稀疏几何集合系统与 Beck-Fiala 猜想

Kunal Dutta * [1] 和 Arijit Ghosh * [2]

[1]华沙大学数学、信息学与力学学院，波兰。电子邮箱：K.Dutta@mimuw.edu.pl[2]印度统计研究所，印度加尔各答。电子邮箱：arijitiitkgpster@gmail.com

arXiv:2301.03329v1 [cs.CG] 9 Jan 2023

摘要

我们研究 Beck-Fiala 情形下浅层单元复杂度有界的几何集合系统的组合差异度，其中每个点至多属于 t 个范围。对于浅层单元复杂度为 $\psi(m,k) = g(m)k^c$ 的集合系统，其中 (i) $g(m) = o(m^\varepsilon)$ 对任意 $\varepsilon \in (0,1]$, (ii) ψ 在 m 中单调不减，且 (iii) $c > 0$ 与 m 和 k 无关，我们得到如下差异界：

$$O\left(\sqrt{\left(\log n + (t^c g(n))^{\frac{1}{1+c}}\right)\log n}\right).$$

对于 $t = \omega(\log^2 n)$，在若干情况下，例如 $\mathbb{R}^2$ 中点与半平面／圆盘／伪圆盘的集合系统，以及 $\mathbb{R}^3$ 等中的点与正交象限，这些界为 $o(\sqrt{t})$，从而验证了 Beck 和 Fiala 的猜想界，并进一步改进了该界（《离散应用数学》，1981年）。

我们的界是通过证明存在交叉数较低的匹配而得到的。证明中结合了 Welzl 的乘法权重更新方法（SoCG，1988年），以及 Mustafa 近期关于集合系统浅层打包的界，该界以其浅层单元复杂度表示。对于浅层单元复杂度为 $\psi(m,k) = m^{c_1} g(m)k^c$ 的集合系统，我们得到交叉数至多为

$$O\left((n^{c_1} g(n) t^c)^{\frac{1}{1+c_1+c}}\right).$$

这些结果本身也具有独立的研究价值。

1 引言

给定一个集合系统 $(X,\mathcal{R})$，其基集为 X，并且包含 X 的子集集合族 $\mathcal{R} \subset 2^X$，该系统的组合差异度定义为

$$\operatorname{disc}(\mathcal{R}) := \min_{\chi: X \to \{\pm 1\}} \max_{S \in \mathcal{R}} \left| \sum_{i \in S} \chi(i) \right|,$$

也就是说，在所有可能的 X 二分划分中取最小值，即集合系统 $\mathcal{R}$ 中所有集合的最大失衡程度。利用系统 $(X,\mathcal{R})$ 的关联矩阵 A，可以将其等价地表述为线性代数问题——向量平衡问题，具体如下：

$$\operatorname{disc}(A) := \min_{x \in \{-1,1\}^n} \|Ax\|_\infty.$$

*由波兰 NCN-SONATA 项目资助，项目编号：2019/35/D/ST6/04525 *(Probabilistic tools for high-dimensional geometric inference, topological data analysis and large-scale networks)*。

对于给定的一类集合系统，差异度最小化旨在根据基集的大小或其他参数，确定该类系统差异度的界。除了本身具有研究价值外，差异度最小化还广泛应用于数学、计算机科学及相关领域 [13, 30]。

Beck-Fiala 猜想与 Komlos 猜想 在开创性研究中，Beck 和 Fiala 首次使用了“差异度”这一术语，并开启了对集合系统差异度的研究。他们 [10]证明，若每个元素至多属于 t 个集合，则该集合系统的差异度至多为 $2t-1$。他们猜想，这一界可以改进为 $O\left(\sqrt{t}\right)$。用线性代数的术语来说，一个密切相关的推广被称为 Komlos 猜想：对于每一列的 ℓ_2 范数至多为 1 的实矩阵，其差异度为 $O(1)$。尽管在过去 4 年里，人们在差异理论的部分结果和理解方面取得了显著进展，这些引人遐思的猜想仍未解决。

早期进展与近期进展 自 Beck 和 Fiala 的开创性成果以来，人们为解决 Beck-Fiala 猜想和 Komlos 猜想取得了许多进展。该领域早期的突破包括 Spencer [39] 和 Gluskin [22] 的工作，他们证明 Komlos 猜想在 $t=\Theta(n)$的情形下成立。此后，Srinivasan [40] 和 Banaszczyk [3] 为一般的 t证明了若~~干界。~~后者给出的$5\sqrt{t}\log\ n$ 界仍是目前关于 Komlos 猜想的最佳结果。

上述差异界都是非构造性的，甚至有人猜测，不存在能够高效构造满足所述差异界的着色的算法。然而，Bansal [4] 取得了重大突破，提出了一种高效算法，能够为 $|\mathcal{R}|=\Theta(n)$ 构造出达到 Spencer 差异界的着色。此后，涌现出大量关于上界和下界的最新成果，既有算法方面的，也有非算法方面的；此外还有许多应用与推广，参见 [28, 5, 6, 16, 27, 37, 7]。

