 $X \sim \mathcal{N}(m, \Sigma)$。对矩阵 A 和向量 b，

$$AX + b \sim \mathcal{N}\left(Am + b, A\Sigma A^\top\right).$$

如果 X 和 Y 是独立高斯，则 $X + Y$ 是高斯，并且

$$\mathrm{Cov}(X + Y) = \mathrm{Cov}(X) + \mathrm{Cov}(Y).$$

特别地，在一维中，如果 X 和 Y 独立且居中，则

$$\mathrm{Var}(aX + bY) = a^2 \,\mathrm{Var}(X) + b^2 \,\mathrm{Var}(Y).$$

证明. 仿射陈述来自高斯特征函数：

$$\mathbb{E}e^{i\langle t, AX+b\rangle} = \exp\left(i\,\langle t, Am + b\rangle - \frac{1}{2} t^\top A\Sigma A^\top t\right).$$

对独立高斯，特征函数相乘，所以均值和协方差相加。一维方差公式就是相应的特例。 $\square$

引理 B.2 (相同协方差高斯之间的 KL)**.** 对 $m, \widehat{m} \in \mathbb{R}^d$ 和正定 Σ，

$$\mathrm{D}_{\mathsf{KL}}(\mathcal{N}(m, \Sigma) \| \mathcal{N}(\widehat{m}, \Sigma)) = \frac{1}{2} \|m - \widehat{m}\|^2_{\Sigma^{-1}}, \qquad \|v\|^2_{\Sigma^{-1}} = v^\top \Sigma^{-1} v.$$

特别地，如果 $\Sigma = \eta I$，则 KL 为 $\|m - \widehat{m}\|^2/(2\eta)$。

证明. 对数密度比为

$$-\frac{1}{2}\left\|x-m\right\|_{\Sigma^{-1}}^2+\frac{1}{2}\left\|x-\widehat{m}\right\|_{\Sigma^{-1}}^2.$$

在 $x\sim\mathcal{N}(m,\Sigma)$ 下取期望，会抵消协方差项并留下 $\frac{1}{2}\left\|m-\widehat{m}\right\|_{\Sigma^{-1}}^2$。 □

引理 B.3 (高斯卷积得分)**.** 令 $X_t=X_0+\sqrt{t}Z$，其中 $Z\sim\mathcal{N}(0,I)$ 与 X_0 独立。如果 p_t 是 X_t 的密度，则

$$\nabla\log p_t(x)=\frac{1}{t}\left(\mathbb{E}[X_0\mid X_t=x]-x\right).$$

证明. 这是连续时间 Tweedie 恒等式 (3.4) 在 $a_t=1$、$\sigma_t^2=t$ 下的特例；该恒等式由对 p_t 的高斯混合表达式求导并除以 $p_t(x)$ 得到。 □

参考文献

[1] M. S. Albergo, N. M. Boffi, and E. Vanden-Eijnden. Stochastic interpolants: a unifying framework for flows and diffusions, 2023. arXiv:2303.08797.

[2] J. M. Altschuler and S. Chewi. Faster high-accuracy log-concave sampling via algorithmic warm starts. In *IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 2169–2176, 2023. arXiv:2302.10249.

[3] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient Flows in Metric Spaces and in the Space of Probability Measures*. Lectures in Mathematics ETH Zürich. Birkhäuser, second edition, 2008.

[4] N. Anari, C. Baronio, CJ Chen, A. Haqi, F. Koehler, A. Li, and T.-D. Vuong. Parallel sampling via autospeculation, 2025. arXiv:2511.07869.

[5] N. Anari, R. Gao, and A. Rubinstein. Parallel sampling via counting, 2024. arXiv:2408.09442.

[6] D. Bakry, I. Gentil, and M. Ledoux. *Analysis and Geometry of Markov Diffusion Operators*. Grundlehren der mathematischen Wissenschaften 348. Springer, 2014.

[7] R. Bauerschmidt, T. Bodineau, and B. Dagallier. Stochastic dynamics and the Polchinski equation: an introduction. *Probability Surveys*, 21:200–290, 2024. arXiv:2307.07619.

