arXiv:2305.03560v1 [math.PR] 5 May 2023

祖先过程转移概率公式的反例

SYLVAIN RUBENTHALER

摘要。我们研究加权粒子系统，其中新世代按与当前粒子权重成正比的概率进行重新抽样。这涵盖了广泛应用于应用统计学的一类序贯蒙特卡洛方法。我们考察嵌入此类粒子系统中的谱系树。时间反向时，粒子系统会诱导出一族取值于划分的过程（谱系树叶节点上的划分）。本文旨在为一个广为人知的公式提供反例，该公式描述了这一过程的转移概率。

1. 引言

1.1. 问题描述。 我们考虑相互作用粒子系统（IPS）。这是一大类随机模型，用于描述物理学、工程学、生物学和金融学等学科中的现象。重要的例子包括粒子滤波器、粒子方法和序贯蒙特卡洛（SMC），它们在非线性滤波的数值逼近方案以及 Feynman-Kac流的平均场近似中都占有重要地位。有关更多背景知识，请读者参阅 [Del04], [CR11]。

这些方法的核心是离散时间下演化的加权粒子系统。粒子之间的相关性源于重采样：这是一种随机选择机制，通常权重较高的粒子会被复制，而权重较低的粒子则会消失，由此形成内嵌的系谱。许多论文都研究了这种系谱。标准做法是将时间倒转：从 n 个粒子开始，它们有一定数量的祖先（少于 n），依此类推。如此继续，直到所有粒子的第一个共同祖先。每一步的谱系结构都可以用起始时的 n 个粒子上的一个划分来表示：也就是说，在给定步骤中，两个粒子若具有相同的祖先，就属于同一个块。有些作者依据转移概率公式来描述划分上的祖先过程（[Möh98]，中的公式 (2)／ [KJJS20] 中的公式 (1)）。本文的贡献是在 SMC 情形下给出这一公式的一个反例。

1.2. 数学模型的描述。 我们考虑 [KJJS20] 中所描述模型的一个子情形。设有一个可测空间 $(E,\mathcal{E})$、一个定义在 $(E,\mathcal{E})$ 上的概率测度 μ、一个马尔可夫转移 K：$E\times\mathcal{E}\to(0,\infty)$，以及一个非负势函数 g：$E\to(0,\infty)$。它们分别对应于相互作用粒子系统的状态空间、初始提议分布、转移核和重要性权重函数。

设 $\zeta_t^{(N)}:=\{(w_t^{(i)},X_t^{(i)})\}_{i=1}^N$ 是时刻 t 下 $\mathbb{N}$ 中的一个加权 N-粒子系统，其中每个 $X_t^{(i)}\in E$，且权重 $w_t^{(i)}$ 非负并满足 $\sum_{i=1}^N w_t^{(i)}=1$。设 S 是一个重采样算子，它作用于 $\zeta_t^{(N)}$，为每个粒子分配一个随机的子代数量。子代总数固定为 N，粒子 $i\in\{1,\dots,N\}$ 的子代平均数量为 $Nw_t^{(i)}$。所有子代均被赋予相同的权重 $1/N$。具体而言

$$S\zeta_t^{(N)}=\{(N^{-1},X_t^{(a_t^{(i)})})\}_{i=1}^N,$$

其中 $a_t^{(i)}=j$ if j in $\zeta_t^{(N)}$ 是 $S\zeta_t^{(N)}$ 中 i 的父粒子。权重较低的粒子会因没有后代而被随机移除，而权重较高的粒子往往会有许多子代。

日期：2023年5月8日。

从时刻 t 到时刻 $t+1$ 的这一步通过将 N-元组 $S\zeta_t^{(N)}$ 中的每个粒子独立地经由转移核 K 传播，得到粒子位置 $X_{t+1}^{(i)}$ $\sim K(X^{(a_t^{(i)})}t$， 。)。最后，根据粒子所在位置处的势函数 g 取值，为每个粒子 $i \in \{1,\dots,N\}$ 赋予一个与之成比例的权重，因此完整的更新为

$$\zeta_t^{(N)} \overset{S}{\mapsto} \{N^{-1}, X_t^{(a_t^{(i)})}\}_{i=1}^N \overset{K}{\mapsto} \left\{ \frac{g(X_{t+1}^{(i)})}{\sum_{j=1}^N g(X_{t+1}^{(j)})}, X_{t+1}^{(i)} \right\}_{i=1}^N .$$

具体说 明见算法1。算法1中嵌入了一个系谱。考虑

算法1 相互作用粒子系统的模拟

```
for i ∈ {1,...,N} do
    Sample X_0^(i) ~ μ.
for i ∈ {1,...,N} do
    Set w_0^(i) ← g(X_0^(i)) / (g(X_0^(1)) + ··· + g(X_0^(N))).
for t ∈ {0,...,T−1} do
    Sample (a_t^(1),...,a_t^(N)) ~ Categorical(w_t^(1),...,w_t^(N)).
    for i ∈ {1,...,N} do
        Sample X_{t+1}^(i) ~ K(X_t^(a_t^(i)), .).
    for i ∈ {1,...,N} do
        Set w_{t+1}^(i) ← g(X_{t+1}^(i)) / (g(X_{t+1}^(1)) + ··· + g(X_{t+1}^(N)))
```

$\zeta_t^{(N)}$ 在固定时刻 t。沿时间向后追踪祖先索引（$a_t^{(1)}$， ....., $a_t^{(N)}$），会得到一个由谱系构成的聚并森林。当到达所有粒子的最近共同祖先（MRCA）时，这片森林便形成一棵树，前提是这一情况发生在到达初始时刻 0 之前。在 [KJJS20], 中，作者证明，在一定条件下，利用适当的时间重标度后，依赖于有限个叶节点（n）的系谱函数泛函会收敛到相应的 Kingman n合并 [Kin82] 泛函，且这种收敛是在 $N \to \infty$ 有限维分布意义下的。特别是，他们证明，从叶节点到最近共同祖先的预期时间步数，对于任意有限数量的叶节点，都随 N 线性增长。其他论文则用更简单的方法估算从叶节点到最近共同祖先所需的时间步数（例如： [DMPR10, JMR15]）。

本文其余部分安排如下。第 2 节介绍系谱过程及其转移公式。第 3 节给出我们的反例。主要结果为定理 3.1 和推论 3.2。本节最后总结所用的记号。

令 $(x)_b := x(x-1)\dots(x-b+1)$ 为下降阶乘。我们采用约定 $\sum_\emptyset = 0, \prod_\emptyset = 1$。对于整数 $n \in \mathbb{N}$，定义 $[n] := \{1,2,\dots,n\}$，其中 $[0] := \emptyset$。对于划分 ξ， $|\xi|$表示 ξ中的块数。

2. 系谱过程及其转移公式

为方便起见，我们用逆时间表示相互作用粒子系统：将算法1中的初始时刻记为 T，将终止时刻记为 0；并用一族以划分为值的过程 $(G_t^{(n,N)})_{t=0}^T$ 来描述系谱，该过程由 $n \le N$ 索引，其中 n 表示在一个含有 N 个粒子的系统中观测到的叶节点（时刻 0 的粒子）数量。过程 $(G_t^{(n,N)})_{t=0}^T$ 根据底层相互作用粒子系统定义，其初始条件为 $G_0^{(n,N)} = \{\{1\},\dots,\{n\}\}$，其动力学则由以下要求驱动：如果叶节点 i 和 j 在时刻 t 有共同祖先，那么 $i \neq j \in [n]$ 必须属于 $G_t^{(n,N)}$ 中的同一个块。