几何集合系统 具有几何解释的集合系统的差异度，例如 $\mathbb{R}^2$ 和 $\mathbb{R}^3$ 中的点与半平面，或 $\mathbb{R}^2$ 中的点与圆盘，其研究在理论和实践上都具有重要意义，因为它与计算几何、统计学习与机器学习、算法分析、数据库理论等多个领域相关。 [30, 25, 26]。自差异理论发展的早期以来，人们便开始研究诸如 Tu´snady 问题这类问题 [9, 11]。其他值得注意的成果包括点与射影平面（Spencer [1]）、VC 维数有界的原系统（Matouˇsek [29]）、对偶 VC 维数有界的系统等（Matouˇsek-Welzl-Wernisch [32]）。

近期，Ezra [21] 提出了尺寸敏感差异度的概念，并进一步探讨了元素度有界的几何集合系统的差异度。结合 Dutta、Ezra 和 Ghosh [17], 的装箱与链式框架，以及 Lovett-Meka 的基于部分着色的算法（或类似算法），相对容易为 $\mathbb{R}^2$中的点与半平面得到 $O\left(t^{1/4}\cdot\log n\right)$ 界。在一定程度上，未发表的工作 [20], 已考虑了这一情形，但其界多出一个 $\sqrt{\log\log\ t}\log\ n$因子（如上例所示）；这是由于采用了部分着色方法以及参数分配并非最优所致。[1]~~据我们~~所知，除上述未发表的草稿外，尚无已发表或公开可得的工作探讨元素度有界的几何集合系统这一问题。然而，这类系统仍有一些值得研究的问题。例如，一个有趣的问题是：在 Beck-Fiala 情形下，能否利用（例如）Bansal 等人的算法框架，将点与半平面的 $O\left(t^{1/4}\log n\right)$ 界改进为 $O\left(t^{1/4}\sqrt{\log n}\right)$ 或更优?

[1]此外，这些界限似乎并未出现在已发表的版本中。

并集与浅层单元复杂度 具有较低并复杂度或浅层单元复杂度的集合系统构成了一个相当一般的类别，其中包括若干重要情形，例如 $\mathbb{R}^2$ 和 $\mathbb{R}^3$中的点与半空间、 $\mathbb{R}^2$ 中的点与（伪）圆盘，以及其他若干情形（例如参见 [36, 第 4]章）。此外，由于低并复杂度的对偶系统，其对应原系统的浅层单元复杂度也较低（例如 [35]），因此已有一系列工作对其进行了研究，例如 [14, 41, 12, 8, 34, 18, 19] ；针对若干值得关注的结构，这些工作得到了更好的界，包括无权和加权 ε 网、ε 近似、相对近似、 ϵ-括号、组合麦克比斯区域等。

我们的贡献

我们研究浅层单元复杂度为线性或近线性的几何集合系统的差异度，其中每个元素最多属于 t 个集合，且 t 是一个与 n无关的给定参数。我们的差异界如下所述。

定理 1. 设$(X, \mathcal{R})$是一个浅层单元复杂度为 $\psi(m,k) = g(m)k^c$的集合系统，其中$g(m) = o(m^\varepsilon)$对每个 $\varepsilon \in (0,1)$均成立，是关于 m 的非递减函数，且 $c > 0$ 是一个与 n 和 k无关的常数；此外， X 的每个元素最多属于 t个 $\mathcal{R}$的集合。那么$(X, \mathcal{R})$的差异度至多为

$$O\left(\sqrt{\left(\log n + (t^c g(n))^{\frac{1}{1+c}}\right)\log n}\right).$$

特别地，对于以下度至多为 t的集合系统族，其差异度界如下。

1. $\mathbb{R}^2$中的点与半平面： $O\left(\sqrt{(\log n + t^{1/2})\log n}\right)$。
2. $\mathbb{R}^2$中的点与凸体的位似像： $O\left(\sqrt{(\log n + t^{1/2})\log n}\right)$。
3. $\mathbb{R}^2$中的点与圆盘： $O\left(\sqrt{(\log n + t^{1/2})\log n}\right)$。
4. $\mathbb{R}^2$中的点与伪圆盘： $O\left(\sqrt{(\log n + t^{1/2})\log n}\right)$。
5. $\mathbb{R}^2$中的点与 α-胖三角形： $O\left(\sqrt{(\log n + (t\cdot \log^* n)^{1/2})\log n}\right)$。
6. 点与具有线性并集复杂度的对象： $O\left(\sqrt{(\log n + t^{1/2})\log n}\right)$。
7. $\mathbb{R}^2$ 中局部 γ-胖的、有界描述复杂度的半代数对象：$O\left(\sqrt{(\log n + t^{1/2}\cdot 2^{O(\log^* n)})\log n}\right)$。
8. $\mathbb{R}^3$中的点与半空间： $O\left(\sqrt{(\log n + t^{2/3})\log n}\right)$。
9. $\mathbb{R}^2$ 和 $\mathbb{R}^3$中的点与正交象限： $O\left(\sqrt{(\log n + t^{1/2})\log n}\right)$。

这些界与已知的有界度集合系统一般界相匹配，或优于这些界，也优于猜想中的一般界。例如，对于 $\mathbb{R}^2$中的点与半平面所构成的集合系统，我们得到的界为

$$O\left(\left(\left(\log n + \sqrt{t}\right)\log n\right)^{1/2}\right),$$

对于 $t = \Omega(\log n)$，其为 $O\left(\sqrt{t}\right)$。对于 $t = \omega\left(\log^2 n\right)$，我们有 $\sqrt{\log n} = o(t^{1/4})$，因此该 $O\left(t^{1/4}\sqrt{\log n}\right)$ 界改进了 Beck 和 Fiala 的猜想 $O\left(\sqrt{t}\right)$ 界。因此，

