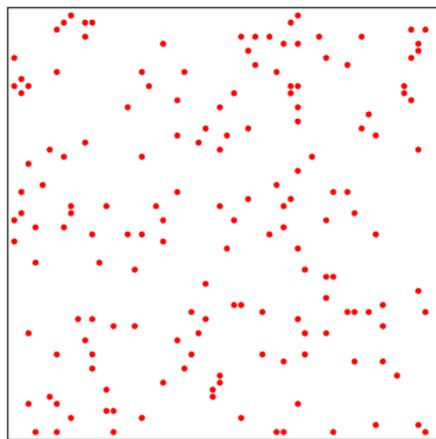
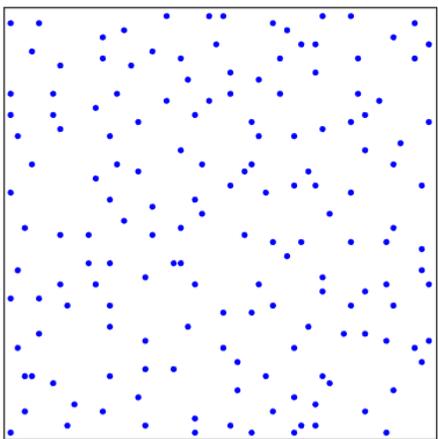


