arXiv:2304.01100v1 [cond-mat.dis-nn] 3 Apr 2023

若干匹配问题系综的能量景观

Till Kahlke 和 Alexander K Hartmann
奥尔登堡卡尔·冯·奥西茨基大学物理研究所，德国奥尔登堡 26111

电子邮件： till.kahlke@uol.de， a.hartmann@uni-oldenburg.de

摘要。 本文通过数值方法研究最大权匹配问题及其能量景观的行为。我们采用一种改编自自旋玻璃分析的微扰方法，从而深入了解不同系综的能量景观的复杂性。本文考察了 Erd˝os-R´enyi 图以及随机添加边的环图，并使用了两种随机边权重分布。对于最大权匹配问题，已有快速且可扩展的算法，因此我们能够研究包含超过10^5 个节点的大型图。结果表明，标准匹配系综的能量景观结构简单，与铁磁体的能量景观相似。不过，对于本文介绍的某些系综，我们的结果表明，可能存在具有副本对称性破缺特征的复杂能量景观。

关键词：匹配、复杂性、微扰技术、随机图、基态、副本对称性破缺、计算机模拟

1. 引言

自旋玻璃是统计物理学中广泛研究的典型复杂系统 [1–6]。Sherrington-Kirkpatrick 自旋玻璃模型的解析解 [7] 引出了副本对称性破缺（RSB）这一概念 [8–13]。这种复杂的 RSB 相通常具有崎岖的能量景观，其中包含许多被高能量势垒分隔开的局部极小值和亚稳态 [13]。这会导致非平凡的平衡态和非平衡行为。

这种复杂性概念可以推广到神经网络、数据分析或优化问题等其他问题 [5, 14–16]。这些问题可能表现出 RSB 或其他类型的复杂性[17]。研究发现，例如，可满足性[18–20], 、旅行商问题 [21–23], 、顶点覆盖 [24–27] 和图着色 [28] ，在某些系综和模型参数下都可能呈现复杂相。这些模型还表现出崎岖的能量景观，以及玻璃态的平衡与非平衡行为。

上述问题都属于非确定性多项式（NP）问题[29]。目前已知的所有用于求解这些问题的精确算法，其最坏情况下的运行时间都会随着问题规模呈指数级增长。这意味着，在对这些问题的平衡行为进行数值研究时，我们受到很大限制，因为只能研究规模相近的小系统。遗憾的是，这些模型的行为非常有趣。

另一方面，在多项式（P）问题这一类中，最坏情况下的运行时间仅呈多项式增长。因此，我们可以研究规模大得多的系统，从而获得数值上更精确的结果和更好的统计数据。遗憾的是，P 类问题通常具有简单的能量景观 [30–35], ，因而呈现相当平凡的行为。这似乎表明，复杂行为与便于数值研究平衡态行为这两者无法兼得。不过需要指出的是，迄今为止的大多数研究都针对模型系综，其中随机性以简单且不相关的方式引入。有趣的是，最近在用于随机介质中的定向高分子模型的系综中，发现了复杂行为。对于该模型，存在一种多项式时间精确抽样算法，因此可以研究多达 10^9 个晶格点的大型系统，并且已观察到有力的迹象表明存在副本对称性破缺 [36]。此外，对于可在多项式时间内求解的递增子序列问题（也称乌拉姆问题），数值证据也显示存在 RSB [37]。最后还应指出，XOR-SAT 问题可以通过高斯消元高效求解 [38, 39], ，但它也表现出一步副本对称性破缺；不过，在热力学极限下，这种破缺在平衡态中变得无关紧要 [40]。

因此，通过选择合适的系综，似乎可以用数值上高效的方式研究一些具有非平凡行为的系统。本文秉承 [36, 37], 的思路，研究 P 类中的另一个问题，以寻找复杂行为的迹象。具体而言，我们考虑最大权图匹配问题。该模型是最早从统计力学[41]视角进行研究的优化问题之一。相关分析后来扩展到了任意图 [30]、有限尺寸效应 [42] 以及欧几里得边权 [43–45]。

图匹配问题之所以受到物理学界关注，还因为其他问题可以通过映射到合适的匹配问题来求解。例如，这种方法已用于二维自旋玻璃[33], 、二聚体覆盖 [46], 、负权渗流[47] 以及动态网络的可控性 [48, 49]。

本文其余部分安排如下：第2节介绍本文所用的图匹配及随机图系综。这些系综包括 Erd˝os-R´enyi 图，以及添加了随机边的环图；两者的边权重均服从高斯分布，或为叠加均匀噪声的恒定边权重。第3 节介绍扰动技术的细节。本文说明了如何通过观察原始图与轻微扰动后的图分别得到的匹配之间的差异，来考察是否存在复杂能量景观。我们还探讨了扰动方法能否用于在统计上正确地采样匹配。分析结果见第4节，首先介绍扰动技术所得的结果。最后，我们在第5节总结并得出结论。

2. 模型

2.1. 匹配

给定一个无向图 $G = (V, E)$，它由节点集合 $i \in V$ 和边集合 $e \in E \subseteq V^{(2)}$构成，其中$n = |V|$ 且 $m = |E|$。若 $e = \{i, j\}$ 是 E的成员，则称两个节点 i 和 j 相邻。称边 e与 i 和 j关联。匹配 M 是 E 的一个子集，其中任意两条边都不与同一节点关联。对于 M中的每条边，其相邻节点称为匹配节点。未匹配的节点称为自由节点。最大基数匹配是使匹配节点数量 $|M|$最大的匹配。在边权重为 $w(e)$的图 G 上，匹配的权重为 $W = \sum_{e\in M} w(e)$。最大权匹配是在 G 上所有可能的匹配中，使权重W 最大的匹配 M。第 3 节将介绍我们用于求给定图的最大权匹配的算法。