即使不使用基于随机游走的框架，也不进行许多近期差异最小化算法所需的复杂分析，它们仍能给出好得多的界，例如 [28, 6, 38]。

更一般地，我们得到定理 1 的以下推论。

推论 2. 对于以下最大度为 t、且满足关于 t 的给定约束的集合系统，Beck-Fiala 猜想成立。

1. $\mathbb{R}^2$ 中的点与半平面：$t = \Omega\left(\log^2 n\right)$。2. $\mathbb{R}^2$ 中的点与凸体的位似像：$t = \Omega\left(\log^2 n\right)$。3. $\mathbb{R}^2$ 中的点与圆盘：$t = \Omega\left(\log^2 n\right)$。4. $\mathbb{R}^2$ 中的点与伪圆盘：$t = \Omega\left(\log^2 n\right)$。5. $\mathbb{R}^2$ 中的点与 α-胖三角形：$t = \Omega\left(\log^2 n \log^* n\right)$。6. 点与具有线性并集复杂度的对象：$t = \Omega\left(\log^2 n\right)$。7. $\mathbb{R}^2$ 中局部 γ-胖且具有有界描述复杂度的半代数对象：$t = \Omega\left(2^{O(\log^* n)} \log^2 n\right)$。8. $\mathbb{R}^3$ 中的点与半空间：$t = \Omega\left(\log^3 n\right)$。9. $\mathbb{R}^2$ 和 $\mathbb{R}^3$中的点与正交象限：$t = \Omega\left(\log^2 n\right)$。

我们的技术基于以下思路：利用 Welzl 的著名乘法权重更新方法 [42], ，结合 Mustafa [33] 以及 Dutta、Ezra 和 Ghosh [17]的浅层打包界，证明存在交叉数较低的匹配。这些交叉数界本身也具有独立价值，具体如下。

定理 3. 设 $(X, \mathcal{R})$ 是一个 n-点集合系统，具有 $|\mathcal{R}| \geq n$，由 X 与一类对象的交集生成，其中该类对象的对偶分裂维数至多为 d，浅层单元复杂度为 $\psi(m, k) = m^{c_1} \cdot g(m) k^c$，其中 c_1，$c \geq 0$ 与 m 和 k无关，并且 $g(m)$ 是关于 m 的非递减函数，使得对每个 $\varepsilon \in (0, 1)$，均有 $g(m) = o(m^{\varepsilon})$。此外，设 $(X, \mathcal{R})$ 满足：X 中的每个点至多属于 t个集合。那么，存在一个由 X 中的点构成的匹配，其交叉数至多为

$$O\left(\left(n^{c_1} t^c g(n)\right)^{\frac{1}{1+c_1+c}}\right).$$

特别地，如果 $c_1 = 0$，那么交叉数至多为

$$O\left(\left(t^c g(n)\right)^{\frac{1}{1+c}}\right).$$

推论 4. 对于以下集合系统 $(X, \mathcal{R})$，其中 X 的每个元素至多属于 $\mathcal{R}$ 的 t 个成员，存在满足以下交叉数界的匹配。

1. $\mathbb{R}^2$中的点与半平面：$O\left(\log n + \sqrt{t}\right)$。2. $\mathbb{R}^2$中的点与凸体的位似像：$O\left(\log n + \sqrt{t}\right)$。3. $\mathbb{R}^2$中的点与圆盘：$O\left(\log n + \sqrt{t}\right)$。4. $\mathbb{R}^2$中的点与伪圆盘：$O\left(\log n + \sqrt{t}\right)$。5. $\mathbb{R}^2$中的点与 α-胖三角形：$O\left(\log n + \sqrt{t \cdot \log^* n}\right)$。

6. 点与具有线性并集复杂度的对象： $O\left(\log n+\sqrt{t}\right)$。

7. 点与局部 γ-胖半代数对象，位于 $\mathbb{R}^2$，具有有界描述复杂度：$O\left(\log\ n+\sqrt{t\cdot 2^{O(\log^* n)}}\right)$。

8. $\mathbb{R}^3$中的点与半空间： $O\left(\log n+t^{2/3}\right)$。

9. $\mathbb{R}^2$ 和 $\mathbb{R}^3$中的点与正交象限： $O\left(\log n+t^{1/2}\right)$。

我们的结果的证明见第3节。虽然我们的差异度界和交叉数界并不难得到，但我们认为它们很重要，因为它们展示了非随机（且非平滑）的自然集合系统的实例，这些系统验证或改进了 Beck-Fiala 猜想。此外，这些界远强于一般有界 VC 维数或浅层单元复杂度的集合系统（没有度数界）的界；已知这些界是紧的 [31]。需要注意的是，仅仅利用低浅层复杂度并不能改进交叉数界，因为 [31] 中的紧性结果适用于二维和三维系统。低交叉数的匹配（或生成路径或树）受到广泛关注，因为它们应用于许多领域 [30, 15]。

