arXiv:2301.00461v1 [math.PR] 1 Jan 2023

稠密图的均匀生成树的 GHP 标度极限

Eleanor Archer* Matan Shalev†

2023年1月3日

摘要

我们考虑收敛到连通图子的稠密图序列，并证明其均匀生成树的 GHP 标度极限是 Aldous 布朗连续随机树。此外，我们还能够从极限图子中提取精确的标度常数。例如，我们可以将此结果应用于 Erd¨os-R´enyi 序列 $(G(n,p))_{n\geq 1}$ 的均匀生成树的标度极限，其中任意固定$p \in (0,1],$，以及稠密扩张图序列。GHP收敛的一个推论是，生成树的若干相关量也会收敛，例如高度、直径和简单随机游走的分布。

1 引言

均匀生成树（UST）是概率论和计算机科学中的基本对象，与数学的许多其他领域密切相关，包括电网络理论、环擦除随机游走和随机交错渗流等。

近期研究表明，在 [7], 基于 [31], 的工作后，某大类图的均匀生成树的普适度量测度空间标度极限是 Aldous 的布朗连续统随机树（连续随机树）。本文旨在将这一结果推广至由图子编码的稠密图序列。由于此前论文中的传递性假设，这些均匀生成树并不适用于 [31] 和 [7], 的结果；但我们在此证明，连续随机树仍然是其标度极限。此外，我们还能够用编码图子表示精确的标度因子，使这一结果比 [7] 中的结果更加精确，并表明图子收敛这一概念足以完全确定均匀生成树的标度极限。

连续随机树（CRT）由 Aldous [1, 2, 3], 引入，是概率论中一个广为人知的对象，或许最广为人知的是它作为临界有限方差 Galton–Watson树的标度极限。我们不打算在此作全面介绍；我们将在第 3 节给出正式定义，并请读者参阅 Le Gall [23]的综述以了解更多背景。

加权图 (G,w) 是一个图 $G=(V,E)$，其中我们为每条边 $e \in E$ 赋予一个非负权重 w_e。本文将研究一系列不含自环或重边的加权图，其中 $w_e \in [0,1]$ 对每个 $e \in E$均成立。当所有边的权重都等于 1 时，我们称该图为简单图。我们将顶点度的定义扩展到加权图，定义 $\deg v$ 为从 v出发的边的权重之和。

加权图 (G,w) 的均匀生成树，是从 G 的所有生成树中随机选取的一棵，其中每棵生成树 t 被选中的概率与 $\prod_{e\in t} w_e$ 成正比。

如果存在 $\delta > 0$ 使得 $\Delta_n := \min_{v\in G_n} \deg(v) \geq \delta \#V(G_n)$ 对所有 n 都成立，我们就称这样的加权图序列 $(G_n)_{n\geq 1}$ 是稠密的。稠密图序列的收敛概念自然可由称为图子的对象来刻画；图子由 Lov´asz 和 Szegedy [25] 以及 Borgs、

* ´巴黎南泰尔大学 Modal'X 团队，法国南泰尔 92000，R´epublique 大道 200 号，G 楼。电子邮件：eleanor.archer@parisnanterre.fr

†特拉维夫大学数学科学学院，以色列特拉维夫。电子邮件： matanshalev@mail.tau.ac.il

Chayes、Lov´asz、S´os 和 Vesztergombi [10] 为此目的而引入。另请参阅 [14]，其中有非常简短的介绍。图子 W 是从 $[0,1]^2$ 到 $[0,1]$ 的对称可测函数，可以将其视为邻接矩阵的（粗略）连续统类比。基于这一观点，离散图与图子之间存在一种自然的距离概念，称为割距离，我们将在第 2.1 节中定义它。这使我们能够考虑收敛到给定图子 W的概念。

图子常用于组合学和计算机科学，以分析大型稠密图。例如，图子已用于极值图论 [12], 平均场博弈 [11], 大型图的分析 [21], 以及研究统计物理系统的热力学极限 [27, 13], ，以上仅举少数例子。

给定一个图子 W，定义常数

$$\alpha_W = \frac{1}{\left(\int_{[0,1]^2} W(x,y)dxdy\right)^2} \cdot \int_{[0,1]} \left(\int_{[0,1]} W(x,y)dy\right)^2 dx. \tag{1}$$

由詹森不等式可立即得出 $\alpha_W \geq 1$，且当且仅当 W 几乎处处为常数时取等号。我们还称图子 W 是连通的，如果对于所有具有正勒贝格测度的 $A \subset [0,1]$，都有

$$\int_A \int_{A^C} W(x,y)dxdy > 0.$$

本文的主要结果如下。下文中，GHP 距离指度量测度空间之间的 Gromov-Hausdorff-Prohorov 距离；我们将在第 2 节中定义它。

定理 1.1. 设 $(G_n)_{n\geq 1}$ 是一个收敛到连通图子 W的确定性加权图稠密序列，其中每个 G_n 都有 n 个顶点。对于每个 $n \geq 1$，令 $\mathcal{T}_n$ 为 G_n的一棵均匀生成树。记 $\mathcal{T}_n$ 上相应的图距离为 $d_{\mathcal{T}_n}$，并记 $\mathcal{T}_n$顶点上的均匀概率测度为 μ_n。那么

$$\left(\mathcal{T}_n, \frac{\sqrt{\alpha_W}}{\sqrt{n}} d_{\mathcal{T}_n}, \mu_n\right) \xrightarrow{(d)} (\mathcal{T}, d_{\mathcal{T}}, \mu)$$

其中，α_W 的定义见 (1)，$(\mathcal{T}, d_{\mathcal{T}}, \mu)$ 是配备典范质量测度 μ 的连续随机树，而$\xrightarrow{(d)}$表示关于度量测度空间之间的 GHP 距离依分布收敛。

单个图子也可以编码随机图序列 $G(k,W)_{k\geq 1}$ 和 $H(k,W)_{k\geq 1}$，它们有 k 个节点，通过从 $[0,1],$中采样 k 个均匀顶点 x_1、.....、x_k 得到；随后以概率 $W(x_i,x_j)$ 在节点 i和 j之间添加一条权重为 1 的边（得到序列 $G(k,W)_{k\geq 1}$），或者改为添加一条权重为 $W(x_i,x_j)$ 的边（得到序列 $H(k,W)_{k\geq 1}$）。我们将推论 1.2 作为定理 1.1 的结果推出。

推论 1.2. 设 W 为连通图子。假设存在 $\delta > 0$，使得当 $n \to \infty$时，$G(n,W)$ 的最小度以趋于 1 的概率至少为 δn。对每个 $n \geq 1$，令 $\mathcal{T}_n$为 $G(n,W)$ 的均匀生成树。记 $d_{\mathcal{T}_n}$为 $\mathcal{T}_n$上的相应图距离，μ_n为 $\mathcal{T}_n$ 顶点上的均匀概率测度。那么

$$\left(\mathcal{T}_n, \frac{\sqrt{\alpha_W}}{\sqrt{n}} d_{\mathcal{T}_n}, \mu_n\right) \xrightarrow{(d)} (\mathcal{T}, d_{\mathcal{T}}, \mu)$$

其中，$(\mathcal{T}, d_{\mathcal{T}}, \mu)$ 是配备典范质量测度 μ 的连续随机树，而 $\xrightarrow{(d)}$ 表示度量测度空间之间关于 GHP 距离的依分布收敛。

此外，将 $G(n,W)$ 替换为 $H(n,W)$ 时，同样的结论仍然成立。

例如，对于任意固定的 $p \in$，Erd¨os-R´enyi 序列 $(G(n,p))_{n\geq 1}$ 都满足这一结论（该序列为 $(G(n,W))_{n\geq 1}$，其中 W 是几乎处处取值为 p 的图子，此时 $\alpha_W = 1$）。

定理 1.1 表明，图子包含了足够的信息，可以确定均匀生成树的标度极限；换言之，GHP 标度极限关于割距离所诱导的拓扑是连续的。在 [16], 中，作者对定理 1.1 中均匀生成树的 Benjamini-Schramm 局部极限证明了类似的结果，并表明该局部极限可以刻画为一个以存活为条件的多类型临界分枝过程，其子代分布由极限图子编码。此外，作者还表明，对 G_n 的生成树总数进行适当归一化后，其连续性也成立。不过，他们也给出了一个例子，说明在较弱的假设下这一结论不再成立。

注意，图序列收敛到一个连通图子，并不自动意味着该图序列必须是稠密图；事实上， [16] 中关于稠密图的均匀生成树所得出的局部极限定理并不要求这一假设。在那里，作者只假设极限图子是非退化的，也就是说

$$\deg_W(x) := \int_{[0,1]} W(x,y)dy > 0 \quad \forall x \in [0,1], \tag{2}$$

并且图序列是连通的。事实上，这意味着"大多数"顶点的度都很高；精确表述见 [16,定理 2.7 和定义 2.6]。这足以证明局部极限结论，因为以高概率而言，局部极限不会看到低阶的例外顶点。另一方面，GHP 标度极限是一个全局结论，因此我们需要对底层图有更一致的控制。通过一个简单的反例即可看出这一点：令 G_n 表示具有$n - n^{2/3}$ 个顶点的完全图，并在完全图的一个顶点上接一条长度为 $n^{2/3}$ 的线段。这些图仍然收敛到处处为 1 的图子，而均匀生成树 UST(G_n) 的局部极限仍是条件为存活的 Poisson(1) Galton–Watson 树。另一方面，唯一非平凡的紧标度极限是一条线段，而不是连续随机树。对于任意序列 $\gamma_n \to \infty$, 也可以构造最小度至少为 $\frac{n}{\gamma_n}$ 的类似反例，这说明线性最小度这一假设确实是必要的。

由于连续随机树的局部极限已知 [1, 第 6] 节$_{n\geq 1}$ 是 Aldous 自相似连续树（SSCRT），因此还可以考察标度极限与局部极限这两个操作是否可交换。一般而言，回答这一问题似乎相当非平凡，因为作为局部极限出现的多类型分枝过程具有很强的非齐次性，且连续世代的子代分布并不独立。然而，当序列（G_n）$_{n\geq 1}$ 为正则时，会出现一个特殊情形。在这种情况下，局部极限是一个条件存活的 Poisson(1) Galton–Watson树，而众所周知，该树经过重新标度后趋于 SSCRT；此外，我们将在注 7.3 中证明，常数 α_W 必定等于 1，这意味着$\frac{1}{\alpha_W}$ 等于 Poisson(1) 子代分布的方差，由此可以推出，在这一情形下这两个操作确实可交换。

对于非正则图序列，这一问题似乎更加微妙。尽管局部极限的根顶点的非骨干邻居数的期望确实为 1，但其方差不一定等于 $\frac{1}{\alpha_W}$。例如，对于完全二分图 $K_{\frac{2n}{3},\frac{n}{3}}$，利用 [16, 定义 1.2] 可以计算出，根顶点的子代数的方差等于 $_{32}$，但$\frac{1}{\alpha_W}$等于 $_{89}$。这并不排除两个操作可交换的可能性，因为后续世代中的方差可能会以适当的方式收敛到$\frac{1}{\alpha_W}$。对于$K_{\frac{2n}{3},\frac{n}{3}}$，我们实际上可以应用 Miermont 的结果 [28]（此时的局部极限实际上是一个具有两种交替类型的 Galton–Watson树：Poi(2) 和 Poi($\frac{1}{2}$)），从而推导出这两个操作确实可交换。然而，在一般情形下，局部极限是一个具有不可数多种类型的 Galton–Watson树；据我们所知，现有的 Galton–Watson树文献尚未涵盖这类树的标度极限。

最后，我们注意到，在 [6], 中，作者考虑了类似的稠密图序列，但没有假设该序列收敛到某个图子。在这一较弱的假设下，他们证明了 UST(G_n) 的直径以高概率为 $\sqrt{n}$ 阶。我们不可能在相同假设下证明缩放极限结果，因为例如可以用一条边连接两个 $K_{n/2}$ 的副本；在这种情况下，直径仍为 $\sqrt{n}$ 阶，但缩放极限并非连续随机树。不过，当这些图具有良好的连通性时，我们可以得到该缩放极限。

本文实际上证明了以下定理。下文中，对于给定的 $\gamma > 0$，若对所有 $U \subset V(G)$， U与 $V(G)$ 之间的边数 $\setminus U$ 至少为$\gamma|U|(|V(G)| - |U|)$，则称图 G 为 γ-扩张图。

定理 1.3。取 $\gamma > 0$ 和 $\delta > 0$，设 $(G_n)_{n\geq 1}$ 是一个由连通的 γ-扩张图组成的稠密序列，其中每个 G_n 有 n 个顶点，且最小度至少为 δn。对于每个 $n \geq 1$，令 $\mathcal{T}_n$ 为 G_n 的一棵均匀生成树。用 $d_{\mathcal{T}_n}$ 表示相应的 $\mathcal{T}_n$ 上的图距离，用 μ_n 表示 $\mathcal{T}_n$ 的顶点上的均匀概率测度。则存在一个序列 $(\alpha_n)_{n\geq 1}$，满足对所有 $n \geq 1$，$1 \leq \alpha_n \leq \delta^{-1}$。

$$\left(\mathcal{T}_n, \frac{\sqrt{\alpha_n}}{\sqrt{n}} d_{\mathcal{T}_n}, \mu_n\right) \xrightarrow{(d)} (\mathcal{T}, d_{\mathcal{T}}, \mu)$$

其中 $\mathcal{T}$、$d_{\mathcal{T}}$、μ 是配备典范质量测度 μ 的连续随机树（CRT），而 $\xrightarrow{(d)}$ 表示度量测度空间之间关于 GHP 距离的依分布收敛。

事实上，定理在稍微更一般的条件下也成立，见注 1.4，不过上述假设使证明更加直接。显然，若不作更强的假设，就无法期望参数 α_n 收敛，因为可以交替选取来自极限值 α_n 不同的序列的图。例如，对于完全图序列 $\alpha_n \to 1$，但若 G_n 换成完全二分图 $K_{\frac{n}{3}, \frac{2n}{3}}$，则 $\alpha_n \to \frac{9}{8}$。

除了重标度直径的收敛之外，由定理 1.3 的 GHP 收敛可直接推出，均匀生成树 UST(G_n) 的重标度高度和重标度简单随机游走也收敛。更正式地说，以下三个收敛均依分布成立。

1. $\frac{\sqrt{\alpha_n}}{\sqrt{n}}$直径$(\mathcal{T}_n) \xrightarrow{(d)}$ 直径$(\mathcal{T})$.
2. $\frac{\sqrt{\alpha_n}}{\sqrt{n}}$高度$(\mathcal{T}_n) \xrightarrow{(d)}$ 高度$(\mathcal{T})$.
3. 若 X_n 是 $\mathcal{T}_n$上的简单随机游走，则 $\left(\frac{\sqrt{\alpha_n}}{\sqrt{n}} X_n(2n^{3/2}\alpha_n^{-1/2}t)\right)_{t\geq 0}$ 的淬火律依分布收敛到连续随机树上布朗运动的淬火律。由此还可推出，相应的混合时间也会在相同的时间尺度上收敛。

有关 GHP 收敛为何能推出这三个性质的更多细节，请参见 [7, 第 1.3]节。在定理 1.1 和推论 1.2 的情形下，上述三个陈述中的 α_n 可以替换为 α_W。

1.1 证明策略

显然，为了证明主要定理，只需先证明定理 1.3，然后证明该图序列是扩张图序列，并证明在定理 1.1 的附加假设下 $\alpha_n \to \alpha_W$。

我们将使用 [8] 的下质量界判据，分两步证明定理 1.3。具体而言，根据[7, 定理 6.5], ，要证明定理 1.3 中的 GHP收敛，只需证明以下两个陈述。

(A) 收敛在有限维意义下成立（这将在定理 3.1 中正式表述）。
(B) 下质量界条件成立；也就是说，若 $m_n(\eta) = \inf_{x\in \mathrm{UST}(G_n)}\left\{\frac{|B(x,\eta\sqrt{n})|}{n}\right\}$，那么对每个$\eta > 0$，序列 $m_n(\eta)^{-1}$ 都是紧的（这将在命题 6.4 中正式表述）。

第二个条件可通过对 [7] 中论证作细微调整，较为直接地得到。本文的主要内容是证明第一个条件。事实上，这个条件等价于：对所有 $k \geq 1$，均匀生成树 UST(G_n) 中均匀随机选取的 $\binom{k}{2}$ 个 k 点之间的距离的联合收敛。

这种收敛此前已在 [31] 中针对一系列高维图的均匀生成树得到证明。这是另一类图，并且假设图具有传递性。他们的证明使用 Wilson 算法，通过运行环擦除随机游走，逐条枝地抽样均匀生成树

（环擦除随机游走）。在他们的证明中，他们将 G_n 上的 Wilson 算法与完全图上的 Wilson 算法耦合起来，并证明两个图上的 $\binom{k}{2}$ 距离集合必定具有相同的标度极限。

相比之下，我们的证明更为直接。我们同样使用 Wilson 算法，但直接处理 UST(G_n)，并利用环擦除随机游走的拉普拉斯随机游走表示来抽样每条分支。通过严格控制环擦除随机游走的容量，我们能够直接计算给定分支超过给定长度的概率，并证明该概率收敛到连续随机树中由 Aldous 的折棍构造得到的对应概率。

注 1.4. 正如定理 1.3 之前的示例和讨论所展示的，为了得到 GHP 拓扑下的收敛，线性最小度这一假设是必要的。为使论述简洁，我们在这些假设下证明上述条件 (A) 和 (B)。不过，条件 (A) 实际上并不需要这一假设。例如，若允许 $o(n)$ 中的顶点度数小于 $\sqrt{n}$，证明仍可原样成立（因为我们在第 5 节分析的环擦除随机游走永远不会到达该集合，高概率成立）。事实上，我们认为，或许可以调整条件 (A) 的证明（定理 3.1），使其适用于 *[31],* 的原始假设；但这需要额外追踪若干繁琐细节，而且不会带来更多见解。

