МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ДЕПАРТАМЕНТ ОСВІТИ І НАУКИ
ВИКОНАВЧОГО ОРГАНУ КИЇВСЬКОЇ МІСЬКОЇ РАДИ
(КИЇВСЬКОЇ МІСЬКОЇ ДЕРЖАВНОЇ АДМІНІСТРАЦІЇ)
КИЇВСЬКЕ ТЕРИТОРІАЛЬНЕ ВІДДІЛЕННЯ МАЛОЇ АКАДЕМІЇ НАУК УКРАЇНИ
(КИЇВСЬКА МАЛА АКАДЕМІЯ НАУК)

Відділення: Математика
Секція: Математичне моделювання

МАТЕМАТИЧНЕ МОДЕЛЮВАННЯ ТА СИМУЛЯЦІЙНИЙ АНАЛІЗ ВЛАСТИВОСТЕЙ ГЕТЕРОГЕННОСТІ МЕТРИК У БАГАТОВИМІРНОМУ ІДЕОЛОГІЧНОМУ ПРОСТОРІ

Роботу виконав:
Прізвище Ім'я По батькові
учень/учениця ___ класу
Назва навчального закладу
Район

Науковий керівник:
Прізвище Ім'я По батькові
Посада, науковий ступінь

Рекомендую до захисту
__________ (підпис)

КИЇВ – 2024

Анотація

Ця робота представляє комплексний математичний та обчислювальний аналіз гетерогенності метрик відстані у багатовимірних просторових моделях голосування. Ми досліджуємо, як різні метрики відстані (L1, L2, Косинусна та Чебишева) впливають на результати голосування, коли виборці використовують гетерогенні функції відстані на основі їх позиції в ідеологічному просторі. Через симуляції Монте-Карло для вимірів 1–10 ми демонструємо, що ефекти гетерогенності метрик зберігаються для всіх протестованих вимірів, з рівнями розбіжностей від 4–18.5% залежно від правила голосування та конфігурації метрики. Ми надаємо теоретичні основи для феномена концентрації міри, показуючи, що центральність виборців збігається до $\sqrt{1/3} \approx 0.577$ у високих вимірах. Наш ключовий результат — це драматична залежна від правил та напрямкова асиметрія: коли косинусна відстань призначається центристським виборцям за методу Борда, розбіжність досягає 60% (проти 8.5% для L2-центру), що становить асиметрію 51.5 процентних пунктів. Ми декомпозуємо розбіжність на сильні (нові результати) та екстремально-вирівняні (ампліфікація) компоненти, виявляючи, що механізм ефектів гетерогенності критично залежить як від правила голосування, так і від напрямку призначення метрики. Ці результати демонструють, що гетерогенність метрик залишається значним фактором навіть у високовимірних просторах, зі складними взаємодіями між вимірністю, властивостями метрик та механізмами агрегації, які не можуть бути пояснені простими аргументами збіжності.

1 Вступ та теоретичні основи геометрії переваг

1.1 Постановка проблеми

Просторові моделі голосування представляють виборців та кандидатів як точки у багатовимірному ідеологічному просторі, де кожен вимір відповідає політичному питанню. Фундаментальне припущення полягає в тому, що виборці віддають перевагу кандидатам, ближчим до їх ідеальної точки, причому сила переваги визначається метрикою відстані. Однак реальні виборці можуть сприймати політичну “відстань” по-різному — деякі можуть зважувати всі питання однаково (L2/Евклідова), інші можуть фокусуватися на екстремізмі одного питання (L1/Манхеттенська або Чебишева), а ще інші можуть пріоритизувати ідеологічне вирівнювання над політичною відстанню (Косинусна).

Вплив високовимірних “просторів питань” на результати голосування залишається погано зрозумілим. Зі збільшенням кількості політичних вимірів геометричні властивості простору переваг змінюються драматично. Прокляття вимірності проявляється кількома способами: (1) більшість об’єму концентрується біля поверхні високовимірних сфер, (2) відстані між випадковими точками стають більш рівномірними, та (3) ефективний радіус для розрізнення “центристських” та “екстремальних” виборців збігається до фіксованого значення.