2 预备知识

集合系统 $(X,\mathcal{R})$ 在基集的子集 $Y\subset X$ 上的投影记作 $\mathcal{R}_{|Y}:=\{R\cap Y\mid R\in\mathcal{R}\}$。

原始集合系统的基集 P 是 $\mathbb{R}^d$ 中的点，其子集由与 $\mathbb{R}^d$ 中一类几何对象 $\mathcal{G}$ 的所有可能交集给出，即集合系统 $(P,\mathcal{G}_{|P})$。给定一个（有限）几何对象集合族 $\mathcal{G}$，其对偶集合系统为 $(\mathcal{G},\mathcal{G}_{|\mathbb{R}^d})$，也就是说，它是 $\mathbb{R}^d$ 中的点按与 $\mathcal{G}$ 中对象相交所形成的等价类。

浅层单元复杂度在计算几何中有许多应用，包括改进 ε-网及相关结构的界（例如参见 [35]）。

定义 5（浅层单元复杂度）。若对任意 $Y\subseteq X$， $\mathcal{R}|_Y$ 中大小为 l 的子集数量至多为 $|Y|\cdot\psi(|Y|,l)$，则集合系统 $(X,\mathcal{R})$ 的浅层单元复杂度为 $\psi(\cdot,\cdot)$。

本文中，我们将始终研究浅层单元复杂度满足如下条件的集合系统

$$\psi(l,k)\ =\ l^{c_1}g(l)k^c, \tag{1}$$

其中 c_1、$c>0$ 是与 l、k无关的常数，并且对每个 $\varepsilon\in(0,1]$， $g(l)=o(l^\varepsilon)$。

定义 6（并复杂度）。若一组 m 几何对象 $\mathcal{O}$ 的并中所有维度的面的数量至多为 $f(m)$，则称其并复杂度为 $f(m)$。

众所周知（例如参见 [35]），由点与并复杂度为 $f(.)$ 的对象相交所形成的集合系统，其浅层单元复杂度至多为 $f(m/k)\cdot k^2$。

对于浅层单元复杂度有界的集合系统，一个核心引理是 [17, 33] 的浅层打包引理。陈述该引理之前，我们先定义浅层打包。

定义 7（浅层打包）。设 $(X,\mathcal{R})$ 是一个集合系统，且 δ 和 k 是正整数。范围 $\mathcal{P}\subseteq\mathcal{R}$ 的一个子集称为 k- 浅层 δ- 打包，或 (k,δ)- 打包，如果其中任意一对范围的对称差都大于 δ，并且对于所有 $R\in\mathcal{P}$，都有 $|R\cap X|\le k$。

接下来是 Mustafa [33], 的浅层打包引理，它根据集合系统的浅层单元复杂度，给出了其浅层打包的最优界。

定理 8（浅层打包引理，Mustafa [33]）。设 $(X,\mathcal{R})$ 是一个具有 $|X|=n$ 和浅层单元复杂度 $\varphi_{\mathcal{R}}$ 的集合系统。若 $(X,\mathcal{R})$ 的 VC 维数至多为 d_0，且 $(X,\mathcal{R})$ 是一个k-浅层 δ-打包，则

$$\frac{24d_0n}{\delta}\cdot\varphi_{\mathcal{R}}\left(\frac{4d_0n}{\delta},\frac{12d_0k}{\delta}\right)\leq C_{d_0}\frac{n}{\delta}\psi\left(\frac{n}{\delta},\frac{k}{\delta}\right),$$

其中 C_{d_0} 是一个与 n 和 δ 无关的常数。

最后，我们定义交叉数和低交叉数划分的概念，它们是本文的核心。

定义 9（交叉数；低交叉数划分）。若范围 $S\in\mathcal{R}$ 满足 $|\{x,y\}\cap X|=1$，则称集合 X, $\{x,y\}$ 中的一个元素对与其相交。给定将全集X 划分为若干对 $\Pi=(\{x_i,y_i\})_{i=1}^{n/2}$，其中对 $i\neq j$ 满足 $\{x_i,y_i\}\cap\{x_j,y_j\}=\emptyset$，且 $\bigcup_{i\geq1}\{x_i,y_i\}=X$，则 Π 的交叉数是 Π 中与任意给定范围 $\mathcal{R}$ 相交的元素对的最大数目。

3 良态对偶系统

本节将证明我们的定理和推论。我们从定理 1 的证明开始。该证明可由定理 3 和推论 4 的陈述直接推出，只需用到以下补充引理；该引理根据元素的完美匹配的交叉数，给出了任意集合系统的差异度上界。引理的第一个陈述是一个简单的观察，而第二个陈述则是众所周知的（例如参见第5 [30]章）。

引理 10. 设 $(X,\mathcal{R})$ 是一个集合系统，$|X|$ 为偶数，其度不超过 t ，且存在一个交叉数至多为 N的匹配。那么 $(X,\mathcal{R})$ 的差异度至多为

$$O\left(\min\left\{N,\sqrt{N\cdot\log|\mathcal{R}|}\right\}\right)=O\left(\sqrt{N\cdot\log|\mathcal{R}|}\right).$$

证明。容易看出，差异度至多为 N：任取一种着色，使每对匹配顶点中的两个顶点颜色相反即可。集合 $S\in\mathcal{R}$的差异度至多为恰有一个顶点属于 S的顶点对数。该数至多为交叉数，即 N。