[8] C. H. L. Beentjes and R. E. Baker. Uniformisation techniques for stochastic simulation of chemical reaction networks. *The Journal of Chemical Physics*, 150:154107, 2019. arXiv:1811.00948.

[9] J. Benton, V. De Bortoli, A. Doucet, and G. Deligiannidis. Nearly d-linear convergence bounds for diffusion models via stochastic localization. In *International Conference on Learning Representations (ICLR)*, 2024. arXiv:2308.03686.

[10] A. Campbell, J. Benton, V. De Bortoli, T. Rainforth, G. Deligiannidis, and A. Doucet. A continuous time framework for discrete denoising models. In *Advances in Neural Information Processing Systems 35*, 2022. arXiv:2205.14987.

[11] F. Chen, S. Chewi, C. Daskalakis, and A. Rakhlin. High-accuracy sampling for diffusion models and log-concave distributions, 2026. arXiv:2602.01338.

[12] H. Chen, H. Lee, and J. Lu. Improved analysis of score-based generative modeling: user-friendly bounds under minimal smoothness assumptions, 2023. arXiv:2211.01916.

[13] H. Chen and L. Ying. Convergence analysis of discrete diffusion model: exact implementation through uniformization, 2024. arXiv:2402.08095.

[14] Y. Chen. An almost constant lower bound of the isoperimetric coefficient in the KLS conjecture. *Geometric and Functional Analysis*, 31:34–61, 2021. arXiv:2011.13661.

[15] Y. Chen. Computational and statistical aspects of diffusion models. Lecture notes, course 401-4634-24L, ETH Zürich, Spring 2026, 2026. `https://metaphor.ethz.ch/x/2026/fs/401-4634-24L/`.

[16] Y. Chen and R. Eldan. Localization schemes: a framework for proving mixing bounds for Markov chains, 2022. arXiv:2203.04163.

[17] Y. Chen and K. Gatmiry. A simple proof of the mixing of Metropolis-adjusted Langevin algorithm under smoothness and isoperimetry, 2023. arXiv:2304.04095.

[18] S. Chewi. Log-concave sampling. Book draft, 2026. `https://chewisinho.github.io/`.

[19] G. Conforti, A. Durmus, and M. Gentiloni Silveri. KL convergence guarantees for score diffusion models under minimal data assumptions, 2024. arXiv:2308.12240.

[20] P. Dhariwal and A. Nichol. Diffusion models beat GANs on image synthesis. In *Advances in Neural Information Processing Systems 34*, 2021. arXiv:2105.05233.

[21] W. E, T. Li, and E. Vanden-Eijnden. *Applied Stochastic Analysis*. Graduate Studies in Mathematics 199. American Mathematical Society, 2019.

[22] R. Eldan. Thin shell implies spectral gap up to polylog via a stochastic localization scheme. *Geometric and Functional Analysis*, 23:532–569, 2013. arXiv:1203.0893.

[23] Z. Geng, M. Deng, X. Bai, J. Z. Kolter, and K. He. Mean flows for one-step generative modeling, 2025. arXiv:2505.13447.

[24] D. T. Gillespie. Approximate accelerated stochastic simulation of chemically reacting systems. *The Journal of Chemical Physics*, 115:1716–1733, 2001.

[25] W. Grassmann. Transient solutions in Markovian queues. *European Journal of Operational Research*, 1(6):396–402, 1977.

[26] J. Ho, A. N. Jain, and P. Abbeel. Denoising diffusion probabilistic models. In *Advances in Neural Information Processing Systems 33*, pages 6840–6851, 2020. arXiv:2006.11239.

[27] J. Ho and T. Salimans. Classifier-free diffusion guidance. NeurIPS 2021 Workshop on Deep Generative Models, 2022. arXiv:2207.12598.

[28] E. Hoogeboom, A. A. Gritsenko, J. Bastings, B. Poole, R. van den Berg, and T. Salimans. Autoregressive diffusion models. In *International Conference on Learning Representations (ICLR)*, 2022. arXiv:2110.02037.

[29] L. P. Kadanoff. Scaling laws for Ising models near T_c. *Physics Physique Fizika*, 2:263–272, 1966.

[30] I. Karatzas and S. E. Shreve. *Brownian Motion and Stochastic Calculus*. Graduate Texts in Mathematics 113. Springer, second edition, 1991.