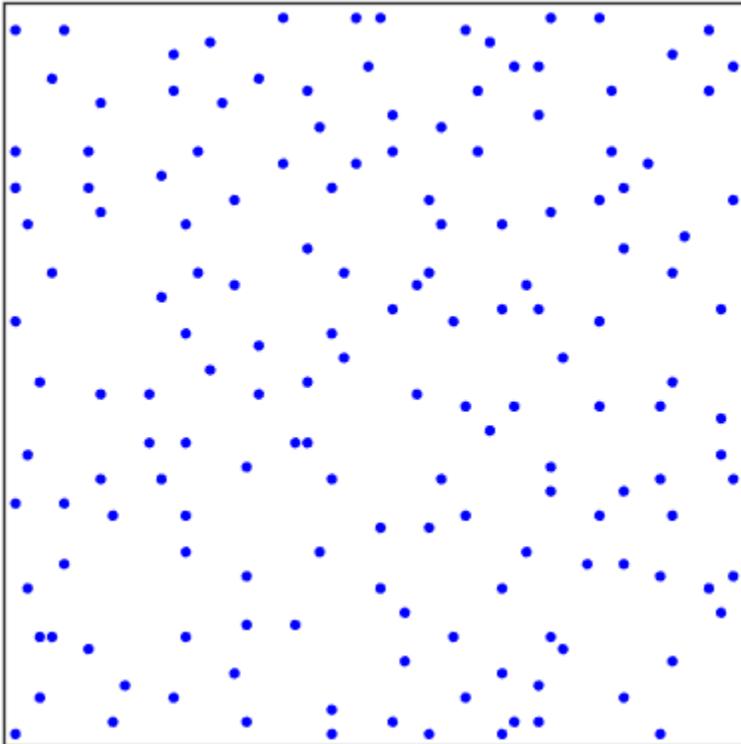
行列式点過程とLアンサンブル

中央大学 工学部 物理学科 香取研究室 小林賢矢



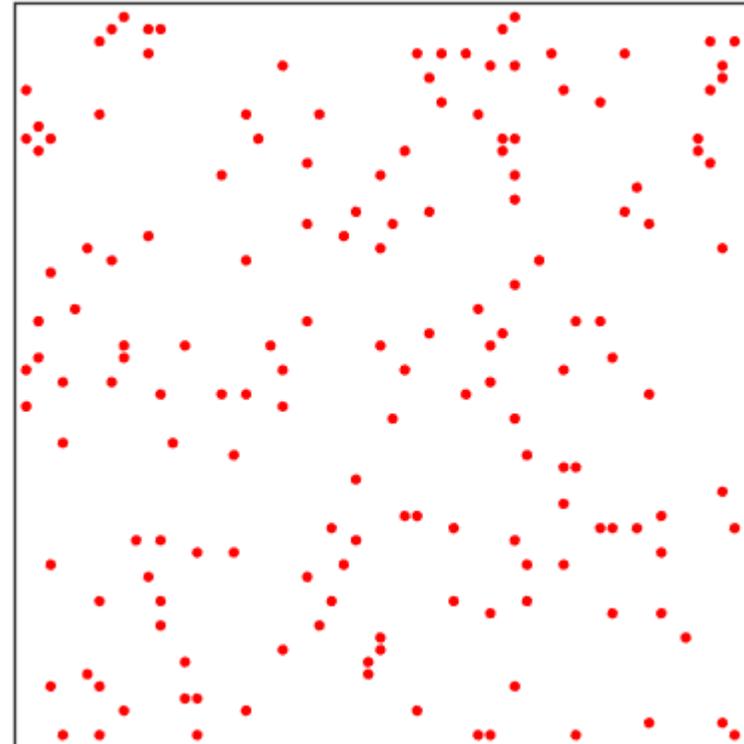
2つの点分布の比較

- ・平面状を一様に分布している。



- 粒子間に斥力が作用しているため、粒子が密集することなく分布している。
- 行列式点過程に従っている。

- ・密になったり疎になったりしていて偏りがある。



- 粒子間に何の力も作用していないため、偶然、粒子の疎密が生じている。

行列式点過程の定義

ランダムな部分集合 Y , 任意の集合 $A \subseteq Y = \{1, \dots, N\}$ (Y : グラウンドセット) に対して

$$P(A \subseteq Y) = \det(K_A)$$

が成り立つとき, 確率 P を行列式点過程と呼ぶ ($\det(K_\emptyset) = 1$ とする).

- $K: N \times N$ 実対称行列, マージナル・カーネル(核), 相関核と呼ばれる.
- K_A : 集合 A の要素に対応する行列 K の成分で構成された $|A| \times |A|$ 行列.
($|A|$ は集合 A の要素数を表す.)

行列式点過程の計算例

$Y = \{1, 2, 3, 4\}$, $A = \{2, 4\}$ とおく. このとき, 行列式点過程に従う集合 Y が集合 A を含むときの確率, つまり, A が生起する確率は

$$\begin{aligned} P(A \subseteq Y) &= \det \begin{pmatrix} K_{11} & K_{12} & K_{13} & K_{14} \\ K_{21} & K_{22} & K_{23} & K_{24} \\ K_{31} & K_{32} & K_{33} & K_{34} \\ K_{41} & K_{42} & K_{43} & K_{44} \end{pmatrix} \\ &= \det \begin{pmatrix} K_{22} & K_{24} \\ K_{42} & K_{44} \end{pmatrix} \\ &= K_{22}K_{44} - K_{24}K_{42}. \end{aligned}$$

行列式点過程の歴史的な背景

- ・ 行列式点過程は、熱平衡状態にあるフェルミ粒子系の分布を与えることから元来は「フェルミオン点過程 (fermion processes)」と呼ばれていた。
- ・ その後、2000年にBorodinとOlshanskiによって「行列式点過程」と呼ばれるようになった。
→ 2人の論文で初めて「**determinantal**」という言葉が使われた。

DISTRIBUTIONS ON PARTITIONS, POINT PROCESSES, AND THE HYPERGEOMETRIC KERNEL

ALEXEI BORODIN AND GRIGORI OLSHANSKI

ABSTRACT. We study a 3-parametric family of stochastic point processes on the one-dimensional lattice originated from a remarkable family of representations of the infinite symmetric group. We prove that the correlation functions of the processes are given by determinantal formulas with a certain kernel. The kernel can be expressed through the Gauss hypergeometric function; we call it the hypergeometric kernel.

※左の論文は[3]の
タイトル, 著者, 及び
ABSTRACTの一部より
引用.

Slater 行列式

フェルミ粒子は Pauli の排他律に従う。Pauli の排他律を満たすような波動関数の記述には、Slater 行列式が用いられる。

$$\begin{aligned}\Psi_{(n_\ell)_{\ell=1}^N}(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) &= \frac{1}{\sqrt{N!}} \det_{1 \leq k, \ell \leq N} (\psi_{n_\ell}(\mathbf{r}_k)) \\ &= \frac{1}{\sqrt{N!}} \begin{pmatrix} \psi_{n_1}(\mathbf{r}_1) & \psi_{n_1}(\mathbf{r}_2) & \cdots & \psi_{n_1}(\mathbf{r}_N) \\ \psi_{n_2}(\mathbf{r}_1) & \psi_{n_2}(\mathbf{r}_2) & \cdots & \psi_{n_2}(\mathbf{r}_N) \\ \vdots & \vdots & \cdots & \vdots \\ \psi_{n_N}(\mathbf{r}_1) & \psi_{n_N}(\mathbf{r}_2) & \cdots & \psi_{n_N}(\mathbf{r}_N) \end{pmatrix}.\end{aligned}$$

