

Schrodinger Bridge 逃课指南

#blog

#Generative_Model

Summary: Schrodinger Bridge 是一种典型的对粒子群体分布的随机最优控制问题，常被用于图像去雨、去雾等计算机视觉任务，或粒子群体的随机插值等科学任务中。其路径空间表述（最小化受控过程与参考过程间的 KL 散度）为我们提供了一种看待问题的视角：从分布匹配层面上，Schrodinger Bridge 等价于一个熵正则化最优传输问题；从单粒子轨迹层面上，它要求取定起点、终点后的单粒子路径是布朗桥。依赖这种视角，先进的 Flow Matching 范式被用于求解 Schrodinger Bridge 问题。然而，对于更多没有路径空间表述的随机最优控制问题，这种“分布匹配-单粒子路径”的两层次分解视角难以推广，为 Flow Matching 系列算法的设计带来了困难。本文在不利用 Schrodinger Bridge 路径空间表述的情况下，提出了一种“单粒子重整化-路径混合”手续，还原了 Schrodinger Bridge 的两层次分解，为更广泛的粒子群体随机最优控制问题的求解提供了备选方案。

Schrodinger Bridge 与 Entropy Regularized OT 的关系

考虑 Schrodinger Bridge 问题中的 KL 分解：

$$\text{KL}(P\|R) = \text{KL}(\pi\|R_{01}) + \mathbb{E}_{\pi}[\text{KL}(P(\cdot|x_0, x_1)\|R(\cdot|x_0, x_1))]$$

其中， P 是待求的路径空间的概率测度， R 是路径空间的参考测度， π 是待求的两个端点处的数据的联合分布， R_{01} 是由参考测度自然诱导出的端点处数据的联合分布。 $P(\cdot|x_0, x_1), R(\cdot|x_0, x_1)$ 是取定两端点后路径空间的概率测度。对于第二项，我们总可以通过使得 $P(\cdot|x_0, x_1)$ 与 $R(\cdot|x_0, x_1)$ 完全相同来将其变成 0，下面只研究第一项。

对于参考过程是布朗运动、初始分布为 μ_0 的情形（下文取 $\sigma(t) \equiv \sigma > 0$ ，并记 $\epsilon = \sigma^2$ ）：

$$R_{01}(x_0, x_1) = \mu_0(x_0)(2\pi\epsilon)^{-d/2} \exp\left(-\frac{\|x_0 - x_1\|^2}{2\epsilon}\right)$$

展开 KL 散度：

$$\begin{aligned} \text{KL}(\pi\|R_{01}) &= \int \pi \log \frac{\pi}{R_{01}} dx_0 dx_1 \\ &= \int \pi \log \pi dx_0 dx_1 + \frac{1}{2\epsilon} \int \pi \|x_0 - x_1\|^2 dx_0 dx_1 + C_R \end{aligned}$$

其中 $C_R = H[\mu_0] + \frac{d}{2} \log(2\pi\epsilon)$ 与 π 无关。因此，我们面对的新优化问题是：

$$\min C[\pi(x_0, x_1)] = \int \|x_0 - x_1\|^2 \pi(x_0, x_1) dx_0 dx_1 - 2\epsilon H[\pi]$$

约束条件是 $\int \pi(x_0, x_1) dx_1 = \mu_0(x_0)$ ， $\int \pi(x_0, x_1) dx_0 = \mu_1(x_1)$ 。这是标准的熵正则化 OT 问题。

单粒子最优控制策略的求解

对于单粒子，我们可给出如下最优控制问题：

$$\min_u \mathbb{E}_{X(0)=x_0} \left[\frac{1}{2} \int_0^1 \|u\|^2 dt + \Phi(x(1)) \right] \quad dx = u(X_t, t)dt + \sigma dW_t$$

由于粒子必须在 $t = 1$ 时到达给定终点 x_1 ，因此有终端代价：

$$\Phi(x(1)) = \begin{cases} 0 & X(1) = x_1 \\ +\infty & X(1) \neq x_1 \end{cases}$$