Ця робота адресує фундаментальне питання: Як гетерогенність метрик відстані впливає на результати голосування у високовимірних ідеологічних просторах, та які математичні механізми лежать в основі цих ефектів?

1.2 Наукова новизна

Наш внесок полягає в порівняльному аналізі метрик Мінковського (L1, L2, Чебишева) проти напрямкових метрик (Косинусна) через призму теорії концентрації міри. Ми демонструємо, що:

1. Центральність виборців у високовимірних гіперкубах концентрується біля $\sqrt{1/3} \approx 0.577$, незалежно від виміру (Теорема 2.3).
2. Ефекти гетерогенності метрик зберігаються для всіх протестованих вимірів (1–10), з рівнями розбіжностей 4–18.5%, які не демонструють просту монотонну збіжність.
3. Правила голосування демонструють драматичні напрямкові асиметрії: метод Борда показує екстремальну чутливість до напрямку призначення метрики, з конфігураціями косинус-центру, що дають 60% розбіжності проти 8.5% для L2-центру (асиметрія 51.5 п.п.).
4. Розбіжність декомпозується на сильні (нові результати) та екстремально-вирівняні (ампліфікація) компоненти, з пропорціями, що варіюються драматично за правилом та конфігурацією (10–93% сильних розбіжностей залежно від напрямку).
5. Ефективний радіус для класифікації центр-екстремальних виборців збігається до порогового значення на основі перцентилів зі збільшенням виміру, дозволяючи стратегії класифікації незалежні від виміру.

Ці результати мають глибокі імплікації для демократичної теорії: гетерогенність метрик залишається значним фактором навіть у високовимірних просторах, з залежними від правил та напрямковими ефектами, які складним чином взаємодіють з механізмами агрегації.

1.3 Просторові моделі та метрики відстані

У просторовій теорії голосування ми моделюємо виборців та кандидатів як точки $\mathbf{x}_v, \mathbf{y}_c \in \mathbb{R}^d$ у d-вимірному просторі політик. Корисність виборця v для кандидата c є спадною функцією відстані: $u_{v,c} = f(d(\mathbf{x}_v, \mathbf{y}_c))$, де d — метрика відстані.

1.3.1 Метрики Мінковського

Сімейство p-норм Мінковського включає:

$$d_p(\mathbf{x}, \mathbf{y}) = \left(\sum_{i=1}^{d} |x_i - y_i|^p \right)^{1/p}, \quad p \geq 1 \tag{1}$$

Особливі випадки:

- L1 (Манхеттенська): $p = 1$, $d_1(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{d} |x_i - y_i|$. Підкреслює координатні різниці, підходить для виборців одного питання.
- L2 (Евклідова): $p = 2$, $d_2(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{i=1}^{d} (x_i - y_i)^2}$. Стандартна геометрична відстань, ставиться до всіх вимірів однаково.
- Чебишева (L∞): $p = \infty$, $d_\infty(\mathbf{x}, \mathbf{y}) = \max_i |x_i - y_i|$. Екстремальний фокус на одному питанні.

1.3.2 Напрямкові метрики: Косинусна відстань

Косинусна відстань вимірює кутове відділення, а не просторову відстань:

$$d_{\cos}(\mathbf{x}, \mathbf{y}) = 1 - \frac{\mathbf{x} \cdot \mathbf{y}}{||\mathbf{x}|| \cdot ||\mathbf{y}||} = 1 - \cos(\theta) \tag{2}$$

де θ — кут між векторами. Ця метрика є інваріантною до масштабу та захоплює ідеологічне вирівнювання: виборці з подібними напрямками (навіть на різних відстанях від початку) мають низьку косинусну відстань.

1.4 Прокляття вимірності

Високовимірні простори демонструють контрінтуїтивні геометричні властивості. Для рівномірного вибірки в гіперкубі $[-1, 1]^d$:

Твердження 1.1 (Концентрація об'єму). Коли $d \to \infty$, майже весь об'єм d-вимірного гіперкуба концентрується біля його поверхні. Частка об'єму на відстані ϵ від межі наближається до 1.