2.2. 随机图

我们考虑两种随机图系综。关于所选模型参数取值的详细信息见第 4 节。第一种系综由 Erdős-Rényi 随机图 [50]构成。生成这类图时，首先从一个包含 n 个节点$V = \{1, 2, \cdots, n\}$的空图开始。然后，对每一对不同的节点，独立地以概率 $p = c/n$ 创建一条边，其中 c 是连通度。由此得到的度分布服从泊松分布。

第二种系综由带有额外随机边的环图构成。首先，创建一个平均度为二的正则环图。这意味着，边 E 集合最初包含边 $\{i, i + 1\}$ 对于$i = 1, \ldots, n - 1$ 以及边 $\{n, 0\}$。接着，以均匀概率随机选取两个节点i，j，并在边 $\{i, j\}$ 尚不存在于 E 中时创建该边。重复此过程，直到以这种方式向图中添加了 m^+ 条边。此时，G 中的边数为 $m = n+m^+$。下文将考虑 $m^+ =$ const、$m^+ \in O(\log n)$ 和 $m^+ \in O(n)$ 这几种情况。

我们研究两种边权重 $w(e)$。首先，考虑 $w(e)$ $\sim \mathcal{N}(1, 0.01)$，即均值为 1、方差为 0.01 的高斯分布随机权重。其次，研究带有额外均匀噪声的常数权重 $w(e) = 1+\varepsilon$。为选择额外噪声 ε 的强度，我们指出，当所有边的权重均为 $w(e) = 1$ 时，最大权匹配等价于最大基数匹配，因为 $W = |M|$ 成立。添加噪声 ε 的目的是使最优解唯一。但除此之外，即使 n 的值很大，它也不应影响最优匹配的大小。为确保这一点，ε 取自一个均匀分布，其取值随 $n^{-\frac{1}{2}}$ 缩放，即 $\varepsilon \sim U(-0.01\, n^{-\frac{1}{2}}, 0.01\, n^{-\frac{1}{2}})$。

3. 方法

为了求解最大权匹配，我们使用在 LEMON 库 [52]中实现的 Edmonds 花收缩算法 [51], 。该算法的最坏情况运行时间为 $O(n^2m)$，即多项式时间，因此可以处理规模相应较大的图。

3.1. 扰动技术

研究能量景观的复杂性时，一种常见的方法是对系统施加弱扰动。据我们所知，这种方法最早用于研究自旋玻璃（例如参见 [53]）。此后，该方法也被用于研究随机场伊辛模型 [54, 55]和旅行商问题 [23]。

基本思路是先计算出一个基态，然后对系统施加轻微扰动，再通过重新计算得到一个新的基态；相对于原系统而言，它是一个激发态。对于匹配问题，我们使用 边翻转 作为一种扰动技术。首先，计算一个总权重为 W_0 的最大权匹配 M_0。然后，随机选择一条边 e 并将其“翻转”。如果 e 属于 M_0，则不允许它出现在扰动图的匹配 M_1 中。另一方面，如果 e 不属于M_0，则强制将其纳入 M_1。实际操作中，可分别将 e 的权重设为一个很小的值（例如 $-2W_0$）或一个很大的值（例如 $2W_0$）。找到 M_1 后，将该边的权重恢复为原值，从而使用原始权重计算 M_1 的权重W_1。下文中，M_0 始终表示相对于原始边权重的最优匹配，而 M_i 和$i > 0$ 表示通过单条边翻转扰动得到的相互独立的匹配，且它们均相对于 M_0。W_0 和 W_i 表示这些匹配对应的权重。

如果W_0 和 W_1 的相对差异满足以下关系，则扰动较弱

$$\frac{W_0 - W_1}{W_0} = \mathrm{O}\left(\frac{1}{n}\right). \tag{1}$$

如果 (1) 成立，M_1 就是准最优。注意，这意味着在正则系综中考察系统时，也就是按照玻尔兹曼分布，匹配 M_0 和 M_1 在热力学极限下的权重相同。

为了比较两个匹配 M_i 和 M_j，我们采用一种相似性度量，也称为重叠度。由于匹配是集合，因此使用 Jaccard

指数 [56], 给出的

$$q_{ij} = \frac{|M_i \cap M_j|}{|M_i \cup M_j|}. \tag{2}$$

两个匹配之间的距离定义为$d_{ij} = 1 - q_{ij}$ [57]。最优匹配 M_0 与通过扰动获得的匹配 M_i 之间的重叠度和距离分别表示为 q_0 和 d_0，其中省略了第二个索引 i 。

复杂能量景观的一个必要条件是，在无限大图 [58] 的热力学极限下，存在具有 $d_0 > 0$的准最优匹配。因此，我们测量了不同 n值下的 $\langle (W_0 - W_i)/W_0 \rangle$和 $\langle d_0 \rangle$ ，并对随机图实现取平均。如果扰动图的匹配是准最优的，并且 $\langle d_0 \rangle$ 随着 n值的增加趋近于零，则可以假定能量景观是简单的。但是，如果 $\langle d_0 \rangle$ 保持有限，则不能排除复杂能量景观。