这里的硬终点约束对应无限控制代价；下文的 HJB 计算按有限终端惩罚趋于无穷的奇异极限理解。系统的 HJB 方程：

$$-\partial_t V(x, t) = \min_u \left[\frac{1}{2} \|u\|^2 + \nabla V(x, t) \cdot u + \frac{1}{2} \sigma^2 \nabla^2 V(x, t) \right]$$

求解 $\min(\cdot)$ 中内容的最小值来获得最优控制律与值函数的关系：

$$u^*(x, t) = -\nabla V(x, t)$$

回代，得到可以求解值函数的 PDE：

$$\partial_t V(x, t) - \frac{1}{2} \|\nabla V\|^2 + \frac{1}{2} \sigma^2 \nabla^2 V = 0$$

有一些方法来求解这个 HJB 方程。首先，注意到这是一个线性动力系统，目标是二次型，因此该问题是典型的 LQR 问题，不妨猜测它有二次型形式的值函数：

$$V(x, t) = \frac{1}{2} P(t) \|x - x_1\|^2 + q(t)$$

$P(t), q(t)$ 是待求函数。计算其对 t 和 x 的各阶导数并回代到 HJB 方程中：

$$\begin{aligned} \partial_t V &= \frac{1}{2} \dot{P}(t) \|x - x_1\|^2 + \dot{q}(t), \nabla V = P(t)(x - x_1), \\ \nabla^2 V &= P(t) \cdot d \left(\frac{1}{2} \dot{P}(t) \|x - x_1\|^2 + \dot{q}(t) \right) - \frac{1}{2} P^2(t) \|x - x_1\|^2 + \frac{1}{2} \sigma^2 P(t) d = 0 \end{aligned}$$

通过常数项和有关 $\|x - x_1\|^2$ 的项前的系数等于 0，得到 Ricatti 方程和另一个方程：

$$\dot{P}(t) - P^2(t) = 0, \quad \dot{q}(t) + \frac{d\sigma^2}{2} P(t) = 0 \Rightarrow P(t) = \frac{-1}{t + C}$$

在终端惩罚趋于无穷的极限下，常数 $C = -1$ ，从而立刻看出极限控制律是布朗桥：

$$dX_t = \frac{x_1 - x_t}{1 - t} dt + \sigma dW_t$$

另一个方法是执行 Cole-Hopf 变换。取：

$$V(x, t) = -\sigma^2 \log \Psi(x, t)$$

那么：

$$\begin{aligned} \partial_t V &= -\sigma^2 \frac{1}{\Psi} \partial_t \Psi \\ \nabla V &= -\sigma^2 \frac{1}{\Psi} \nabla \Psi, \quad \|\nabla V\|^2 = \sigma^4 \frac{\|\nabla \Psi\|^2}{\Psi^2} \\ \nabla^2 V &= \nabla \cdot \left(-\sigma^2 \frac{1}{\Psi} \nabla \Psi \right) = \sigma^2 \frac{\|\nabla \Psi\|^2}{\Psi^2} - \sigma^2 \frac{\nabla^2 \Psi}{\Psi} \end{aligned}$$

代入原方程化简后得到倒向热方程：

$$\partial_t \Psi + \frac{\sigma^2}{2} \nabla^2 \Psi = 0$$

归一化奇异极限下的终端条件 $\Psi(x, 1) = \delta(x - x_1)$ ，从而直接读出其解，即其格林函数：

$$\Psi(x, t) = (2\pi\sigma^2(1-t))^{-d/2} \exp \left(-\frac{\|x - x_1\|^2}{2\sigma^2(1-t)} \right)$$

进而可得到和之前（使用二次型假设）完全相同的最优控制律。

提取单粒子运行代价中的有限项

然而，直接将这个控制律代入控制目标并不会得到有限代价。尝试进行这个计算：我们需要的是

$$\frac{1}{2} \mathbb{E}_{X(0)=x_0} \left[\int_0^1 \|u_{BB}(X_t)\|^2 dt \right] = \frac{1}{2} \mathbb{E}_{X(0)=x_0} \left[\int_0^1 \frac{\|x_1 - X_t\|^2}{(1-t)^2} dt \right] dX_t = \frac{x_1 - X_t}{1-t} dt + \sigma dW_t$$