1.2 论文结构

本文的组织结构如下。第 2 节介绍必要的背景知识，包括图子、均匀生成树以及所关注拓扑的简介。第 3 节介绍树的折棍过程的一般框架，并陈述 Aldous 的连续随机树折棍构造。第 4 节给出一些精确的随机游走估计；在第 5 节中，我们利用拉普拉斯随机游走表示，将这些估计应用于 Wilson 算法的前几步，从而得到相应的估计。第 6 节中，我们利用这些估计，将连续随机树上的折棍过程与 Wilson 算法耦合起来，并证明当 n 足够大时，这两个过程非常相似。这证明了上述条件 (A)。我们还解释了如何根据 [7] 的结果推出条件 (B)；事实上，该结果建立了定理 1.3。最后，在第 7 节中，我们证明定理 1.1 和推论 1.2。

1.3 致谢

我们感谢 Asaf Nachmias 和 Jan Hladky 建议研究图子，并提供了许多有益的意见。本研究得到 ERC 整合者基金 101001124（UniversalMap）和 ISF 基金 1294/19 的支持。EA 部分得到了 ANR ProGraM 基金的支持。

2 背景

2.1 图子

图子是一个对称可测函数 $[0,1]^2 \to [0,1]$。如引言中所述，图子由 Borgs、Chayes、Lov´asz、S´os、Szegedy 和 Vesztergombi 引入 [25, 10]，旨在刻画稠密图极限。为了理解这个定义为何自然，我们如下定义离散图 G 的图子表示。假设 G 是一个有 n 个顶点的简单图。将顶点从 v_1 到 v_n 编号，并将区间 $[0,1]$ 划分为一列区间 $(I_i)_{i=1}^n$，其中对于每个 $1 \le i \le n$，$I_i = \left[\frac{i-1}{n}, \frac{i}{n}\right]$。我们定义图子 W_G: $[0,1]^2 \to [0,1]$ 如下（例如，参见 [25, 第 7 节。1]）

$$W_G((x,y)) = \mathbb{1}\{v_{\lceil nx \rceil \vee 1} \sim v_{\lceil ny \rceil \vee 1}\} \qquad \forall\ (x,y) \in [0,1]^2.$$

如果 G 是一个加权图，我们则定义

$$W_G((x,y)) = w(v_{\lceil nx \rceil \vee 1}, v_{\lceil ny \rceil \vee 1}) \qquad \forall\ (x,y) \in [0,1]^2,$$

其中，$w(v_i, v_j)$ 表示连接 v_i 和 v_j 的边的权重（如果不存在这样的边，则为零）。

注意，仅给定 G时， W_G 的这个定义并不唯一，因为它取决于顶点的排序。因此，为了在图子的空间上定义一个度量，我们将改为考虑图子的等价类。具体而言，给定两个图子 W_1 和 W_2，它们之间的割距离定义为（例如，参见 [25, 公式（8.16）]）

$$\delta_{\square}(W_1, W_2) = \inf_{\varphi} ||W_1^{\varphi} - W_2||_{\square},$$

其中，下确界取遍 $[0,1]$, 的所有保测自同构，其中 W^{φ} 由 $W^{\varphi}(x,y) = W(\varphi(x)\ \varphi(y))$ 定义，而可测函数 U: $[0,1]^2 \to [-1,1]$ 的割范数定义为

$$||U||_{\square} = \sup_{S,T\in\mathcal{B}([0,1])} \left| \int_{x\in S} \int_{y\in T} U(x,y)dxdy \right|.$$

因此，我们称确定性图序列 $(G_n)_{n\geq 1}$ 收敛到图子 W，若$\delta_{\square}(W_{G_n}, W) \to 0$，当 $n \to \infty$ 时。

注记 2.1. 事实上，图子可以定义为从 $\Omega^2 \to [0,1]$ 到某处的函数，其中 Ω 是任意概率空间，见 *[25,* 第 *13],* 章；但由于所有概率空间都同构，这并不能带来多大的泛化。

我们将用到以下引理。

引理 2.2. *[9,* 引理 *7]*。设 W 是一个连通图子。那么，对于每个 $\alpha \leq 1/2$，都存在某个常数 $\beta = \beta(W, \alpha)$, 使得对于每个满足 $\alpha \leq \mu(A) \leq 1/2$ 的集合 A，我们有

$$\int_A \int_{A^C} W(x,y)dxdy > \beta.$$

2.1.1 随机图与图子

图子 W 可通过两种方式定义一个含有 n 个顶点的随机图。

1. 在 $[0,1]$上均匀地抽取 $x_1, \ldots, x_n$ 个独立同分布的样本。通过以概率 $W(x_i, x_j)$连接节点 i 和 j，定义一个以 $\{1, \ldots, n\}$ 为顶点集的随机简单图；对每一对（无序）节点 (i, j)，独立进行此操作。我们将得到的随机图记为 $G(n, W)$。
2. 在 $[0,1]$上均匀地抽取 $x_1, \ldots, x_n$ 个独立同分布的样本。对于每一对（无序）节点 (i, j)，在 i 和 j 之间添加一条权重为 $W(x_i, x_j)$ 的边，从而定义一个顶点集为 $\{1, \ldots, n\}$ 的随机加权图。我们将得到的随机图记为 $H(n, W)$。

在这两种构造中，请注意，我们可以用单个图子定义一列随机图。下面的引理指出，无论采用哪种方式，随机抽样 $G(k, W)$ 或 $H(k, W)$ 与 W 之间的割距离都以高概率趋于零，当 $k \to \infty$ 时。

引理 2.3. *[25,* 引理 10.*16].* 固定一个图子 W，对于 $k \geq 1$，令 $G(k, W)$ 和 $H(k, W)$ 如上定义。那么，当 $k \to \infty$ 时， $\delta_{\square}(W_{G(k,W)}, W)$ 和 $\delta_{\square}(W_{H(k,W)}\, W)$ 都依概率趋于 0。

特别地，这意味着我们针对确定性图序列的均匀生成树所证明的结果，会自动推广到满足推论 1.2 假设的 $G(k, W)_{k\geq 1}$ 或 $H(k, W)_{k\geq 1}$ 形式的序列。

例如，经典的 Erd¨os-R´enyi 图 $\mathcal{G}(n, p)$（ $n \geq 1$, $p \in [0,1]$）对应于图 $G(n, W_p)$，其中 W_p 是处处等于 p 的图子。

关于图子的更多背景和应用，请参阅 [25, 第 3]部分。

2.2 混合时间

设 G 为一个具有 n 个顶点的连通加权图，其权重为 $(w(x,y))_{x,\ y\in V(G)}$，且没有自环或重边。在 G 上的随机游走是马尔可夫链 $(X_m)_{m\geq 0}$，满足对所有顶点 $x,y\in V(G)$ 以及所有 $m\geq 1$,

$$\mathbb{P}(X_m=y\mid X_{m-1}=x)=\frac{w(x,y)}{\sum_{z\sim x}w(x,z)},$$

其中 $z\sim x$ 表示 z 是 x的一个邻点。考虑到周期性，有时使用惰性随机游走的概念会更方便。其定义如下:

$$\mathbb{P}(X_m=y\mid X_{m-1}=x)=\frac{w(x,y)}{2\sum_{z\sim x}w(x,z)}\ \forall y\sim x\ \text{ and }\ \mathbb{P}(X_m=x\mid X_{m-1}=x)=\frac{1}{2}$$

对所有 $m\geq 1$均成立。

对于每个 $t\geq 0$，令 p_t 表示惰性随机游走的 t 步转移密度，即对所有 x、$y\in V(G)$，$p_t(x,y)=\mathbb{P}(X_t=y\mid X_0=x)$。我们将 G 的混合时间定义为

$$t_{\mathrm{mix}}(G)=\min\left\{t\geq 0:\max_{x,y\in G}|p_t(x,y)-\pi(x)|\leq\frac{1}{4}\right\},\tag{3}$$

（见 [24, 公式 (4.31)]），其中 π 表示 G上的平稳测度。

我们还需要定义两个概率测度之间的全变差距离，其中一个定义在 μ 上，另一个定义在有限子集 $X\subset V(G)$ 上的 ν 上。其定义为

$$d_{\mathrm{TV}}(\mu,\nu)=\max_{A\subset X}|\mu(A)-\nu(A)|.$$

此外，根据 [24, 第 4 节。5],，对于任意 $k\geq 1$、任意 $t\geq kt_{\mathrm{mix}}$ 以及任意顶点 x，有 t

$$d_{\mathrm{TV}}(p_t(x,\cdot),\pi(\cdot))\leq 2^{-k}.\tag{4}$$

2.3 扩张图

我们将采用以下扩张图的定义。

定义 2.4. (*[16,* 定义 2.*1]*)。对于任意 $\gamma>0$，无自环加权图 G 若满足以下条件，则称为 γ-扩张图：对于所有 $U\subset V(G)$，都有 $w(U,V(G)\setminus U)\geq\gamma|U|(V(G)-|U|)$，其中 $w(A,B)=\sum_{v\in A,u\in B}w(v,u)$。

虽然我们给出了无自环图的定义，但请注意，给图添加自环不会改变其均匀生成树的分布，因为自环不可能出现在均匀生成树中。还要注意，文献中通常采用略有不同的扩展图定义，其中会用到切格常数。我们采用上述定义，因为它更自然地适用于稠密图的框架（正如我们稍后将在命题 7.1 中所示），并且这也是研究 [16] 中局部极限时所采用的定义。

我们将使用的扩张图的主要性质如下。

断言 2.5。设 $\gamma>0$，且设 G 是一个具有 $n\geq 2$ 个顶点的 γ-扩展图。那么，只要 n 足够大（仅取决于 γ），就有

$$t_{\mathrm{mix}}(G)\leq\frac{64}{\gamma^4}\log n.$$

证明。注意，由定义 2.4 可知，G 的最小度至少为 $\frac{\gamma}{2}n$。首先注意，根据[24, 定理 12.4],

$$t_{\mathrm{mix}}(G)\leq t_{\mathrm{rel}}\log\left(\frac{8n}{\gamma}\right),$$

其中 t_{rel} 是 G的弛豫时间。根据切格不等式（各种证明见 [4, 5, 19, 22] ）， $\frac{1}{t_{\mathrm{rel}}}$有下界 $\Phi(G)^2/2$，其中

$$\Phi(G)=\min_{S\subset V(G),\pi(S)\le 1/2}\frac{w(S,V(G)\setminus S)}{\sum_{v\in S}\deg v}.$$

注意， $\pi(S)\le 1/2$ 意味着 $(|V(G)|-|S|)\ge\frac{n\gamma}{4}$。由于 G 是一个 γ-扩展图，因此

$$\Phi(G)\ge\frac{\gamma|S|(|V(G)|-|S|)}{\sum_{v\in S}\deg v}\ge\frac{\gamma|S|(|V(G)|-|S|)}{|S|n}\ge\frac{\gamma^2}{4}.$$

综合所有不等式即可得到结论。 □

2.4 环擦除随机游走与 Wilson 算法

现在我们介绍 Wilson 算法 [32] ，这是一种广泛用于抽样均匀生成树的算法。游走$X=(X_0,\dots X_L)$若长度为 $L\in\mathbb{N}$，则是一个顶点序列，其中对每个 $0\le i\le L-1$，$(X_i,X_{i+1})\in E(G)$。对于满足 a,b为整数的区间 $J=[a,b]\subset[0,L]$，我们用 $X[J]$表示 $\{X_i\}_{i=a}^{b}$。给定游走 X，我们按如下方式归纳定义其环擦除 $Y=\mathrm{LE}(X)$ $=\mathrm{LE}(X[0,L])$。令 $Y_0=X_0$，并令 $\lambda_0=0$。接着，对每个 $i\ge 1$，令 $\lambda_i=1+\max\{t\mid X_t=Y_{\lambda_{i-1}}\}$；若 $\lambda_i\le L$，则令 $Y_i=X_{\lambda_i}$。当 $\lambda_i>L$时，我们终止这一过程。当 X 是加权图 G 上的随机游走，从某个顶点v出发，并在到达另一个顶点 u 时终止（此时L 是随机的），我们称 $\mathrm{LE}(X)$ 是一条从 v 到 u的环擦除随机游走（LERW）。

要从有限的连通加权图 G 中抽样得到一棵均匀生成树，我们首先固定顶点 $V=(v_1,\dots,v_n)$ 的一个排列。先令 T_1 为包含 v_1 且不含任何边的树。然后，对每个 $i>1$，从 v_i 到 T_{i-1} 抽样一条循环擦除随机游走，并令 T_i 为 T_{i-1} 与刚刚抽样得到的循环擦除随机游走的并集。算法在 T_n 处终止。Wilson [32] 证明了 T_n 服从 $\mathrm{UST}(G)$。由此立即可得，$\mathrm{UST}(G)$ 中任意两个顶点之间的路径，服从这两个顶点之间的循环擦除随机游走分布。Pemantle [30] 最先证明了这一结论。

2.5 拉普拉斯随机游走

这里我们概述循环擦除随机游走的拉普拉斯随机游走表示（完整细节见 [26, 第 4 节。1] ）及其在 Wilson 算法中的应用。取一个有限的加权连通图 G，并假设我们已按上述 Wilson 算法抽样得到某个 $j\ge 1$ 对应的 T_j。现在我们从 v_{j+1}到 T_j 抽样一条循环擦除随机游走。将这条循环擦除随机游走记为 $(Y_m)_{m\ge 0}$。再令 X 表示在 G 上的一次随机游走。对于集合 $A\subset G$，令 τ_A 表示 X 命中 A 的时间，令 τ_A^+ 表示 X 首次返回 A 的时间。Y 的拉普拉斯随机游走表示表明，在给定 T_j 且事件 $\{(Y_m)_{m=0}^{i}\cap T_j=\emptyset\}$ 发生的条件下，对任意 $i\ge 0$，有

$$\mathbb{P}\big(Y_{i+1}=v\;\big|\;(Y_m)_{m=0}^{i}\big)=\mathbb{P}_{Y_i}\Big(X_1=v\;\Big|\;\tau_{T_j}<\tau^+_{\cup_{m=0}^{i}\{Y_m\}}\Big)=\frac{\mathbb{P}_{Y_i}(X_1=v)\mathbb{P}_v\Big(\tau_{T_j}<\tau_{\cup_{m=0}^{i}\{Y_m\}}\Big)}{\mathbb{P}_{Y_i}\Big(\tau_{T_j}<\tau^+_{\cup_{m=0}^{i}\{Y_m\}}\Big)}.$$

显然，只有当 $v\notin\bigcup_{m=0}^{i}\{Y_m\}$ 时，这才非零。现在我们可以进一步探究：给定 $(Y_m)_{m=0}^{i}$，对于某个 $H\ge 1$，$(Y_m)_{m=i+1}^{i+H}$ 的分布是什么？特别地，如果 $u_0,u_1,\dots,u_H$ 是 G_n 中的一条简单路径，其中 $\{u_1,\dots,u_{H-1}\}$ 与 $\bigcup_{m=0}^{i}\{Y_m\}\cup T_j$ 和 $u_0=Y_i$ 不相交，那么

$$\mathbb{P}\big((Y_m)_{m=i+1}^{i+H}=(u_m)_{m=1}^{H}\;\big|\;(Y_m)_{m=0}^{i}\big)=\mathbb{P}_{u_0}\big((X_m)_{m=1}^{H}=(u_m)_{m=1}^{H}\big)C\big((Y_m)_{m=0}^{i},T_j,(u_m)_{m=1}^{H}\big),$$

其中

$$C\big((Y_m)_{m=0}^{i},T_j,(u_m)_{m=1}^{H}\big)=\prod_{h=1}^{H}\frac{\mathbb{P}_{u_h}\Big(\tau_{T_j}<\tau_{\cup_{m=0}^{i}\{Y_m\}\cup\,\cup_{m=1}^{h-1}\{u_m\}}\Big)}{\mathbb{P}_{u_{h-1}}\Big(\tau_{T_j}<\tau^+_{\cup_{m=0}^{i}\{Y_m\}\cup\,\cup_{m=1}^{h-1}\{u_m\}}\Big)}.$$

2.6 容量与接近性

回忆一下， G 是一个有 n 个顶点且最小度至少为 δn的连通加权图。 G 中顶点集的容量量化了随机游走到达该集合的难度。设$(X_i)_{i\geq 0}$ 是 G 上的随机游走，对于 $U \subset V(G)$，令 $\tau_U = \inf\{i \geq 0 : X_i \in U\}$。给定 $k \geq 0$，我们将 U 的 k-容量定义为 $\mathrm{Cap}_k(U) = \mathbb{P}_\pi(\{v18)$。

这里汇总一些关于容量的有用事实。

引理 2.6. 设 $A \subset V(G)$ 和 $k \geq 1$。那么

$$\mathrm{Cap}_k(A) \leq k\pi(A) \leq \frac{k|A|}{\delta n}. \tag{5}$$

此外，若 $k|A| \leq \frac{\delta^3 n}{2}$，则

$$\mathrm{Cap}_k(A) \geq \frac{k\pi(A)}{2} \geq \frac{\delta k|A|}{2n}. \tag{6}$$

证明. 上界由并集界得到。下界由 Bonferroni 不等式以及度的下界推出，它们意味着

$$\mathrm{Cap}_k(A) \geq k\pi(A) - \left(\frac{k|A|}{\delta n}\right)^2 \geq \frac{\delta k|A|}{2n}. \qquad \square$$

我们还将用到以下断言。

断言 2.7. 设 $t_{\mathrm{mix}} = t_{\mathrm{mix}}(G)$。设 $A \subset V(G)$，令 $M \geq (\log n)^2 t_{\mathrm{mix}}$，并假设 $(\log n)^2 \cdot t_{\mathrm{mix}}|A| \leq n$。那么，只要 n 足够大，

$$\sup_{u \in V(G)\setminus A} |\mathbb{P}_u(\tau_A \leq M) - \mathrm{Cap}_M(A)| \leq \frac{3\log n \cdot t_{\mathrm{mix}}|A|}{\delta n}.$$

证明。设 X 是一条从 $u \in G$ 出发的随机游走。显然，对于任意 $t \geq 0$， X 的前 t 步可以与一条惰性随机游走 X 的前 t 个非重复步进行耦合。因此，先运行一条从 u 出发的惰性随机游走，直到时间 $T = 2\log n \cdot t_{\mathrm{mix}}$。令 N 表示这条˜惰性随机游走的非重复跳跃总数。根据 (4)， X_t 的分布几乎是平稳分布。此外，我们确定有 $0 \leq N \leq T$。要抽样 $(X_t)_{t=0}^M$，我们先按照上述方法将其与 $(\tilde{X}_t)_{t=0}^T$ 的前 N 步进行耦合，然后再运行 X，继续走 $M - N$ 步。在这种耦合下，由并集界可得

$$\mathbb{P}_u(\tau_A \leq M) \leq \frac{2\log n \cdot t_{\mathrm{mix}}|A|}{\delta n} + \mathbb{P}_\pi(\tau_A \leq M) + 2^{-2\log n} \leq \mathrm{Cap}_M(A) + \frac{3\log n \cdot t_{\mathrm{mix}}|A|}{\delta n}.$$