如 [KJJS20], 所述，系谱过程 $(G_t^{(n,N)})_{t=0}^T$ 在一个追踪观测粒子祖先关系但不追踪其状态的空间上演化。该过程是 [DM01], 历史过程时间反演的投影，其中粒子位置已被边缘化。这意味着，$(G_t^{(n,N)})_{t=0}^T$ 通常不具有马尔可夫性。

令 $\nu_t^{(i)}$ 表示粒子 i 在时间 t 拥有的子代数量（时间为 $t-1$，其中$\boldsymbol{\nu}_t = (\nu_t^{(i)})_{1\le i\le N}$）。令 $\boldsymbol{a}_t = (a_t^{(i)})_{1\le i\le N}$ 对所有 t 均成立。我们采用与 [KJJS20] 相同的基本假设（“基本假设”，第 5 页）。

假设 1. 给定子代数量 $\boldsymbol{a}_t|\boldsymbol{\nu}_t$ 后，亲代索引的条件分布在所有满足 $\nu_t^{(i)} = \#\{j \in [N] : a_t^{(j)} = i\}$（对每个 $i \in [N]$）的向量上均匀分布。

备注 2.1. 上述假设涉及不含粒子位置的亲代分配边缘分布。该假设的一个充分条件是算法 2 第 6 行中的分类变量具有可交换性（确实如此）。

在 [Möh98]，中，还有以下附加假设（第 439 页）。

假设 2。子代数量（$\boldsymbol{\nu}_0$、 $\boldsymbol{\nu}_1$、.....）相互独立。

我们不作此假设，因为它对于我们的相互作用粒子系统而言很可能不成立。

设 ξ 和 η 是 $[n]$ 的分划，各块按块中最小元素排序，并且 η 由 ξ 通过合并若干块的子集得到。对于 $i \in [|\eta|]$，设 b_i 为 ξ中合并后形成 η中的块 i的块数，因此 $b_1 + \cdots + b_{|\eta|} = |\xi|$，并定义

$$p_{\xi,\eta}(t) := \frac{1}{(N)_{|\xi|}} \sum_{\substack{i_1 \neq \ldots \neq i_{|\eta|} = 1 \\ \text{all distincts}}} (\nu_t^{(i_1)})_{b_1} \ldots (\nu_t^{(i_{|\nu|})})_{b_{|\nu|}} \tag{2.1}$$

（这是 [KJJS20]中的公式（1）以及 [Möh98]中的公式（2））。 [KJJS20 和 Möh98] 声称，上式是在给定 $\boldsymbol{\nu}_t$（记号中省略）的情况下，从时刻 $t-1$ 的状态 ξ 转移到时刻 t的状态 η 的转移概率。两篇文章（[KJJS20、Möh98]）给出的论证相同：我们可以把从第 $t-1$ 世代到第 t 世代的过程看作一项实验，将 $|\xi|$ 个球（子代）投进 $|\nu|$ 个盒子，使得实验结束时，第 i个盒子中恰好有 $\nu_t^{(i)}$ 个球。

[KJJS20 中的这套推理有一个隐含假设，Möh98] ；我们将在这里揭示它。假设在时刻 $t-1$有一个划分 ξ。在图 2.1 中，我们展示了粒子过程的谱系结构。划分 ξ 由三个粒子表示（对于每一个粒子，

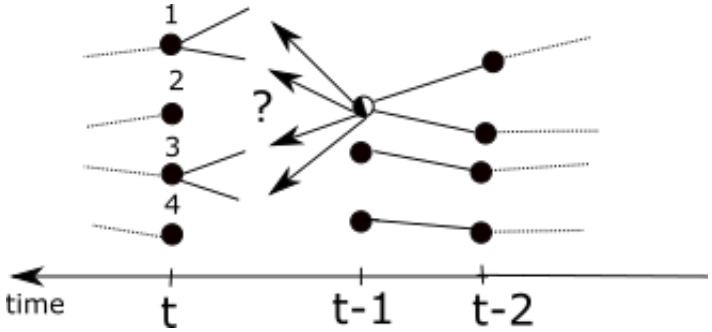

图 2.1. 祖先的选择

它在时间 0 的后代构成一个块 ξ）。我们想在已知 $\boldsymbol{\nu}_t$、已知 ξ 的条件下进行计算。在我们的例子中，我们知道时间 t 有两个粒子各有两个子代，时间 t 有两个粒子没有子代。在图中，这些由“悬垂”分支表示。现在，假设我们先为黑白相间的粒子选择一个父代（它是构成块 ξ 的叶节点的共同祖先）。我们必须选择连接到其中一条悬垂分支，才能得到一个父代。在 [KJJS20 中，Möh98]，假设选择从粒子 1 出发的某一条分支的概率为

$$\frac{\nu_t^{(1)}}{N} \tag{2.2}$$

（我们图中的 $= 2/4$ ）。这使得证明方程 (2.1) 成为可能。我们的目标是表明，在假设 1 下，这一基本假设并不成立。

3. 反例

3.1. 一个简单情形下的系谱过程。 我们关注一个非常简单的相互作用粒子系统。状态空间由两点组成： a 和 b。势函数 g 满足 $g(a) = p_a > 0$, $g(b) = p_b > 0$, 其中

$$p_b < p_a .$$

终止时刻为 $T = 2$ （起始时刻为 0）。概率分布 μ 满足 $\mu(\{a\}) = \alpha \in (0,1), \mu(\{b\}) = 1 - \alpha$。核 K 满足 $K(\{a\},\{a\}) = K(\{b\},\{b\}) = 1$。时刻 T 的子代数满足

$$\nu_T^{(1)} = 2 .$$

请记住，这个终止时刻 T 是相互作用粒子系统（IPS）的起始时刻。我们假设 $t = 2$。我们假设 $\xi = \{\{1,2\},\dots\}$。这一情形总结于图 3.1 中。每个黑点代表一个

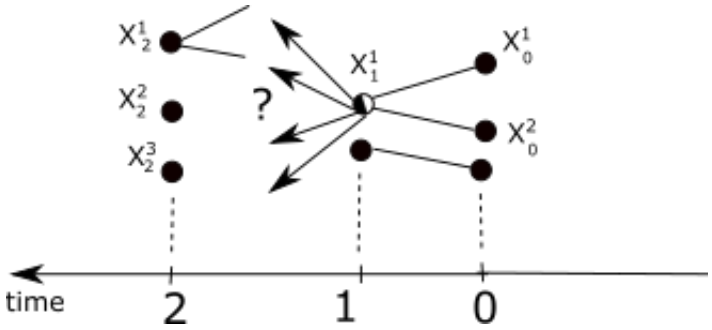

图 3.1. 简单情况

粒子。对于所有 i, s, $X_s^{(i)}$，粒子在时间 s 的位置为 i ，我们将该粒子记为$i,\ s$。在图中，作为粒子 1，0 和 2，0 的父粒子是粒子 1，1。

在下文中，我们将用 $\boldsymbol{\nu}$ 表示 $\boldsymbol{\nu}_2 = (\nu_2^{(1)},\ \nu_2^{(2)}、\dots\dots、\ \nu_2^{(N)})$，即时间 2 的子代数。这样可以让记号简洁一些。