Твердження 1.2 (Рівномірність відстані). Для двох випадкових точок $\mathbf{x}, \mathbf{y} \sim \text{Uniform}([-$ розподіл $||\mathbf{x} - \mathbf{y}||_2$ стає все більш концентрованим навколо середнього зі збільшенням d.

1.4.1 Аналітичне наближення середньої відстані

Хоча точні формули існують для середньої відстані у низьких вимірах (до 4-го та 5-го вимірів), складність зростає експоненційно з виміром, що робить це непрактичним для вищих вимірів. Замість цього ми використовуємо асимптотичне наближення. Для нормалізованих відстаней (поділених на $\sqrt{d}$), середнє збігається до $\sqrt{2/3}$ при $d \to \infty$, але для скінченних вимірів більш точне наближення:

$$\mathbb{E}\left[\frac{||\mathbf{x}-\mathbf{y}||_2}{\sqrt{d}}\right] \approx \sqrt{\frac{2}{3}}\left(1-\frac{7}{40d}\right) \tag{3}$$

Ця формула враховує скінченновимірні корекції до асимптотичної межі. Зі збільшенням d, корекційний член $7/(40d)$ зникає, і середнє наближається до $\sqrt{2/3} \approx 0.816497$. Помилка наближення становить максимум 1% у вимірі 1, де точне значення $2/3 \approx 0.666667$, і швидко зменшується для вищих вимірів.

Рисунок 1 надає емпіричне підтвердження Твердження 2.2, показуючи, як розподіл нормалізованих відстаней (поділених на $\sqrt{d}$) концентрується навколо аналітичного середнього зі збільшенням виміру. Гістограми демонструють, що хоча відстані можуть ще варіюватися у низьких вимірах (вимір 1 показує широкий, асиметричний розподіл), до виміру 100 розподіл стає надзвичайно вузьким та піковим, зі стандартним відхиленням, що зменшується від 0.471 (вимір 1) до 0.049 (вимір 100).

Рисунок 2 кількісно визначає цю збіжність, побудовуючи коефіцієнт варіації (стандартне відхилення поділене на середнє) та абсолютне стандартне відхилення як функції виміру. Обидві міри зменшуються з виміром, підтверджуючи, що відносна дисперсія зменшується як $1/\sqrt{d}$, тоді як абсолютна дисперсія також зменшується, демонструючи феномен концентрації міри.

Ці феномени створюють "порожні кути" — області простору, які геометрично можливі, але статистично малоймовірні для міщення виборців або кандидатів. Це має глибокі імплікації для голосування: зі збільшенням виміру ефективний простір конфігурації зменшується, призводячи до більш передбачуваних результатів.

1.5 Концентрація міри та ефективний радіус

Наш ключовий теоретичний результат стосується концентрації центральності виборців у високих вимірах.

Визначення 1.3 (Нормалізована L2 центральність). Для виборця в позиції $\mathbf{x} \in [-1,1]^d$ з центром $\mathbf{c} = \mathbf{0}$ (геометричний центр), нормалізована L2 центральність:

$$c(\mathbf{x}) = \frac{||\mathbf{x}||_2}{\sqrt{d}\cdot\frac{\text{range}}{2}} = \frac{||\mathbf{x}||_2}{\sqrt{d}} \tag{4}$$

де знаменник — половина діагоналі гіперкуба.

Теорема 1.4 (Концентрація центральності). Для $\mathbf{X} \sim \text{Uniform}([-1,1]^d)$, коли $d \to$