同样，

$$\mathbb{P}_u(\tau_A \leq M) \geq \mathbb{P}_\pi(\tau_A \leq M - T) - 2^{-2\log n} \geq \mathrm{Cap}_M(A) - \frac{3\log n \cdot t_{\mathrm{mix}}|A|}{\delta n}.$$

$\square$

为了 获得关于容量的下界，我们定义集合 U 和 W b 的 k-接近度 y

$$\mathrm{Close}_k(U, W) = \mathbb{P}_\pi(\tau_U < k, \tau_W < k). \tag{7}$$

推论 2.8. 对任意不相交集合 U、 $W \subset G$，有

$$\sup_{v \in G\setminus(U\cup W)} \mathbb{P}_v(\tau_U < k, \tau_W < k) \leq \frac{2k^2|U||W|}{\delta^2 n^2}.$$

特别地，$\mathrm{Close}_k(U, W) \leq \frac{2k^2|U||W|}{\delta^2 n^2}$。

证明。注意到

$$\begin{aligned}\sup_{v\in G\setminus(U\cup W)}\mathbb{P}_v(\tau_U<k,\tau_W<k)&\le\sup_{v\in G\setminus(U\cup W)}\{\mathbb{P}_v(\tau_U<\tau_W<k)+\mathbb{P}_v(\tau_W<\tau_U<k)\}\\&\le\sup_{v\in G\setminus(U\cup W),u\in U,w\in W}\{\mathbb{P}_v(\tau_U<k)\mathbb{P}_u(\tau_W<k)+\mathbb{P}_v(\tau_W<k)\mathbb{P}_w(\tau_U<k)\}\\&\le\frac{2k^2|U||W|}{(\delta n)^2}.\end{aligned}$$

□

2.7 随机变量

在这里，我们介绍两个将在第 6 节中发挥作用的基本结果。

命题2.9. 设 $\varepsilon>0$，并设 $0<a<b$，其中 $b-a\le\varepsilon$。令 $X_a\sim U([0,a])$ 和 $X_b\sim U([0,b])$。那么，我们可以将 X_a 和 X_b 耦合，使得 $\mathbb{P}(|X_a-X_b|>\varepsilon)<\varepsilon$。

证明. 我们取 $X_b=\frac{b}{a}X_a$。于是，$|X_b-X_a|=|\frac{b-a}{a}\cdot X_a|\le|b-a|\le\varepsilon$。 □

引理 2.10. 对于任意 $L>0$，令 X_L 为定义在 $(0,\infty)$ 上且满足以下条件的随机变量：

$$\mathbb{P}(X_L>x)=\exp\left\{-\frac{(x+L)^2-L^2}{2}\right\}.$$

那么，对于任意 $\delta>0$，存在 $\eta=\eta(\delta,L)>0$，使得以下结论成立。令 Y 为定义在 $(0,\infty)$ 上的另一个随机变量，并假设对于所有 $x>0$，

$$|\mathbb{P}(X_L>x)-\mathbb{P}(Y>x)|<\eta. \tag{8}$$

于是，这意味着我们可以将 X_L 和 Y 耦合起来，使得 $\mathbb{P}(|X_L-Y|>\delta)<\delta$。

此外，对于任意满足 $L_1<L_2$ 的 δ、L_1 和 L_2，存在 $\eta=\eta(\delta,L_1,L_2)$，使得对于每个 $L\in[L_1,L_2]$，我们都可以按照上述方式将 X_L 和 Y 耦合起来。

证明。注意，我们可以通过先抽取 $U\sim\text{Uniform}([0,1])$，再令

$$X_L(\omega)=\sup_{x\ge0}\{\mathbb{P}(X_L\ge x)\ge U(\omega)\},\qquad Y(\omega)=\sup_{x\ge0}\{\mathbb{P}(Y\ge x)\ge U(\omega)\}.$$

现在选择$K_{\delta,L}<\infty$，使得 $\mathbb{P}(X_L\ge K_{\delta,L})<\delta$。不失一般性，假设 $\delta<1$ 和 $K_{\delta,L}>1$；否则，必要时将它们减小或增大。注意，对于所有 $0\le x<K_{\delta,L}$，我们有

$$\exp\left\{-\frac{(x+L)^2-L^2}{2}\right\}-\exp\left\{-\frac{(x+\delta+L)^2-L^2}{2}\right\}\ge\delta(x+L)\exp\left\{-\frac{(x+\delta+L)^2-L^2}{2}\right\}\ge M_{\delta,L},$$

其中 $M_{\delta,L}=\delta L\exp\left\{-\frac{(2K_{\delta,L}+L)_2^2-L^2}{}\right\}>0$。

现在假设 (8) 成立且 $\eta<M_{\delta,L}$。那么，对于任意 $0\le x<K_{\delta,L}$，我们有

$$\mathbb{P}(Y\ge x+\delta)\le\mathbb{P}(X_L\ge x+\delta)+\eta\le\mathbb{P}(X_L\ge x)-M_{\delta,L}+\eta\le\mathbb{P}(X_L\ge x).$$

因此，在该耦合下，对于任意 $x<K_{\delta,L}$，我们有

$$\{X_L\le x\}\Leftrightarrow\{\mathbb{P}(X_L\ge x)\le U\}\Rightarrow\{\mathbb{P}(Y\ge x+\delta)\le U\}\Leftrightarrow\{Y\le x+\delta\}.$$

类似地，$\{X_L\ge x\}\Rightarrow\{Y\ge x-\delta\}$。因此，在这种耦合下，我们有

$$\mathbb{P}(|X_L-Y|>\delta)\le\mathbb{P}(X_L\ge K_{\delta,L})<\delta,$$

这正是所需的。

对于第二个断言，注意对每个 $L'>L$，我们也有 $\mathbb{P}(X_{L'}\ge K_{\delta,L})<\delta$。因此，对于区间 $[L_1,L_2]$，我们只需在整个区间上使用 K_{δ,L_1} 和 M_{δ,L_1}。 □

2.8 GHP 拓扑

这里我们定义 GHP 拓扑。我们采用 [29, 第 1.3 节和 6] 中的框架，并在空间 $\mathbb{X}_c$中讨论度量测度空间（mm-空间）的等价类 (X, d, μ)，其中 (X, d) 是紧致度量空间，且 μ 是其上的博雷尔概率测度。若存在双射等距映射 $\phi\colon X \to X'$ 使得 $\phi_*\mu = \mu'$（这里 $\phi_*\mu$ 是 μ 在 ϕ下的推前测度），则称 (X, d, μ) 与 (X', d', μ') 等价。为简化记号，我们用 $\mathbb{X}_c$ 中的一个等价类代表元来表示该等价类。

首先回顾一下，若 (X, d) 是一个度量空间，则两个集合 $A, A' \subset X$ 之间的豪斯多夫距离 d_H 定义为

$$d_H(A, A') = \max\{\sup_{a\in A} d(a, A'),\ \sup_{a'\in A'} d(a', A)\}.$$

对于 $\varepsilon > 0$ 和 $A \subset X$，我们还记 $A^\varepsilon = \{x \in X\colon d(x, A) < \varepsilon\}$ 为 X 中 A 的 ε-膨胀。若 μ 和 ν 是 X 上的两个测度，则它们之间的普罗霍罗夫距离定义为

$$d_P(\mu, \nu) = \inf\{\varepsilon > 0 : \mu(A) \le \nu(A^\varepsilon) + \varepsilon \text{ and } \nu(A) \le \mu(A^\varepsilon) + \varepsilon \text{ for any closed set } A \subset X\}.$$

定义 2.11。设 (X, d, μ) 和 (X', d', μ') 是 $\mathbb{X}_c$ 的元素。(X, d, μ) 与 (X', d', μ') 之间的格罗莫夫-豪斯多夫-普罗霍罗夫（GHP）距离定义为

$$d_{\mathrm{GHP}}((X, d, \mu), (X', d', \mu')) = \inf\{d_H(\phi(X), \phi'(X')) \vee d_P(\phi_*\mu, \phi'_*\mu')\},$$

其中，下确界取遍所有等距嵌入 $\phi\colon X \to F$, $\phi'\colon X' \to F$ 到某个公共度量空间 F 中的情形。

回顾一下，本文的目标是证明关于 GHP 拓扑的分布收敛。给定一个 mm-空间（X、d、μ）和一个固定的 $m \in \mathbb{N}$，我们将 $\nu_m((X、d、\mu))$ 定义为 $\mathbb{R}^{\binom{m}{2}}$ 上的一个测度，即按照 μ 抽取 m 个独立同分布点后，这些点之间 $\binom{m}{2}$ 两两距离的分布。因而，每个 $\mathbb{P}$ 在 $\mathbb{X}_c$ 上定义的分布都会在 $\mathbb{R}^{\binom{m}{2}}$ 上定义随机测度 $(\nu_m)_{m\ge 2}$ 和退火测度 $(\tilde{\nu}_m)_{m\ge 2}$，其定义为

$$\tilde{\nu}_m(\mathbb{P}) := \int_{\mathbb{X}_c} \nu_m((X, d, \mu))\, d\mathbb{P}.$$

在 [7] 中，我们将 [8, 定理 6.1] 的一个结果改写为分布意义下的表述，从而用测度 $(\tilde{\nu}_m)_{m\ge 2}$ 的收敛和一个体积条件来刻画 GHP 收敛。为陈述本文将使用的版本，给定 $c > 0$ 和一个 mm-空间（X、d、μ），我们定义

$$m_c((X, d, \mu)) = \inf_{x\in X}\{\mu(B(x, c))\}$$

（参见 [8, 节 3]）。

在下一个命题的证明中，我们还将提及格罗莫夫–普罗霍罗夫拓扑（它更粗糙），其定义如下。

定义 2.12. 设 (X, d, μ) 和 (X', d', μ') 是 $\mathbb{X}_c$ 的元素。(X, d, μ) 与 (X', d', μ') 之间的 Gromov–Prohorov（GP）距离定义为

$$d_{\mathrm{GP}}((X, d, \mu), (X', d', \mu')) = \inf\{d_P(\phi_*\mu, \phi'_*\mu')\},$$

其中下确界取遍所有等距嵌入 $\phi\colon X \to F$, $\phi'\colon X' \to F$，使其嵌入到某个共同的度量空间 F 中。

关键结果如下。

命题 2.13. 设 (X, d, μ) 是 $\mathbb{X}_c$ 中的一个元素，其分布为 $\mathbb{P}$，且 μ 几乎必然具有全支撑。设 $((X_n, d_n, \mu_n))_{n\ge 1}$ 是 $\mathbb{X}_c$ 中的一个序列，其相应的分布为 $(\mathbb{P}_n)_{n\ge 1}$，并假设：

(a) 对所有 $m \geq 0$, $\tilde{\nu}_m(\mathbb{P}_n) \to \tilde{\nu}_m(\mathbb{P})$，当 $n \to \infty$ 时均成立。(b) 对任意 $c > 0$，序列 $\left(m_c((X_n, d_n, \mu_n))^{-1}\right)_{n\geq 1}$ 是紧的。

那么，关于 GHP 拓扑，$(X_n, d_n, \mu_n) \xrightarrow{(d)} (X, d, \mu)$。

证明。首先，我们通过验证 [15, 推论 3.1]中的两个条件，说明 (a) 和 (b) 一起蕴含关于 GP 拓扑的 $(X_n, d_n, \mu_n) \xrightarrow{(d)} (X, d, \mu)$。[15, 推论 3.1] 的第二个条件恰好就是 (a)。为验证第一个条件，我们还要用到 [15, 定理 3]（回顾一下，根据普罗霍罗夫定理，测度的相对紧性等价于其紧性），并验证其中的条件 (i) 和 (ii)（另见 [15]中的命题 8.1）。条件 (i) 仅仅是说，$\tilde{\nu}_2$ 是 $\mathbb{R}$上的一个紧测度序列，这由 (a) 得出。最后，(b) 直接蕴含条件 (ii)。

Th因此，由 [7, 定理 6.5], ，这些空间关于 GHP 拓扑收敛 . □

3 树的断棍构造

我们的首要目标是 证明命题 2.13 的条件 (a)，它等价于以下陈述 .

定理 3.1. 取 $\gamma > 0$ 和 $\delta > 0$，并设 $(G_n)_{n\geq 1}$ 是一个 γ-扩张图的稠密序列，其中每个 G_n 有 n 个顶点，且最小度至少为 δn。用 $d_{\mathcal{T}_n}$ 表示 $\mathcal{T}_n$ 上的图距离，并用 $(\mathcal{T}, d, \mu)$ 表示连续随机树。则存在一个序列 $(\beta_n)_{n\geq 1}$，满足对所有 $n \geq 1$ 均有 $\sqrt{\delta} \leq \beta_n \leq 1$；对于任意固定的 $k \geq 1$，若从 G_n 中均匀选取 $\{x_1, \ldots, x_k\}$ 个相互独立的顶点，则距离

$$\frac{d_{\mathcal{T}_n}(x_i, x_j)}{\beta_n \sqrt{n}}$$

联合依分布收敛到 $\mathcal{T}$ 中 $\binom{k}{2}$ 距离的分布，其中距离取自按照 μ 抽取的 k 独立同分布点之间。

为证明这一 k 定理，我们将使用 Aldous 的连续随机树（CRT）折棍构造；该构造尤其适用于处理一组 k 均匀点之间的两两距离。我们的策略是证明，当 n 较大时，Wilson 算法在 G_n 上的前 k 步与这一断棍过程十分接近。本节将简要回顾连续随机树的断棍构造及其若干关键性质。

我们先更一般地介绍如何利用实数轴上的线段构造树序列。

定义 3.2。（树序列的折棍构造）。令 $y_0 = z_0 = 0$，并假设有一列点 y_1、y_2、$\ldots \in [0, \infty)$以及 z_1、z_2、$\ldots \in [0, \infty)$，满足对所有$i \geq 1$都有 $y_{i-1} < y_i$ 且 $z_i \leq y_i$。按如下方式构造树。从时刻 1 开始，取线段 $[y_0, y_1]$。这就是 $T^{(2)}$（因为它包含两个标记点）。接下来进行归纳构造。在时刻 $i \geq 2$，取区间 $[y_{i-1}, y_i]$，并将区间 $[y_{i-1}, y_i]$的底端连接到 $T^{(i)}$ 上与 z_{i-1}对应的点。由此得到一棵新的树，它有 $i+1$个标记点（与集合$(y_j)_{j=0}^{i}$一一对应），记作 $T^{(i+1)}$。

给定两个这样的序列和任意一个 $k \geq 2$，我们将 $\mathrm{SB}^{(k)}((y_0, y_1, y_2, \ldots), (z_0, z_1, z_2, \ldots))$，或等价地将 $\mathrm{SB}^{(k)}((y_0, y_1, y_2, \ldots, y_{k-1}), (z_0, z_1, z_2, \ldots, z_{k-2}))$，定义为等于树 $T^{(k)}$。

一般来说，上述方式构造出的树序列未必收敛，但 Aldous 指出，只要以恰当的方式选择这些点，我们实际上就能通过折棍过程构造出连续随机树。

命题 3.3. *[1,过程 3]*。令 $Y_0 = Z_0 = 0$，设 $(Y_1, Y_2, \ldots)$ 表示区间 $[0, \infty)$ 上强度为 $t\ dt$ 的非齐次泊松过程的有序点集，并对每个 $i \geq 1$，令 Z_i 在区间 $[0, Y_i)$ 上均匀选取。按定义 3.2 构造序列 $(T_{(k)})_{k=2}^{\infty}$。则 $T^{(k)}$ 的极限的闭包依分布等于连续随机树。此外，若在 $k-1$ 步后停止该过程，则所得的树 $T^{(k)}$ 与连续随机树中由 k 个均匀点张成的子树具有相同的分布，而与集合 $(Y_i)_{i=0}^{k-1}$ 对应的点可视为连续随机树中的 k 个均匀点。

特别地，对应于$(Y_i)_{i=1}^k$ 的点之间的 $\binom{k}{2}$ 个两两距离所成的集合，依分布等于连续随机树中 k 个均匀点之间的 $\binom{k}{2}$ 个两两距离所成的集合。

以下命题对于稍后与 Wilson 算法进行比较至关重要。它可以通过直接计算验证。

命题 3.4。定义序列 $(Y_1, Y_2, \ldots)$ 如命题 3.3 中所述。那么，对于任意 $k \geq 1$ 和任意 $x \geq 0$,

$$\mathbb{P}\big(Y_{k+1} - Y_k \geq x \;\big|\; (Y_i)_{i=0}^k\big) = \exp\left\{-\frac{1}{2}\big((Y_k + x)^2 - Y_k^2\big)\right\}.$$

下面的引理也会很有用。

引理 3.5。存在函数 $f\colon [0,\infty) \times \mathbb{N} \to [0,1]$，使得对于每个 $k \in \mathbb{N}$，都有 $\lim_{C\to\infty} f(C,k) \to 0$；并且若 Y_k 如命题 3.3 中所述，则

$$\mathbb{P}\big(C^{-1} \leq Y_k \leq C\big) \geq 1 - f(C,k).$$

命题 3.6。设 $(y_0, y_1, y_2, \ldots)$、$(z_0, z_1, z_2, \ldots)$ 和 $(y'_0, y'_1, y'_2, \ldots)$、$(z'_0, z'_1, z'_2, \ldots)$ 分别是按照定义 3.2 所定义的两个独立折棍过程的输入。任取 $k \geq 1$，并令 $T^{(k+1)}$ 和 $T^{(k+1)'}$ 分别为两个过程经过 k 步后形成的树。令 d 和 d' 分别表示 $T^{(k+1)}$ 和 $T_{(k+1)'}$ 上的距离。

固定某个 $\varepsilon > 0$，并假设以下条件成立。

(i) 对所有 $i \leq k$ 和所有 $i \leq k-1$，$|y_i - y'_i| \leq \varepsilon$ 和 $|z_i - z'_i| \leq \varepsilon$ 均成立，