3.2. 反例的思路。 我们考虑粒子 1， 1。我们知道 $\xi = \{\{1,2\},\dots\}$，1， 1 是 1， 0 和 2， 0 的父粒子，并且 $\nu_2^{(1)} = 2$。根据这些信息（“1，1 有很多子代”），且由于 $p_a > p_b$， $X_1^{(1)}$ 更可能等于 a 而非 b。但只有当 1，1 的父粒子位于位置 a时，它才可能等于a 。因此，悬挂分支的选择（这将决定父粒子的选择）在所有这些悬挂分支中可能并非均匀的。

这种启发式方法相当不可靠，也可以用同类论据为相反的观点辩护。因此，我们在下一节采用精确计算。

3.3. **主要结果。** 我们设

$$f : x \in \mathbb{R} \mapsto \frac{x^2}{2} e^{-x}$$

$$q_a = \frac{p_a}{\alpha p_a + (1-\alpha) p_b},\ q_b = \frac{p_b}{\alpha p_a + (1-\alpha) p_b},$$

$$q'_a = \frac{p_a}{\alpha p_a q_a + (1-\alpha) p_b q_b},\ q'_b = \frac{p_b}{\alpha p_a q_a + (1-\alpha) p_b q_b} .$$

定理 3.1。 我们有 （当 n 固定时）

$$\liminf_{N \to +\infty} \frac{N}{2} \times \mathbb{P}(a_2^{(1)} = 1 | \nu_2^{(1)} = 2, \nu_1^{(1)} = 2) \tag{3.1}$$

$$\geq \underbrace{\frac{1}{\alpha q_a} \times \frac{\alpha f(q_a)}{\alpha f(q_a) + (1-\alpha) f(q_a)} \times \frac{\alpha q_a f(q'_a)}{\alpha q_a f(q'_a) + (1-\alpha) q_b f(q'_b)}}_{=: R(\alpha, p_a, p_b)} .$$

（我们提醒读者， $a_2^{(1)}$ 是粒子 j的父代索引，1。）

定理 *3.1* 的证明。我们有

$$\begin{aligned}(3.2)\quad & \mathbb{P}(a_2^{(1)} = 1 | \nu_2^{(1)} = 2, \nu_1^{(1)} = 2) \geq \mathbb{P}(X_2^{(1)} = a, X_1^{(1)} = a, a_2^{(1)} = 1 | \nu_2^{(1)} = 2, \nu_1^{(1)} = 2) \\ & = \mathbb{P}(a_2^{(1)} = 1 | \nu_2^{(1)} = 2, \nu_1^{(1)} = 2, X_2^{(1)} = a, X_1^{(1)} = a) \times \mathbb{P}(X_2^{(1)} = a, X_1^{(1)} = a | \nu_2^{(1)} = 2, \nu_1^{(1)} = 2) \\ & = \underbrace{\mathbb{P}(a_2^{(1)} = 1 | \nu_2^{(1)} = 2, \nu_1^{(1)} = 2, X_2^{(1)} = a, X_1^{(1)} = a)}_{=:(1)} \\ & \times \underbrace{\mathbb{P}(X_2^{(1)} = a | X_1^{(1)} = a, \nu_2^{(1)} = 2, \nu_1^{(1)} = 2)}_{=:(2)} \times \underbrace{\mathbb{P}(X_1^{(1)} = a | \nu_2^{(1)} = 2, \nu_1^{(1)} = 2)}_{=:(3)}\end{aligned}$$

对于 $s = 1$, 2，我们令

$$N_s^a = \#\{k : X_s^{(k)} = a\}, N_s^b = \#\{k : X_s^{(k)} = b\}.$$

由于在时刻 1 等于 a 的唯一方式，是成为时刻 2 等于 a 的粒子的后代，因此我们有

$$\sum_{k\,:\, X_2^{(k)} = a} \nu_2^{(k)} = N_1^a ,$$

因此

$$(1) = \mathbb{E}\left(\frac{\nu_2^{(1)}}{N_1^a} \middle| \nu_2^{(1)} = 2, \nu_1^{(1)} = 2, X_2^{(1)} = a, X_1^{(1)} = a \right) .$$

对于所有 $N \geq 0$，我们令

$$A_N = \frac{p_a N_2^a}{p_a N_2^a + p_b N_2^b} = \frac{p_a (N_2^a / N)}{p_a (N_2^a / N) + p_b (N_2^b / N)}.$$

我们有

$$A_N \xrightarrow[N \to +\infty]{\text{a.s.}} \frac{p_a \alpha}{p_a \alpha + p_b (1-\alpha)} =: A_\infty .$$

此外,

$$x \in [0,1] \mapsto \frac{p_a x}{p_a x + p_b (1-x)}$$

是（p_a/p_b）-利普希茨的（因为 $p_b < p_a$），因此，对所有 N,

$$|A_N - A_\infty| \leq \frac{p_a}{p_b} \times \left| \frac{N_2^a}{N} - \alpha \right| ,$$

这意味着（因为 N_2^a 是参数为 α 的 N 个伯努利变量之和）

$$\mathbb{E}((A_N - A_\infty)^2) \leq \left(\frac{p_a}{p_b} \right)^2 \times \frac{\alpha(1-\alpha)}{N} .$$

我们有

$$\frac{N_1^a}{N} = \frac{1}{N} \sum_{i=1}^{N} \mathbb{1}_{U_i \leq A_N}$$

对于某些独立同分布的 U_i，它们独立于 A_N（按照分类分布抽样时便会出现这种情况）。由引理 4.1，我们得到

$$(3.3)\qquad \frac{N_1^a}{N} \xrightarrow[N \to +\infty]{\text{a.s.}} \frac{p_a \alpha}{p_a \alpha + p_b (1-\alpha)} = \alpha q_a ,$$

这意味着（由于 $N_1^a + N_1^b = N$）

$$\frac{N_1^b}{N} \xrightarrow[N\to+\infty]{\text{a.s.}} \frac{p_b(1-\alpha)}{p_a\alpha + p_b(1-\alpha)} = (1-\alpha)q_b\,. \tag{3.4}$$

因此，根据Fatou引理,

$$\liminf_{N\to+\infty} N \times (1) \geq 2 \times \frac{p_a\alpha + p_b(1-\alpha)}{p_a\alpha} = \frac{2}{\alpha q_a}\,. \tag{3.5}$$

我们有

$$\begin{aligned}(2) &= \mathbb{P}(X_2^{(1)} = a | \nu_2^{(1)} = 2, X_1^{(1)} = a, \nu_1^{(1)} = 2) \\ &= \frac{\mathbb{P}(X_2^{(1)} = a, \nu_2^{(1)} = 2, X_1^{(1)} = a)}{\mathbb{P}(\nu_2^{(1)} = 2, X_1^{(1)} = a)}\,.\end{aligned}$$

现在，作为混沌传播的推论，我们很容易相信，当 $N \to +\infty$ 时，1，2及其后代会与1，1及其后代相互独立（完整证明见第4.2节）。因此

$$(2) = \frac{\mathbb{P}(X_2^{(1)} = a, \nu_2^{(1)} = 2) \times \mathbb{P}(X_1^{(1)} = a, \nu_1^{(1)} = 2) + O(1/\sqrt{N})}{\mathbb{P}(\nu_2^{(1)} = 2) \times \mathbb{P}(X_1^{(1)} = a, \nu_1^{(1)} = 2) + O(1/\sqrt{N})}\,. \tag{3.6}$$

根据大数定律,

$$\frac{N_2^a}{N} \xrightarrow[N\to+\infty]{\text{a.s.}} \alpha\,,\ \frac{N_2^b}{N} \xrightarrow[N\to+\infty]{\text{a.s.}} 1-\alpha\,, \tag{3.7}$$