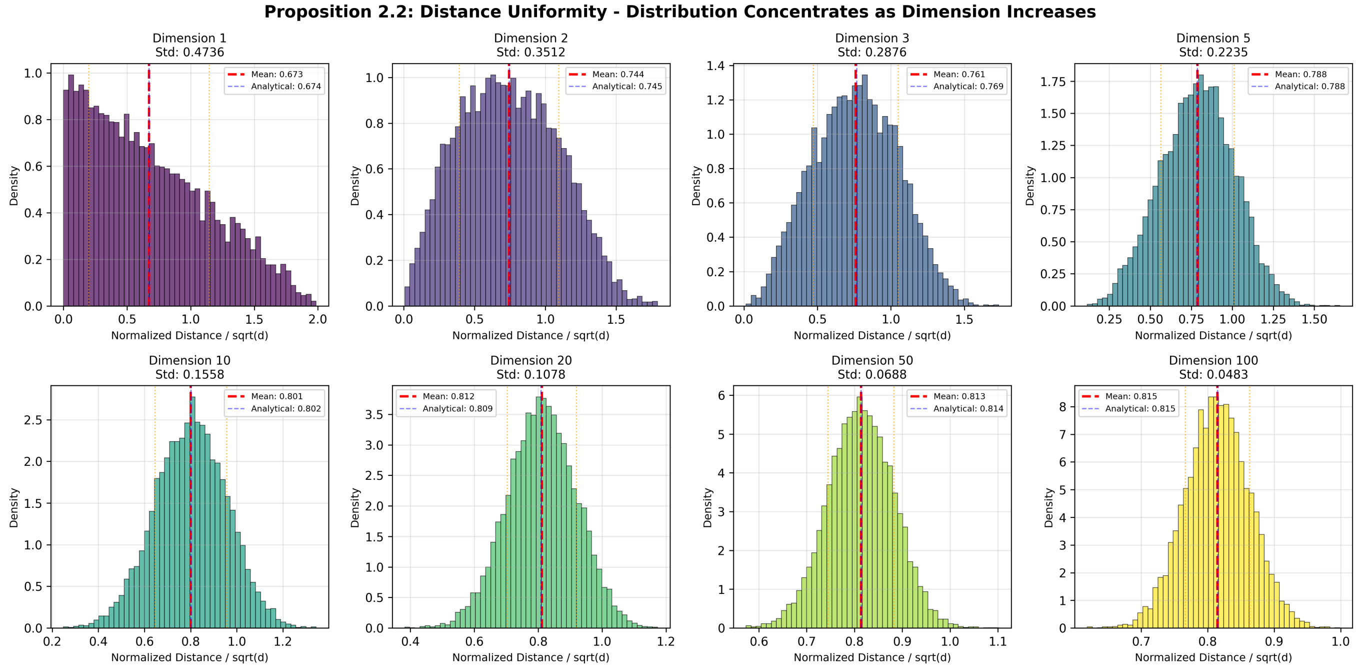

Рис. 1: Емпіричне підтвердження Твердження 2.2: Розподіл нормалізованих Евклідових відстаней між випадковими парами точок у $[-1,1]^d$ для вимірів 1, 2, 3, 5, 10, 20, 50 та 100. Кожна гістограма показує щільність нормалізованих відстаней (поділених на $\sqrt{d}$), з червоними пунктирними лініями, що вказують емпіричне середнє, синіми пунктирними лініями, що показують аналітичне середнє $\sqrt{2/3}(1-7/(40d))$, та помаранчевими пунктирними лініями, що позначають ± 1 стандартне відхилення. Зі збільшенням виміру розподіл стає все більш концентрованим навколо середнього, зі стандартним відхиленням, що зменшується від 0.471 (вимір 1) до 0.049 (вимір 100).

∞:

$$\frac{||\mathbf{X}||_2}{\sqrt{d}} \xrightarrow{p} \sqrt{\mathbb{E}[X_1^2]} = \sqrt{\frac{1}{3}} \approx 0.577 \tag{5}$$

Доведення. За законом великих чисел:

$$\frac{||\mathbf{X}||_2^2}{d} = \frac{1}{d}\sum_{i=1}^{d} X_i^2 \xrightarrow{p} \mathbb{E}[X_1^2] = \int_{-1}^{1} \frac{x^2}{2} dx = \frac{1}{3} \tag{6}$$

Взявши квадратні корені та застосувавши неперервність функції квадратного кореня, отримуємо результат. □

Ця теорема (Результат 5) пояснює, чому наша стратегія класифікації центр-екстремальних виборців стає незалежною від виміру: у високих вимірах майже всі виборці мають центральність біля 0.577, роблячи порогові значення на основі перцентилів (а не абсолютні радіальні пороги) природним вибором.