3.2. 抽样偏差检验

为了进一步探究给定图系综上的匹配问题是否确实呈现复杂能量景观，理想情况下应能高效地对匹配进行抽样。要获得正确的统计结果，总权重相同的匹配 W 应以相同的概率被抽取。因此，抽样方法需要知道整个问题及所遇到的子问题中不同匹配的数量。不幸的是，匹配问题属于#P完全问题 [59, 60]。这意味着，使用所有已知算法，都需要枚举全部匹配，而匹配的数量通常呈指数级增长。因此，这种方法在计算上代价高昂。

原则上，我们可以将通过微扰方法得到的匹配用作样本。不幸的是，边翻转会在抽样中引入未知偏差，因为某些构型会被禁止，而具体哪些构型被禁止取决于翻转的边。尽管如此，偏差误差仍有可能相互抵消，尤其是在大型图中。为评估这一点，我们进行了以下研究。

对于给定的图，我们在此研究 Erd˝os-R´enyi 图，其中抽取了一组带噪声的常数权重边权重 $w = 1 + \varepsilon$ ，参见第 2.2节。随后，根据这组权重计算一个最优匹配。接下来，通过随机边翻转多次获得扰动图，并再次计算匹配。然后针对初始噪声的多次实现重复这一过程。

将获得的匹配汇总到样本直方图中。这里，我们仅纳入与最优匹配具有相同基数的匹配。随后，利用所得直方图比较相对频率 $Q(M)$ 与所有匹配均以相等概率抽取时的情况。即 $P(M) = 1/n_{\mathrm{M}}$，其中 n_{M} 是不同匹配的数量。由于匹配的数量会随着边数呈指数增长，因此该过程仅适用于小型图。

为了定量比较 P 和 Q ，使用了两种距离度量。第一种是*Kullback-Leibler* 散度 D_{KL} ，定义为

$$D_{\mathrm{KL}}(P,Q) = \sum_{i=1}^{n_{\mathrm{M}}} P(M_i) \ln\left(\frac{P(M_i)}{Q(M_i)}\right). \tag{3}$$

第二种度量是均值相对误差 MRE，定义为

$$\mathrm{MRE}(P,Q) = \sum_{i=1}^{n_{\mathrm{M}}} P(M_i) \frac{\sqrt{(P(M_i) - Q(M_i))^2}}{P(M_i)} \tag{4}$$

$$= \sum_{i=1}^{n_{\mathrm{M}}} \sqrt{(P(M_i) - Q(M_i))^2}, \tag{5}$$

它表示每个匹配的平均相对误差。

如果对于更大的图， P 和 Q 变得更加相等，则 D_{KL} 和 MRE 的值会变小。特别是，如果抽样趋近于均匀抽样，则距离度量应收敛于零。若情况确实如此，则可以有意义地研究重叠分布 $P(q)$。如果在大图极限下，重叠分布仍然很宽，这将表明复杂行为在平衡态中具有重要作用。另一方面，如果 D_{KL} 和 MRE 的值随图规模增大而增加，则不能忽略抽样偏差误差。在这种情况下，距离 d_0 的非零极限值只能表明存在差异显著的匹配，但采用当前方法无法判断它们的权重是否足够大到影响平衡态行为。

4. 结果

以下结果通过对两种图系综、不同参数取值以及图规模范围从$n = 7$ 到 131072 进行模拟 [61] 获得。除非另有说明，所有结果均对 100 个不同的随机图实现，以及每个图实现对应的 100 个不同边翻转扰动及其产生的匹配取平均值。部分模拟通过 GNU Parallel 工具分配到多个 CPU 上执行 [62]。所有数据拟合均按照参考文献 [63]中的建议进行。

4.1. 扰动

我们首先验证扰动图的匹配是准最优的。结果表明，对于

两种随机图，Erd˝os-R´enyi 图和环图，以及两种边权重，扰动图的匹配均为准最优，式（1）成立。图1展示了连通度为˝的 Erd´os-Rényi 图的结果。图中还包括使用以下方法进行拟合的结果：

$c \in \{0.5, 1, 2, 4\}$。

$$\langle (W_0 - W_i)/W_0 \rangle = \alpha \frac{1}{n}. \tag{6}$$

在其他研究案例中，结果看起来相似，因此这里不再展示。请注意，对于环图，小规模时的数据偏离了拟合结果 $n \leq 2^8$。由于（1）只需渐近成立，因此该案例也满足准最优性条件。

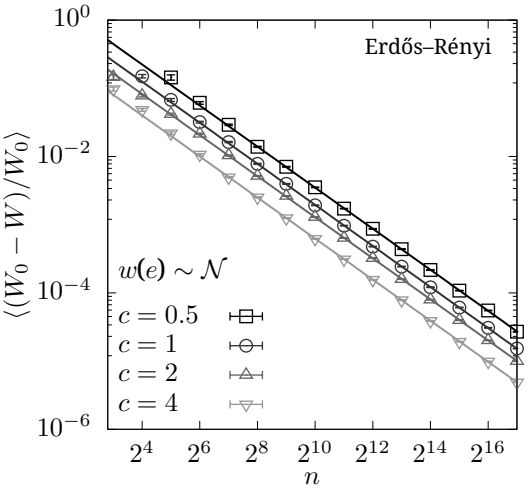

图 1。 验证扰动图上的匹配是准最优的，即 $\langle (W_0 - W)/W_0 \rangle$ 作为 n 的函数，其表现为 O(1/n）。这里展示的是具有高斯分布随机权重和连通度 ˝的 Erd´os-Rényi 图的结果。直线表示采用（6）形式的拟合。