我们知道，布朗桥的均值和方差是：

$$X_t \sim \mathcal{N}((1-t)x_0 + tx_1, \sigma^2 t(1-t)I_d)$$

从而：

$$\begin{aligned} \mathbb{E}\|x_1 - X_t\|^2 &= \|x_1 - \mathbb{E}[X_t]\|^2 + \text{TrCov}(X_t) \\ &= (1-t)^2 \|x_1 - x_0\|^2 + d\sigma^2 t(1-t) \end{aligned}$$

其中 d 是数据的维度。即时代价：

$$\mathbb{E}_{X(0)=x_0} \|u_{BB}(X_t)\|^2 = \|x_1 - x_0\|^2 + d\sigma^2 \frac{t}{1-t}$$

这个东西直接积分是要发散的，注意到发散是在 $t = 1$ 附近（控制律 u_{BB} 需要抵抗恒定的扩散，将粒子拉向终点 x_1 处的 $\delta(\cdot)$ 分布，这需要无穷大的控制律）考虑通过将终点时间截断到 $1 - \delta$ 的方式把发散项提取出来：

$$\begin{aligned} C_\delta^{BB}(x_0, x_1) &= \frac{1}{2} \mathbb{E} \int_0^{1-\delta} \|u_t^{BB}(X_t)\|^2 dt \\ &= \frac{1-\delta}{2} \|x_1 - x_0\|^2 + \frac{d\sigma^2}{2} \int_0^{1-\delta} \frac{t}{1-t} dt \\ &= \frac{1-\delta}{2} \|x_1 - x_0\|^2 + \frac{d\sigma^2}{2} \int_0^{1-\delta} \left(\frac{1}{1-t} - 1 \right) dt \\ &= \frac{1}{2} \|x_1 - x_0\|^2 - \frac{d\sigma^2}{2} + \frac{\delta}{2} (d\sigma^2 - \|x_1 - x_0\|^2) + \frac{d\sigma^2}{2} \log \frac{1}{\delta} \end{aligned}$$

这里最后一项就是对数发散了。扣除对数发散项（下文中它会在混合代价中抵消）再令 $\delta \rightarrow 0$ 得到单条布朗桥代价的有限部分 $\frac{1}{2} \|x_1 - x_0\|^2 - \frac{d\sigma^2}{2}$ 。

通过混合路径获得熵项

实际上，最优控制律 $u(X_t, t)$ 并不会将一个粒子移动到一个预先指定的固定终点。上面的发散在固定端点问题中是真实的，但在混合代价有限时可以被抵消。此外，上面这种取定端点的研究方式无法找出熵正则化项。要想做到这两件事，我们需要混合通向不同终点的布朗桥。考虑从 x_0 处出发的粒子，记 Transport Plan $\pi(x_0, x_1)$ ，那么，传播子是 $q(x_1|x_0) = \frac{\pi(x_0, x_1)}{\mu_0(x_0)}$ 。如果将从 x_0 出发的布朗桥进行混合，则这些布朗桥的密度：

$$p_t(x|x_0) = \int q(x_1|x_0) \rho_t^{x_0, x_1}(x) dx_1$$

$\rho_t^{x_0, x_1}$ 代表一条布朗桥的概率密度。取定当前位置 $X_t = x$ ，终点后验：

$$\omega_t(x_1|x, x_0) = \frac{q(x_1|x_0) \rho_t^{x_0, x_1}(x)}{p_t(x|x_0)}$$

使用这个概率对控制律进行加权，得到只对起点取条件的控制律：

$$u_\pi(t, x|x_0) = \int \frac{x_1 - x}{1-t} \omega_t(x_1|x, x_0) dx_1$$

这个控制律可以正确地把从 x_0 出发的粒子运送到终点分布 $q(\cdot | x_0)$ ；再对起点按 μ_0 混合，得到终点分布 μ_1 。注意到：

$$\begin{aligned}
\int \omega_t(x_1|x, x_0) \nabla_x \log \rho_t^{x_0, x_1}(x) dx_1 &= \int \frac{q(x_1|x_0) \rho_t^{x_0, x_1}(x)}{p_t(x|x_0)} \nabla_x \log \rho_t^{x_0, x_1}(x) dx_1 \\
&= \int \frac{q(x_1|x_0) \nabla_x \rho_t^{x_0, x_1}(x)}{p_t(x|x_0)} dx_1 \\
&= \frac{\nabla_x p_t(x|x_0)}{p_t(x|x_0)} \\
&= \nabla_x \log p_t(x|x_0)
\end{aligned}$$