(ii) 对所有 $i \leq k-1$、$j \leq k$，$|z_i - y_j| \geq 3\varepsilon$ 均成立。

那么，对所有 $0 \leq i$、$j \leq k$，有

$$|d(y_i, y_j) - d'(y'_i, y'_j)| \leq 2k\varepsilon.$$

证明。当条件 (i) 和 (ii) 成立时，对所有 $i \leq k-1$、$j \leq k$，$y_j \leq z_i \leq y_{j+1}$ 当且仅当 $y'_j \leq z'_i \leq y'_{j+1}$。我们断言，这意味着对所有 i、$j \leq k+1$，$|d(y_i, y_j) - d'(y'_i, y'_j)| \leq 2k\varepsilon$。事实上，根据构造可知，$d(y_i, y_j)$ 是 $T^{(k+1)}$ 中至多 k 条分支线段的长度之和，且它们的长度都可以写成 $|y_j - y_{j-1}|$、$|z_j - y_\ell|$ 或 $|z_j - z_\ell|$。此外，根据构造，当条件 (i) 和 (ii) 成立时，$d'(y'_i, y'_j)$ 可以写成相同的和，只需将每个 z_j 替换为 z'_j，并将每个 y_j 替换为 y'_j。因此，由三角不等式可得 $|d(y_i, y_j) - d'(y'_i, y'_j)| \leq 2k\varepsilon$。 □

4 随机游走性质

本节证明一些关于随机游走命中概率和容量的结果，之后我们将利用第 2.5 节中的拉普拉斯随机游走表示，将这些结果推广到循环擦除随机游走的线段上。

在本节中，我们固定一个较小的 $\kappa \in (0, 1/32)$，并对 $n \geq 1$，令 $M_n = n^\kappa$。下文中，我们将直接写作 M，而不写 M_n。

记号说明。对于本节中的陈述，我们取一个满足定理 1.3 假设的图序列，因此它与两个正常数 $\gamma > 0$ 和 $\delta > 0$ 相关联。本节中，我们将这些常数视为固定值，因此 $o(\cdot)$ 和 $O(\cdot)$ 量也可能依赖于 γ 和 δ。

4.1 命中概率

我们先介绍一些关于命中概率的结果。设 X 是在 G_n 上的一个（非惰性）随机游走，其中 $n \geq 1$。对于集合 $A \subset V(G_n)$，我们定义

$$\tau_A = \inf\{t \geq 0 : X_t \in A\}.$$

主要引理如下。

引理 4.1. 取 $\gamma > 0$ 和 $\delta > 0$，并设 $(G_n)_{n\geq 1}$ 为一个由 γ-扩张图组成的稠密序列，其中每个 G_n 有n 个顶点，且最小度至少为 δn。取 κ和 M 为第 4 节开头所定义的对象。则存在一个序列 $(\eta_n)_{n\geq 1}$，满足 $\eta_n \to 0$，且仅依赖于 δ 和 γ，使得对于任意满足 $|A| + |B| \leq \frac{\delta^3}{2} n^{\frac{1}{2}+2\kappa}$的互不相交的 A、$B \subset G_n$:

$$\left| \mathbb{P}_\pi(\tau_A < \tau_B) - \frac{\mathrm{Cap}_M(A)}{\mathrm{Cap}_M(A) + \mathrm{Cap}_M(B)} \right| \leq \frac{\mathrm{Cap}_M(A)\eta_n}{\mathrm{Cap}_M(A) + \mathrm{Cap}_M(B)}.$$

证明. 设 $(X_i)_{i=1}^M$ 为长度为 M的随机游走。那么，由贝叶斯公式、推论 2.8 以及引理 2.6 中的下界可得，

$$\begin{aligned}
\mathbb{P}_\pi(\tau_A < M \mid \tau_A \wedge \tau_B < M) &= \frac{\mathrm{Cap}_M(A)}{\mathrm{Cap}_M(A) + \mathrm{Cap}_M(B) - \mathbb{P}_\pi(\tau_A \vee \tau_B < M)} \\
&= \frac{\mathrm{Cap}_M(A)}{\mathrm{Cap}_M(A) + \mathrm{Cap}_M(B)} \left(1 + O\left(\frac{\delta^{-3}M|B|}{n}\right)\right). \\
\mathbb{P}_\pi(\tau_A \vee \tau_B < M \mid \tau_A \wedge \tau_B < M) &= \frac{\mathbb{P}_\pi(\tau_A \vee \tau_B < M)}{\mathrm{Cap}_M(A) + \mathrm{Cap}_M(B) - \mathbb{P}_\pi(\tau_A \vee \tau_B < M)} \\
&\leq \mathbb{P}_\pi(\tau_A < M \mid \tau_A \wedge \tau_B < M) O\left(\frac{\delta^{-3}M|B|}{n}\right).
\end{aligned}$$

因此，综合以上结果并应用引理 2.6，可得:

$$\begin{aligned}
\mathbb{P}_\pi(\tau_A < \tau_B \mid \tau_A \wedge \tau_B < M) &= \mathbb{P}_\pi(\tau_A < M \mid \tau_A \wedge \tau_B < M) + O(\mathbb{P}_\pi(\tau_A \vee \tau_B < M \mid \tau_A \wedge \tau_B < M)) \\
&= \frac{\mathrm{Cap}_M(A)}{\mathrm{Cap}_M(A) + \mathrm{Cap}_M(B)} \left(1 + O\left(\frac{\delta^{-3}M|B|}{n}\right)\right).
\end{aligned}$$

类似地，由断言 2.7 和引理 2.6 中的下界可知，对于所有 $u \in G_n \backslash (A \cup B)$，该结论均成立:

$$\mathbb{P}_u(\tau_A < \tau_B \mid \tau_A \wedge \tau_B < M) = \frac{\mathrm{Cap}_M(A)}{\mathrm{Cap}_M(A) + \mathrm{Cap}_M(B)} \left(1 + O\left(\frac{\delta^{-3}M|B|}{n} + \frac{t_{\mathrm{mix}} \cdot \log n}{\delta^2 M}\right)\right).$$

现在我们将时间分解为长度为 M 的区间。对于每个 $i \geq 1$，定义区间 A_i 为

$$A_i = [iM, (i+1)M].$$

于是，根据推论 2.8，我们有:

$$\begin{aligned}
\mathbb{P}_\pi(\tau_A < \tau_B) &\geq \sum_{i=0}^{\infty} \mathbb{P}_\pi(\tau_A < \tau_B \mid \tau_{A\cup B} \in A_i) \mathbb{P}_\pi(\tau_{A\cup B} \in A_i) \\
&\geq \sum_{i=0}^{\infty} \inf_{u \in G_n \backslash (A\cup B)} \mathbb{P}_u(\tau_A < \tau_B \mid \tau_{A\cup B} \in A_0) \mathbb{P}_\pi(\tau_{A\cup B} \in A_i) \\
&\geq \frac{\mathrm{Cap}_M(A)}{\mathrm{Cap}_M(A) + \mathrm{Cap}_M(B)} \left(1 + O\left(\frac{\delta^{-3}|B|M}{n} + \frac{t_{\mathrm{mix}} \cdot \log n}{\delta^2 M}\right)\right)
\end{aligned}$$

我们推得，对于所有允许的 A 和 B，均匀地有:

$$\mathbb{P}_\pi(\tau_A < \tau_B) \geq \frac{\mathrm{Cap}_M(A)}{\mathrm{Cap}_M(A) + \mathrm{Cap}_M(B)} (1 - o_{\delta,\gamma}(1)), \tag{9}$$

其中 $o_{\delta,\gamma}(1)$ 项对所有 A 和 B 均一致，但可能依赖于 δ 和 γ。类似地，要给出 $\mathbb{P}_\pi(\tau_A<\tau_B)$ 的上界，只需交换 A 和 B 的角色。我们推得，对于所有允许的 A和 B，均匀地有：

$$\mathbb{P}_\pi(\tau_A<\tau_B)=\frac{\mathrm{Cap}_M(A)}{\mathrm{Cap}_M(A)+\mathrm{Cap}_M(B)}(1-o_{\delta,\gamma}(1)). \tag{10}$$

□

我们还需要以下一个小的改动。

引理 4.2. 取 $\gamma>0$ 和 $\delta>0$，并设 $(G_n)_{n\geq 1}$ 是一个由 γ-扩展图组成的稠密序列，其中每个 G_n 都有 n 个顶点，且最小度至少为 δn。取 κ 和 M，其定义见第 4 节开头。那么，对于任意满足 $|A|+|B|\leq \frac{\delta^3}{2}n^{\frac{1}{2}+2\kappa}$ 的不相交的 A、$B\subset G_n$，每个 $u\in G_n\setminus(A\cup B)$ 以及每个 $v\in G_n\setminus A$，我们都有

$$\mathbb{P}_u(\tau_A<\tau_B)=\mathbb{P}_\pi(\tau_A<\tau_B)(1+o(n^{3\kappa-1/2}t^+_{\mathrm{mix}}))\quad and\quad \mathbb{P}_v\big(\tau_A<\tau_B^+\big)=\mathbb{P}_\pi(\tau_A<\tau_B)(1+o(n^{3\kappa-1/2}t^+_{\mathrm{mix}})).$$

证明. 我们先对惰性随机游走证明第一个结论，因为这与原结论等价，并将这种惰性随机游走记为 X。在整个证明过程中，我们还将使用以下记号。对于集合 $C\subset G_n$ 和某个时刻 $t\geq 0$，我们用 $\tau(C,t)$ 表示严格大于 t 且满足 $X_s\in C$ 的最早时刻 s。此外，记 $t^+{}_{\text{混合}}$ 为 $\log_2^2(n)t_{\mathrm{mix}}$，因此由 (4) 可知，对于每个 $u\in G_n$，

$$d_{\mathrm{TV}}(p_{t^+_{\mathrm{mix}}}(u,\cdot),\pi(\cdot))\leq n^{-\log(2)\log(n)}.$$

现在令 $u\in G_n\setminus A$。我们先给出 $\mathbb{P}_u(\tau_A<\tau_B)$ 的一个下界。我们有

$$\begin{aligned}\mathbb{P}_u(\tau_A<\tau_B)&\geq \mathbb{P}_u\big(t^+_{\mathrm{mix}}<\tau_A<\tau_B\big)\\&\geq \mathbb{P}_u\big(\tau(A,t^+_{\mathrm{mix}})<\tau(B,t^+_{\mathrm{mix}})\big)-\mathbb{P}_u\big(\tau_{A\cup B}<t^+_{\mathrm{mix}}<\tau(A,t^+_{\mathrm{mix}})<\tau(B,t^+_{\mathrm{mix}})\big)\end{aligned} \tag{11}$$

注意，由 (4) 可知，第一项可以下界为 $\mathbb{P}_\pi(\tau_A<\tau_B)-n^{-\log(2)\log(n)}$。对于第二项，我们来给事件 $\{\tau_{A\cup B}<t^+{}_{\text{混合}}<\tau(A,t_{\mathrm{mix}})<\tau(B,t^+{}_{\text{混合}})\}$ 的概率取一个上界。利用并集界，我们得到

$$\begin{aligned}&\mathbb{P}_u\big(\tau_{A\cup B}<t^+_{\mathrm{mix}}<\tau(A,t^+_{\mathrm{mix}})<\tau(B,t^+_{\mathrm{mix}})\big)\\&\leq \mathbb{P}_u\big(\tau_{A\cup B}<t^+_{\mathrm{mix}}<\tau(A,t^+_{\mathrm{mix}})<2t^+_{\mathrm{mix}}\big)+\mathbb{P}_u\big(\tau_{A\cup B}<t^+_{\mathrm{mix}}\text{ and }\tau(A,2t^+_{\mathrm{mix}})<\tau(B,2t^+_{\mathrm{mix}})\big).\\&\leq \frac{|A|t^+_{\mathrm{mix}}}{\delta n}+\frac{(|A|+|B|)t^+_{\mathrm{mix}}}{\delta n}\cdot(\mathbb{P}_\pi(\tau_A<\tau_B)+n^{-\log(2)\log(n)}).\end{aligned}$$

注意，由引理 2.6 和引理 4.1 可知

$$\frac{|A|t^+_{\mathrm{mix}}}{\mathbb{P}_\pi(\tau_A<\tau_B)\delta n}\leq\frac{|A|t^+_{\mathrm{mix}}}{\delta n}\cdot\frac{2(|A|+|B|)}{\delta^2|A|}\leq\frac{2(|A|+|B|)t^+_{\mathrm{mix}}}{\delta^3 n},$$

因此，由断言 2.5 可知

$$\frac{|A|t^+_{\mathrm{mix}}}{\delta n}=\mathbb{P}_\pi(\tau_A<\tau_B)\cdot O_\gamma(n^{2\kappa-1/2}t^+_{\mathrm{mix}})=\mathbb{P}_\pi(\tau_A<\tau_B)\cdot o_\gamma(n^{3\kappa-1/2}t^+_{\mathrm{mix}}) \tag{12}$$

将所有结果代回 (11)，因此可推出

$$\mathbb{P}_u(\tau_A<\tau_B)\geq\mathbb{P}_\pi(\tau_A<\tau_B)(1+o(n^{3\kappa-1/2}t^+_{\mathrm{mix}})).$$

对于 $\mathbb{P}_u(\tau_A<\tau_B)$ 的一个上界，我们只需写出

$$\mathbb{P}_u(\tau_A<\tau_B)\leq\mathbb{P}_u\big(\tau_A<t^+_{\mathrm{mix}}\big)+\mathbb{P}_u\big(t^+_{\mathrm{mix}}<\tau(A,t^+_{\mathrm{mix}})<\tau(B,t^+_{\mathrm{mix}})\big)\leq\frac{|A|t^+_{\mathrm{mix}}}{\delta n}+\mathbb{P}_\pi(\tau_A<\tau_B)+n^{-\log(2)\log(n)}.$$

再次利用（12），我们得到

$$\mathbb{P}_u(\tau_A<\tau_B)=\mathbb{P}_\pi(\tau_A<\tau_B)(1+o(n^{3\kappa-1/2}t^+_{\mathrm{mix}})).$$

对于第二个陈述，同样只需证明惰性随机游走的情形：采用完全相同的证明，将 τ_B^+ 替换为至少进行一次非惰性步之后首次命中 B 的时间。 □

4.2 容量

这里我们证明随机游走的容量与接近度的一些类似性质。 .

本节还可以引入定理 1.3 中出现的序列 $(\alpha_n)_{n\geq 1}$ 。给定图序列 $(G_n)_{n\geq 1}$，取第 4 节开头定义的 $M = n^{\kappa}$ ，令 X 为 G_n 上的随机游走，并对每个 $n \geq 1$ ，定义

$$\alpha_n = \frac{n\mathbb{E}_{\pi}\big[\mathrm{Cap}_M(X[0, n^{\kappa/2}))\big]}{Mn^{\kappa/2}}.$$

命题 4.3。取 $\gamma > 0$ 和 $\delta > 0$ ，令 $(G_n)_{n\geq 1}$ 为稠密的 γ-扩展图序列，其中每个 G_n 有 n 个顶点，且最小度至少为 δn。取 $u \in G_n$ ，并令 $(X_i)_{i\geq 0}$ 表示从 u 出发、在 G_n 上进行的随机游走。取第 4 节开头定义的 $M = n^{\kappa}$ 。那么对所有充分大的 n,

$$\mathbb{P}\left(\left|\mathrm{Cap}_M(X[0, M)) - \frac{\alpha_n M^2}{n}\right| \geq \frac{\alpha_n M^2}{n} n^{-\kappa/16}\right) \leq \frac{2M^2}{\delta n}.$$

证明。证明是对引理 5.3] 的证明的简化版本。首先回顾引理 4.2 中的结论：$t^{+}_{\text{混合}} = \log_2^2(n) t_{\text{mix}}$。令 $(T_j)_{j=1}^{n^{\kappa/2}}$ 为一列独立同分布、服从 $\mathrm{Bin}(t^{+}_{\text{混合}}, 1/2)$ 分布的随机变量。那么，对所有 $1 \leq j \leq n^{\kappa/2}$ ，定义

$$B_j = [(j-1)n^{\kappa/2} + T_j, jn^{\kappa/2} - t^{+}_{\text{mix}} + T_j].$$

注意，由 (4) 可知，对于所有 $j \leq n^{\kappa/2}$，给定 $X[0, jn^{\kappa/2}]$ 后， B_{j+1} 的起点几乎服从平稳分布。另外，令 $(X^{\text{ind},\, j})_{j=1}^{n^{\kappa/2}}$ 表示一列独立随机游走片段，每段长度为 $n^{\kappa/2} - t^{+}_{\text{mix}}$，且每段均从平稳分布出发。注意，由 (4) 可知，片段 $(X_{B_j})_{j=1}^{n^{\kappa/2}}$ 可以与片段 $(X^{\text{ind},\, j})_{j=1}^{n^{\kappa/2}}$ 耦合，使得对于所有 $j \leq n^{\kappa/2}$ ，这些片段以至少...
...的概率重合

$$1 - n^{\kappa/2} n^{-\log_2(n)}. \tag{13}$$

注意，片段 $(X^{\text{ind},\, j})_j$ 是独立同分布的，并且根据定义，

$$\mathbb{E}\big[\mathrm{Cap}_M(X^{\text{ind},j})\big] = \mathbb{E}\Big[\mathrm{Cap}_M(X^{\text{ind},j}_{[0,n^{\kappa/2})})\Big] + O\left(\frac{Mt^{+}_{\text{mix}}}{\delta n}\right) = \frac{\alpha_n M n^{\kappa/2}}{n}\left(1 + O\left(\frac{t^{+}_{\text{mix}}}{\delta n^{\kappa/2}}\right)\right). \tag{14}$$

此外，由并集界可得确定性界

$$\mathrm{Cap}_M(X^{\text{ind},j}) \leq M\pi(X^{\text{ind},j}) \leq \frac{Mn^{\kappa/2}}{\delta n}. \tag{15}$$