ここで、 $\sqrt{N!}$ は規格化定数である（スピン自由度は考えないものとした）。

Slater 行列式と Ginibre 点過程①

$\mathbf{r}_k = (x_k, y_k)$ とし，波動関数を

$$\psi_{\ell-1}(x_k, y_k) = \frac{e^{-(x_k^2 + y_k^2)/4\sigma^2}}{\sqrt{2\pi(\ell-1)!}\sigma} \left(\frac{x_k + iy_k}{\sqrt{2}\sigma} \right)^{\ell-1} \cdots (1) \text{式}$$

とおく．これは $2\sigma^2 = \xi^2$ のとき，量子力学的調和振動子の解として得られる（ $\xi = \sqrt{\hbar/m\omega}$ ， $\hbar := h/2\pi$ ， h ：プランク定数， m ：振動子の質量， ω ：角振動数）．フォノンをフェルミ粒子系と見なし，この基底状態での存在確率 $|\Psi_{(\ell-1)_{\ell=1}^N}(x_k, y_k)_{k=1}^N|^2$ を計算する．2つの行列式の多重線形性より，まず， $(\ell-1)!$ 部分の係数と規格化定数を合わせて

$$\frac{1}{N!} \cdot \frac{1}{(N-1)!} \cdots \frac{1}{1!} \cdot \frac{1}{0!} = \frac{1}{\prod_{n=1}^N n!}$$

を得る．続いて，(1)式の括弧内の分母部分 $(\sqrt{2}\sigma)^{\ell-1}$ に注目し，多重線形性を2つの行列式に用いると以下が示せる．

$$\prod_{n=0}^{N-1} \left(\frac{1}{2\sigma^2} \right)^n = \left(\frac{1}{2\sigma^2} \right)^{N(N-1)/2} .$$

Slater 行列式と Ginibre 点過程②

同様に多重線形性を指数部分と分母部分 $\sqrt{2\pi\sigma}$ にも適用し、最後に2つ行列式の中にある $(x_k + iy_k)^{\ell-1}$ と $(x_k - iy_k)^{\ell-1}$, $(k, \ell = 1, \dots, N)$ の部分に Vandermonde の行列式を用いると先の存在確率密度は以下のようになり、下の式の右辺は Ginibre 点過程の確率密度に一致する。

$$\left| \Psi_{(\ell-1)_{\ell=1}^N} (x_k, y_k)_{k=1}^N \right|^2 = c_N(\sigma^2) \prod_{1 \leq k < \ell \leq N} \left\{ (x_\ell - x_k)^2 + (y_\ell - y_k)^2 \right\} \\ \times \prod_{j=1}^N \left\{ p(0, \sigma^2)(x_j) p(0, \sigma^2)(y_j) \right\}.$$