并且直接计算得到：

$$\nabla_x \log \rho_t^{x_0, x_1}(x) = -\frac{x - (1-t)x_0 - tx_1}{\sigma^2 t(1-t)}$$

从而推出：

$$\frac{x_1 - x}{1-t} = \frac{x - x_0}{t} + \sigma^2 \nabla_x \log \rho_t^{x_0, x_1}(x)$$

也就是：

$$u_\pi(t, x|x_0) = \frac{x - x_0}{t} + \sigma^2 \nabla_x \log p_t(x|x_0)$$

此时，再考虑计算从 x_0 出发的一束到达不同终点的轨迹的控制代价：

$$\begin{aligned}
\mathbb{E}_{X(0)=x_0} \|u_{BB}(X_t)\|^2 &= \mathbb{E}_{X(0)=x_0} \|u_\pi(t, X_t|X_0)\|^2 + \mathbb{E}_{X(0)=x_0} \|u_{BB}(X_t) - u_\pi(t, X_t|X_0)\|^2 \\
dX_t &= u_\pi(t, X_t|X_0)dt + \sigma dW_t
\end{aligned}$$

我们需要拿到的是第二项，我们看看其余两项是什么东西。首先，第一项：

$$\begin{aligned}
\frac{1}{2} \int_0^{1-\delta} \mathbb{E}_{X(0)=x_0} \|u_{BB}(t)\|^2 dt &= \frac{1}{2} \int_0^{1-\delta} dt \int q(x_1|x_0) \mathbb{E}_{X(0)=x_0, X(1)=x_1} \|u_{BB}(t)\|^2 \\
&= \mathbb{E}_{q(x_1|x_0)} [C_\delta^{BB}(X_0, X_1)]
\end{aligned}$$

就是上文中求出的发散的布朗桥代价。通过直接计算，第三项可以写成：

$$u_{BB}(X_t) - u_\pi(t, X_t|X_0) = \sigma^2 [\nabla \log \rho_t^{X_0, X_1}(X_t) - \nabla \log p_t(X_t|X_0)]$$

继续计算这个差的平方项（以 III 代表这一项）：

$$\text{III} = \sigma^4 \int dx_1 q(x_1|x_0) \int dx \rho_t^{x_0, x_1}(x) (\|\nabla \log \rho_t^{x_0, x_1}(x)\|^2 + \|\nabla \log p_t(x|x_0)\|^2 - 2 \nabla \log \rho_t^{x_0, x_1}(x) \cdot \nabla \log p_t(x|x_0))$$

重点看交叉项变成了什么：

$$\begin{aligned}
&\mathbb{E}_{X(0)=x_0} \left[\nabla \log \rho_t^{x_0, X_1}(X_t) \cdot \nabla \log p_t(X_t|x_0) \right] \\
&= \int dx \int dx_1 q(x_1|x_0) \rho_t^{x_0, x_1}(x) \nabla \log \rho_t^{x_0, x_1}(x) \cdot \nabla \log p_t(x|x_0) \\
&= \int dx \int dx_1 q(x_1|x_0) \nabla \rho_t^{x_0, x_1}(x) \cdot \nabla \log p_t(x|x_0) \\
&= \int dx \nabla p_t(x|x_0) \cdot \nabla \log p_t(x|x_0) \\
&= \int dx p_t(x|x_0) \|\nabla \log p_t(x|x_0)\|^2
\end{aligned}$$

所以我们知道这一项的计算结果：

$$\begin{aligned}\text{III} &= \sigma^4 \left[\int dx_1 q(x_1|x_0) \int dx \rho_t^{x_0, x_1}(x) \|\nabla \log \rho_t^{x_0, x_1}(x)\|^2 - \int dx p_t(x|x_0) \|\nabla \log p_t(x|x_0)\|^2 \right] \\ &= \frac{d\sigma^2}{t(1-t)} - \sigma^4 \int p_t(x|x_0) \|\nabla \log p_t(x|x_0)\|^2 dx\end{aligned}$$