因此，由霍夫丁界 [17, 定理 1] 可知，存在 $C < \infty$、$c > 0$ ，使得对于任意 $t > 0$,

$$\begin{aligned}\mathbb{P}\left(\left|\sum_{j=1}^{n^{\kappa/2}} \mathrm{Cap}_M(X^{\text{ind},j}) - n^{\kappa/2}\mathbb{E}\big[\mathrm{Cap}_M(X^{\text{ind},1})\big]\right| \geq \frac{\alpha_n M^2 t^{+}_{\text{mix}}}{2n^{1+\kappa/8}}\right) &\leq 2\exp\left(-2n^{\kappa/2}\left(\frac{\frac{\alpha_n M^2 t^{+}_{\text{mix}}}{2n^{1+5\kappa/8}}}{\frac{Mn^{\kappa/2}}{\delta n}}\right)^2\right) \\ &= 2\exp\left(-n^{\kappa/4}(t^{+}_{\text{mix}})^2 \frac{\alpha_n^2 \delta^2}{2}\right).\end{aligned}$$

特别地，由（14）可知

$$\left|n^{\kappa/2}\mathbb{E}\big[\mathrm{Cap}_M(X^{\text{ind},1})\big] - \frac{\alpha_n M^2}{n}\right| \leq O\left(\frac{\alpha_n M n^{\kappa/2}}{n}\frac{t^{+}_{\text{mix}}}{\delta}\right) \ll \frac{\alpha_n M^2 t^{+}_{\text{mix}}}{n^{1+\kappa/8}},$$

我们推导出

$$\mathbb{P}\left(\left|\sum_{j=1}^{n^{\kappa/2}} \mathrm{Cap}_M(X^{\text{ind},j}) - \frac{\alpha_n M^2}{n}\right| \geq \frac{\alpha_n M^2 t^{+}_{\text{mix}}}{n^{1+\kappa/8}}\right) \leq 2\exp\left(-n^{\kappa/4}(t^{+}_{\text{mix}})^2 \frac{\alpha_n^2 \delta^2}{2}\right). \tag{16}$$

我们希望用各个较小线段的容量之和来近似整个线段 $X[iM,(i+1)M)$ 的容量，但这样可能会略微高估，因为我们重复计算了命中多个较小线段的随机游走轨迹。为此，我们使用第 2.6 节中定义的接近性概念。对于每个 $J \le n^{\kappa/2}$，注意，在条件 $(X^{\mathrm{ind},\,j})_{j\le J}$ 全部互不相交时（其发生概率至少为

$$1-\frac{M^2}{\delta n}, \tag{17}$$

由推论 2.8 可得

$$\mathrm{Close}_M(X^{\mathrm{ind},J},\cup_{j<J}(X^{\mathrm{ind},j})_j) \le \frac{2M^3 n^{\kappa/2}}{\delta^2 n^2}.$$

同样，用 $\sum_{j=1}^{n^{\frac{\kappa}{2}}}\mathsf{Cap}_M(X_{B_j})$ 近似 $\mathsf{Cap}_M(X[0,M)$，可能会略微低估，因为来自集合 $X[0,M)\setminus\left(\cup_{j<n^{\kappa/2}}X_{B_j}\right)$ 的容量贡献也需要考虑。

现在我们将上述估计合并如下。注意，在事件 $X_{B_j}=X^{\mathrm{ind},\,j}$ 对所有 $j\le J$ 均成立的条件下（同时利用 (15) 和并集界），对所有充分大的 n，我们有：

$$\begin{aligned}\left|\mathrm{Cap}_M(X[0,M))-\sum_{j=1}^{n^{\frac{\kappa}{2}}}\mathrm{Cap}_M(X^{\mathrm{ind},j})\right| &\le \sum_{J=1}^{n^{\kappa/2}}\mathrm{Close}_M(X_{B_j},\cup_{j<J}X_{B_j})+\mathbb{P}\Big(\tau_{X[0,M)\setminus\left(\cup_{j<n^{\kappa/2}}X_{B_j}\right)}\le M\Big)\\ &\le \frac{2M^4}{\delta^2n^2}+\frac{Mn^{\kappa/2}t^+_{\mathrm{mix}}}{\delta n}\le n^{-\kappa/8}\frac{\alpha_n M^2 t^+_{\mathrm{mix}}}{n^{1+\kappa/8}}.\end{aligned} \tag{18}$$

因此，将 (13)、(16)、(17) 的估计通过并集界合并，应用 (18)，并利用 $t^+_{\mathrm{mix}}\ll n^{\kappa/32}$，可知以至少 $1-\frac{2M^2}{\delta n}$ 的概率有

$$\left|\mathrm{Cap}_M(X[0,M))-\frac{\alpha_n M^2}{n}\right|\le o\Big(\frac{\alpha_n M^2}{n}\cdot n^{-\kappa/16}\Big).$$ □

5 拉普拉斯随机游走表示与 Wilson 算法

在本节中，我们始终令 $(G_n)_{n\ge1}$ 为满足定理 1.3 中参数为 $\delta>0$ 和 $\gamma>0$ 的假设的图序列。由断言 2.5 可知，这意味着 $t_{\mathrm{mix}}=O(\log n)$。对于每个 n，$k\ge1$ 将表示在顶点集 $(v_1,\dots,v_{k-1})$ 上对 G_n 运行 Wilson 算法后得到的树。给定这样的序列 $(G_n)_{n\ge0}$，我们令

$$\alpha_n=\frac{n\mathbb{E}_\pi\big[\mathrm{Cap}_M(X[0,n^{\kappa/2}))\big]}{Mn^{\kappa/2}},\qquad \beta_n=\frac{1}{\sqrt{\alpha_n}}, \tag{19}$$

其中，X 是在 G_n 上的随机游走，κ 如第 4 节开头所定义。最后，如果 $A\subset G_n$，我们将使用记号 τ_A 表示在 G_n 上进行随机游走时到达 A 的命中时间。

本节的目标是证明针对这类图序列的下述命题 5.2，为此我们需要以下定义。

定义 5.1. 如果满足以下条件，则称子图 $T\subset G_n$ 是良好的：

1. T 是一棵树。
2. $|T|\le n^{1/2+\kappa}$。
3. 对于每个具有 $|A|\ge n^{3\kappa}$ 的开连通子集 $A\subset T$，我们都有 $|\mathsf{Cap}_M(A)-\alpha_nM|A|/n|\le\frac{\alpha_nM|A|}{n}\cdot n^{-\kappa/16}$。

注意， $T_n^{(1)}$（由第一个单顶点构成的树）显然是良好的。

命题 5.2. 取任意良子图 $T \subset G_n$。取任意 $u \in V(G_n) \setminus T$，并令 Y 为一条从 u 出发、在到达 T 时终止的循环擦除随机游走。

(1) 取任意 $C \in (0,\infty)$ 和任意 $B \in (0,\infty)$。另外，假设 $|T| \le B\sqrt{n}$。那么

$$\mathbb{P}_u\big(Y[0, C\beta_n s\sqrt{n}] \cap T = \emptyset\big) = \exp\left\{-\frac{(C + |T|/\beta_n\sqrt{n})^2 - (|T|/\beta_n\sqrt{n})^2}{2}\right\}(1 + o_B(1)).$$

(2) 令 H_T 为 Y 到达 T 的时间。那么对于任意满足 $|A| \ge n^{3\kappa}$ 的连通 $A \subset T$，以及所有 $n^{2\kappa}/M \le i \le n^{1/2+\kappa}/M$，

$$\mathbb{P}_u(Y_{H_T} \in A \mid H_T \in [iM, (i+1)M)) = \frac{|A|}{|T|}(1 + o(1)),$$

其中 $o(1)$ 在所有 $n^{2\kappa}/M \le i \le n^{1/2+\kappa}/M$ 上均匀分布。

(3) 对任意 $k \ge 1$ 和 $B \in (0,\infty)$，若 $|T| \le B\sqrt{n}$，则

$$\mathbb{P}\Big(T_n^{(k)} \textit{ is good} \;\Big|\; T_n^{(k-1)} = T\Big) = 1 - o_B(1).$$

命题 5.2 将使我们能够比较 Wilson 算法与第 6 节中的连续随机树折棍过程，并证明定理 1.3。

5.1 路径概率的比较

我们将使用第 2.5 节中概述的 LERW 拉普拉斯随机游走表示，比较不同环擦除随机游走轨迹的概率。首先固定一些 $n \ge 1$，并在 G_n 上运行 Wilson 算法。令 $\{v_1, \dots, v_n\}$ 表示此过程中顶点的排列顺序。

现在固定一些 $k \ge 2$，假设我们已运行 Wilson 算法 $k-2$ 步，构造出由顶点 $(v_1, \dots, v_{k-1})$ 张成的树，记为 $T_n^{(k-1)}$。令 X 表示在 G_n 上的随机游走，并在到达 $T_n^{(k-1)}$ 时停止；令 $(Y_m)_{m\ge 0}$ 表示其环擦除。回顾第 2.5 节，若 $u_0, u_1, \dots, u_H$ 是 G_n 中的一条简单路径，其中 $\{u_0, u_1, \dots, u_{H-1}\}$ 与 $\bigcup_{m=0}^{L}\{Y_m\} \cup T_n^{(k-1)}$ 和 $u_0 = Y_L$ 不相交，则

$$\mathbb{P}\big((Y_m)_{m=L+1}^{L+H} = (u_m)_{m=1}^{H} \;\big|\; (Y_m)_{m=0}^{L}\big) = \mathbb{P}_{u_0}\big((X_m)_{m=1}^{H} = (u_m)_{m=1}^{H}\big) C\big((Y_m)_{m=0}^{L}, T_n^{(k-1)}, (u_m)_{m=1}^{H}\big), \quad (20)$$

其中

$$C\big((Y_m)_{m=0}^{L}, T_n^{(k-1)}, (u_m)_{m=1}^{H}\big) = \prod_{h=1}^{H} \frac{\mathbb{P}_{u_h}\Big(\tau_{T_n^{(k-1)}} < \tau_{\{Y_m\}_{m=0}^{L} \cup \{u_m\}_{m=0}^{h-1}}\Big)}{\mathbb{P}_{u_{h-1}}\Big(\tau_{T_n^{(k-1)}} < \tau^{+}_{\{Y_m\}_{m=0}^{L} \cup \{u_m\}_{m=0}^{h-1}}\Big)}. \quad (21)$$

在本节的某些地方，我们将以形如 $T_n^{(k-1)} = T$ 的事件为条件，其中 X 和 Y 分别表示一条随机游走和一条循环擦除随机游走，二者均在到达 T 时停止。为方便记号，若 H_T 是 Y 到达 T 的时间，则对所有 $m \ge H_T$，我们令 $Y_m = Y_{H_T}$。

注 5.3。为证明定理 3.1，我们应在 G_n 上均匀地选择顶点 $\{v_1, \dots, v_k\}$。事实上，我们将证明一个对任意选取的互异 $\{v_1, \dots, v_k\}$ 都成立的结果。

证明命题 5.2 的策略大致如下。首先，引理 5.4 使我们能够控制 Y的首个较短线段的行为。这使我们能够对由 (21) 定义的常数 C给出紧确估计，我们将在引理 5.5 中完成此事。在推论 5.6 中，我们结合这一结果、命题 4.3 和 (20)，对循环擦除随机游走线段的容量进行紧确控制。在引理 5.7 中，我们估计良好树 T 上的随机游走命中测度；在推论 5.8 中，我们将这一结果与关于常数 C 的估计结合起来，代入 (20)，从而得到循环擦除随机游走的类似估计。

引理 5.4。存在 $N<\infty$，使得对所有 $n\ge N$，以下结论成立。设 $T\subset G_n$是一个满足 $|T|\le n^{\frac{1}{2}+\kappa}$的子图。任取 $u\in V(G_n)\setminus T$，并设 Y是一个从 u出发、在命中T时终止的循环擦除随机游走。那么

$$\mathbb{P}_u\big(|Y|\le n^{2\kappa}\big)\le\frac{4n^{2\kappa}(n^{2\kappa}+|T|)}{\delta^3 n}.$$

证明。对于每个 $K<n^{2\kappa}$，我们将使用拉普拉斯随机游走表示来界定 $\mathbb{P}(|Y|=K+1\mid |Y|>K)$ 的概率。为此，由 (20) 可知，对每个 $v\in T$以及每条长度为 K、满足 $\varphi\cap T=\emptyset$ 的简单路径 φ，有

$$\mathbb{P}_u\big(Y_{K+1}=v\;\big|\;(Y_m)_{m=0}^K=\varphi\big)\le\frac{1}{\delta n}\cdot\frac{1}{\mathbb{P}_{Y_K}\big(\tau_T<\tau_\varphi^+\big)}.$$

由引理 4.1 和引理 4.2，第二项可以通过估计 T 和φ的容量来界定。我们断言，除依赖于最小度的某些常数因子外，它们都可以用各自的大小来估计。事实上，对于任意大小小于 $n^{1/2+\kappa}$的 A 集合，当 n足够大时，由引理 2.6 可得 $\frac{\delta M|A|}{2n}\le \mathrm{Cap}_M(A)\le\frac{|A|M}{\delta n}$。

Th因此，由于 φ 的大小为 $K<n^{2\kappa}<n^{1/2+\kappa}$，而 T的大小至多为 $n^{1/2+\kappa}$，我们有

$$\frac{\mathrm{Cap}_M(T)}{\mathrm{Cap}_M(T)+\mathrm{Cap}_M(\varphi)}\ge\frac{\delta^2|T|}{2(n^{2\kappa}+|T|)}.$$

因此，根据引理 4.1、引理 4.2，并对所有 $v\in T$ 求和可知，对所有充分大的 n，有

$$\mathbb{P}_u\big(Y_{K+1}\in T\;\big|\;(Y_m)_{m=0}^K=\varphi\big)\le\frac{4(n^{2\kappa}+|T|)}{\delta^3 n}.$$

由并集界可得，$\mathbb{P}_u\big(|Y|\le n^{2\kappa}\big)\le {}^{4n^{2\kappa}}{}_{(n^{2\kappa}+|T|)\ \delta^3 n}$，如所需。 ——— □

我们还需要以下结果来控制由（21）定义的常数 C。读者应注意，我们最终将把该结果应用于 $T=T_n^{(k-1)}$。

引理 5.5。取 $i\le\frac{n^{1/2+\kappa}}{M}$。那么，对于所有简单路径 $\{Y_m\}_{m=0}^{iM}$、所有满足$u_0=Y_{iM}$且 $H\le M$的简单路径 $\{u_m\}_{m=0}^H$，以及所有满足 $\{u_m\}_{m=1}^{H-1}$、$\{Y_m\}_{m=0}^{iM}$ 与 T 互不相交且满足 $|T|\le n^{1/2+\kappa}$ 的连通子图 $T\subset G_n$，我们有如下结论

(a) 若 $u_H\notin T$，则

$$\big|C((Y_m)_{m=0}^{iM},T,(u_m)_{m=1}^H)-1\big|=o(n^{5\kappa-1/2}).$$

(b) 若 $u_H\in T$和 $i\ge n^{2\kappa}/M$，则

$$C((Y_m)_{m=0}^{iM},T,(u_m)_{m=1}^H)=\frac{[\mathrm{Cap}_M(T)+\mathrm{Cap}((Y_m)_{m=0}^{iM})](1+o(1))}{\mathrm{Cap}_M(T)}.$$

证明。固定某个 $1\le h\le H$。在情形 (a) 中，为了给 (21) 中乘积里出现的一项作界，我们希望比较以下概率

$$\mathbb{P}_{u_h}\Big(\tau_T<\tau_{\{Y_m\}_{m=0}^{iM}\cup\ \{u_m\}_{m=0}^{h-1}}\Big)\ \text{ and }\ \mathbb{P}_{u_{h-1}}\Big(\tau_T<\tau^+_{\{Y_m\}_{m=0}^{iM}\cup\ \{u_m\}_{m=0}^{h-1}}\Big).\tag{22}$$

由引理 4.2 和三角不等式可得

$$\begin{aligned}&\Big|\mathbb{P}_{u_h}\Big(\tau_T<\tau_{\{Y_m\}_{m=0}^{iM}\cup\ \{u_m\}_{m=0}^{h-1}}\Big)-\mathbb{P}_{u_{h-1}}\Big(\tau_T<\tau^+_{\{Y_m\}_{m=0}^{iM}\cup\ \{u_m\}_{m=0}^{h-1}}\Big)\Big|\\&\le\mathbb{P}_{u_h}\Big(\tau_T<\tau_{\{Y_m\}_{m=0}^{iM}\cup\ \{u_m\}_{m=0}^{h-1}}\Big)(o(n^{3\kappa-1/2}t^+_{\mathrm{mix}})).\end{aligned}$$

换句话说，

$$\frac{\mathbb{P}_{u_h}\Big(\tau_T < \tau_{\{Y_m\}_{m=0}^{iM} \cup \ \{u_m\}_{m=0}^{h-1}}\Big)}{\mathbb{P}_{u_{h-1}}\Big(\tau_T < \tau^+_{\{Y_m\}_{m=0}^{iM} \cup \ \{u_m\}_{m=0}^{h-1}}\Big)} = 1 + o(n^{3\kappa-1/2} t^+_{\mathrm{mix}}).$$

因此，(21) 中的乘积有如下界

$$(1 + o(n^{3\kappa-1/2} t^+_{\mathrm{mix}}))^{n^\kappa} = 1 + o(n^{4\kappa-1/2} t^+_{\mathrm{mix}}).$$

最后，我们利用 $t_{\mathrm{mix}} = o(n^\kappa)$ 如 $n \to \infty$，这由断言 2.5 得出。

对于 (b)，若 $u_H \in T$，则 $C((Y_m)_{m=0}^{iM}, T, (u_m)_{m=1}^{H})$ 反而等于

$$\frac{1}{\mathbb{P}_{u_{H-1}}\Big(\tau_T < \tau^+_{\{Y_m\}_{m=0}^{iM} \cup \ \{u_m\}_{m=0}^{h-1}}\Big)} \prod_{h=1}^{H-1} \frac{\mathbb{P}_{u_h}\Big(\tau_T < \tau_{\{Y_m\}_{m=0}^{iM} \cup \ \{u_m\}_{m=0}^{h-1}}\Big)}{\mathbb{P}_{u_{h-1}}\Big(\tau_T < \tau^+_{\{Y_m\}_{m=0}^{iM} \cup \ \{u_m\}_{m=0}^{h-1}}\Big)},$$