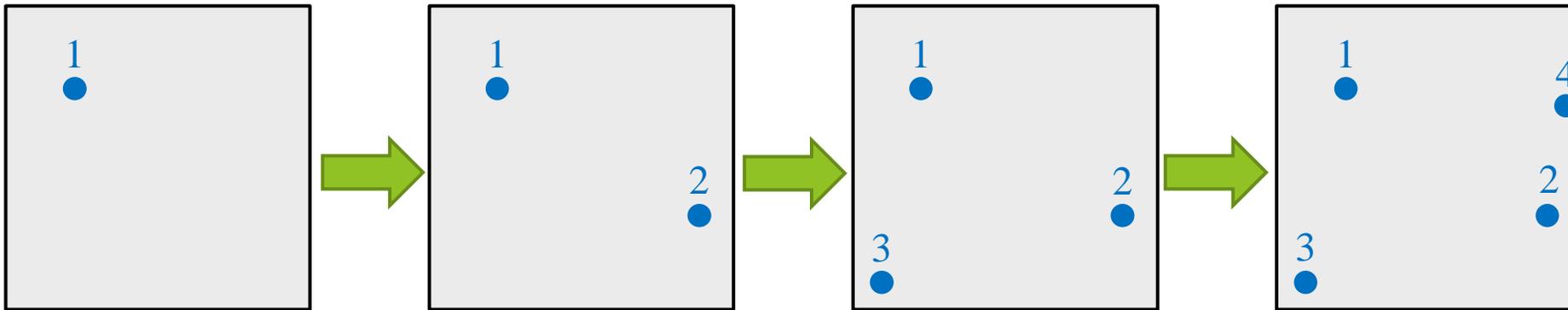
ここで,

$$c_N(\sigma^2) = 1 / \left\{ (2\sigma^2)^{N(N-1)/2} \prod_{n=1}^N n! \right\}$$

であり、 $p(0, \sigma^2)(x_j)$, $p(0, \sigma^2)(y_j)$ は平均0, 分散 σ^2 の正規分布を表す。

ランダム行列の
複素固有値の
2次元実平面上
での確率密度

点分布のプロットの過程



- ・ 2回目のプロットは1回目のプロットに影響を受ける。
- ・ 3回目のプロットは1回目と2回目のプロットに影響を受ける。
- ・ 4回目のプロットは1回目と2回目，そして3回目のプロットに影響を受ける。

→行列式点過程に従う点分布のプロットの過程では，マルコフ過程と異なる性質が見られる。

行列式点過程の有用な性質

ランダムな部分集合 \mathbf{Y} がマージナルカーネル \mathbf{K} の行列式点過程に従っているとき、その補集合 \mathbf{Y}^c もまた \mathbf{K} の行列式点過程に従う。

つまり、以下の2つの不等式

$$\begin{cases} P(\{i, j\} \subseteq \mathbf{Y}) \leq P(\{i\} \subseteq \mathbf{Y})P(\{j\} \subseteq \mathbf{Y}) \\ P(\{i, j\} \not\subseteq \mathbf{Y}) \leq P(\{i\} \not\subseteq \mathbf{Y})P(\{j\} \not\subseteq \mathbf{Y}) \end{cases}$$

が同時に成り立つ。

→行列式点過程に従う集合の要素のサンプリングを考える。着目する集合の要素数が大きく計算が困難な場合、補集合を用いてサンプリングを行えば計算コストを削減できる。

Lアンサンブルによる行列式点過程の定義

半正定値性を有する $N \times N$ 実対称行列 L について

$$P_L(\mathbf{Y} = A) = \frac{\det(L_A)}{\det(L + I)}$$

が成り立つとき、**確率 P_L** は**行列式点過程**である (I は単位行列)。

- ・規格化定数は以下の等式で現れる集合 B が空集合のときに相当する。

$$\sum_{Y: B \subseteq Y \subseteq \mathcal{Y}} \det(L_Y) = \det(L + I_{\bar{B}}).$$

- ・ $I_{\bar{B}}$: $\bar{B} = \mathcal{Y} - B$ の要素に対応する対角成分が1で、それ以外の成分が全て0である単位行列。

カーネル L の非対角成分

集合AがA={i, j}で与えられる例を考えると、その確率は以下のように与えられる。

$$P_L(\mathbf{Y} = \{i, j\}) = \det(L + I) \left\{ P_L(\mathbf{Y} = i)P_L(\mathbf{Y} = j) - \left(\frac{L_{ij}}{\det(L + I)} \right)^2 \right\}.$$

よって L_{ij} が大きくなると、集合Yが2つの要素を持つ確率は小さくなり、**カーネル L の非対角成分は負の相関を決定している**ことがわかる。カーネル L の各成分が類似度を表すものとする、

無相関のとき、すなわち $L_{ij} = 0$ のとき、

$$P_L(\mathbf{Y} = \{i, j\}) \propto P_L(\mathbf{Y} = i)P_L(\mathbf{Y} = j)$$

であるので、このときの点分布は独立になる (**ポアソン点過程**)。

他方、 $L_{ij} = \sqrt{L_{ii}L_{jj}}$ のとき、 $P_L(\mathbf{Y} = \{i, j\}) = 0$ が示される。

Lアンサンブルの有用な性質①

2つのカーネル K と L には以下の関係式が成り立つ。

$$K = L(L + I)^{-1}, \quad L = K(I - K)^{-1}.$$

Lアンサンブルの定義式から、ランダムな部分集合 \mathbf{Y} が空集合 \emptyset である確率は

$$P_L(\mathbf{Y} = \emptyset) = \frac{1}{\det(L + I)}.$$

ここで、先ほどの2つのカーネル行列 K と L の関係式を使うと

$$\begin{aligned}(L + I) &= K^{-1}L \\ &= K^{-1}K(I - K)^{-1} \\ &= (I - K)^{-1}.\end{aligned}$$

$$\therefore P_L(\mathbf{Y} = \emptyset) = \det(I - K).$$

Lアンサンブルの有用な性質②

ここで、

$$(I - K)^{-1} \text{が存在する} \iff \det(I - K) \neq 0$$

であるので、空集合の確率について $P_L(\mathbf{Y} = \emptyset) \neq 0$ が示される。
これはLアンサンブルを応用したモデリングの際に、有用な性質を与える。