因此我们有（使用控制收敛定理）

$$\mathbb{P}(X_1^{(1)} = a) = \mathbb{E}\left(\frac{p_a N_2^a}{p_a N_2^a + p_b N_2^b}\right) \xrightarrow[N\to+\infty]{} \frac{\alpha p_a}{\alpha p_a + (1-\alpha)p_b} = \alpha q_a\,. \tag{3.8}$$

同理:

$$\mathbb{P}(X_1^{(1)} = b) \xrightarrow[N\to+\infty]{} \frac{(1-\alpha)p_b}{\alpha p_a + (1-\alpha)p_b} = (1-\alpha)q_b\,. \tag{3.9}$$

我们有

$$\begin{aligned}\mathbb{P}(\nu_1^{(1)} = 2 | X_1^{(1)} = a) &= \mathbb{E}\left(\frac{N(N-1)}{2}\left(1 - \frac{p_a}{N_1^a p_a + N_1^b p_b}\right)^{N-2}\left(\frac{p_a}{N_1^a p_a + N_1^b p_b}\right)^2\right) \\ &= \mathbb{E}\left(\exp\left((N-2)\log\left(1 - \frac{p_a}{N_1^a p_a + N_1^b p_b}\right)\right) \times \frac{N(N-1)}{2}\left(\frac{p_a}{N_1^a p_a + N_1^b p_b}\right)^2\right)\,.\end{aligned}$$

利用等式 (3.3)、(3.4) 和控制收敛定理，可得

$$\mathbb{P}(\nu_1^{(1)} = 2 | X_1^{(1)} = a) \xrightarrow[N\to+\infty]{} \frac{1}{2}\left(q_a'\right)^2 e^{-q_a'} = f(q_a')\,. \tag{3.10}$$

同理:

$$\mathbb{P}(\nu_1^{(1)} = 2 | X_1^{(1)} = b) \xrightarrow[N\to+\infty]{} f(q_b')\,. \tag{3.11}$$

因此，我们可以将项 (2) 化简为

$$(2) = \frac{\mathbb{P}(X_2^{(1)} = a, \nu_2^{(1)} = 2) + O(1/\sqrt{N})}{\mathbb{P}(\nu_2^{(1)} = 2) + O(1/\sqrt{N})} \tag{3.12}$$

利用贝叶斯公式和全概率公式，我们得到

$$\mathbb{P}(X_2^{(1)} = a | \nu_2^{(1)} = 2) \quad = \quad \mathbb{P}(\nu_2^{(1)} = 2 | X_2^{(1)} = a)\frac{\mathbb{P}(X_2^{(1)} = a)}{\mathbb{P}(\nu_2^{(1)} = 2)}$$

$$\begin{aligned} &= \frac{\mathbb{P}(\nu_2^{(1)}=2|X_2^{(1)}=a)\mathbb{P}(X_2^{(1)}=a)}{\mathbb{P}(\nu_2^{(1)}=2|X_2^{(1)}=a)\mathbb{P}(X_2^{(1)}=a)+\mathbb{P}(\nu_2^{(1)}=2|X_2^{(1)}=b)\mathbb{P}(X_2^{(1)}=b)} \\ &= \frac{\mathbb{P}(\nu_2^{(1)}=2|X_2^{(1)}=a)\times\alpha}{\mathbb{P}(\nu_2^{(1)}=2|X_2^{(1)}=a)\times\alpha+\mathbb{P}(\nu_2^{(1)}=2|X_2^{(1)}=b)\times(1-\alpha)}\,. \end{aligned}$$

我们有

$$\begin{aligned} \mathbb{P}(\nu_2^{(1)}=2|X_2^{(1)}=a) &= \mathbb{E}\left(\frac{N(N-1)}{2}\left(1-\frac{p_a}{N_2^a p_a+N_2^b p_b}\right)^{N-2}\left(\frac{p_a}{N_2^a p_a+N_2^b p_b}\right)^2\right) \\ &= \mathbb{E}\left(\exp\left((N-2)\log\left(1-\frac{p_a}{N_2^a p_a+N_2^b p_b}\right)\right)\times\frac{N(N-1)}{2}\left(\frac{p_a}{N_2^a p_a+N_2^b p_b}\right)^2\right). \end{aligned}$$

利用公式（3.7）和控制收敛定理，我们得到

$$\mathbb{P}(\nu_2^{(1)}=2|X_2^{(1)}=a)\times\alpha \underset{N\to+\infty}{\longrightarrow} \frac{\alpha}{2}q_a^2 e^{-q_a}=\alpha f(q_a)\,. \tag{3.13}$$

同理，我们得到

$$\mathbb{P}(\nu_2^{(1)}=2|X_2^{(1)}=b)\times(1-\alpha) \underset{N\to+\infty}{\longrightarrow} \frac{(1-\alpha)}{2}q_b^2 e^{-q_b}=(1-\alpha)f(q_b)\,. \tag{3.14}$$

因此我们得到

$$(2) \underset{N\to+\infty}{\longrightarrow} \frac{\alpha f(q_a)}{\alpha f(q_a)+(1-\alpha)f(q_b)}\,. \tag{3.15}$$

我们有

$$(3)=\frac{\mathbb{P}(X_1^{(1)}=a,\nu_2^{(1)}=2,\nu_2^{(1)}=2)}{\mathbb{P}(\nu_2^{(1)}=2,\nu_2^{(1)}=2)}\,.$$

同样，作为混沌传播的结果，我们很容易相信，当 $N\to+\infty$ 时，1，2及其后代将与1，1及其后代相互独立。因此

$$\begin{aligned} (3) &= \frac{\mathbb{P}(X_1^{(1)}=a,\nu_1^{(1)}=2)\mathbb{P}(\nu_2^{(1)}=2)+O(1/\sqrt{N})}{\mathbb{P}(\nu_2^{(1)}=2,)\mathbb{P}(\nu_2^{(1)}=2)+O(1/\sqrt{N})} \\ &\quad \text{(by Equations (3.7), (3.13), (3.14), we can symplify)} \\ &= \frac{\mathbb{P}(X_1^{(1)}=a,\nu_1^{(1)}=2)+O(1/\sqrt{N})}{\mathbb{P}(\nu_2^{(1)}=2)+O(1/\sqrt{N})} \end{aligned}$$

（上述公式的证明与公式（3.12）的证明非常相似，因此略去）。利用贝叶斯公式和全概率公式，

$$\begin{aligned} \mathbb{P}(X_1^{(1)}=a|\nu_1^{(1)}=2) &= \mathbb{P}(\nu_1^{(1)}=2|X_1^{(1)}=a)\times\frac{\mathbb{P}(X_1^{(1)}=a)}{\mathbb{P}(\nu_1^{(1)}=2)} \\ &= \frac{\mathbb{P}(\nu_1^{(1)}=2|X_1^{(1)}=a)\mathbb{P}(X_1^{(1)}=a)}{\mathbb{P}(\nu_1^{(1)}=2|X_1^{(1)}=a)\mathbb{P}(X_1^{(1)}=a)+\mathbb{P}(\nu_1^{(1)}=2|X_1^{(1)}=b)\mathbb{P}(X_1^{(1)}=b)}\,. \end{aligned}$$

由公式（3.8）、（3.9）、（3.10）、（3.11），可得

$$(3) \underset{N\to+\infty}{\longrightarrow} \frac{\alpha q_a f(q'_a)}{\alpha q_a f(q'_a)+(1-\alpha)q_b f(q_{b'})}\,. \tag{3.16}$$