接下来我们考察 d_0 的行为，首先从高斯分布权重的情况开始。Erd˝os-Rényi 图的结果如图 ´所示。可以观察到，距离 $\langle d_0 \rangle$ 随着 n的增加而趋近于零。为系统地确认这一行为，采用带偏移量的幂律拟合 d_∞ ，

$$\langle d_0 \rangle (n) = d_\infty + \beta_1 n^{\beta_2}, \tag{7}$$

从而得到良好的拟合。所有研究案例的拟合参数均列于表1中。

对于高斯分布权重，在所有情况下，d_∞ 的值都被设为固定值 0，以获得合理的结果。自由估计d_∞ 的拟合会导致该参数取负值，即产生非物理结果。观察到的行为表明，扰动图的匹配与原始图的匹配共享大部分边，并且随着系统规模的增大，这种共享程度越来越高。这可以通过采用非均匀权重分

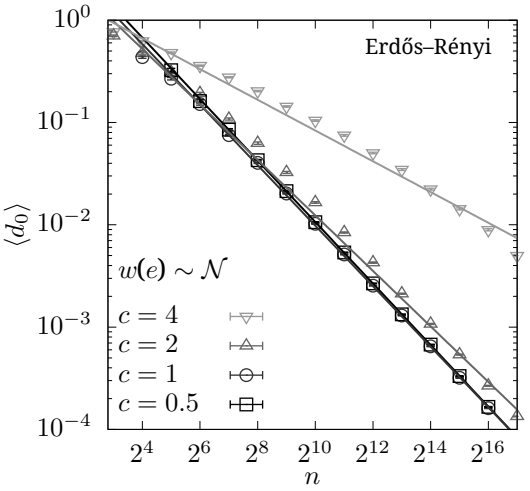

图 2。 扰动图中的匹配与最优匹配之间的平均距离 $\langle d_0 \rangle$随 n 变化的关系，其中图为 Erd˝os-R´enyi 图，采用高斯分布随机权重和连通度 $c \in \{0.5, 1, 2, 4\}$。直线表示采用（7）形式的幂律拟合。

布来解释。为了最大化 W_0，匹配算法会集中于少数具有较大权重的边。现在，对于具有扰动权重的匹配，如果一条边发生翻转，通过使用大部分与 M_0中相同的高权重边，仍然可以获得较大的总权重。因此，与 M_0 相比，只有少数边会发生变化，以补偿这次边翻转。这导致最优匹配与扰动图的匹配之间存在很小的差异，并且当图的规模趋于无穷大时，该差异会消失。

总体而言，高斯权重的结果表明能量景观具有简单的结构。环图的结果与此类似，因此未在此展示。由于 m^+ 的恒定值会导致 $d_\infty = 0$，我们省略了对高斯分布权重下 $m^+ \in O(\log n)$ 或 $m^+ \in O(n)$ 值的研究。

接下来，我们考察带有额外噪声的恒定边权重。由于此时所有边权重彼此更加相似，因此可以预期存在权重非常相近但差异很大的匹配，即更复杂的能量景观。对于 Erd˝os-R´enyi 图，$\langle d_0 \rangle$ 的结果与上述高斯分布权重的情况类似，因此这里不再展示。再次强调，需要固定 $d_\infty = 0$ 的值，才能根据（7）获得物理上有意义的拟合结果。因此，在这种情况下同样找不到复杂能量景观的迹象。

对于高斯权重，前述论证在这里不再成立，因为每条边对于总权重几乎同等重要。不过，这里的行为可以用 Erd˝os-R´enyi 图的拓扑约束来解释：对于 Erd˝os-R´enyi 图，环路的长度为 log(n) 量级，因此匹配的局部变化不会扩散太远，最终使得原图与扰动图的

表 1. 以（7)形式进行拟合所用的拟合参数值，以确定 d_∞。固定值以 * 标记。结果列出了埃尔德什-雷尼图和具有偶数或奇数条边的环图在不同参数值（第一列）和边权重（第二列）下的情况。值 $d_\infty = 0$ 对应简单的能量景观。但对于 $d_\infty > 0$，不能排除复杂行为。给出 β_1 和 β_2 的值以供参考。

埃尔德什-雷尼图				
c	w	d_∞	β_1	β_2
0.50	$\sim\mathcal{N}$	0*	10.7(1)	0。997(1)
1	$\sim\mathcal{N}$	0*	8.8(1)	0.982(1)
2	$\sim\mathcal{N}$	0*	6。27(7)	0.898(1)
4	$\sim\mathcal{N}$	0*	2.18(2)	0.446(1)
0.50	$1+\varepsilon$	0*	10.8(2)	0.998(2)
1	$1+\varepsilon$	0*	8。9(1)	0.983(1)
2	$1+\varepsilon$	0*	6.36(7)	0。901(1)
4	$1+\varepsilon$	0*	2.65(2)	0.499(1)
环图，n 偶数				
$m^+(n)$	w	d_∞	β_1	β_2
0	$\sim\mathcal{N}$	0*	83(3)	0.989(3)
1	$\sim\mathcal{N}$	0*	81(3)	0.986(3)
10	$\sim\mathcal{N}$	0*	80(2)	0.985(3)
0	$1+\varepsilon$	1*	0*	0*
1	$1+\varepsilon$	0.947(2)	0*	0*
10	$1+\varepsilon$	0.583(3)	2.5(5)	1。16(9)
$2\ln(n)$	$1+\varepsilon$	0.345(5)	1.00(3)	0.35(1)
$4\ln(n)$	$1+\varepsilon$	0.255(3)	1。31(3)	0.42(1)
环图，n 奇数				
$m^+(n)$	w	d_∞	β_1	β_2
0	$\sim\mathcal{N}$	0*	81(3)	0.987(3)
1	$\sim\mathcal{N}$	0*	79(3)	0.984(3)
10	$\sim\mathcal{N}$	0*	84(2)	0.991(3)
0	$1+\varepsilon$	0.544(3)	1.5(5)	1。1（2）
1	$1+\varepsilon$	0.523(3)	2.0(5)	1。1(1)
10	$1+\varepsilon$	0.397(3)	2.2(2)	0.92(4)
$2\ln(n)$	$1+\varepsilon$	0.295(3)	1.42(6)	0.56(2)
$4\ln(n)$	$1+\varepsilon$	0.222(3)	1.50(4)	0。51(1)