→実データをモデリングするとき、稀なイベントやノイズのあるイベントに対し、ゼロでない確率を割り当てなければならないが、Lアンサンブルではその手続きがすでに行われている。

Lアンサンブルの最適化問題①

アルゴリズムにおける問題が属するクラスの定義[4]

- **P**:問題の入力のサイズの多項式で表せる時間で解決できる問題が属するクラス。
- **NP**:問題に対して提案された解が与えられたら,問題の入力のサイズとその提案された解のサイズの多項式で表せる時間で,その提案された解が問題の解だということが確かめられる問題が属するクラス。
- **NP困難**:問題を多項式時間で解決できるアルゴリズムがあれば, NPのすべての問題をこの問題に変換し,多項式時間で解けるようにする方法があるという条件を満たす問題が属するクラス。
- **NP完全**:問題がクラスNPに属し,その問題を多項式時間で解くアルゴリズムが存在すれば, NPのすべての問題をこの問題に変換し,多項式時間で解くことができる方法があるという条件を満たす問題が属するクラス。

Lアンサンブルの最適化問題②

Lアンサンブルで定義される $P_L(Y = A) \propto \det(L_A)$ の最大値を見つける問題はNP困難である。

証明. この証明では, X3C問題がNP完全であることを採用する.

X3C問題の定義

$|S|$ が3の倍数である集合 S を考える. C を S の中から3つの要素を持つ集合 C_n で作られる集合とする. このとき, S の要素を重複なく与える C の部分集合 C' は存在するか.

例

$S = \{1, 2, 3, 4, 5, 6\}$, $C_1 = \{1, 5, 6\}$, $C_2 = \{2, 3, 5\}$, $C_3 = \{4, 5, 6\}$, $C_4 = \{2, 3, 4\}$
 $C = \{C_1, C_2, C_3, C_4\}$. このとき, $C' = \{C_1, C_4\}$ が解となる.

Lアンサンブルの最適化問題③

$Y = \{1, \dots, |C|\}$, $|S| \times |C|$ 行列 B を

$$B = (B_1, \dots, B_i, \dots, B_{|C|})$$

と置き, 行列 B の各成分を

$$B_{si} = \begin{cases} 1/\sqrt{3} & (C_i \text{ が要素 } s \text{ を含むとき}) \\ 0 & (\text{その他}) \end{cases}$$

とする. このとき, 以下の $|C| \times |C|$ 行列を得る.

$$B_i^\top B_j = \begin{cases} 0 & (C_i \cap C_j = \emptyset) \\ 1/3 & (|C_i \cap C_j| = 1) \\ 2/3 & (|C_i \cap C_j| = 2) \\ 1 & (|C_i \cap C_j| = 3). \end{cases}$$

Lアンサンブルの最適化問題④

カーネル L は半正定値行列であるから以下で定義できる。

$$L = \gamma B^\top B, \quad (1 < \gamma \leq 9/8).$$

このとき、以下の命題を示す。

$\det(L_A)$ の最大値が γ^{k-1} より大きい $\iff C$ が C' を含む

(\Leftarrow の証明)

$|S|=3k$ と置くと、集合 S の要素を重複することなく含む集合 C' が存在するとき、 C' に属する集合に改めて添え字を付けると

$$C' = \{C_1, \dots, C_k\}$$

と表され、 $C_i, C_j \subset C'$ に対し $i \neq j$ のとき $|C_i \cap C_j| = 0$ 、 $i = j$ のとき $|C_i \cap C_j| = 3$ となる。今、 $\det(L_A)$ の最大値を求めたいので、 A を C' の添え字の集合とすると以下を得る。

$$\det(L_A) = \det(\gamma I) = \gamma^k > \gamma^{k-1}.$$

Lアンサンブルの最適化問題⑤

(\Rightarrow の証明)

C' が存在しないと仮定する。

(i) $|A| < k$ のとき, $|A| \leq k-1$ であるので

$$\det(L_A) \leq \prod_{i \in A} L_{ii} = \gamma^{|A|} \leq \gamma^{k-1} < \gamma^k.$$