为证明差异度至多为 $O\left(\sqrt{N\log|\mathcal{R}|}\right)$，考虑一种随机着色，使每个匹配点对中的顶点都被赋予相反的颜色。对于任意集合 $S\in\mathcal{R}$， S 的期望差异度为零，而 S 的差异度方差至多为与 S相交的点对数，即至多为交叉数 N。现在应用 Chernoff 界，允许最大偏差为 $O\left(\sqrt{N\log|\mathcal{R}|}\right)$，并对所有 $\mathcal{R}$ 的集合应用并集界，即可得到该引理的结论。 ——————

□

为证明定理 3，我们需要以下引理，它保证存在一对点，使其与少量范围相交。

引理 11（短边引理）。假设满足定理 3 的条件。则对于任意范围集合或范围多重集 $T\subseteq\mathcal{R}$，都存在点 x、$y\in X$，使得边 $\{x,y\}$ 至多与 δ 个范围相交，其中 δ 是满足以下条件的最大值

$$n\ \leq\ C_{d_0}\frac{|T|}{\delta}\cdot\psi\left(\frac{|T|}{\delta},\frac{t}{\delta}\right),$$

其中 C_{d_0} 是一个仅取决于 d_0 的常数。

证明。对于每个点 $x \in X$，定义范围集合 $D_x := \{S \in T : x \in S\}$。令$\mathcal{D} := \{D_x : x \in X\}$ 为由此得到的范围集合族。于是，对于任意点对$x, y \in X$，对称差 $D_x \Delta D_y$ 是被点对$\{x, y\}$穿过的范围集合。系统 $\mathcal{D}$ 的浅层单元复杂度至多为由 T给出的对偶集合系统的并复杂度。令 δ 表示 $\mathcal{D}$中任意两集合之间的最小对称差。若存在一对 $\{x, y\}$ 满足 $D_x = D_y$，则结论成立，因为对于这一对，有 $\delta = 0$。否则，注意到根据定义， $\mathcal{D}$ 是一个 δ-打包。由于 $(X, \mathcal{R})$中的每个元素至多属于 t 个范围， $\mathcal{D}$ 中任意集合的最大大小为 t。因此，我们可以应用浅层装填引理 8 来界定 δ-打包的大小。我们得到

$$n = |\mathcal{D}| \leq C_{d_0} \frac{|T|}{\delta} \cdot \psi\left(\frac{|T|}{\delta}, \frac{t}{\delta}\right).$$

其中 C_{d_0} 是一个仅取决于 d_0的常数。 □

现在我们给出定理 3 的证明。

定理 12（定理 3 的重述）。设 $(X, \mathcal{R})$ 是一个 n-点集合系统，其中 $|\mathcal{R}| \geq n$由 X 与一类对象的交集生成，该类对象的对偶分裂维数至多为 d，浅层单元复杂度为 $\psi(m, k) = m^{c_1} \cdot g(m)k^c$，其中 c_1、$c \geq 0$与 m和k无关，且 $g(m)$ 是关于 m 的非递减函数，并满足对每个 $\varepsilon \in (0, 1)$，都有 $g(m) = o(m^{\varepsilon})$。此外，设 $(X, \mathcal{R})$ 满足： X 中的每个点至多属于 t 个集合。则存在一个由 X中的点构成的匹配，其交叉数至多为

$$O\left((n^{c_1} t^c g(n))^{\frac{1}{1+c_1+c}}\right).$$

特别地，若 $c_1 = 0$，则交叉数至多为

$$O\left((t^c g(n))^{\frac{1}{1+c}}\right).$$

证明。我们将通过依次以贪心方式选取点对来构造匹配。每次都选取一对与最少范围相交的点。不过，我们不直接考察给定点对与多少个范围相交，而是进行加权计数：为每个范围赋予权重，并选择使其所相交范围的权重最小的点对。范围 $S \in \mathcal{R}$ 的权重将是一个关于已选点对数量的函数，这些点对与 S相交。这样一来，所选的边将不会与那些已多次相交的范围相交，因此，我们的策略倾向于选择相对于已选边而言交叉数较低的边。

初始时，我们为每个范围赋予权重 1。选出第 i 对之后，对于给定的范围 $S \in \mathcal{R}$，令 $c_S(i)$ 表示已选出的、与 i交叉的对数。于是， S 的权重为 $w_i(S) = 2^{c_S(i)}$。范围子集 $\mathcal{S} \subseteq \mathcal{R}$ 的权重是其中各范围权重之和。每一步中，我们贪心地选取一对，使其与 $\mathcal{R}$中权重最小的范围集合相交；这些权重已在上一步结束时更新。第 i步结束时，令 T_i 表示所得的多重集，它由将每个范围 $S \in \mathcal{R}$ 替换为 $w_i(S)$ 个 S而得到。注意， $w_i(\mathcal{R}) = |T_i|$，其中 T_i的基数指加权基数。

根据短边引理 11，我们有

$$\begin{aligned} n &\leq \frac{|T_i|^{1+c_1}}{\delta^{1+c_1+c}} g(|T_i|/\delta) t^c \\ &\leq \frac{w_i(\mathcal{R})^{1+c_1}}{\delta^{1+c_1+c}} g(w_i(\mathcal{R})) t^c. \end{aligned}$$