由公式（3.2）、（3.5）、（3.15）、（3.16），我们得到

$$\liminf_{N\to+\infty} N\mathbb{P}(a_2^{(1)}=1|\nu_2^{(1)}=2,\nu_1^{(1)}=2)\geq 2\times\frac{1}{\alpha q_a}\times\frac{\alpha f(q_a)}{\alpha f(q_a)+(1-\alpha)f(q_b)}$$

$$\times \frac{\alpha q_a f(q'_a)}{\alpha q_a f(q'_a) + (1-\alpha) q_b f(q_{b'})}$$

□

推论 3.2. 对于经过适当选择的 α、 p_a、 p_b 的值，不等式 *3.1*（定理 *3.1*）右侧的 $R(\alpha, p_a, p_b)$ 严格大于1。这与公式（*2.2*）的断言相矛盾。

证明。关键是选择 α、 p_a、 p_b的取值，并计算 $R(\alpha, p_a, p_b)$。下面给出用于计算 R 的 Python 代码。我们先从 q_a 和 q_b的定义开始。

```
import numpy a s np
import matplotlib.pyplot as plt
定义 qa（al, pa, pb）:
    返回（pa/(al *pa+(1−al)*pb)）
定义 qb（al, pa, pb）:
    return(pb/(al*pa+(1−al)*pb))
```

这里定义 q'_a、 q'_b。

```
定义 qap（al, pa, pb）:
    return(pa/(al*qa(al,pa,pb)*pa+(1−al)*qb(al,pa,pb)*pb))
定义 qbp(a l,pa, p: b)
    return(pb/(al*qa(al,pa,pb)*pa+(1−al)*qb(al,pa,pb)*pb))
```

这里，我们定义 f，以及出现在 $R(\alpha, p_a, p_b)$ 中的第一项（$1/(\alpha q_a)$）和第二项（$\alpha f(q_a)/(\alpha f(q_a)+(1-\alpha) f(q_b))$）。

```
定义 f(t):
    return(0.5*t**2*np.exp(−t))
定义 t1(al, pa, pb):
    return(1/(al*qa(al,pa,pb)))
定义 t2(al, pa, pb):
    return((al*f(qa(al,pa,pb)))/
        (al*f(qa(al,pa,pb))+(1−al)*f(qb(al,pa,pb))))
```

在此，我们定义出现在 $R(\alpha, p_a, p_b)$ 中的第三项（$\alpha q_a f(q'_a)/(\alpha q_a f(q'_a) + (1-\alpha) q_b f(q'_b))$）。

```
def t3(al, pa, pb):
    return(al*qa(al,pa,pb)*f(qap(al,pa,pb))/
        (al*qa(al,pa,pb)*f(qap(al,pa,pb))
        +(1−al)*qb(al,pa,pb))*f(qbp(al,pa,pb))))
```

现在， R 定义如下。

```
def R(al, pa, pb):
    return(t1(al,pa,pb)*t2(al,pa,pb)*t3(al,pa,pb))
```

我们为 p_b绘制 $R(0.5, 1, p_b)$，其中 $[0, 0.2]$。

```
def trace(x): return (R(0.5,
1, x)) x=np.linspace(0, 0.2, 500)
plt.plot(x, trace(x)) plt.xlabel('x')
plt.ylabel('R(0.5, 1, x)')
plt.show()
```

我们得到图3.2。1 上方的直线。 此外，如果我们询问 R(0。5，1，0。075)，得到的答案是 1。04104942 ……，严格大于 1。

我们有

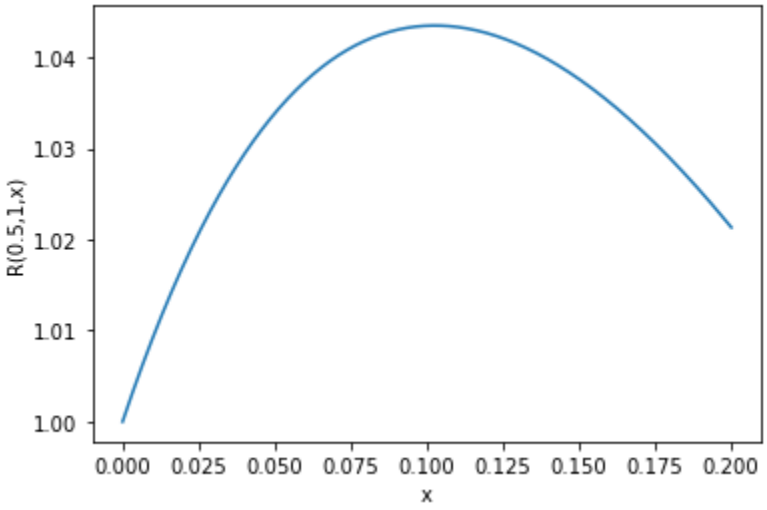

图3.2。项 $R_\circ$

$$\begin{aligned}\mathbb{E}(\mathbb{P}(a_2^{(1)}=1|\boldsymbol{\nu},\xi)|\nu_2^{(1)}=2,\nu_1^{(1)}=2)&=\mathbb{E}(\mathbb{E}(\mathbb{1}_{a_2^{(1)}=1}|\boldsymbol{\nu},\xi)|\nu_2^{(1)}=2,\nu_1^{(1)}=2))\\&=\mathbb{E}(\mathbb{1}_{a_2^{(1)}=1}|\nu_2^{(1)}=2,\nu_1^{(1)}=2)=\mathbb{P}(a_2^{(1)}=1|\nu_2^{(1)}=2,\nu_1^{(1)}=2)\end{aligned}$$

因此，如果取 $\alpha=0.5$、 $p_a=1$、 $p_b=0.075$，则存在 N 和 $\boldsymbol{\nu}$、 ξ （其中 $\nu_2^{(1)}=2$、 $\nu_1^{(1)}=2$），使得

$$\mathbb{P}(a_2^{(1)}=1|\boldsymbol{\nu},\xi)>\frac{\nu_2^{(1)}}{N}.$$

这与公式（2.2）矛盾。 □

4. Appendix

4.1. 概率技术引理。

引理 4.1. 若有一个 L^2 随机变量序列 $(A_N)^{N\geq 0}$ 在 $[0,1]$ 中几乎必然收敛到常数 A_∞，当 $N\to+\infty$ 时，且

$$\mathbb{E}((A_N-A_\infty)^2)\leq\frac{C}{N}$$

（对某个常数 C）并且若有一个由概率分布为 $\mathcal{U}([0,1])$的变量构成的独立同分布序列 $(U_i)^{i\geq 0}$ （与 A_n独立），则

$$\frac{1}{N}\sum_{i=1}^{N}\mathbb{1}_{U_i\leq A_N}\xrightarrow[N\to+\infty]{a.s.}A_\infty\,.$$

证明。 我们有（大数定律）

$$\frac{1}{N}\sum_{i=1}^{N}\mathbb{1}_{U_i\leq A_\infty}\xrightarrow[N\to+\infty]{\text{a.s.}}A_\infty\,.$$

因此，我们考察这个差值

$$\begin{aligned}\left|\frac{1}{N}\sum_{i=1}^{N}\mathbb{1}_{U_i\leq A_N}-\frac{1}{N}\sum_{i=1}^{N}\mathbb{1}_{U_i\leq A_\infty}\right|&=\left|\frac{1}{N}\sum_{i=1}^{N}u(U_i,A_N,A_\infty)\right|\\&\leq\left|\frac{1}{N}\sum_{i=1}^{N}u(U_i,A_N,A_\infty)-(A_N-A_\infty)\right|+\frac{1}{N}\sum_{i=1}^{N}|A_N-A_\infty|\,,\end{aligned}$$