この対偶をとると, 「 $\det(L_A) > \gamma^{k-1} \Rightarrow C'$ が存在する」となる。

(ii) $|A| \geq k$ のとき

2×2 行列の行列式は平行四辺形の面積を, 3×3 行列の行列式は平行六面体の体積を与えることを思い出すと, $\det(L_A)$ は以下のように書ける。ただし, $\text{Vol}(\{\mathbf{B}\}_{i \in A})$ は列ベクトル $\{\mathbf{B}\}_{i \in A}$ が張る体積を表す。

$$\det(L_A) = \gamma^{|A|} \text{Vol}(\{\mathbf{B}\}_{i \in A})^2.$$

Lアンサンブルの最適化問題⑥

$\text{Proj}_{\perp B_i}$ を B_i に直交する部分空間への射影演算子とすると

$$\det(L_A) = \gamma^{|A|} \left\{ \underbrace{\|B_1\| \| \text{Proj}_{\perp B_1} B_2 \| \cdots \| \text{Proj}_{\perp B_1, \dots, B_{k-2}} B_{k-1} \|}_{\textcircled{1}} \right.$$

$$\left. \times \underbrace{\| \text{Proj}_{\perp B_1, \dots, B_{k-1}} B_k \| \cdots \| \text{Proj}_{\perp B_1, \dots, B_{|A|-1}} B_{|A|} \|}_{\textcircled{2}} \right\}^2$$

と表せる。Cは存在しないと仮定すると、 $\{B\}_{i \in A}$ のうちk個の列ベクトルは直交していないことになる。したがって

$$\begin{cases} \textcircled{1} \text{の各項} \leq 1 \\ \textcircled{2} \text{の各項} < 1 \end{cases}$$

が成り立つ。また、 B_i と B_j が直交していないとき、

$$\| \text{Proj}_{\perp B_j} B_i \| = \| B_i - (B_i^\top B_j) B_j \| \leq \| B_i - \frac{1}{3} B_j \| \leq \sqrt{\frac{10}{9} - \frac{2}{9}} = \sqrt{\frac{8}{9}}$$

Lアンサンブルの最適化問題⑦

よって、行列式 $\det(L_A)$ は

$$\det(L_A) = \gamma^{|A|} \left\{ \underbrace{\|B_1\| \| \text{Proj}_{\perp B_1} B_2 \| \cdots \| \text{Proj}_{\perp B_1, \dots, B_{k-2}} B_{k-1} \|}_{\text{blue}} \right. \\ \left. \underbrace{\times \| \text{Proj}_{\perp B_1, \dots, B_{k-1}} B_k \| \cdots \| \text{Proj}_{\perp B_1, \dots, B_{|A|-1}} B_{|A|} \|}_{\text{red}} \right\}^2 \\ \leq \gamma^{|A|} \left\{ \underbrace{1^{k-1} \cdot \left(\frac{8}{9} \right)^{|A|-(k-1)}}_{\text{blue}} \right\} \leq \gamma^{|A|} \left(\frac{1}{\gamma} \right)^{|A|-(k-1)} = \gamma^{k-1}.$$

したがって、 $\det(L_A) \leq \gamma^{k-1}$ を得る。以上の議論の対偶をとると
「 $\det(L_A) > \gamma^{k-1} \Rightarrow C'$ が存在する」となる。

→ 以上より、 $\det(L_A)$ の最大化問題はNP困難である。

参考文献

2つの点分布の比較, Slater 行列式, Slater 行列式と Ginibre 点過程に関するスライドは[1]を参考にした. 行列式点過程の歴史的な背景に関するスライドと, その他のスライドの有限かつ離散からなる集合の点過程に関連する議論は[2]を参考にした.

[1]香取眞理「偶然を科学する 統計力学と確率論」数理科学, 2016年4月号,
No.634, pp.30-36, サイエンス社

[2] Alex Kulesza and Ben Taskar, “Determinantal Point Processes for Machine Learning,” Foundations and Trends in Machine Learning, Vol. 5, no. 2-3, pp. 123-286, 2012.

[3] A. Borodin and G. Olshanski, “Distributions on partitions, point processes, and the hypergeometric kernel,” Communications in Mathematical Physics, Vol. 211, no. 2, pp. 335-358, 2000.

[4] トーマス・H・コルメン (著) 長尾高弘 (訳) 「アルゴリズムの基本」
日経BP社, 2016.