最优匹配之间的整体差异很小。因此，对于 $n\to\infty$，差值 $\langle d_0\rangle$ 消失。

另一方面，对于环图，其行为有所不同，如图3中所示。在所有研究案例中，要获得良好拟合，都需要取 $d_\infty > 0$ 。

为理解这一行为，考虑具有偶数个 n 节点且没有随机边的环图，即 $m^+ = 0$。由于沿路径的匹配呈现为匹配边与未匹配边交替出现，简单环图的结构要求所有满足 M_i with$i > 0$ 的匹配均由 $E \setminus M_0$给出，也就是说，所有匹配边

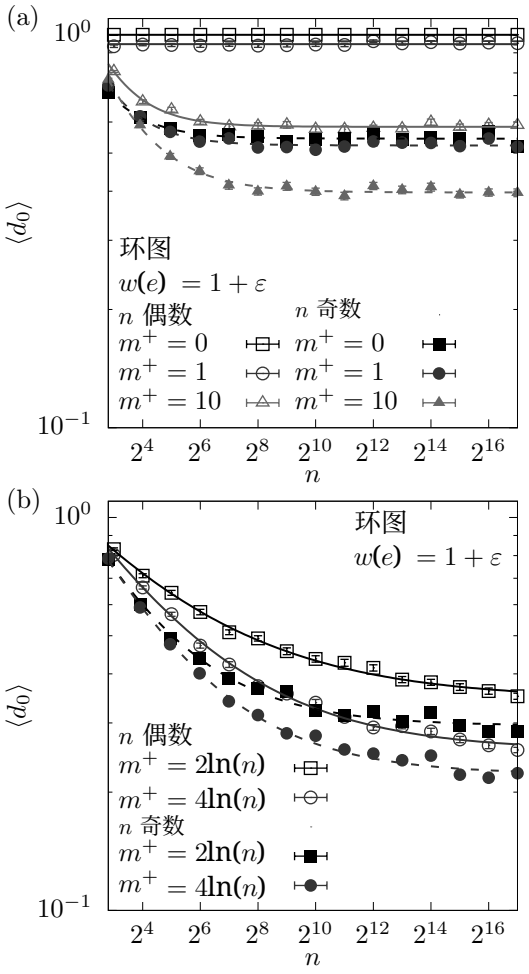

图3。环图中扰动图上的匹配与最优匹配之间的平均距离 $\langle d_0\rangle$随 n 变化的关系。（a）展示了恒定边权重加噪声时、取值为 $m^+ \in \{0, 1, 10\}$ 的结果；（b）展示了 $m^+ = l^+\ln(n)$且 $l^+ \in \{2, 4\}$的结果。空心符号表示节点数为偶数 n，实心符号则用于 n的奇数值。直线或虚线表示带偏移量 d_∞ 的幂律拟合，其形式分别为（7）。

都被替换为未匹配边，反之亦然。因此，显然$d_0 = 1$ 成立。

如果 n 为奇数，就会有一个节点空闲，并且存在 n 个最大基数匹配。 M_0 由总权重最大的最大基数匹配给出。对扰动图使用 $E \setminus M_0$ 会使匹配的基数降低，因而总权重也会降低。然而， $n-1$ 仍有其他匹配可以弥补边翻转，每个匹配都会使另一个节点空闲。因此，图受到扰动时，最优匹配中平均约有一半的边会发生变化。这可以与一个简单的自旋系统作比较：对于铁磁体，引入反周期边界条件时会产生两个畴壁，一个位于新边界条件所在之处，另一个则可能位于系统中的任意位置。

这也会产生许多不同的构型，但系统实际上仍不会表现出复杂行为。总之，对于简单环图，$m^+ = 0$ 时 d_∞ 保持有限，并非源于复杂的能量景观，而是因为图的结构简单。

加入随机边后，情况变得更加有趣。对于 $m^+ =$ const 和 $m^+ \in O(\ln n)$ 这两种情况，扰动会产生多种匹配，并使匹配之间出现更复杂的重叠。仅凭图结构已无法解释 d_∞ 的有限值。因此，我们观察到准最优解与最优解之间相差大量边。这是复杂性的必要条件[58]。因此，这个系综可能存在复杂能量景观。