其中，乘积等于 $C((Y_m)_{m=0}^{iM}, T, (u_m)_{m=1}^{H-1})$，因此由 (a) 可知它是 $1 + o(1)$。接着注意到，由引理 4.1 和引理 4.2 可得

$$\begin{aligned}\mathbb{P}_{u_{H-1}}\Big(\tau_T < \tau^+_{\{Y_m\}_{m=0}^{iM} \cup \ \{u_m\}_{m=0}^{h-1}}\Big) &= \frac{\mathrm{Cap}_M(T)(1+o(1))}{\mathrm{Cap}_M(T) + \mathrm{Cap}_M((Y_m)_{m=0}^{iM} \cup \ \{u_m\}_{m=0}^{h-1})} \\ &= \frac{\mathrm{Cap}_M(T)(1+o(1))}{\mathrm{Cap}_M(T) + \mathrm{Cap}_M((Y_m)_{m=0}^{iM})},\end{aligned}$$

其中，最后一个等号成立是因为 $i \geq n^{2\kappa}/M$，因此 $\mathsf{Cap}_M(\{u_m\}_{m=0}^{h-1}) \leq \frac{M^2}{\delta n} = o(\frac{\delta i M^2}{2n})$，而根据引理 2.6，这是 $\mathsf{Cap}_M((Y_m)_{m=0}^{iM})$ 的一个下界。这就证明了该断言。 □

利用拉普拉斯随机游走表示，我们现在可以精确控制下一段循环擦除随机游走具有良好容量的概率。

推论 5.6. 设 $T \subset G_n$ 是一个子图，使得 $|T| \leq n^{1/2+\kappa}$，并设 Y 是一条从某个 $u \in V(G_n) \setminus T$ 出发、命中 T 时终止的循环擦除随机游走轨迹。那么对每个 $0 \leq i \leq \frac{n^{1/2+\kappa}}{M}$，以及任意一条与 T 不相交的循环擦除随机游走轨迹 $\{Y_m\}_{m=0}^{iM} = \{y_m\}_{m=0}^{iM}$，我们有

$$\mathbb{P}\bigg((Y_m)_{m=iM}^{(i+1)M} \cap T = \emptyset \text{ and } \Big|\mathrm{Cap}_M(Y[iM,(i+1)M)) - \frac{\alpha_n M^2}{n}\Big| \geq \frac{\alpha_n M^2}{n} n^{-\kappa/16} \ \Big|\ \{Y_m\}_{m=0}^{iM} = \{y_m\}_{m=0}^{iM}\bigg) \leq \frac{4M^2}{\delta n}.$$

证明. 设 A 为不与 T 相交的简单路径 $(u_m)_{m=0}^{M}$ 所构成的集合，其中 $\Big|\mathsf{Cap}_M(u[0,M)) - \frac{\alpha_n M^2}{n}\Big| \geq \frac{\alpha_n M^2}{n} n^{-\kappa/16}$，且满足 $u_0 = Y_{iM}$。由 (20) 可知，对任意 $i \geq 1$,

$$\mathbb{P}\big((Y_m)_{m=iM+1}^{iM+M} \in A \ \big|\ (Y_m)_{m=0}^{iM} = (y_m)_{m=0}^{iM}\big) \leq \mathbb{P}_{Y_{iM}}\big((X_m)_{m=1}^{H} \in A\big) \sup_{(u_m)_{m=1}^{M} \in A} \big\{C((y_m)_{m=0}^{iM}, T, (u_m)_{m=1}^{M}))\big\}. \tag{23}$$

由于 $(u_m)_{m=1}^{M}$ 与 T 不相交，根据引理 5.5，对于所有充分大的 n，上确界至多为 2。因此，根据命题 4.3，我们有

$$\mathbb{P}\bigg(\Big|\mathrm{Cap}_M(Y[iM,(i+1)M)) - \frac{\alpha_n M^2}{n}\Big| \geq \frac{\alpha_n M^2}{n} n^{-\kappa/16} \text{ and } Y[iM,(i+1)M) \cap T = \emptyset \ \Big|\ (Y_m)_{m=0}^{iM} = (y_m)_{m=0}^{iM}\bigg) \leq \frac{4M^2}{\delta n}.$$

□

在下一个引理中，我们计算随机游走的一些命中概率。

引理 5.7。设 X 是 G_n 上的随机游走。设 $T \subset G_n$ 是一个满足 $|T| \le n^{1/2+\kappa}$ 的子图，并设 A 是 T 的一个连通子集。那么对于任意 $0 \le i \le \frac{n^{1/2+\kappa}}{M}$、任意一条在 G_n 上且与 T 不相交的简单路径 $y[0, iM]$，以及任意 $u \notin T \cup y[0, iM]$

$$\begin{aligned}&\mathbb{P}_u(X[0,M] \cap A \neq \emptyset \text{ and } \nexists 0 < j < \ell \le \tau_A : X_j = X_\ell \text{ and } \nexists j \le \tau_A : X_j \in y[0,iM] \cup (T \setminus A))\\ &= \mathrm{Cap}_M(A)(1 + o(1)).\end{aligned}$$

证明。上界。回顾一下，$t^+_{\text{混合}} = \log_2^2(n) t_{\text{mix}}$。首先注意，根据断言 2.7，我们有

$$\begin{aligned}&\mathbb{P}_u(X[0,M] \cap A \neq \emptyset \text{ and } \nexists 0 < j < \ell \le \tau_A : X_j = X_\ell \text{ and } \nexists j \le \tau_A : X_j \in y[0,iM] \cup (T \setminus A))\\ &\le \mathbb{P}_u(X[0,M] \cap A \neq \emptyset)\\ &\le \mathrm{Cap}_M(A) + \frac{3|A| \log n \cdot t_{\text{mix}}}{\delta n} \le \mathrm{Cap}_M(A)\left(1 + \frac{t^+_{\text{mix}}}{\delta^2 M}\right).\end{aligned}$$

这里最后一行由引理 2.6 得出，它意味着在$|T| \le n^{1/2+\kappa}$事件发生时，$\mathrm{Cap}_M(A) \ge \frac{\delta M|A|}{2n}$。

下界。我们首先注意到

$$\begin{aligned}&\mathbb{P}_u(X[0,M] \cap A \neq \emptyset) - \mathbb{P}_u(X[0,M] \cap A \neq \emptyset \text{ and } \exists 0 < j < \ell < \tau_A : X_j = X_\ell)\\ &- \mathbb{P}_u(X[0,M] \cap A \neq \emptyset \text{ and } \exists j < M : X_j \in y[0,iM] \cup (T \setminus A)).\\ &\le \mathbb{P}_u(X[0,M] \cap A \neq \emptyset \text{ and } \nexists 0 < j < \ell \le \tau_A : X_j = X_\ell \text{ and } \nexists j \le \tau_A : X_j \in y[0,iM] \cup (T \setminus A))\end{aligned}$$

现在我们将对左边的三个项分别进行界定。首先，根据断言 2.7，我们用 $\mathrm{Cap}_M(A)\left(1 - \frac{3t^+_{\text{混合}}}{\delta^2 M}\right)$ 对 $\mathbb{P}_u(X[0,M] \cap A \neq \emptyset)$取下界。对于第二项，我们用 X 在 M 步内自相交的概率与随后在另外 M 步内命中 A 的概率之积作为上界。该项的上界为

$$\frac{M^2}{\delta n} \cdot \frac{M|A|}{\delta n}.$$

第三项可以用至多 M 步内命中 $y[0, iM] \cup (T \setminus A)$、随后至多再用 M 步命中 A 的概率来界定，其上界为

$$\frac{M|A|}{\delta n} \cdot \frac{M(iM + |T|)}{\delta n}.$$

由此得出

$$\begin{aligned}&\mathrm{Cap}_M(A)\left(1 + O\left(\frac{t^+_{\text{mix}}}{M}\right)\right) \le \mathrm{Cap}_M(A)\left(1 - \frac{t^+_{\text{mix}}}{\delta^2 M}\right) - 2 \cdot \frac{M|A|}{\delta n} \cdot \frac{M(iM + |T|)}{\delta n}\\ &\le \mathbb{P}_u(X[0,M] \cap A \neq \emptyset \text{ and } \nexists 0 < j < \ell \le \tau_A : X_j = X_\ell \text{ and } \nexists j \le \tau_A : X_j \in y[0,iM] \cup (T \setminus A)).\end{aligned}$$ □

现在，我们已对式 (20) 中出现的所有量得到了估计。我们将在下一个推论中将这些估计结合起来。

推论 5.8. 设 $T \subset G_n$ 是一个满足 $|T| \le n^{\kappa+1/2}$ 的子图，并设 $A \subseteq T$。令 Y 是一条定义在 G_n 上的循环擦除随机游走，在其命中 T 时停止。对于每个 $\frac{n^{2\kappa}}{M} \le i \le \frac{n^{1/2+\kappa}}{M}$，以及任意一条不与 T 相交的简单路径 $(y_m)_{m=0}^{iM}$，满足

$$\left|\mathrm{Cap}_M(y[jM,(j+1)M)) - \frac{\alpha_n M^2}{n}\right| \le \frac{\alpha_n M^2}{n} n^{-\kappa/16} \text{ for all } \frac{n^{2\kappa}}{M} \le j < i,$$

有

$$\begin{aligned}&\mathbb{P}\big(Y \text{ hits } T \text{ in time interval } [iM,(i+1)M) \text{ in set } A \;\big|\; (Y_m)_{m=0}^{iM} = (y_m)_{m=0}^{iM}\big)\\ &= \frac{\mathrm{Cap}_M(A)}{\mathrm{Cap}_M(T)}\left(\mathrm{Cap}_M(T) + \frac{\alpha_n i M^2}{n}\right)(1 + o(1)).\end{aligned}$$

证明。首先注意到，由推论 2.8 可知

$$\mathrm{Cap}_M(y[0,iM)) = \frac{\alpha_n i M^2(1+o(1))}{n}.$$

给定 $1 \le H < M$、$(y_m)_{m=0}^{iM}$和 T，令 $\Gamma_{y_{iM}\to A,T,H}$表示这样一组简单路径：它们在时刻 H 首次到达集合 A 中的 T，路径从 $u_0 = y_{iM}$ 出发，并且在此之前避开 $(y_m)_{m=0}^{iM}$。于是，利用引理 5.5(b) 和引理 5.7，可写为：

$$\begin{aligned}
&\mathbb{P}\big(Y \text{ hits } T \text{ in time interval } [iM,(i+1)M) \text{ in set } A \,\big|\, (Y_m)_{m=0}^{iM} = (y_m)_{m=0}^{iM}\big)\\
&= \sum_{H<M} \sum_{\substack{(u_m)_{m=0}^{H}\in\\ \Gamma_{y_{iM}\to A,T,H}}} \mathbb{P}\big(Y[iM,iM+H] = (u_m)_{m=0}^{H} \,\big|\, (Y_m)_{m=0}^{iM} = (y_m)_{m=0}^{iM}\big)\\
&= \sum_{H<M} \sum_{\substack{(u_m)_{m=0}^{H}\in\\ \Gamma_{y_{iM}\to A,T,H}}} C((y_m)_{m=1}^{iM}, T, (u_m)_{m=1}^{H})\mathbb{P}_{y_{iM}}\big(X[0,H] = (u_m)_{m=0}^{H}\big)\\
&= \sum_{H<M} \sum_{\substack{(u_m)_{m=0}^{H}\in\\ \Gamma_{y_{iM}\to A,T,H}}} \frac{[\mathrm{Cap}_M(T) + \mathrm{Cap}((Y_m)_{m=0}^{iM})](1+o(1))}{\mathrm{Cap}_M(T)}\mathbb{P}_{y_{iM}}\big(X[0,H] = (u_m)_{m=0}^{H}\big)\\
&= \frac{[\mathrm{Cap}_M(T) + \mathrm{Cap}((Y_m)_{m=0}^{iM})](1+o(1))}{\mathrm{Cap}_M(T)} \sum_{H<M} \sum_{\substack{(u_m)_{m=0}^{H}\in\\ \Gamma_{y_{iM}\to A,T,H}}} \mathbb{P}_{y_{iM}}\big(X[0,H] = (u_m)_{m=0}^{H}\big)\\
&= \frac{[\mathrm{Cap}_M(T) + \mathrm{Cap}((Y_m)_{m=0}^{iM})]}{\mathrm{Cap}_M(T)}\mathrm{Cap}_M(A)(1+o(1)).
\end{aligned}$$

最后一行由断言 2.7、推论 2.8 和引理 5.7 得出。 □

命题 5.2 的证明。（1）固定一个满足良好条件的 T，再固定一个 $C>0$。对于每个 $i \le C\beta_n\sqrt{n}/M$，令$E_{C,i}$ 表示如下事件：

$$\left|\mathrm{Cap}_M(Y[iM,(i+1)M)) - \frac{\alpha_n M^2}{n}\right| \le \frac{\alpha_n M^2}{n} n^{-\kappa/16} \text{ and } Y[iM,(i+1)M)\cap T = \emptyset.$$

记 E_{prefix} 为事件$\cap_{i\le n^{2\kappa}/M} E_{C,i}$。注意，由引理 5.4 和推论 5.6 可得

$$\mathbb{P}(E_{prefix}) = 1 - o(1).$$

注意，由推论 5.6 和推论 5.8，对于任意 $i \ge n^{2\kappa}/M$，在给定 $E_{prefix}\cap_{n^{2\kappa}/M\le j\le i} E_{C,j}$ 的条件下，利用推论 5.8 可得

$$\begin{aligned}
&\mathbb{P}\big(E_{C,i} \,\big|\, E_{prefix} \text{ and } \cap_{n^{2\kappa}/M\le j\le i} E_{C,j}\big)\\
&= 1 - \left(\mathrm{Cap}_M(T) + \frac{\alpha_n i M^2}{n}\right)(1+o(1)) - O\left(\frac{4M^2}{\delta n}\right) = 1 - \left(\frac{\alpha_n M|T|}{n} + \frac{\alpha_n i M^2}{n}\right)(1+o(1)).
\end{aligned}$$

这里最后一行成立，因为 T 是好的， $i \ge \frac{n^{2\kappa}}{M}$ 并且我们是在$E_{prefix}\cap\cap_{n^{2\kappa}/M\le j\le i}E_{C,j}$的条件下进行的。接下来，记 E_C 为以下事件：

$$E_{prefix} \text{ and } \{Y[0,C\beta_n\sqrt{n}]\cap T = \emptyset\} \text{ and } \left(\cap_{i=n^{2\kappa}/M}^{C\beta_n\sqrt{n}/M} E_{C,i}\right),$$

我们有，

$$\begin{aligned}
&\mathbb{P}\Big(E_C \,\Big|\, T_n^{(k-1)} = T\Big) =\\
&\mathbb{P}\Big(E_{prefix} \,\Big|\, T_n^{(k-1)} = T\Big).
\end{aligned}$$

$$
\begin{aligned}
&\prod_{i=n^{2\kappa}/M}^{C\beta_n\sqrt{n}/M} \mathbb{P}\big(E_{C,i} \,\big|\, E_{prefix} \text{ and } \cap_{n^{2\kappa}/M\le j<i} E_{C,j}\big)\\
&= (1-o(1)) \prod_{i=n^{2\kappa}/M}^{C\beta_n\sqrt{n}/M} \left(1-\left(\frac{\alpha_n M|T|}{n}+\frac{\alpha_n i M^2}{n}\right)(1+o(1))\right)\\
&= (1-o(1))\exp\left(-\sum_{i=n^{2\kappa}/M}^{C\beta_n\sqrt{n}/M}(1-o(1))\alpha_n(M|T|/n+iM^2/n)\right)\\
&= (1-o(1))\exp\left(-\alpha_n\frac{C\beta_n\sqrt{n}}{2M}(2M|T|/n+C\beta_n M/\sqrt{n})\right)\\
&= (1-o(1))\exp(-(C^2/2+C|T|/(\beta_n\sqrt{n})).
\end{aligned}
$$

（最后一行中，我们根据定义使用了 $\beta_n^2=\frac{1}{\alpha_n}$。）最后，请注意

$$
\mathbb{P}\Big(Y[0,C\beta_n\sqrt{n}]\cap T_n^{(k-1)}=\emptyset \,\Big|\, T_n^{(k-1)}=T\Big)-\mathbb{P}\Big(E_C \,\Big|\, T_n^{(k-1)}=T\Big)\le\frac{C\beta_n\sqrt{n}}{M}\cdot\frac{4M^2}{\delta n}=o(1) \tag{24}
$$

由推论 5.6。在事件 $|T|\le B\sqrt{n}$ 上，这可以写成 $o_B(1)\mathbb{P}\Big(E_C \,\Big|\, T_n^{(k-1)}=T\Big)$的形式，其中使用了上述估计。

(2) 由推论 5.8 和定义 5.1 可直接得出，对于任意 $i>n^{2\kappa}/M$，在给定 $H_{T_n^{(k-1)}}\in(iM,(i+1)M]$的条件下，我们有

$$
\mathbb{P}\Big(Y_{H_{T_n^{(k-1)}}}\in A \,\Big|\, H_{T_n^{(k-1)}}\in(iM,(i+1)M]\Big)=\frac{|A|(1+o(1))}{|T_n^{(k-1)}|},
$$

正如所需。

(3) 给定 $\varepsilon>0$，先选择 $C<\infty$，使第 (1) 部分中出现的概率至多为 ε。然后，在事件 $Y[0,C\beta_n\sqrt{n}]\cap T\neq\emptyset$ 上，由 (24) 可知，$T_n^{(k)}$ 不好的概率以上界 $\varepsilon+o(1)$ 控制。由于 $\varepsilon>0$ 是任意的，因此得到所需结果。

□

6 定理 1.3 的证明

本节证明定理 1.3。我们首先利用上一节的估计证明定理 3.1。在本节末尾，我们将讨论下质量界条件，从而完成定理 1.3 的证明。

6.1 定理 3.1 的证明

在定义 3.2 中，我们定义了如何通过断棍过程构造树序列。接下来，我们将概述对于任意 $k\ge1$，如何在 G_n 上使用 Wilson 算法得到两个序列 $(Y_i)_{i=0}^n$ 和 $(Z_i)_{i=0}^{n-1}$，使得 $\mathrm{SB}^{(k)}((Y_0,Y_1,\dots,Y_{k-1}),(Z_0,Z_1,\dots,Z_{k-2}))$ 等于 Wilson 算法前 $k-1$ 步后得到的子树，并且定理 3.1 中出现的 $\binom{k}{2}$ 距离因此与断棍构造中点 $(Y_0,Y_1,\dots,Y_{k-1})$ 之间的距离相对应。