[31] H. Lavenant and G. Zanella. Error bounds and optimal schedules for masked diffusions with factorized approximations, 2025. arXiv:2510.25544.

[32] Y. T. Lee and S. S. Vempala. Eldan's stochastic localization and the KLS conjecture: isoperimetry, concentration and mixing, 2016. arXiv:1612.01507.

[33] Y. Liang, Y. Liang, L. Lai, and N. Shroff. Discrete diffusion models: novel analysis and new sampler guarantees, 2025. arXiv:2509.16756.

[34] Y. Lipman, R. T. Q. Chen, H. Ben-Hamu, M. Nickel, and M. Le. Flow matching for generative modeling. In *International Conference on Learning Representations (ICLR)*, 2023. arXiv:2210.02747.

[35] X. Liu, C. Gong, and Q. Liu. Flow straight and fast: learning to generate and transfer data with rectified flow, 2022. arXiv:2209.03003.

[36] A. Lou, C. Meng, and S. Ermon. Discrete diffusion modeling by estimating the ratios of the data distribution. In *International Conference on Machine Learning (ICML), PMLR 235*, 2024. arXiv:2310.16834.

[37] S. P. Meyn and R. L. Tweedie. *Markov Chains and Stochastic Stability*. Cambridge University Press, second edition, 2009.

[38] A. Montanari. Sampling, diffusions, and stochastic localization, 2023. arXiv:2305.10690.

[39] H. Nisonoff, J. Xiong, S. Allenspach, and J. Listgarten. Unlocking guidance for discrete state-space diffusion and flow models. In *International Conference on Learning Representations (ICLR)*, 2025. arXiv:2406.01572.

[40] B. Øksendal. *Stochastic Differential Equations: An Introduction with Applications*. Springer, sixth edition, 2003.

[41] J. Polchinski. Renormalization and effective lagrangians. *Nuclear Physics B*, 231:269–295, 1984.

[42] J. Song, C. Meng, and S. Ermon. Denoising diffusion implicit models. In *International Conference on Learning Representations (ICLR)*, 2021. arXiv:2010.02502.

[43] Y. Song, P. Dhariwal, M. Chen, and I. Sutskever. Consistency models. In *International Conference on Machine Learning (ICML), PMLR 202*, 2023. arXiv:2303.01469.

[44] Y. Song, J. Sohl-Dickstein, D. P. Kingma, A. Kumar, S. Ermon, and B. Poole. Score-based generative modeling through stochastic differential equations. In *International Conference on Learning Representations (ICLR)*, 2021. arXiv:2011.13456.

[45] B. Uria, I. Murray, and H. Larochelle. A deep and tractable density estimator. In *International Conference on Machine Learning (ICML)*, 2014. arXiv:1310.1757.

[46] S. S. Vempala and A. Wibisono. Rapid convergence of the unadjusted Langevin algorithm: isoperimetry suffices. In *Advances in Neural Information Processing Systems 32*, 2019. arXiv:1903.08568.

[47] C. Villani. *Optimal transport: old and new*. Springer, 2009.

[48] K. G. Wilson. Renormalization group and critical phenomena. I. Renormalization group and the Kadanoff scaling picture. *Physical Review B*, 4:3174–3183, 1971.

[49] K. G. Wilson. Renormalization group and critical phenomena. II. Phase-space cell analysis of critical behavior. *Physical Review B*, 4:3184–3205, 1971.

[50] K. G. Wilson and J. Kogut. The renormalization group and the ϵ expansion. *Physics Reports*, 12:75–199, 1974.

[51] K. Wu, S. Schmidler, and Y. Chen. Minimax mixing time of the Metropolis-adjusted Langevin algorithm for log-concave sampling. *Journal of Machine Learning Research*, 23(270):1–63, 2022. arXiv:2109.13055.

[52] L. Wu, B. L. Trippe, C. A. Naesseth, D. M. Blei, and J. P. Cunningham. Practical and asymptotically exact conditional sampling in diffusion models. In *Advances in Neural Information Processing Systems 36*, 2023. arXiv:2306.17775.

Mathematics Department, Duke University, Box 90320, Durham, NC 27705 USA.
Email address: jianfeng@math.duke.edu