因此，存在一对点位于

$$X \setminus \{x_1, y_1, x_2, y_2, \ldots, x_i, y_i\},$$

其交叉的范围至多为 δ，其中 δ 是满足以下条件的最大值

$$\delta \leq \left(\frac{|T_i|^{1+c_1}}{n} g(|T|) t^c \right)^{1/(1+c_1+c)}.$$

我们选择这对 $\{x_{i+1}, y_{i+1}\}$ 作为匹配中的第 $(i+1)$ 条边。令 $\mathcal{R}_{i+1}$ 表示被这对 $\{x_{i+1}, y_{i+1}\}$ 交叉的 $\mathcal{R}$ 的范围。因此，经过第 $(i+1)$ 步后，系统 $\mathcal{R}$ 的权重为

$$w_{i+1}(\mathcal{R}) \quad = \quad w_i(\mathcal{R}) - w_i(\mathcal{R}_{i+1}) + w_{i+1}(\mathcal{R}_{i+1}).$$

现在，由于选择 $\{x, y\}$ 后，$\mathcal{R}$ 中范围的交叉数增加了 1，因此根据权重函数的定义，它们的权重在第 $(i+1)$ 步后翻倍，也就是说，

$$w_{i+1}(\mathcal{R}_{i+1}) = 2w_i(\mathcal{R}_{i+1}).$$

因此我们得到

$$\begin{aligned} w_{i+1}(\mathcal{R}) &= w_i(\mathcal{R}) - w_i(\mathcal{R}_{i+1}) + 2w_i(\mathcal{R}_{i+1}) \\ &= w_i(\mathcal{R}) \left(1 + \frac{w_i(\mathcal{R}_{i+1})}{w_i(\mathcal{R})} \right). \end{aligned} \tag{2}$$

注意到，根据我们对这对 $\{x_{i+1}, y_{i+1}\}$ 的选择，$|\mathcal{R}_{i+1}| \leq \delta$，我们得到

$$w_i(\mathcal{R}_{i+1}) \quad \leq \quad C \cdot \left(w_i(\mathcal{R})^{1+c_1} t^c \cdot \left(\frac{g(w_i(\mathcal{R}))}{n-2i} \right) \right)^{1/(1+c_1+c)}.$$

现在将 $\mathcal{R}$ 的权重代入递推关系 (2)，得到

$$\begin{aligned} w_{i+1}(\mathcal{R}) &\leq w_i(\mathcal{R}) \left(1 + \frac{w_i(\mathcal{R}_{i+1})}{w_i(\mathcal{R})} \right) \\ &= w_i(\mathcal{R}) \left(1 + \left(\frac{t^c}{(n-2i)} \cdot \frac{g(w_i(\mathcal{R}))}{w_i(\mathcal{R})^c} \right)^{\frac{1}{1+c_1+c}} \right). \end{aligned}$$

根据定义，对于每个 $\varepsilon \in (0, 1]$，都有 $g(n) = o(n^\varepsilon)$；又由于我们假设 c 是常数，因此比值 $\frac{g(w_i(\mathcal{R}))}{w_i(\mathcal{R})^c}$ 在 $w_i(\mathcal{R})$ 取最小值时达到最大，即 $w_0(\mathcal{R})$，或 $|\mathcal{R}|$。此外，根据我们的假设 $|\mathcal{R}| \geq n$，该比值至多为 $\frac{g(n)}{n^c}$。由此可得

$$\begin{aligned} w_{n/2}(\mathcal{R}) &\leq w_0(\mathcal{R}) \prod_{i=1}^{n/2-1} \left(1 + \left(\frac{t^c g(n))}{(n-2i)n^c} \right)^{\frac{1}{1+c_1+c}} \right) \\ &\leq |\mathcal{R}| \prod_{i=1}^{n/2-1} \left(1 + \left(\frac{t^c g(n)}{(n-2i)n^c} \right)^{\frac{1}{1+c_1+c}} \right). \end{aligned}$$

对两边取对数，得到

$$\begin{aligned} \log w_{n/2}(\mathcal{R}) &\leq \log |\mathcal{R}| + O\left(\left(\left(\frac{t}{n} \right)^c g(n) \right)^{\frac{1}{1+c_1+c}} \sum_{i=1}^{n/2-1} \frac{1}{(n-2i)^{1/(1+c_1+c)}} \right) \\ &\leq \log |\mathcal{R}| + O\left((t^c g(n))^{\frac{1}{1+c_1+c}} n^{c_1/(1+c_1+c)} \right) \\ &= \log |\mathcal{R}| + O\left((n^{c_1} t^c g(n))^{\frac{1}{1+c_1+c}} \right). \end{aligned} \tag{3}$$

这里，在上面的第二个不等式中，我们用到了这一事实：

$$\sum_{i=1}^{n/2-1} \frac{1}{(n-2i)^{\frac{1}{1+c_1+c}}} \leq n^{1-\frac{1}{1+c_1+c}}.$$

最后，令 $c_{\max} := \max_{S\in\mathcal{R}} c_S(n/2)$ 表示所有集合$S \in \mathcal{R}$中的最大交叉数。我们有

$$c_{\max} = \max_{S\in\mathcal{R}} \log w_{n/2}(S) \leq \log w_{n/2}(\mathcal{R}).$$