我们还对环图进行了模拟，其中随机边的数量 m^+ 随 n线性增长。在这里，当n取较大值时，环图会越来越接近 Erd˝os-R´enyi 图，环结构也变得不那么重要。因此，我们得到了与 Erd˝os-R´enyi 图类似的发现，即对于 $n \to \infty$，我们观察到 $\langle d_0 \rangle \to 0$。因此，我们不再报告这一情况的更多结果。

4.2. 抽样

我们先前的结果表明，存在一些用于匹配问题的系综，其中存在一些最优匹配，它们相差 $O(n)$个变量，但能量几乎相同，也就是说，它们是准最优的。这是复杂能量景观存在的必要条件。如上所述，若能对这些匹配进行抽样并相互比较，将有助于更好地理解能量景观，尤其是在热力学极限下。为此，需要以相同的概率进行抽样，也就是说，采用无偏的方式。

在抽样偏差检验中，采用了边权重为 $1+\epsilon$、$c = 2$ 和 n 介于4至32之间的Erd˝os-R´enyi图。需要注意的是，只有对于如此小的规模，才能对所有解进行完整枚举并以良好的统计效果进行抽样。对于每个图，使用1000组不同的边权重，并针对每组权重计算扰动图的1000个匹配，因此每个图共得到 10^6 个匹配。

这些匹配出现的频率 $Q(M_i)$ 按照第3.2节所述进行测量，并与要求的均匀分布进行比较。所得的 D_{KL} 和均方误差的数值在100个不同的随机图实现上取平均。该平均过程
记作 $\langle . \rangle_{\mathrm{G}}$。结果如图所示4。从中可以看出，$\langle D_{\mathrm{KL}} \rangle_{\mathrm{G}}$ 和 $\langle$均方误差$\rangle_{\mathrm{G}}$ 均随 n的增大而增大。因此，抽样偏差不会相互抵消。还可以看出，

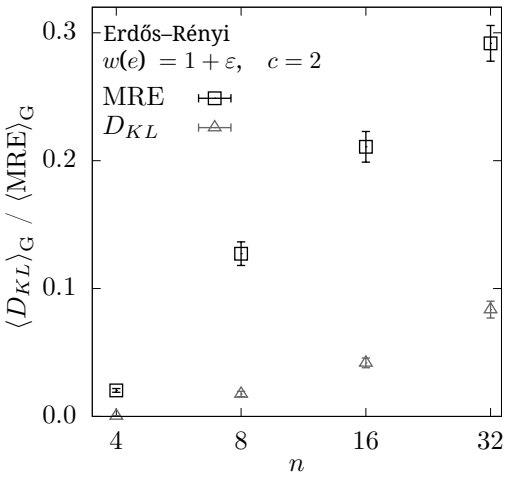

图4。抽样偏差检验的结果。平均 Kullback- Leibler 散度 $\langle D_{\mathrm{KL}} \rangle_{\mathrm{G}}$ 和平均相对误差 $\langle \mathrm{MRE} \rangle_{\mathrm{G}}$ 均随 n取值增大而系统性增加。结果来自 Erd˝os-R´enyi 图，其连通度为 $c = 2$，并采用带有额外噪声的恒定边权重。

误差棒随 n的增大而增大。这是因为可能的匹配数量会随着节点数 n 的增加而增加，从而导致方差增大。

该结果表明，抽样误差不会被平均掉。因此，直接使用从扰动过程抽取的样本来研究抽样重叠分布 $P(q)$ 是没有意义的。

5. 结论

我们通过对原始图的最优解进行“边翻转”扰动，研究了匹配的能量景观行为。分析表明，对于适当的系综，存在与基态相差大量边的准最优态。因此，不能排除这里存在复杂能量景观。这些系综包括带有恒定数量或对数增长数量额外边的环图，以及具有恒定边权重和均匀噪声的情况。在这些情况下，匹配之间外推得到的相对距离 d_∞ 保持有限。对于其他观测到的系综——额外边数量呈线性增长的环图以及 Erd˝os-R´enyi 图——结果没有显示出存在复杂能量景观的证据。

研究表明，直接从微扰方法中获取的样本存在偏差。这妨碍了进一步开展更细致的研究，例如求得多个最优匹配之间重叠度的分布。因此，尚不清楚准最优态在热力学极限下是否具有重要意义。如果并非如此，那么此处观察到的能量景观行为就类似于“弱破缺

副本对称性” [64]。需要指出的是，一般而言，基于对给定最优解施加扰动的方法，只能表明是否存在与之相差 $O(n)$ 个变量的其他解，却无法判断这些解在热力学极限下是否具有重要意义。不过，如果不存在这些差异极大的其他准最优解，那么显然能量景观是简单的。

为开展进一步研究，需要指出，对于具有恒定数量或按对数增长的额外边 $m^+(n)$ 的环图，观察到了有限的 $d_\infty > 0$ 。但如果 $m^+(n)$ 随 n 线性增长，我们发现 $d_\infty = 0$。因此，在 $m^+(n)$ 的取值之间存在一个转变：一类取值下 d_∞ 保持有限，另一类取值下则趋于零。为进一步分析这一问题，可以用幂律 $m^+ \in O(n^\lambda)$ 对额外边的数量建模。这样便可考察是否存在指数的临界值 λ_c ，使得 $d_\infty > 0$ 与 $d_\infty = 0$ 之间发生突变。