其中

$$u:(v,x,y)\in[0,1]\times\mathbb{R}\times\mathbb{R}\mapsto u(v,x,y)=\begin{cases}-1 & \text{if } x<y \text{ and } v\in(x,y],\\ 1 & \text{if } y<x \text{ and } v\in(y,x],\\ 0 & \text{otherwise.}\end{cases}$$

我们有

$$\frac{1}{N}\sum_{i=1}^{N}|A_N-A_\infty|=|A_N-A_\infty|\xrightarrow[N\to+\infty]{\text{a.s.}}0\,.$$

并且我们有（当 $\mathbb{E}(u(U_i,\ A_N,\ A_\infty)|A_N,\ A_\infty)=A_N-A_\infty$ 对所有 $i\geq 0, N\geq 0$）

$$\begin{aligned}\mathbb{E}\left(\left(\frac{1}{N}\sum_{i=1}^{N}u(U_i,A_N,A_\infty)-(A_N-A_\infty)\right)^2\right)&=\mathbb{E}\left(\frac{1}{N^2}\sum_{i=1}^{N}(u(U_i,A_N,A_\infty)-(A_N-A_\infty))^2\right)\\&=\mathbb{E}\left(\frac{1}{N^2}\sum_{i=1}^{N}\mathbb{E}((u(U_i,A_N,A_\infty)-(A_N-A_\infty))^2|A_N,A_\infty)\right)\\&=\mathbb{E}\left(\frac{1}{N^2}\sum_{i=1}^{N}\mathbb{E}(u(U_i,A_N,A_\infty)^2|A_N,A_\infty)-(A_N-A_\infty)^2\right)\\&=\mathbb{E}\left(\frac{1}{N^2}\sum_{i=1}^{N}|A_N-A_\infty|-(A_N-A_\infty)^2\right)\\&\leq\mathbb{E}\left(\frac{1}{N^2}\sum_{i=1}^{N}|A_N-A_\infty|\right)\\&\text{(Jensen's inequality)}\leq\frac{N}{N^2}\mathbb{E}((A_N-A_\infty)^2)^{1/2}\leq\frac{\sqrt{C}}{N^{3/2}}\,.\end{aligned}$$

根据 Borell-Cantelli 引理:

$$\frac{1}{N}\sum_{i=1}^{N}u(U_i,A_N,A_\infty)-(A_N-A_\infty)\xrightarrow[N\to+\infty]{\text{a.s.}}0\,,$$

由此完成证明。 □

4.2. 公式（3.6）的证明。 要证明公式（3.6），只需证明以下引理。

引理 4.2。 我们有

(4.1)

$$\mathbb{P}(X_2^{(1)}=a,\nu_2^{(1)}=2,X_1^{(1)}=a,\nu_1^{(1)}=2)=\mathbb{P}(X_2^{(1)}=a,\nu_2^{(1)}=2)\mathbb{P}(X_1^{(1)}=a,\nu_1^{(1)}=2)+O(1/\sqrt{N})\,,$$

(4.2)

$$\mathbb{P}(\nu_2^{(1)}=1,X_1^{(1)}=a,\nu_1^{(1)}=2)=\mathbb{P}(\nu_2^{(1)}=1)\mathbb{P}(X_1^{(1)}=a,\nu_1^{(1)}=2)+O(1/\sqrt{N})\,.$$

上述引理的证明将通过耦合完成。

4.2.1. 耦合的描述。从权重为 $(w_2^{(1)},\dots,w_2^{(N)})$ 的分类分布中抽样时（这些是时刻 2 粒子的权重），我们假设使用概率分布为 $\mathcal{U}([0,1])$ 的独立同分布变量 $(U_2^{(1)},\dots,U_{(N)\,2})$，具体如下

$$a_2^{(i)}=k \text{ if and only if } w_2^{(1)}+\cdots+w_2^{(k-1)}\leq U_2^{(i)}<w_2^{(1)}+\cdots+w_2^{(k)}\ (k\in[N]).$$

提醒读者，对于所有 i 属于 $[N]$ 的情形，

$$w_2^{(i)}=\frac{g(X_2^{(i)})}{g(X_2^{(1)})+\cdots+g(X_2^{(N)})}\,.$$

我们引入

$$\widehat{w}_2^{(1)} = \frac{g(X_2^{(1)})}{N(\alpha p_a + (1-\alpha)p_b)}\,,$$

$$\widehat{\nu}_2^{(1)} = \#\{i : U_2^{(i)} \le \widehat{w}_2^{(1)}\}\,.$$

随机数 $\widehat{\nu}_2^{(1)}$ 扮演 1,2 的"后代数"，且与 $X_2^{(2)},\dots,X_2^{(N)}$ 独立。

我们引入额外的独立同分布变量 $(U_{(N+1)\,2}, U_2^{(N+2)},\dots)$，其概率分布为 $\mathcal{U}([0,1])$。随后，对所有 i 属于 $\{N+1, N+2,\dots\}$、k 属于 $[N]$，我们令

$$a_2^{(i)} = k \text{ if and only if } w_2^{(1)}+\dots+w_2^{(k-1)} \le U_2^{(i)} < w_2^{(1)}+\dots+w_2^{(k)}\,.$$

对所有 i 属于 $\{1,2,\dots,N,N+1,\dots\}$，我们令

$$X_1^{(i)} = X_2^{a_2^{(i)}}$$

（由于我们的 K 过于简单，这相当于抽样 $X_1^{(i)} \sim K(X_2^{a_2^{(i)}},.)$ 。）

我们以递归方式定义 $\mathcal{I} = \{i_1, i_2, \dots, i_N\}$，其中 $i_1 < i_2 < \dots < i_N$：

- $i_1 = \inf\{i : a_2^{(i)} \neq 1\}$
- for $k \in \{2,\dots,N\}$, $i_k = \inf\{i > i_{k-1} : a_2^{(i_k)} \neq 1\}$.

N元组

$$(X_1^{(i_1)},\dots,X_1^{(i_N)}) = (X_2^{(a_2^{(i_1)})},\dots,X_2^{(a_2^{(i_N)})}) \text{ is independent of } (X_2^{(1)}, \widehat{\nu}_2^{(1)})\,. \tag{4.3}$$

事实上，序列 i_1，i_2，..... 是通过接受-拒绝方案构造的：每个 $U^{(i_k)\,2}$ 服从概率分布 $\mathcal{U}([w_2^{(1)},1])$，且对所有 j 属于 $\{2,\dots N\}$的情形，$a^{(i_k)}{}_2 \quad = j$ 以如下概率发生

$$\frac{g(X_2^{(j)})}{\sum_{r=2}^N g(X_2^{(r)})}\,.$$

我们设

$$I = i_1\,,$$

and

$$\widetilde{X}_1^{(1)} = X_1^{(I)}\,.$$

假设我们使用服从概率分布 $\mathcal{U}([0,1])$ 的独立同分布变量 $(U_1^{(1)}$，.....，$U_{(N)\,1})$ 来计算亲代索引（对所有 i 属于 $[N]$的情形）：

$$a_1^{(i)} = k \text{ if and only if } w_1^{(1)}+\dots+w_1^{(k-1)} \le U_1^{(i)} < w_1^{(1)}+\dots+w_1^{(k)} \ (k \in [N]),$$