设 G_n 为一个具有 n 个顶点的图，并回顾（19）中 β_n 的定义。我们将定义一个折棍过程 $(Y_i^n)_{i=0}^n$、$(Z_i^n)_{i=0}^{n-1}$，它源自对 G_n应用 Wilson 算法。为简化记号，我们将省略上标，并从 $Y_0=0$、$Z_0=0$开始。我们选取 G的顶点排列，记为$\{v_1,\dots,v_n\}$。然后，在第一步中，我们从 v_2到 v_1抽样均匀生成树路径（使用 Wilson 算法）。我们

将这条路径记为 $T_n^{(2)}$，并令 Y_1 为该路径长度除以 $\beta_n\sqrt{n}$所得的值。对于路径上的每个顶点 z，我们称 z是在第一步添加的。令 $k \geq 2$，并假设我们抽样得到了 $T_n^{(k)}$和 Z_0、……、Z_{k-2}和 Y_0、……、Y_{k-1}。对于第 k^{th}步，取 v_{k+1}，并抽样（再次使用 Wilson 算法）其到 $T_n^{(k)}$的路径。将该路径记为 P_k，并令 $T_n^{(k+1)} = P_k \cup T_n^{(k)}$。对于 $P_k \setminus T^{(k)}{}_n$中的每个顶点，我们称其是在第 k^{th} 步添加的。令 $Y_k = Y_{k-1} + \frac{|P_k|}{\beta_n\sqrt{n}}$。为定义 Z_{k-1}，首先令 z 为 P_k 与$T_n^{(k)}$相交的顶点。如果 z 的形式为 v_m （其中 m），则令 $Z_{k-1} = Y_{m-1}$。否则，令 $m < k$ 为 z 被添加的步骤。于是， $Y_{m-1} \leq Z_{k-1} \leq Y_m$，而 Z_{k-1} 的确切值为

$$Z_{k-1} = Y_m - \frac{d(z, v_{m+1})}{\beta_n\sqrt{n}}.$$

此外，通过这种方式，我们可以定义一个函数 I，将每个 $v \in T_n^{(k)}$与一个点对应起来 $[0, \infty)$。如果v 是在第 m^{th} 步添加的，那么我们令 $I(v) = Y_m - {}^{d}{}_{(v,\ v_{m+1})\ \beta_n\sqrt{n}}$。

在本节中，我们也令 $(Y_i')_{i\geq 0}$ 和 $(Z_i')_{i\geq 0}$ 分别表示连续随机树折棍过程中的对应量，其抽样方式如命题 3.3 所述。

我们将使用以下断言。回顾定义 5.1 中“good”的定义。

断言 6.1. 假设 $T^{(k-1)}{}_n$是 good，且 $|T^{(k-1)}{}_n| \in [C^{-1}\sqrt{n}, C\sqrt{n}]$。令 $I_{\max} = \max_{v\in T_n^{(k-1)}} I(v)$。设 Y为从 v_k出发的循环擦除随机游走，并令 $H_{T_n^{(k-1)}}$为 Y命中 $T^{(k-1)}{}_n$的时刻。令 $j \leq n^{1/2+\kappa}/M$，并令 $\mathbb{P}_{d,\ j}$ 为定义在 $[0, I_{\max}]$ 上的测度，其定义如下

$$\mathbb{P}_{d,j}(I(v)) = \mathbb{P}\Big(Y_{H_{T_n^{(k-1)}}} = v \;\Big|\; H_{T_n^{(k-1)}} \in [j\varepsilon\beta_n\sqrt{n}/2, (j+1)\varepsilon\beta_n\sqrt{n}/2)\Big) \quad \forall v \in T_n^{(k-1)}.$$

于是，对每个 $\varepsilon > 0$，都存在 $N \in \mathbb{N}$，使得对所有 $n > N$ 和所有 $j \leq n^{1/2+\varepsilon}/M$，测度 $\mathbb{P}_{d,\ j}$ 与 $[0, I_{\max}]$ 上的均匀概率测度之间的普罗霍罗夫距离小于 ε。

证明。我们假设 C/β_n 大于 ε；否则 $I_{\max}$ 小于 ε，无需证明。我们还不失一般性地假设 $\varepsilon < 1$。将 $[0, I_{\max}]$ 分解为大小为 ε 的区间，写作 $[0, I_{\max}] = \cup_{i\leq \lfloor I_{\max}/\varepsilon \rfloor}[i\varepsilon, \min\{I_{\max}, (i+1)\varepsilon\}]$。固定 j，用 $\mathbb{P}_d$代替 $\mathbb{P}_{d,\ j}$，并记 $\mathbb{P}_u$ 为 $[0, I_{\max}]$上的均匀测度。注意，每个长度为 ε的区间 $I \subset [0, I_{\max}]$ 都可以视为至多 k 个 $T_n^{(k-1)}$的连通子集之并，且这些子集的长度之和为 $\varepsilon\beta_n\sqrt{n}$ （远大于 $n^{3\varepsilon}$）。舍去长度小于 $n^{3\varepsilon}$ 的子集后，我们可以对剩余子集应用命题 5.2(2)（必要时，将它们及 I 分解为更小的区间），从而推出

$$\mathbb{P}_d(I) = \frac{\varepsilon}{I_{\max}}(1 + o(1)), \quad \mathbb{P}_u(I) = \frac{\varepsilon}{I_{\max}}.$$

现在取足够大的 N，使得 $o(1)$ 误差不超过 ε。然后，取某个集合 A，使其属于 $[0, I_{\max}]$，并令 I_A 为与 A 相交的形如 $[i\varepsilon, (i+1)\varepsilon)$ 的区间集合。现在我们有

$$|I_A|\frac{\varepsilon}{I_{\max}} - \varepsilon \leq |I_A|\frac{\varepsilon}{I_{\max}}(1 + o(1)) \leq \mathbb{P}_d(I_A) \leq |I_A|\frac{\varepsilon}{I_{\max}}(1 + o(1)) \leq |I_A|\frac{\varepsilon}{I_{\max}} + \varepsilon$$

and

$$\mathbb{P}_d(A) \leq \mathbb{P}_d(I_A) \leq |I_A|\frac{\varepsilon}{I_{\max}} + \varepsilon \leq \mathbb{P}_u(I_A) + \varepsilon \leq \mathbb{P}_u(A^\varepsilon) + \varepsilon,$$

$$\mathbb{P}_u(A) \leq \mathbb{P}_u(I_A) = |I_A|\frac{\varepsilon}{I_{\max}} \leq \mathbb{P}_d(I_A) + \varepsilon \leq \mathbb{P}_d(A^\varepsilon) + \varepsilon.$$

因此，这两个测度之间的普罗霍罗夫距离至多为 ε。 □

本节的主要断言如下。

命题 6.2. 对于每个 $\varepsilon > 0$ 和 $k \geq 1$，存在 N，使得对于所有 $n > N$，我们都可以将连续随机树的折棍过程与均匀生成树的折棍过程耦合，使得对于所有 $0 \leq i \leq k-1$，$|Y_i - Y_i'| \leq \varepsilon$，并且对于每个$0 \leq i \leq k-2$，$|Z_i - Z_i'| \leq \varepsilon$的概率至少为 $1-\varepsilon$。

证明。我们通过归纳法证明该断言。显然，当 $k = 1$（即对于 $T_n^{(1)}$）时，该陈述显然成立，因为树是一个单点，且 $Y_0 = Y_0' = 0$ 根据构造得出。此外，由于由单个顶点组成的树总是良好的，并且由于 $Z_0 = Z_0' = 0$，根据引理 2.10、命题 3.4 和命题 5.2 可直接推出，该陈述对于 $k = 2$ 也成立。

现在固定 $k \geq 3$，并假设该断言对所有 $m < k$ 都成立。我们现在将证明该断言对 k 也成立。也就是说，假设对于每个 $\varepsilon > 0$，都存在足够大的 N，使得对于所有 $n > N$，我们都可以成功地将 $T_n^{(k-1)}$ 与连续随机树耦合。只需证明，对于每个 $\varepsilon > 0$，都存在 $0 < \zeta < \varepsilon/8$，使得在以参数 ζ 为条件、上一步成功耦合的情况下，我们可以将 (Y_{k-1}, Z_{k-2}) 与 (Y_{k-1}', Z_{k-2}') 耦合，使得 $|Y_{k-1} - Y_{k-1}'| < \varepsilon$ 且 $|Z_{k-2} - Z_{k-2}'| < \varepsilon$ 的概率至少为$1 - \varepsilon/2$。

为此，令 $\zeta > 0$（其具体取值稍后确定），并假设我们已成功地按照断言中参数为 ζ 的陈述，将 $(Y_i)_{i \leq k-2}$ 和 $(Z_i)_{i \leq k-3}$ 与 $(Y_i')_{i \leq k-2}$ 和 $(Z_i')_{i \leq k-3}$ 耦合。注意，直接迭代命题 5.2 的第 3 点可知，对于所有充分大的 n，$T_n^{(k-1)}$ 为好的概率至少为 $1 - \varepsilon/3$。此外，由引理 3.5 还可知，对某些函数 f 和 g，以至少 $1 - \varepsilon/3$ 的概率有 $0 < g(\varepsilon) \leq Y_{k-2} \leq f(\varepsilon)$，其中 $g(\varepsilon) > 0$ 且 $f(\varepsilon) < \infty$。因此，我们可以假设已将 $T_{(k-1)n}$ 与连续随机树耦合，$T_{(k-1)n}$ 是好的，并且 $g(\varepsilon) \leq Y_{k-2} \leq f(\varepsilon)$。

在该耦合下，我们可以写成 $\frac{|T_n^{(k-1)}|}{\beta_n \sqrt{n}} = Y_{k-2} = Y_{k-2}' + \varepsilon'$，其中 $\varepsilon' \in [-\zeta, \zeta]$。因此，由于 $T_{(k-1)n}$是良好的，根据命题 5.2(1)及 $B = f(\varepsilon)$，可知对任意 $C \in (0, \infty)$，

$$\begin{aligned}\mathbb{P}\Big(Y_{k-1} - Y_{k-2} > C \;\Big|\; T_n^{(k-1)}\Big) &= \exp\left\{-\frac{(C + Y_{k-2}' + \varepsilon')^2 - (Y_{k-2}' + \varepsilon')^2}{2}\right\} + o(1) \\ &= \exp\left\{-\frac{(C + Y_{k-2}')^2 - (Y_{k-2}')^2}{2}\right\}\Big(1 + e^{-C\varepsilon'} - 1\Big) + o(1),\end{aligned}$$

so

$$\left|\mathbb{P}\Big(Y_{k-1} - Y_{k-2} > C \;\Big|\; T_n^{(k-1)}\Big) - \exp\left\{-\frac{(C + Y_{k-2}')^2 - (Y_{k-2}')^2}{2}\right\}\right| \leq |1 - e^{-C\varepsilon'}| e^{\frac{-C^2}{2}} + o(1) \leq C\zeta e^{\frac{-C^2}{2}} + o(1). \tag{25}$$

右侧第一项随着 $\zeta \to 0$ 趋于 0，且对 $C > 0$ 均匀成立。根据引理 2.10，存在一个依赖于 ε 的 η，使得若 $f(\varepsilon) \leq Y_{k-2} \leq g(\varepsilon)$，且 (25) 右侧小于 η，那么我们可以将 $Y_{k-1} - Y_{k-2}$ 与 $Y_{k-1}' - Y_{k-2}'$ 耦合，使它们彼此 $\varepsilon/4$ 接近的概率至少为 $1 - \varepsilon/4$。发生这种情况时，根据三角不等式，我们有 $|Y_{k-1} - Y_{k-1}'| < \varepsilon/2$。因此，我们选取足够小的 ζ（且小于 $\varepsilon/8$），并选取足够大的 n，使得右侧小于这个 η。

然而，我们注意到 Z_{k-2} 与 Y_{k-1} 并不独立，因此需要将 (Y_{k-1}, Z_{k-2}) 与 (Y_{k-1}', Z_{k-2}') 这对变量进行耦合。为此，我们将 $\mathbb{R}^+$ 分解为长度为 $\varepsilon/2$的区间，也就是说，我们写作 $\mathbb{R}^+ = \bigcup_{j=0}^{\infty} I_j$，其中 $I_j = [j\varepsilon/2, (j+1)\varepsilon/2)$。令 M_{k-1}（分别令 M_{k-1}'）为满足 $Y_{k-1} \in I_j$（分别满足$Y_{k-1}' \in I_j$）的唯一 j。根据引理 2.10 及上述讨论，存在 M_{k-1} 与 M_{k-1}' 之间的一种耦合，使得它们之差至多为 1 的概率为 $1 - \varepsilon/8$。根据引理 3.5，当 n 足够大时（此时 M_{k-1} 也足够大），以至少 $1 - \varepsilon/8$ 的概率有 $M_{k-1}' \leq n^{1/2+\varepsilon} - 1$。

然后，给定 M_{k-1}，我们根据其条件分布抽取 Z_{k-2}。由断言 6.1，当 n 足够大时，对每个 $j \leq n^{1/2+\kappa}$，在给定 $M_{k-1} = j$ 的条件下，Z_{k-2} 与 $[0, Y_{k-2}]$ 上的均匀随机变量之间的普罗霍罗夫距离至多为 ζ。由命题 2.9，$[0, Y_{k-2}]$ 上的均匀随机变量与 Z_{k-2}' 之间的普罗霍罗夫距离至多为 ζ（回顾一下，给定 Y_{k-2}'，Z_{k-2}' 与 Y_{k-1}' 独立，因此也与 M_{k-1}' 独立）。因此，在给定 $M_{k-1} = j$ 的条件下，Z_{k-2}' 与 Z_{k-2} 之间的普罗霍罗夫距离至多为 2ζ。由于 $\zeta < \varepsilon/8$，可知我们可以对数对 (Y_{k-1}, Z_{k-2}) 和 (Y_{k-1}', Z_{k-2}') 进行耦合，使得 $|Y_{k-1} - Y_{k-1}'| < \varepsilon$ 和 $|Z_{k-2}' - Z_{k-2}| < \varepsilon$ 以至少 $1 - \varepsilon/2$ 的概率成立，正如所需。 □

推论 6.3. 对每个 $\varepsilon > 0$ 和 $k \geq 1$，存在 N，使得对所有 $n > N$，我们都可以耦合连续随机树的折棍过程与均匀生成树的折棍过程，使得以至少 $1-\varepsilon$ 的概率，对所有 $0 \leq i$、$j \leq k$，均有

$$|d(y_i, y_j) - d'(y'_i, y'_j)| \leq \varepsilon.$$

证明。取 $\eta > 0$ 和 $k \geq 1$。我们验证存在一种耦合，使得命题 3.6 的每个条件都以高概率成立。

对于第一个条件，注意到由命题 6.2 可知，我们可以对连续随机树和均匀生成树的断棍过程进行耦合，使得对于所有 $0 \leq i \leq k$ 和每个 $0 \leq i \leq k-1$，$|Y_i - Y'_i| \leq \eta$ 和 $|Z_i - Z'_i| \leq \eta$，其概率对于所有充分大的 n 至少为 $1-\eta$。对于第二个条件，注意到由命题 3.3 可知，我们可以选择 $\delta = \delta(\eta, k) > 0$，使得对于所有 $i \leq k-1$，$|Z'_i - Y'_j| \geq 3\eta$ 和 $j \leq k$ 的概率至少为 $1-\delta$，并且当 $\eta \downarrow 0$ 时，$\delta \downarrow 0$。

因此，由命题 3.6 可知，在这种耦合下，$1-\eta-\delta$ 的概率至少为 $_{1\leq i,\ j\leq k}\, |d(y_i, y_j) - d'(y'_i, y'_j)| \leq 2k\eta$。给定 $\varepsilon > 0$，因此我们可以将 $\eta > 0$ 取得足够小，使得 $2k\eta + \delta < \varepsilon$，从而得出所述断言。 □

定理 3.1 的证明。对于 $k \geq 1$，令 $D_n^{(k)}$ 表示 UST(G_n) 中 k 个均匀点之间的距离矩阵。令 $D^{(k)}$ 表示连续随机树中相应的距离矩阵。

我们证明了，对于任意 $k \geq 1$ 和任意 $\varepsilon > 0$，都可以将 $T_n^{(k)}$ 与连续随机树耦合，使得对于所有充分大的 n，以至少 $1-\varepsilon$ 的概率有 $||D_n^{(k)} - D^{(k)}||_\infty < \varepsilon$。因此，$||D_n^{(k)} - D^{(k)}||_\infty$ 依概率收敛到 0，进而 $D_n^{(k)}$ 依分布收敛到 $D^{(k)}$，这等价于定理 3.1 的陈述。 □

6.2 下质量界

为了将定理 3.1 中得到的收敛加强为 GHP收敛（从而证明定理 1.3），只需验证命题 2.13(b) 成立。因此，在我们的情形下，只需证明以下结论。

命题 6.4（下质量界）。设 $(G_n)_{n\geq 1}$ 是满足定理 1.3 假设的确定性图稠密序列。对每个 $n \geq 1$，令 $\mathcal{T}_n$ 为从 G_n 中均匀抽取的生成树。记 $d_{\mathcal{T}_n}$ 为 $\mathcal{T}_n$ 上相应的图距离，记 μ_n 为 $\mathcal{T}_n$ 的顶点上的均匀概率测度。那么，对于任意 $c > 0$ 和任意 $\eta > 0$，存在某个 $\varepsilon > 0$，使得对所有 $n \in \mathbb{N}$

$$\mathbb{P}\big(\exists v \in \mathcal{T}_n : |B_{\mathcal{T}_n}(v, c\sqrt{n})| \leq \varepsilon n\big) \leq \eta.$$

[7] 的结果为序列 $(G_n)_{n\geq 1}$ 建立了下质量界，其中对所有 n 均有 $|G_n| = n$，且该序列满足以下三个条件。

$\sqrt{n}$1. 存在 $\theta < \infty$，使得 $\sup_n \sup_{x\in G_n} \sum_{t=0} (t+1)p_t(x,x) \leq \theta$。

2. 存在 $\alpha > 0$，使得当 $n \to \infty$ 时，$t_{\text{mix}}(G_n) = o(n^{\frac{1}{2}-\alpha})$。

3. 对所有 n，G_n 都是传递的。

对于满足定理 1.3 假设的图序列，注意到第二个条件由断言 2.5 可立即得到满足。由于对于所有 $x \in G_n$ 和所有 $t \geq 1$，$p_t(x,x) \leq \frac{1}{\delta n}$，第一个条件也得到满足。

然而，我们希望放宽 G_n 是传递的这一条件，转而只要求图是平衡的；也就是说，存在一个常数 $D < \infty$，使得

$$\frac{\max_{v\in G_n} \deg v}{\min_{v\in G_n} \deg v} \leq D$$

对所有 n 都成立。正如 [7], 末尾所述，只需在所有计算中保留常数 D，就可以轻松地将下质量界的证明推广到这种情形；由于这些细节并无启发性，我们不再赘述。在定理 1.3 的假设下，我们可以取 $D = \delta^{-1}$，因此可轻松验证命题 6.4，进而验证命题 2.13(b)。此外，定理 3.1 保证命题 2.13(a) 也成立。因此，定理 1.3 可直接得出。

7 定理 1.1 和推论 1.2 的证明

回顾引言中的内容，图子 W 是非退化的，当且仅当函数

$$\deg_W(x) := \int_{[0,1]} W(x,y)dy$$

对每个 $x \in [0,1]$都有定义且严格为正；非退化图子 W 是连通的，当且仅当对每个可测的 $A \subset [0,1]$，都有

$$\int_A \int_{A^C} W(x,y)dxdy > 0.$$

为了根据定理 1.3 验证定理 1.1，我们需要验证在定理 1.1 的假设下，该图序列是一个扩展图序列，且 $\alpha_n \to \alpha_W$。

我们先从第 一个性质开始。回顾一下， γ-扩展图序列的定义见定义 2.4 .