此外，若能找到一种高效生成无偏样本的方法，也会很有意义。可以采用数值开销更大的抽样方法，例如蒙特卡罗马尔可夫链模拟 [65] ，或穷举。在现有算法下，这会大幅限制可研究的系统规模。另一方面，也可以在之后利用弹道搜索 方法 [66], 校正抽样偏差；该方法已成功用于校正进化算法应用于自旋玻璃时观察到的抽样偏差 [67]。这种做法在算法上相当复杂，但未来的项目完全可以实现。采用这类方法后，便可研究抽样得到的重叠分布 $P(q)$，从而更深入地了解此处提出的系综的能量景观。这样也可以检验这些差异很大的准最优解在热力学极限下是否具有重要意义。

除匹配和定向高分子之外，其他 P类问题在适当的系综中也可能呈现复杂能量景观的迹象。本研究或其他研究中所描述的方法 [36–39]可以针对这些问题进行调整，以类似的方式对其进行研究。

致谢

模拟部分在奥尔登堡大学（德国）的 HPC 集群 CARL 上完成，该集群由德国研究联合会通过其大型科研仪器计划（INST 184/157- 1 FUGG）以及下萨克森州科学与文化部资助。

参考文献

[1]Binder K 和 Young A P 1986年 《现代物理评论》58(4) 801–976 网址 https://link.aps.org/doi/10.1103/RevModPhys.58.801[2] M´ezard M、Parisi G 和 Virasoro M A 1987年 《自旋玻璃理论及其后续发展：副本方法及其应用导论》第9卷（世界科技出版公司，新加坡）[3] Fischer K H 和 Hertz J A 1993年《自旋玻璃》第1卷（剑桥大学出版社）[4] Young A P 1998年《自旋玻璃与随机场》第12卷（世界科技出版公司，新加坡）[5] Nishimori H 2001年《自旋玻璃与信息处理的统计物理：导论》第111卷（牛津大学出版社，牛津）[6] Kawashima N 和 Rieger H 2013年《受挫自旋系统》509–614[7] Sherrington D 和 Kirkpatrick S 1975年《物理评论快报》35(26) 1792–1796 网址 https://link.aps.org/doi/10.1103/PhysRevLett.35.1792[8] Parisi G 1979年《物理快报A》73 203–205 ISSN 0375-9601 网址 https://www.sciencedirect.com/science/article/pii/0375960179907084[9] Parisi G 1979年《物理评论快报》43(23) 1754–1756 网址 https://link.aps.org/doi/10.1103/PhysRevLett.43.1754[10] Parisi G 1980年《物理学杂志A：数学与一般物理》13 1887–1895 网址 https://doi. org/10.1088/0305-4470/13/5/047[11] Parisi G 1980年《物理学杂志A：数学与一般物理》13 L115–L121 网址 https://doi.org/10.1088/0305-4470/13/4/009[12] Parisi G 1980年《物理学杂志A：数学与一般物理》13 1101–1112 网址 https://doi. org/10.1088/0305-4470/13/3/042[13] Parisi G 1983年《物理评论快报》50(24) 1946–1948 网址 https://link.aps.org/doi/10.1103/PhysRevLett.50.1946[14] Hartmann A K 和 Weigt M 2005年《组合优化问题中的相变：基础、算法与统计力学》（德国魏因海姆：Wiley-VCH）ISBN 3527404732[15] Mezard M 和 Montanari A 2009年《信息、物理学与计算》（牛津：牛津大学出版社）[16]Moore C 和 Mertens S 2016年《计算的本质》（2016年重印，含勘误版）（牛津：牛津大学出版社）ISBN 9780199233212[17]Franchini S 2022年 arXiv 网址 https://arxiv. org/abs/1610.03941v12

[18]Kirkpatrick S 和 Selman B 1994年《科学》264 1297–1301 网址 https://www.science.org/doi/abs/10.1126/science.264.5163.1297[19]Hayes B 1997年《美国科学家》85 108–112 ISSN 00030996 网址 http://www.jstor.org/ stable/27856726[20] Cocco S 和 Monasson R 2001年《物理评论快报》86(8) 1654–1657 网址 https://link.aps.org/doi/10.1103/PhysRevLett.86.1654[21]Gent I P 和 Walsh T 1996年《人工智能》88 349–358 ISSN 0004-3702 网址 https://www.sciencedirect.com/science/article/pii/S0004370296000306[22]Schawe H 和 Hartmann A K 2016年 EPL 113 30004 网址 https://doi.org/10.1209/ 0295-5075/113/30004[23] Schawe H、Jha J K 和 Hartmann A K 2019年《物理评论E辑》100(3) 032135 网址 https://link.aps.org/doi/10.1103/PhysRevE.100.032135[24] Weigt M 和 Hartmann A K 2000 年《物理评论快报》 84(26) 6118–6121 网址 https://link.aps.org/doi/10.1103/PhysRevLett.84.6118[25]Hartmann A K 和 Weigt M 2001年《理论计算机科学》265 199–225 ISSN 0304-3975 组合问题中的相变 网址 https://www.sciencedirect.com/science/article/pii/S0304397501001633[26] Weigt M 和 Hartmann A K 2001年《物理评论快报》86(8) 1658–1661 网址 https://link.aps.org/doi/10.1103/PhysRevLett.86.1658[27] Weigt M 和 Hartmann A K 2001年《物理评论E辑》63(5) 056127 网址 https://link.aps.org/doi/10.1103/PhysRevE.63.056127[28]Mulet R、Pagnani A、Weigt M 和 Zecchina R 2002年《物理评论快报》89(26) 268701 网址 https://link.aps.org/doi/10.1103/PhysRevLett.89.268701[29]Garey M R 和 Johnson D S 1979年《计算机与难解性》第174卷（旧金山: W.H. Freeman）[30] M´ezard M 和 Parisi G 1987年《物理学杂志》48 1451–1459[31] Middleton A A 和 Fisher D S 2002年《物理评论B辑》65(13) 134411 网址 https://link.aps.org/doi/10.1103/PhysRevB.65.134411[32]Hartmann A K 和 Moore M A 2003 年《物理评论快报》 90(12) 127201 网址 https://link.aps.org/doi/10.1103/PhysRevLett.90.127201[33] Hartmann A K 和 Moore M A 2004年《物理评论B辑》69(10) 104409 网址 https://link.aps.org/doi/10.1103/PhysRevB.69.104409