根据公式（3），这至多为

$$\log\ |\mathcal{R}| + O\left((n^{c_1}t^c g(n))^{1/(1+c_1+c)}\right).$$

由于该系统的 VC 维数是有界的（事实上，将 $k = n$ 代入浅层单元复杂度后，我们看到 VC 维数至多为 $2 + c_1 + c$），因此我们得到 $\log\ |\mathcal{R}| = O(\log\ n)$，因为范围的数量关于全集 X 的大小是多项式的。

定理的证明至此完成。 □

现在只需证明推论 4。该推论直接源自各种几何集合系统的已知浅层单元复杂度和并复杂度界。这些界列于下表中。

推论 4 的证明。推论的证明是将下表中的浅层单元复杂度函数应用于定理 3 的结果得出的。这些界可见于 [18]和 [36][，参见第 4 章中的列表。3] 及其中的参考文献。我们也推荐感兴趣的读者参阅综述 [2]。正交象限的界已在 [24] 中得到证明。

对象	$m\psi(m,k)$
$\mathbb{R}$中的区间	m.
$\mathbb{R}^2$中的半空间	mk.
$\mathbb{R}^2$中的凸体的位似像	mk.
$\mathbb{R}^2$中的圆盘	mk.
$\mathbb{R}^2$中的伪圆盘	mk.
α中的 $\mathbb{R}^2$-胖三角形	mk对数$^*\ \frac{m}{k} + mk$ 对数$^2\ \alpha$。
局部 γ-胖半代数对象	$mk2^{O(\log^* m)}$。
在 $\mathbb{R}^2$ 中具有有界描述复杂度并复杂度为 $m\phi(m)$ 的对象	$m\phi(m/k)k$。
$\mathbb{R}^3$ 中的半空间	mk.
$\mathbb{R}^2$ 中的正交象限和 $\mathbb{R}^3$	mk.

表1：不同几何集合系统的浅层单元复杂度界。

□

4 结论

我们证明了，在浅层单元复杂度接近线性的集合系统中，只要每个点属于有界数量的集合，就能得到更优的差异界。虽然算法方面的考量并非本文重点，但我们仍在此简要讨论。要获得低差异着色，我们需要计算交叉数较低的匹配。多位作者研究过此类算法，例如 [42, 43, 23, 15]。就我们的

目的而言，使用 Welzl 最初提出的算法就足够了；在我们的情形下，该算法的运行时间至多为 $O(n^3 t)$。值得注意的是，Csikos 和 Mustafa [15] 以及 Har-Peled [23] 提出的更快算法所得匹配的交叉数界略差（在n中相差一个对数因子），因此差异度也会相应增加。

参考文献

[1] 射影平面的着色。《离散数学》，73(1):213–220，1988年。[2] P. K. Agarwal、J. Pach 和 M. Sharir。几何对象的现状：一篇综述。载于 J. Goodman、J. Pach 和 R. Pollack 编，《计算几何：二十年之后》，第 9–48 页。美国数学会，2008年。[3] W. Banaszczyk。向量平衡与 n 维凸体的高斯测度。《Random Struct. Algorithms》，12(4):351–360，1998年。[4] N. Bansal。差异度最小化的构造性算法。载于 FOCS，第 3–10 页，2010年。[5] N. Bansal、D. Dadush 和 S. Garg。一个匹配 Banaszczyk 界的 Koml´os 猜想算法。载于 FOCS，第 788–799 页，2016年。[6] N. Bansal 和 S. Garg。超越部分着色的算法差异度。载于 STOC 2017：第 49 届年度 ACM SIGACT 计算理论研讨会论文集，第 914–926 页，2017年。[7] Nikhil Bansal、Daniel Dadush、Shashwat Garg 和 Shachar Lovett。Gram-Schmidt 游走：治愈 Banaszczyk 的烦恼。载于第 50 届年度 ACM SIGACT 计算理论研讨会论文集，STOC，第 587–597 页，2018年。[8]Nikhil Bansal 和 Kirk Pruhs。通过准均匀抽样求解加权几何集合多重覆盖。《J. Comput. Geom.》，7(1):221–236，2016年。网址：https://doi.org/10.20382/jocg.v7i1a11，doi:10.20382/jocg.v7i1a11。[9] J. Beck。正方形中有限集的平衡二染色 I。《Combinatorica》，1(4):327–335，1981年。[10] J. Beck 和 T. Fiala。“整数化”定理。《离散应用数学》，3(1):1–8，1981年。[11] J´ozsef Beck。立方体中有限集的平衡二染色。《Discret. Math.》，73(1-2):13–25，1989年。[12] T. M. Chan、E. Grant、J. K¨onemann 和 M. Sharpe。通过改进的准均匀抽样求解加权容量约束、优先级及几何集合覆盖。载于第 23 届 ACM-SIAM 离散算法研讨会论文集，SODA，第 1576–1585 页，2012年。[13] Bernard Chazelle。《差异度方法——随机性与复杂性》。剑桥大学出版社，2001年。[14] K. L. Clarkson 和 K. R. Varadarajan。改进的几何集合覆盖近似算法。《Discrete & Computational Geometry》，37(1):43–58，2007年。[15] M. Csik´os 和 N. H. Mustafa。摆脱空间划分的困境：低交叉数匹配及其应用。载于第 37 届国际计算几何研讨会论文集，SoCG，第 189 卷，第 28:1–28:17 页，2021年。