考虑将这一项与数据的熵

$$H[p] = - \int p(x) \log p(x) dx$$

建立联系。考虑从 x_0 出发的传播子的熵的演化：

$$\begin{aligned}\frac{d}{dt} H[p_t(x|x_0)] &= - \int (\partial_t p_t)(1 + \log p_t) dx \\ &= - \int (-\nabla \cdot (p_t u_\pi) + \frac{\sigma^2}{2} \nabla^2 p_t)(1 + \log p_t) dx\end{aligned}$$

处理一下里面的各种东西： $\int \nabla \cdot (p_t u_\pi)$ 可以用分部积分扔掉；第二项用分部积分把梯度符号转移到 $\log p_t$ 上，获得 $-\int p_t u_\pi \cdot \nabla \log p_t dx$ ，第三项 $\int \nabla^2 p_t$ 同样可以分部积分扔掉；第四项分部积分转移一个梯度符号，得到 $\int \nabla p_t \cdot \nabla \log p_t = \int p_t \|\nabla \log p_t\|^2$ ，从而继续上面的计算：

$$\begin{aligned}\frac{d}{dt} H[p_t(x|x_0)] &= - \int p_t u_\pi \cdot \nabla \log p_t dx + \frac{\sigma^2}{2} \int p_t(x|x_0) \|\nabla \log p_t\|^2 dx \\ &= - \frac{1}{t} \int (x - x_0) \cdot \nabla p_t dx - \frac{\sigma^2}{2} \int p_t(x|x_0) \|\nabla \log p_t\|^2 dx \\ &= \frac{d}{dt} - \frac{\sigma^2}{2} \int p_t(x|x_0) \|\nabla \log p_t\|^2 dx\end{aligned}$$

另外，对于布朗桥有：

$$G(t) := H[\rho_t^{x_0, x_1}] = \frac{d}{2} \log(2\pi e \sigma^2 t(1-t)) \Rightarrow G'(t) = \frac{d}{2} \left(\frac{1}{t} - \frac{1}{1-t} \right)$$

从而，刚才的 $\mathbb{E} \|u_{BB}(X_t) - u_\pi(t, X_t|X_0)\|^2$ 其实可以写成混合路径和布朗桥的熵差，直接计算得到：

$$\frac{d}{dt} [H - G] = \frac{1}{2\sigma^2} \mathbb{E}_{X(0)=x_0} \|u_{BB}(X_t) - u_\pi(t, X_t|X_0)\|^2$$

从而这一项（这里使用了 $t \downarrow 0$ 时 $H[p_t(\cdot|x_0)] - G(t) \rightarrow 0$ ，并假设相应熵极限及分部积分成立）

$$\frac{1}{2} \int_0^{1-\delta} \text{III}(t; x_0) dt = \sigma^2 H(X_{1-\delta}|X_0 = x_0) - \frac{d\sigma^2}{2} \log(2\pi e \sigma^2 (1-\delta)) + \frac{d\sigma^2}{2} \log \frac{1}{\delta}$$

这里，我们遇到了与从条件住 x_0, x_1 的控制代价中相同的发散项。两者相减，得到控制代价（保留与 coupling 无关的常数，并取 $\delta \rightarrow 0$ ）：

$$\frac{1}{2} \int_0^1 \mathbb{E}_{X(0)=x_0} \|u_\pi(t, X_t|X_0)\|^2 dt = \frac{1}{2} \int dx_1 q(x_1|x_0) \|x_1 - x_0\|^2 - \sigma^2 H(X_1|X_0 = x_0) + C_\sigma$$

其中 $C_\sigma = \frac{d\sigma^2}{2} \log(2\pi\sigma^2)$ 。最后，注意到：

$$\begin{aligned}H(X_1|X_0 = x_0) &= - \int dx_1 q(x_1|x_0) \log q(x_1|x_0) \\ &= - \int dx_1 \frac{\pi(x_0, x_1)}{\mu_0(x_0)} [\log \pi(x_0, x_1) - \log \mu_0(x_0)]\end{aligned}$$

两侧再对 x_0 取期望：

$$-\mathbb{E}_{x_0 \sim \mu_0} H(X_1 | X_0 = x_0) = \int dx_0 dx_1 \pi(x_0, x_1) \log \pi(x_0, x_1) + H[\mu_0] = -H[\pi(x_0, x_1)] + H[\mu_0]$$