[34] Ahrens B and Hartmann A K 2011 《物理评论B辑》84(14) 144202 URL https://link.aps.org/doi/10.1103/PhysRevB.84.144202[35]M¨unster L, Norrenbrock C, Hartmann A K and Young A P 2021 《物理评论E》103(4) 042117 URL https://link.aps.org/doi/10.1103/PhysRevE.103.042117[36]Hartmann A K 2022 《欧洲物理快报》137 41002 URL https://dx.doi.org/10.1209/ 0295-5075/ac5226[37] Krabbe P, Schawe H and Hartmann A K 2023 《物理评论B辑》107(6) 064208 URL https://link.aps.org/doi/10.1103/PhysRevB.107.064208[38] Ricci-Tersenghi F, Weigt M and Zecchina R 2001 《物理评论E》63(2) 026702[39] M´ezard M, Ricci-Tersenghi F and Zecchina R 2003 J. Stat. Phys. 111 505[40] Ricci-Tersenghi F 2010 《科学》330 1639–1640[41] M´ezard M and Parisi G 1985 《物理学快报》46 771–778[42]Parisi G and Rati´eville M 2002 《欧洲物理学杂志B：凝聚态与复杂系统》29 457–468[43] M´ezard M and Parisi G 1988 《物理学杂志》49 2019–2025[44]Houdayer J, De Monvel J B and Martin O 1998 《欧洲物理学杂志B：凝聚态与复杂系统》6 383–393[45] Caracciolo S, Lucibello C, Parisi G and Sicuro G 2014 Physical Review E 90 012118[46]Kenyon R 2009 关于二聚体的讲座 arXiv:0910.3129[47]Melchert O, Norrenbrock C and Hartmann A 2014 《物理学程序》57 58 – 67 ISSN 1875-3892 凝聚态物理计算机模拟研究研讨会（CSP2014）论文集 URL http://www.sciencedirect.com/science/article/pii/S1875389214002776[48] Liu Y Y, Slotine J J and Barab´asi A L 2011 《自然》473 167[49]Menichetti G, Dall'Asta L and Bianconi G 2014 《物理评论快报》113(7) 078701 URL https://link.aps.org/doi/10.1103/PhysRevLett.113.078701[50] Erd˝os P and R´enyi A 1959 Publicationes Mathe- maticae 6 290–297 [51] Cook W J, Cunningham W H, Pulleyblank W R and Schrijver A 1998 组合优化（纽约, NY, USA：John Wiley & Sons, Inc.）ISBN 0-471-55894-X

[52] Egerv´ary Research Group on Co mbinatorial Opti- mization LEMON 是网络高效建模与优化库的缩写。它是一个开源 C++ 库。http://lemon.cs. elte.hu/pub/doc/1.2.3/index.html[53] Palassini M and Young A P 2000 《物理评论快报》85(14) 3017–3020 URL https://link.aps.org/doi/10.1103/PhysRevLett.85.3017[54]Zumsande M, Alava M J and Hartmann A K 2008 《统计力学杂志：理论与实验》2008 P02012[55]Zumsande M and Hartmann A K 2009 《欧洲物理学杂志B》72 619–627[56] Jaccard P 1912 《新植物学家》11 37–50[57] Levandowsky M and Winter D 1971 《自然》234 34–35[58] M´ezard M and Parisi G 1986 《物理学杂志》47 1285–1296[59] 在 [68] is 中证明，二分图上的完美匹配是#P完全问题。每个图都含有一个二分子图，且该子图的匹配数至少与完美匹配数一样多。因此可知，匹配问题同样是#P完全问题。[60] Vazirani V V 2002 《近似算法》（柏林 海德堡：施普林格）第28章[61] Hartmann A K 2015 《计算机模拟实用指南》（新加坡：世界科技出版公司）[62] Tange O 2020 Gnu parallel 20210322（'2002-01-06'）GNU Parallel 是一种通用并行化工具，无需修改多个串行命令行程序，就能并行运行它们。URL https://doi.org/ 10.5281/zenodo.4628277[63] Young P 2014 关于数据分析与拟合，你想知道的一切，却又不敢问 arXiv:1210.3781v3[64] Parisi G 1990 《法国物理学杂志》51 1595 – 1606[65] Newman M E J and Barkema G T 1999 《统计物理中的蒙特卡罗方法》（牛津：克拉伦登出版社）[66] Hartmann A K 2000 《物理学杂志A辑》33 657[67] Hartmann A K and Ricci-Tersenghi F 2002 《物理评论B辑》66 224419[68]Valiant L 1979 Theoretical Computer Sci- ence 8 189–201 ISSN 0304-3975 URL https://www.sciencedirect.com/science/article/pii/0304397579900446