命题 7.1. 设 $W\colon [0,1]^2 \to [0,1]$ 是一个连通图子，且 G_n 是一个最小度至少为 δn、在割距离下收敛到 W 的加权图序列。那么存在 $\gamma = \gamma(W,\delta) > 0$，使得 $(G_n)_{n\geq 1}$是一个 γ-扩展图序列。

证明. 取 $U \subset G_n$。根据 $|U| \geq \frac{1}{2}\delta n$ 是否成立，我们将证明分为两种情况。

情况 1： $|U| \leq \frac{1}{2}\delta n$。由于 G_n的最小度至少为 δn，且每条边的最大权重为 1，因此从每个顶点出发、通向 $V(G) \setminus U$的边的总权重为 $\frac{1}{2}\delta n$，于是

$$w(U, V(G)\setminus U) \geq \frac{1}{2}|U|\delta n \geq \frac{1}{2}\delta |U|(V(G) - |U|).$$

情况 2： $|U| > \frac{1}{2}\delta n$。由引理 2.2，存在一个常数 $\beta = \beta(W,\delta)$，使得对于每个满足$\frac{\delta}{2} \leq \mu(U) \leq 1/2$ 的集合 U，我们有

$$\int_U \int_{U^C} W(x,y) > \beta.$$

特别地，由于 G_n 收敛到 W，这意味着存在 $N > \infty$，使得对于所有这样的 U 以及所有 $n \geq N$，

$$\frac{1}{n^2} w(U, V(G)\setminus U) = \int_U \int_{U^C} W_n(x,y) > \frac{\beta}{2}.$$

由于 $|U|(V(G) - |U|) \leq n^2$ 显然成立，这意味着 $e(U, V(G) \setminus U) \geq \frac{\beta}{2}|U|(V(G) - |U|)$。 □

现在我们来验证收敛 $\alpha_n \to \alpha_W$。回想一下，在第 4 节中，我们定义了

$$\alpha_n = \frac{n\mathbb{E}_\pi\big[\mathrm{Cap}_M(X[0, n^{\kappa/2}))\big]}{Mn^{\kappa/2}}.$$

其中 X 是在 G_n 从平稳分布出发，并证明在某些假设下，序列

$$\left(\mathrm{UST}(G_n), \frac{\sqrt{\alpha_n} d_n}{\sqrt{n}}, \mu_n\right) \xrightarrow{(d)} (\mathcal{T}, d, \mu) \tag{26}$$

关于 t GHP 拓扑。我们还令 U_{π_n} 表示 G_n的一个随机平稳顶点，并定义 e

$$\tilde{\alpha}_n = n\mathbb{E}[\pi_n(U_{\pi_n})].$$

事实上，处理 $\tilde{\alpha}_n$ 比处理 α_n更方便。正如下述断言所示，这已经足够（我们为完整起见写出证明，但实际上只需利用期望的线性性和推论 2.8 即可直接得到）。

论断 7.2. 设 $(G_n)_{n\geq 0}$ 是一列定义在 n 个顶点上的加权图，且最小度为 δn。设 α_n 和 $\tilde{\alpha}_n$ 如上定义。则当 $n \to \infty$ 时， $\alpha_n = \tilde{\alpha}_n(1 + o(1))$。

证明. 根据 Bonferroni 不等式和期望的线性性，令 Z 表示一个从平稳分布出发的独立随机游走，并回顾引理 2.6 中 $\mathrm{Cap}_M(U_\pi) \geq \frac{M\delta}{2n}$ 确定成立，可得：

$$\begin{aligned}\left|\mathbb{E}\Big[\mathrm{Cap}_M(X[0,n^{\varepsilon/2}))\Big] - n^{\varepsilon/2}\mathbb{E}[\mathrm{Cap}_M(U_\pi)]\right| &= \left|\mathbb{E}\Big[\mathrm{Cap}_M(X[0,n^{\varepsilon/2}))\Big] - \sum_{i=0}^{n^{\varepsilon/2}}\mathbb{E}[\mathrm{Cap}_M(X_i)]\right| \\ &\leq \mathbb{P}_\pi\Big(\exists 0 \leq t_1 < t_2 \leq M-1 : Z_{t_1} \cap X[0,n^{\varepsilon/2}) \neq \emptyset \text{ and } Z_{t_2} \cap X[0,n^{\varepsilon/2}) \neq \emptyset\Big) \\ &\leq \left(\frac{Mn^{\varepsilon/2}}{\delta n}\right)^2 \leq \frac{2Mn^{\varepsilon/2}}{\delta^3 n} \cdot n^{\varepsilon/2}\mathbb{E}[\mathrm{Cap}_M(U_\pi)].\end{aligned}$$

类似地，注意到对于所有 $v \in G_n$， $\pi(v) \geq \frac{\delta}{n}$ 都确定成立，因此

$$\begin{aligned}|\mathbb{E}[\mathrm{Cap}_M(U_\pi)] - M\mathbb{E}[\pi(U_\pi)]| &= \left|\mathbb{E}[\mathrm{Cap}_M(U_\pi)] - \sum_{i=0}^{M-1}\mathbb{P}_\pi(Z_i = U_\pi)\right| \\ &\leq \mathbb{P}_\pi(\exists 0 \leq t_1 < t_2 \leq M-1 : Z_{t_1} = Z_{t_2} = U_\pi) \leq \left(\frac{M}{\delta n}\right)^2 \leq \frac{M}{\delta^3 n} \cdot M\mathbb{E}[\pi(U_\pi)].\end{aligned}$$

综上，我们合并可得

$$\alpha_n = \frac{n}{Mn^{\varepsilon/2}}\mathbb{E}\Big[\mathrm{Cap}_M(X[0,n^{\varepsilon/2}))\Big] = \frac{n}{M}\mathbb{E}[\mathrm{Cap}_M(U_\pi)](1+o(1)) = n\mathbb{E}[\pi(U_\pi)](1+o(1)) = \tilde{\alpha}_n(1+o(1)),$$

如所要求。 □

因此，(26) 的收敛性对于用 $(\tilde{\alpha}_n)_{n\geq 1}$ 替代 $(\alpha_n)_{n\geq 1}$ 的序列同样成立。为证明主要收敛定理，只需证明在定理 1.1 的假设下，

$$\tilde{\alpha}_n \to \alpha_W, \tag{27}$$

其中 α_W 如 (1) 所示。

注 7.3。注意，当 G_n 为正则时， $\tilde{\alpha}_n$ 为 1，因此显然 (27) 将推出，对于一个正则图序列， $\alpha_W = 1$ 。

我们的下一个目标是证明以下结论。

命题 7.4. 设 $W:\ [0,1]^2 \to [0,1]$ 为一个连通图子，且设 G_n 为一个带有平稳测度 π_n 、在割距离下收敛至 W的图序列。那么，

$$\tilde{\alpha}_n := n\mathbb{E}_{\pi_n}[\pi_n(v)] \to \frac{1}{\left(\int_{[0,1]^2} W(x,y)dxdy\right)^2} \cdot \int_{[0,1]}\left(\int_{[0,1]} W(x,y)dy\right)^2 dx.$$

证明. 我们先来证明

$$\frac{1}{n^2}\sum_{v\in V}\deg_{G_n}(v)\to\left(\int_{[0,1]^2}W(x,y)dxdy\right). \tag{28}$$

事实上，由于 $G_n\to W$ 在割距离下，存在 ϕ_n 保测自同构，使得 G_n 的图子表示 W_n 满足

$$\sup_{S,T\in\mathcal{B}([0,1])}\left|\int_S\int_T W_n^{\phi_n}(x,y)-W(x,y)dxdy\right|\to 0.$$

因此我们 今后用 W_n 代替 $W_n^{\phi_n}$。特别地，取 $S=T=[0,1]$ 可得 t

$$\int_{[0,1]}\int_{[0,1]}W_n(x,y)dxdy\to\int_{[0,1]}\int_{[0,1]}W(x,y).$$

然而，根据第 2.1 节中给出的图的图子表示的定义，我们还有

$$\int_{[0,1]}\int_{[0,1]}W_n(x,y)dxdy=\frac{1}{n^2}\sum_{v\in V}\deg_{G_n}(v),$$

由此得到式 (28)。接下来，我们将证明

$$\frac{1}{n^3}\sum_{v\in G_n}\deg_{G_n}(v)^2\to\int_{[0,1]}\left(\int_{[0,1]}W(x,y)dy\right)^2dx \tag{29}$$

注意，在 G_n 的每个图子表示 W_n 中，我们都有

$$\frac{1}{n^3}\sum_{v\in G_n}\deg_{G_n}(v)^2=\frac{1}{n}\sum_{v\in G_n}\left(\frac{\deg_{G_n}(v)}{n}\right)^2=\frac{1}{n}\sum_{v\in G_n}\left(\int_{[0,1]}W_n(x_v,y)dy\right)^2,$$

其中， x_v 是 $[0,1]$ 中的某个点，对应于 v。此外，按照第 2.1 节中的记号，由那里给出的构造可知,

$$\frac{1}{n}\sum_{v\in G_n}\left(\int_{[0,1]}W_n(x_v,y)dy\right)^2=\sum_{i=1}^n\int_{I_i}\left(\int_{[0,1]}W_n(x,y)dy\right)^2dx=\int_{[0,1]}\left(\int_{[0,1]}W_n(x,y)dy\right)^2dx.$$

因此，要证明 (29)，只需证明

$$\int_{[0,1]}\left(\int_{[0,1]}W_n(x,y)dy\right)^2dx\to\int_{[0,1]}\left(\int_{[0,1]}W(x,y)dy\right)^2dx.$$

换言之，记 deg_W 和 deg_{W_n} 分别为图子W 和 W_n 对应的归一化度函数（定义见 (2)），我们需要证明

$$\int_{[0,1]}\big(\deg_{W_n}(x)^2-\deg_W(x)^2\big)dx\to 0. \tag{30}$$

由于 deg_{W_n} 和 deg_W 是可测函数，我们有集合 $\{x\in[0,1]\colon \mathsf{deg}_{W_n}(x)\ >\mathsf{deg}_W(x)\}$ 是可测的。将此集合记为 S。我们有

$$\begin{aligned}\int_S\deg_{W_n}(x)^2-\deg_W(x)^2dx&=\int_S(\deg_{W_n}(x)-\deg_W(x))(\deg_{W_n}(x)+\deg_W(x))dx\\&\le 2\int_S(\deg_{W_n}(x)-\deg_W(x))dx=2\int_S\int_{[0,1]}W_n(x,y)-W(x,y)dydx\to 0.\end{aligned}$$

由对称性并考虑 S^c，我们同样有

$$\int_{S^c} \deg_{W_n}(x)^2 - \deg_W(x)^2 dx \to 0,$$

由此　　我们得出 (30) 成立，因而 (29) 也成立。最后，由 (28) 和 (29) 可知

$$\begin{aligned}\tilde{\alpha}_n &= n \sum_{v \in G_n} \left(\frac{\deg(v)}{\sum_{v \in G_n} \deg(v)} \right)^2 = n \cdot \frac{1}{\left(\sum_{v \in G_n} \deg(v)\right)^2} \sum_{v \in G_n} (\deg(v)^2) \\ &= \left(n^2 \cdot \frac{1}{\sum_{v \in G_n} \deg(v)} \right)^2 \cdot \left(\sum_{v \in G_n} \frac{1}{n^3} \deg(v)^2 \right) \to \frac{1}{\left(\int_{[0,1]^2} W(x,y) dx dy \right)^2} \cdot \int_{[0,1]} \left(\int_{[0,1]} W(x,y) dy \right)^2 dx,\end{aligned}$$

正如所需。 □

定理 1.1 的证明。我们在命题 7.1 中证明，在定理 1.1 的假设下，所讨论的图序列是扩展图序列，因此定理 1.3 适用。在论断 7.2 和命题 7.4 中，我们证明了定理 1.3 结论中出现的序列 α_n 按 $n \to \infty$ 收敛到 α_W，正如所需。

□

推论 1.2 直接由定理 1.1 和引理 2.3 得出。

参考文献

[1] D. Aldous. 连续统随机树 I. Ann. Probab., 19(1):1–28, 1991.[2] D. Aldous. 连续统随机树 II：概述. 167:23–70, 1991.[3] D. Aldous. 连续统随机树 III. 概率年刊, pages 248–289, 1993.[4] N. Alon. 特征值与扩张图. 组合数学, 6(2):83–96, 1986.[5] N. Alon and V. Milman. $\lambda 1$、图的等周不等式与超集中器. 组合理论杂志 B 辑, 38(1):73–88, 1985.[6] N. Alon, A. Nachmias, and M. Shalev. 稠密图的均匀生成树直径. 组合学、概率与计算, 31(6):1010–1030, 2022.[7] E. Archer, A. Nachmias, and M. Shalev. 高维均匀生成树的 GHP 标度极限. arXiv 预印本 arXiv:2112.01203, 2021.[8] S. Athreya, W. L¨ohr, and A. Winter. Gromov-模糊拓扑与 Gromov–Hausdorff-模糊拓扑之间的差距. 随机过程及其应用, 126(9):2527–2553, 2016.[9] B. Bollob´as, C. Borgs, J. Chayes, and O. Riordan. 稠密图序列上的渗流. 概率年刊, 38(1):150–183, 2010.[10] C. Borgs, J. Chayes, L. Lov´asz, V. S´os, and K. Vesztergombi. 稠密图的收敛序列 I：子图频率、度量性质与测试. 数学进展, 219(6):1801–1851, 2008.[11] P. Caines and M. Huang. graphon 均值场博弈与 GMFG 方程： ε-纳什均衡. 载于《2019 IEEE 第 58 届决策与控制会议（CDC）论文集》，第 286–292 页。IEEE，2019.[12] J. Cooper, T. Martins, and D. Kr´al. 有限可强制图极限具有普适性. 数学进展, 340:819–854, 2018.

[13] C. Giardin`a, C. Giberti, and E. Magnanini. Erd¨os–r´enyi 随机图中三角形累积量生成函数的近似. 统计物理学杂志, 182(2):1–22, 2021.[14] D. Glasscock. 什么是……图子. 美国数学学会通报, 62(1), 2015. [15] A. Greven, P. Pfaffelhuber, and A. Winter. 随机度量测度空间依分布收敛（Λ-合并测度树）. 概率论及相关领域, 145(1-2):285–322, 2009.[16] J. Hladk`y, A. Nachmias, and T. Tran. 稠密图上均匀生成树的局部极限. 统计物理学杂志, 173(3):502–545, 2018.[17] W. Hoeffding. 有界随机变量之和的概率不等式. 载于《Wassily Hoeffding 文集》，第 409–426 页。施普林格，1994.[18] T. Hutchcroft. 交错渗流与接线均匀生成森林. 概率年刊, 46(2):1170–1200, 2018.[19] M. Jerrum, J.-B. Son, P. Tetali, and E. Vigoda. 可分解马尔可夫链的 Poincar´é 常数与对数索博列夫常数的初等界. 应用概率年刊, 14(4):1741–1765, 2004.[20] G. Kirchhoff. ¨关于求解线性分布研究中所导出的方程. 物理学年鉴, 148(12):497–508, 1847.[21] D. Krioukov. 聚类意味着网络具有几何结构. 物理评论快报, 116(20):208302, 2016.[22] G. Lawler and A. Sokal. 马尔可夫链与马尔可夫过程的谱隙界：切格尔不等式的推广. 美国数学会会刊, 309(2):557– 580, 1988.[23] J.-F. Le Gall. 随机树及其应用. 概率综述, 2:245–311, 2005.[24] D. Levin, Y. Peres, and W. E. 马尔可夫链与混合时间，第 107 卷。美国数学会，2017.[25] L. Lov´asz and B. Szegedy. 稠密图序列的极限. 组合理论杂志 B 辑, 96(6):933–957, 2006.[26] R. Lyons and Y. Peres. 树与网络上的概率，第 42 卷。剑桥大学出版社，2017.[27] G. Medvedev. 稠密图与图极限上的非线性热方程. 工业与应用数学学会数学分析杂志, 46(4):2743–2766, 2014.[28] G. Miermont. 空间多类型 Galton–Watson 树的不变性原理. Annales de l’Institut Henri Poincar´e, Probabilit´es et Statistiques, 44(6):1128 – 1161, 2008.[29] G. Miermont. 任意亏格随机地图的镶嵌. Ann. Sci. ´Ec. Norm. Sup´er. (4), 42(5):725–781, 2009.[30] R. Pemantle. 均匀地选取整数格上的一棵生成树. 概率年刊, 19(4):1559–1574, 1991.[31] Y. Peres and D. Revelle. 有限图上均匀生成树与环擦随机游走的标度极限. arXiv 数学电子预印本, page math/0410430, 2004 年 10 月.[32] D. Wilson. 以快于覆盖时间的速度生成随机生成树. 载于《第二十八届 ACM 计算理论年会论文集》（费城，PA，1996），第 296–303 页。美国计算机协会，纽约，1996.