所以再对 x_0 取期望：

$$\begin{aligned} \frac{1}{2} \int_0^1 \mathbb{E}_{X_0 \sim \mu_0} \|u_\pi(t, X_t | X_0)\|^2 dt &= \frac{1}{2} \int \mu_0(x_0) dx_0 \int_0^1 \mathbb{E}_{X(0)=x_0} \|u_\pi(t, X_t | X_0)\|^2 dt \\ &= \frac{1}{2} \int dx_0 dx_1 \pi(x_0, x_1) \|x_1 - x_0\|^2 - \sigma^2 H[\pi(x_0, x_1)] + C \end{aligned}$$

其中 $C = \sigma^2 H[\mu_0] + C_\sigma$ ，优化 coupling 时可以省略。

注意，在这里，我们得到的并不是原问题中的最优控制代价 $\frac{1}{2} \int \mathbb{E} \|u(t, X_t)\|^2$ ，如果你希望得到原问题的控制代价，还需要模仿上文中的过程对起点进行一次混合。但是我们发现：对终点进行混合后，**我们已经得到了正确的熵正则化最优传输问题**，这意味着，再进行一次起点的混合，我们其实会得到一个错误的结果（这里的“错误”是针对 Flow Matching 算法设计来说的，应该使用第一次混合后的结果优化 Coupling）。

证明终点混合之后的控制代价是一个正确的 Surrogate Cost

上面，我们只对终点进行了混合，得到了控制律 $u_\pi(t, x | x_0)$ ，但是 Schrodinger Bridge 的原问题要求我们给出一个不依赖于起点的控制律 $\bar{u}_\pi(t, x)$ 。我们继续完成这个混合，并且说明上面对终点混合后得到正确控制代价的巧合是如何出现的。考虑：

$$p_t(x) = \int \mu_0(x_0) p_t(x | x_0) dx_0$$

给定位置 x ，起点的后验：

$$\kappa_t(x_0 | x) = \frac{\mu_0(x_0) p_t(x | x_0)}{p_t(x)}$$

对起点混合后的速度：

$$\bar{u}_\pi(t, x) = \int \kappa_t(x_0 | x) u_\pi(x, t | x_0) dx_0$$

对起点进行混合后，瞬时代价的降低：

$$\mathbb{E}_{x_0 \sim \mu_0, x \sim p_t(x | x_0)} \|u_\pi(t, x | x_0)\|^2 - \mathbb{E}_{x \sim p_t(x)} \|\bar{u}_\pi(t, x)\|^2 = \mathbb{E}_{x_0 \sim \mu_0, x \sim p_t(x | x_0)} \|u_\pi(t, x | x_0) - \bar{u}_\pi(t, x)\|^2$$

下面我们要说明一个重要性质：在 π 刚好取得上述熵正则化最优传输问题的最优解 π_* 时， $u_\pi(t, x | x_0)$ 对起点的依赖刚好消失，也就是说此时使用混合起点的控制律或是不混合起点的控制律会得到相同代价。

具体而言，考虑乘子法求解上述熵正则化最优传输问题：

$$\begin{aligned} \mathcal{J} &= \int \left(\frac{1}{2} \|x_0 - x_1\|^2 \pi + \sigma^2 \pi \log \pi \right) dx_0 dx_1 \\ &\quad - \int dx_0 \lambda_0(x_0) \left(\int dx_1 \pi(x_0, x_1) - \mu_0(x_0) \right) - \int dx_1 \lambda_1(x_1) \left(\int dx_0 \pi(x_0, x_1) - \mu_1(x_1) \right) \\ &= \int \left(\frac{1}{2} \|x_0 - x_1\|^2 \pi + \sigma^2 \pi \log \pi - \lambda_0(x_0) \pi - \lambda_1(x_1) \pi \right) dx_0 dx_1 + \dots \end{aligned}$$

对 Transport Plan π 变分，得到：

$$\frac{\delta \mathcal{J}}{\delta \pi} = \frac{1}{2} \|x_0 - x_1\|^2 + \sigma^2 \log \pi + \sigma^2 - \lambda_0(x_0) - \lambda_1(x_1)$$