其中，$(w_1^{(1)}$，.....，$w_1^{(N)})$ 是时刻 1 粒子的权重。对所有 i 属于 $[N]$,

$$w_1^{(i)} = \frac{g(X_1^{(i)})}{\sum_{r=1}^N g(X_1^{(r)})}\,.$$

我们引入替代权重

$$\widetilde{w}_1^{(i)} = \frac{g(X_1^{(i)})}{\sum_{k\in\mathcal{I}} g(X_1^{(k)})} \ (i \in \mathcal{I}).$$

由此，我们计算出其他亲代索引（适用于所有 i、$k \in [N]$）：

$$\widetilde{a}_1^{(i)} = k \text{ if and only if } \widetilde{w}_1^{(i_1)}+\dots+\widetilde{w}_1^{(i_k)} \le U_1^{(i)} \le \widetilde{w}_1^{(i_1)}+\dots+\widetilde{w}_1^{(i_k)}\,.$$

并且我们设

$$\widetilde{\nu}_1^{(1)} = \#\{i \in [N] : \widetilde{a}_1^{(i)} = I\}\,.$$

4.2.2. 技术引理。

引理 4.3。 变量 $(\widetilde{X}_1^{(1)}, \widetilde{\nu}_1^{(1)})$ 与 $(X_2^{(1)}, \widehat{\nu}_2^{(1)})$ 独立。

证明。我们想要证明，对于所有 $u \in \{a,b\}$ 和 $k \in \mathbb{N}$，$\mathbb{P}((\widetilde{X}_1^{(1)}, \widetilde{\nu}_1^{(1)}) = (u, k) | X_2^{(1)}, \widehat{\nu}_2^{(1)})$ 是一个常数。我们设定

$$\beta(u) = \begin{cases} \alpha & \text{if } u = a, \\ 1-\alpha & \text{if } u = b. \end{cases}$$

对于所有 $u \in \{a,b\}$,

$$\begin{aligned}
\mathbb{P}(\widetilde{X}_1^{(1)} = u | X_2^{(1)}, \widehat{\nu}_2^{(1)}) &= \sum_{j=2}^{N} \mathbb{P}(X_2^{(j)} = u, a_2^I = j | X_2^{(1)}, \widehat{\nu}_2^{(1)}) \\
&= \sum_{j=2}^{N} \mathbb{P}(a_2^I = j | X_2^{(1)}, \widehat{\nu}_2^{(1)}, X_2^{(j)} = u) \mathbb{P}(X_2^{(j)} = u | X_2^{(1)}, \widehat{\nu}_2^{(1)}) \\
&= \sum_{j=2}^{N} \mathbb{E}\left(\mathbb{E}\left(\mathbb{1}_{a_2^I = j} \middle| X_2^{(1)}, \widehat{\nu}_2^{(1)}, (X_2^{(i)})_{2 \le i \le N} \right) \middle| X_2^{(1)}, \widehat{\nu}_2^{(1)}, X_2^{(j)} = u \right) \times \mathbb{P}(X_2^{(j)} = u | X_2^{(1)}, \widehat{\nu}_2^{(1)}) \\
&= \sum_{j=2}^{N} \mathbb{E}\left(\frac{g(u)}{\sum_{i=2}^{N} g(X_2^{(i)})} \middle| X_2^{(1)}, \widehat{\nu}_2^{(1)}, X_2^{(j)} = u \right) \times \beta(u) \\
&= \sum_{j=2}^{N} \mathbb{E}\left(\frac{g(u)}{g(u) + \sum_{i \in \{2,\dots,N\} \setminus \{j\}} g(X_1^{(i)})} \right) \times \beta(u).
\end{aligned}$$

因此 $\mathbb{P}(\widetilde{X}_1^{(1)} = u | X_2^{(1)}, \widehat{\nu}_2^{(1)})$ 是常数。

接下来（根据公式（4.3）），对于所有 $k \in \{0, 1, \dots, N\}$,

$$\begin{aligned}
\mathbb{P}(\widetilde{\nu}_1^{(1)} = k | \mathcal{I}, (X_1^{(i)})_{i \in \mathcal{I}}, X_2^{(1)}, \widehat{\nu}_2^{(1)}) &= \binom{N}{k} (\widetilde{w}_1^{(I)})^k (1 - \widetilde{w}_1^{(I)})^{N-k} \\
&= \binom{N}{k} \left(\frac{g(\widetilde{X}_1^{(1)})}{\sum_{i \in \mathcal{I}} g(X_1^{(i)})} \right)^k \left(1 - \frac{g(\widetilde{X}_1^{(1)})}{\sum_{i \in \mathcal{I}} g(X_1^{(i)})} \right)^{N-k}.
\end{aligned}$$

So

$$\begin{aligned}
&\mathbb{P}(\widetilde{\nu}_1^{(1)} = k | \widetilde{X}_1^{(1)} = u, X_2^{(1)}, \widehat{\nu}_2^{(1)}) \\
&\quad = \mathbb{E}\left(\binom{N}{k} \left(\frac{g(u)}{g(u) + \sum_{i \in \mathcal{I} \setminus \{i_1\}} g(X_1^{(i)})} \right)^k \left(1 - \frac{g(u)}{g(u) + \sum_{i \in \mathcal{I} \setminus \{i_1\}} g(X_1^{(i)})} \right)^{N-k} \right)
\end{aligned}$$

是常数。

因此，对于所有 k,

$$\mathbb{P}(\widetilde{X}_1^{(1)} = u, \widetilde{\nu}_1^{(1)} = k | X_2^{(1)}, \widehat{\nu}_2^{(1)}) = \mathbb{P}(\widetilde{\nu}_1^{(1)} = k | X_2^{(1)}, \widehat{\nu}_2^{(1)}, \widetilde{X}_1^{(1)} = u) \mathbb{P}(\widetilde{X}_1^{(1)} = u | X_2^{(1)}, \widehat{\nu}_2^{(1)})$$

是一个常数。至此，证明完毕。 □

引理 4.4。 $\mathbb{P}((\widetilde{X}_1^{(1)}, \widetilde{\nu}_1^{(1)}) \neq (X_1^{(1)}, \nu_1^{(1)})) = O(1/N)$

证明。我们有

$$\begin{aligned}
\mathbb{P}((\widetilde{X}_1^{(1)}, \widetilde{\nu}_1^{(1)}) = (X_1^{(1)}, \nu_1^{(1)})) &\ge \mathbb{P}(I = 1, \widetilde{\nu}_1^{(1)} = \nu_1^{(1)}) = \mathbb{P}(a_2^{(1)} \neq 1, \widetilde{\nu}_1^{(1)} = \nu_1^{(1)}) \\
&= \mathbb{P}(a_2^{(1)} \neq 1) \mathbb{P}(\nexists i \in [N] : U_1^{(i)} \in [w_1^{(1)}, \widetilde{w}_1^{(1)}] | a_2^{(1)} \neq 1) \\
&= \mathbb{E}(1 - w_2^{(1)}) \times \mathbb{E}\left((1 - |w_1^{(1)} - \widetilde{w}_1^{(1)}|)^N \middle| a_2^{(1)} \neq 1 \right)
\end{aligned}$$

$$\text{(for } N \text{ such that } \frac{p_a}{Np_b} < 1)$$

$$\geq \left(1 - \frac{p_a}{Np_b}\right) \times \mathbb{E}\left(1 - N|w_1^{(1)} - \widetilde{w}_1^{(1)}| \middle| a_2^{(1)} \neq 1\right)$$