[16] D. Dadush、S. Garg、S. Lovett 和 A. Nikolov。迈向 Banaszczyk 向量平衡定理的构造性版本。载于 RANDOM，第 28:1–28:12 页，2016年。[17] K. Dutta、E. Ezra 和 A. Ghosh。浅层打包的两种证明。《Discrete & Computational Geometry》，56(4):910–939，2016年。[18] K. Dutta、A. Ghosh、B. Jartoux 和 N. H. Mustafa。浅层打包、半代数集合系统、麦克比斯区域与多项式划分。《Discrete & Computational Geometry》，61(4):756–777，2019年。[19] K. Dutta、A. Ghosh 和 S. Moran。一致括号、容器与组合麦克比斯区域。载于第 13 届理论计算机科学创新会议论文集，ITCS，第 215 卷，第 59:1–59:10 页，2022年。[20] K. Dutta 和 G. Ghosh。汉明立方体的尺寸敏感打包数及其推论。2014年。网址：https://arxiv.org/abs/1412.3922。[21] E. Ezra。具有有界原始粉碎维数的集合系统的尺寸敏感差异界。《SIAM 计算杂志》，45(1):84–101，2016年。[22] E. D. Gluskin。正交平行六面体的极值性质及其在巴拿赫空间几何中的应用。《数学文集》，64(1):85–96，1989年。[23] S. Har-Peled。近似计算低交叉数生成树。2009年。网址：https://arxiv.org/abs/0907.1131。[24] Haim Kaplan、Natan Rubin、Micha Sharir 和 Elad Verbin。盒子中的颜色计数。载于第 18 届年度 ACM-SIAM 离散算法研讨会论文集，SODA '07，第 785–794 页，美国，2007年。工业与应用数学学会。[25] K. G. Larsen。群模型中的范围查询与组合差异。《SIAM J. Comput.》，43(2):673–686，2014年。[26] Kasper Green Larsen。具有遗传 L2 保证的构造性差异度最小化。载于 Rolf Niedermeier 和 Christophe Paul 编，第 36 届理论计算机科学专题国际研讨会，STACS 2019，2019年3月13–16日，德国柏林，第 126 卷 LIPIcs，第 48:1–48:13 页。Schloss Dagstuhl - Leibniz 信息学中心 f¨ur Informatik，2019年。网址：https://doi.org/10.4230/LIPIcs.STACS.2019.48，doi:10.4230/LIPIcs.STACS.2019.48。[27] A. Levy、H. Ramadas 和 T. Rothvoss。通过乘法权重更新方法进行确定性差异度最小化。第 1 页，待刊，2017 年。[28] S. Lovett 和 R. Meka。通过沿边行走进行构造性差异度最小化。载于 FOCS，第 61–67 页，2012年。[29] J. Matouˇsek。《几何差异：图解指南》。施普林格，1999年。[30] J. Matouˇsek。《几何差异：图解指南》。施普林格，2009年。[31] J. Matouˇsek。基于对偶粉碎函数的差异界。《Mathematika》，44(1):42–49，1997年。[32] J. Matouˇsek、E. Welzl 和 L. Wernisch。有限 VC 维数下的差异度与近似。《Combinatorica》，13(4):455–466，1993年。[33] N. H. Mustafa。浅层打包引理的一个简单证明。《Discrete & Computational Geometry》，55(3):739–743，2016年。

[34] N. H. Mustafa、K. Dutta 和 A. Ghosh。《最优 ε-网的一个简单证明》。《Combinatorica》，2017年。[35] N. H. Mustafa 和 K. Varadarajan。ε-逼近与 ε-网。载于 J. E. Goodman、J. O'Rourke 和 C. D. T´oth 编，《离散与计算几何手册》。CRC Press LLC，2017年。[36] Nabil H. Mustafa。《组合与几何系统中的抽样》。美国数学学会，2022年。[37] Aleksandar Nikolov。《盒与多胞体差异度的更紧界》。《Mathematika》，63(3):1091–1113，2017年。[38] T. Rothvoss。《通过差异理论改进装箱逼近》。《SIAM J. Comput.》，45(3):930–946，2016年。[39] J. Spencer。《六个标准差足矣》。《Trans. Amer. Math. Soc.》，289(2):679–706，1985年。[40] A. Srinivasan。《改进稀疏矩阵的差异界：稀疏格逼近问题的更优逼近》。载于 SODA，第692–701页，1997年。[41] K. R. Varadarajan。《通过准均匀抽样实现加权几何集合覆盖》。载于第42届 ACM 计算理论研讨会论文集，STOC，第641–648页，2010年。[42] E. Welzl。《用于三角形计数及其他范围搜索问题的划分树》。载于第四届年度计算几何研讨会论文集，SCG '88，第23–33页，美国纽约州纽约市，1988年。计算机协会。网址：https://doi.org/10.1145/73393.73397，doi:10.1145/73393.73397。[43] Emo Welzl。《关于低交叉数生成树》，第233–249页。Springer Berlin Heidelberg，德国柏林，1992年。