推出 Optimal Transport Plan 满足：

$$\pi_{\star} = \exp \left(\frac{\lambda_0(x_0) + \lambda_1(x_1)}{\sigma^2} - 1 - \frac{\|x_1 - x_0\|^2}{2\sigma^2} \right) := A(x_0)k_1(x_0, x_1)B(x_1)$$

其中

$$k_s(x, y) = (2\pi\sigma^2 s)^{-d/2} \exp \left(-\frac{\|y - x\|^2}{2\sigma^2 s} \right)$$

是布朗运动传播子。为求出 $u_{\pi}(x, t|x_0)$ ，先求出 $p_t(x|x_0)$ 。考虑：

$$\begin{aligned} q_{\star}(x_1|x_0) &= \frac{\pi_{\star}(x_0, x_1)}{\mu_0(x_0)} \\ &= \frac{A(x_0)k_1(x_0, x_1)B(x_1)}{\mu_0(x_0)} \\ &= \frac{A(x_0)k_1(x_0, x_1)B(x_1)}{A(x_0) \int k_1(x_0, x_1)B(x_1)dx_1} \end{aligned}$$

布朗桥的密度可以写成：

$$\rho_t^{x_0, x_1} = \frac{k_t(x_0, x)k_{1-t}(x, x_1)}{k_1(x_0, x_1)}$$

那么：

$$\begin{aligned} p_t(x|x_0) &= \int q_{\star}(x_1|x_0)\rho_t^{x_0, x_1}(x)dx_1 \\ &= \frac{1}{\int k_1(x_0, x_1)B(x_1)dx_1} \int k_1(x_0, x_1)B(x_1) \frac{k_t(x_0, x)k_{1-t}(x, x_1)}{k_1(x_0, x_1)} dx_1 \\ &= \frac{k_t(x_0, x)}{\int k_1(x_0, x_1)B(x_1)dx_1} \int k_{1-t}(x, x_1)B(x_1)dx_1 \end{aligned}$$

考虑到上文 $p_t(x|x_0)$ 和 $u_{\pi}(t, x|x_0)$ 的关系：

$$\begin{aligned} u_{\pi}(t, x|x_0) &= \frac{x - x_0}{t} + \sigma^2 \nabla_x \log p_t(x|x_0) \\ &= \frac{x - x_0}{t} + \sigma^2 \nabla_x \log k_t(x_0, x) + \sigma^2 \nabla_x \log \left(\int k_{1-t}(x, x_1)B(x_1)dx_1 \right) \\ &= \sigma^2 \nabla_x \log \left(\int k_{1-t}(x, x_1)B(x_1)dx_1 \right) \end{aligned}$$

于是，我们得到了一个重要的结论：**在最优 Coupling $\pi_{\star}$ 处，对起点取条件的控制律 $u_{\pi}(t, x|x_0)$ 的起点依赖自动取消。因此，此时对起点混合也不会引起控制代价的下降。**这意味着我们对终点混合得到的控制代价是一个对真正控制代价的 Surrogate Model：除了在最优 Coupling $\pi_{\star}$ + 布朗桥诱导的控制律处之外，对终点混合得到的控制代价都比真正的控制代价高，因此优化上文中对终点混合的控制代价，就会得到最优 Coupling $\pi_{\star}$ 和最优控制代价！

为什么会出现这个现象

我们发现了奇妙的现象：只混合终点，我们得到了与 KL 分解相同的结果：将 Schrodinger Bridge 问题表述成了熵正则化最优传输问题；而混合两次之后控制代价再次下降，我们无法得到正确的结果。这是为什么呢？直观上，第一次混合只是允许起点到达不同的终点，从而消除“将粒子的末时刻分布强行拉到 $\delta(\cdot)$ 分布”这一行为带来的无穷大控制代价，但是如果取定终点，控制律又回到布朗桥。然而第二次混合后，如果取定起点和终点，控制律一般不再是布朗桥（除非 Coupling 取为最优 Coupling $\pi_{\star}$ ，此时取定起点的控制律自动取消对起点的依赖）。以上说法可以通过对第一次混合后的控制律做 Doob's h Transform 来验证。