$$\text{(because } a_2^{(1)} \neq 1 \Rightarrow i_1 = 1)$$

$$= \left(1 - \frac{p_a}{Np_b}\right) \mathbb{E}\left(1 - N\left|\frac{g(X_1^{(1)})}{\sum_{i\in[N]} g(X_1^{(i)})} - \frac{g(X_1^{(1)})}{\sum_{i\in\mathcal{I}} g(X_1^{(i)})}\right| \middle| a_2^{(1)} \neq 1\right)$$

$$= \left(1 - \frac{p_a}{Np_b}\right) \mathbb{E}\left(1 - N\left|\frac{g(X_1^{(1)})(\sum_{i\in[N]} g(X_1^{(i)}) - \sum_{i\in\mathcal{I}} g(X_1^{(i)}))}{\sum_{i\in[N]} g(X_1^{(i)}) \times \sum_{i\in\mathcal{I}} g(X_1^{(i)})}\right| \middle| a_2^{(1)} \neq 1\right)$$

$$\text{(as } \#([N]\backslash\mathcal{I}) = \#(\mathcal{I}\backslash[N]) = \nu_1^{(2)}) \geq \left(1 - \frac{p_a}{Np_b}\right) \mathbb{E}\left(1 - \frac{\nu_2^{(1)} p_a(p_a - p_b)}{Np_b^2} \middle| a_2^{(1)} \neq 1\right)$$

$$= \left(1 - \frac{p_a}{Np_b}\right)\left(1 - \frac{\mathbb{E}(\nu_2^{(1)}|a_2^{(1)} \neq 1)p_a^2}{Np_b^2}\right)$$

$$\text{(as } \nu_2^{(1)} = \sum_{i=1}^{N} \mathbb{1}_{\{1\}}(a_2^{(i)})) \geq \left(1 - \frac{p_a}{Np_b}\right)\left(1 - \frac{\mathbb{E}(\nu_2^{(1)})p_a^2}{Np_b^2}\right)$$

$$\geq \left(1 - \frac{p_a}{Np_b}\right)\left(1 - \frac{\left(\frac{p_a}{p_b}\right) \times p_a^2}{Np_b^2}\right)$$

□

引理 4.5。 $\mathbb{P}((X_2^{(1)},\ \widehat{\nu}_2^{(1)}) \neq (X_2^{(1)},\ \nu_2^{(1)})) = O(1/\sqrt{N})$

证明。我们有

$$\mathbb{P}((X_2^{(1)}, \widehat{\nu}_2^{(1)}) \neq (X_2^{(1)}, \nu_2^{(1)}) = \mathbb{P}(\exists i \in [N] : U_2^{(i)} \in [w_2^{(1)}, \widehat{w}_2^{(1)}])$$

$$= 1 - \mathbb{P}\left(\forall i \in [N], U_2^{(i)} \notin \left[\frac{g(X_2^{(1)})}{\sum_{i=1}^N g(X_2^{(i)})}, \frac{g(X_2^{(1)})}{N(\alpha p_a + (1-\alpha)p_b)}\right]\right)$$

$$= 1 - \mathbb{E}\left(\left(1 - \left|\frac{g(X_2^{(1)})}{N(\alpha p_a + (1-\alpha)p_b)} - \frac{g(X_2^{(1)})}{\sum_{i=1}^N g(X_2^{(i)})}\right|\right)^N\right)$$

$$\leq N\mathbb{E}\left(\left|\frac{g(X_2^{(1)})}{N(\alpha p_a + (1-\alpha)p_b)} - \frac{g(X_2^{(1)})}{\sum_{i=1}^N g(X_2^{(i)})}\right|\right)$$

$$= N\mathbb{E}\left(\left|g(X_2^{(1)}) \times \frac{\sum_{i=1}^N g(X_2^{(i)}) - N(\alpha p_a + (1-\alpha)p_b)}{N(\alpha p_a + (1-\alpha)p_b) \times \sum_{i=1}^N g(X_2^{(i)})}\right|\right)$$

$$\text{(Jensen's inequality)} \leq \frac{Np_a}{N^2 p_b^2}\mathbb{E}\left(\left(\sum_{i=1}^N (g(X_2^{(i)}) - (\alpha p_a + (1-\alpha)p_b)\right)^2\right)^{1/2}$$

$$\text{(with } \sigma \geq 0\,,\, \sigma^2 = \alpha p_a^2 + (1-\alpha)p_b^2 - (\alpha p_a + (1-\alpha)p_b)^2) \ = \frac{p_a}{Np_b^2}\sqrt{N}\sigma$$

□

4.3. **引理 4.2 的证明**。

证明。 我们计算

$$\mathbb{P}(X_2^{(1)} = a, \nu_2^{(1)} = 2, X_1^{(1)} = a, \nu_1^{(1)} = 2)$$

$$\text{(by Lemma 4.4 and Lemma 4.5)}$$
$$= \mathbb{P}(X_2^{(1)} = a, \widehat{\nu}_2^{(1)} = 2, \widetilde{X}_1^{(1)} = a, \widetilde{\nu}_1^{(1)} = 2) + O(1/\sqrt{N})$$
$$\text{(by Lemma 4.3)}$$
$$= \mathbb{P}(X_2^{(1)} = a, \widehat{\nu}_2^{(1)} = 2) \times \mathbb{P}(\widetilde{X}_1^{(1)} = a, \widetilde{\nu}_1^{(1)} = 2) + O(1/\sqrt{N})$$
$$\text{(by Lemma 4.4 and Lemma 4.5)}$$
$$= \mathbb{P}(X_2^{(1)} = a, \nu_2^{(1)} = 2) \times \mathbb{P}(X_1^{(1)} = a, \nu_1^{(1)} = 2) + O(1/\sqrt{N})$$

这就证明了等式（4.1）。等式（4.2）的证明非常类似。 □

参考文献

[CR11] Dan Crisan 和 Boris Rozovskiĭ（编），《牛津非线性滤波手册》，牛津大学出版社，牛津，2011。MR 2882749[Del04] Pierre Del Moral，《费曼-卡茨公式》，《概率及其应用》（纽约），Springer-Verlag，纽约，2004，附有应用的谱系与相互作用粒子系统。MR 2044973[DM01] P. Del Moral 和 L. Miclo，《费曼-卡茨模型与遗传模型的谱系及混沌传播的增强》，《应用概率年刊》**11** 2001 4, 1166 1198 1878294 ( )，第 – 期。MR DMPR P. Del Moral、L. Miclo、F. Patras 和 S. Rubenthaler，《中性遗传模型的平衡态收敛》，《随机分析与应用》**28** 2010 1, 123 143 2597983 ( )，第 – 期。MR JMR Pierre E. Jacob、Lawrence M. Murray 和 Sylvain Rubenthaler，《粒子滤波中的路径存储》，《统计计算》**25** 2015 2, 487 496 3306720 ( )，第 – 期。MR Kin J. F. C. Kingman，《合并过程》，《随机过程及其应用》**13** 1982 3, 235 248 671034 ( )，第 – 期。MR KJJS Jere Koskela、Paul A. Jenkins、Adam M. Johansen 和 Dario Spanò，《相互作用粒子系统的渐近谱系及其在序贯蒙特卡洛中的应用》，《统计年刊》**48** 1，560 2020 583 4065174 ( )，第 – 期。MR [Möh98] M. Möhle，《合并过程的稳健性结果》，《应用概率杂志》**35** (1998)，第2期，438–447。MR 1641829