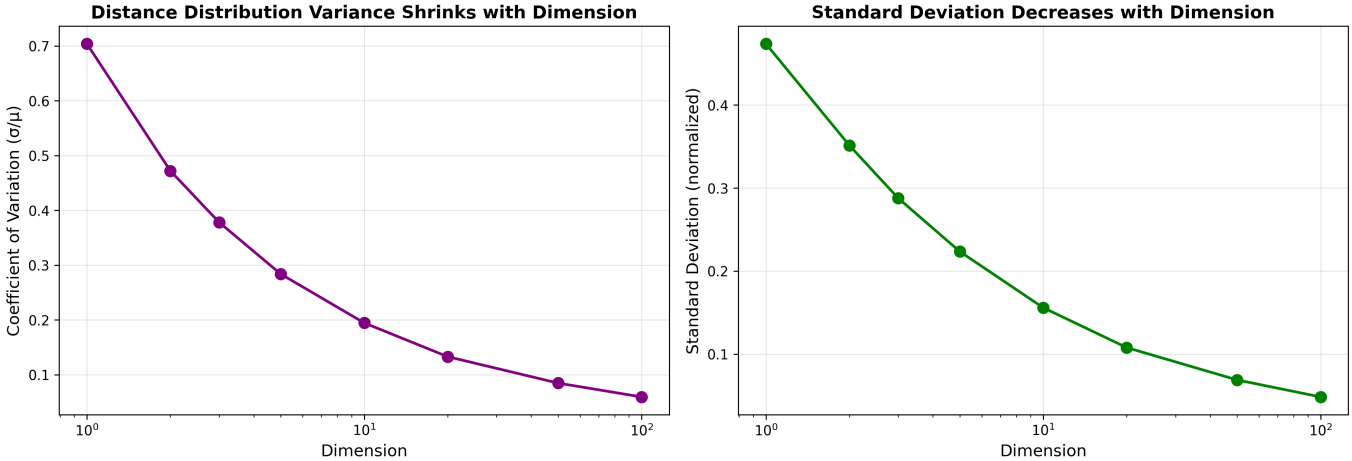

Рис. 2: Збіжність дисперсії розподілу відстані з виміром. Зліва: Коефіцієнт варіації (σ/μ) зменшується з виміром, показуючи, що відносна дисперсія зменшується як $1/\sqrt{d}$. Справа: Абсолютне стандартне відхилення (нормалізоване на $\sqrt{d}$) також зменшується з виміром. Обидва графіки використовують логарифмічну вісь x для показу збіжності через порядки величини.

1.5.1 Емпіричне підтвердження

Наші симуляції підтверджують це теоретичне передбачення. Таблиця 1 показує середню центральність за виміром для рівномірної вибірки гіперкуба:

Табл. 1: Емпірична збіжність центральності до $\sqrt{1/3} \approx 0.577$

Вимір	Середня центральність	Відхилення від теорії
1	0.4995	0.0778
2	0.5410	0.0363
3	0.5541	0.0232
5	0.5645	0.0128
10	0.5712	0.0062
20	0.5745	0.0029
50	0.5762	0.0012
100	0.5768	0.0006
200	0.5770	0.0003

Збіжність очевидна: до виміру 200 середня центральність (0.5770) відхиляється від теоретичного значення (0.5774) лише на 0.0003. Для вимірів ≥ 10 середня центральність становить 0.5751, з відхиленням 0.0022. Це підтверджує, що у високих вимірах позиції виборців концентруються біля ефективного радіуса, роблячи розрізнення центр-екстремальних все більш значущим лише через класифікацію на основі перцентилів.

2 Методологія симуляцій, нормалізація корисності та аналіз гетерогенності метрик

2.1 Рамки Монте-Карло для електоральних симуляцій

Наша система симуляцій реалізує комплексний підхід Монте-Карло до електорального моделювання:

- Генерація профілів: Для кожної експериментальної конфігурації ми генеруємо $N = 200$ незалежних електоральних профілів.
- Просторова вибірка: Виборці та кандидати вибіркові рівномірно з $[-1, 1]^d$ для забезпечення повного кутового діапазону для косинусної відстані.
- Кількість виборців: Базові експерименти використовують $n = 100$ виборців на профіль, з тестами масштабування від 10 до 500 виборців.
- Кількість кандидатів: Фіксована на $M = 5$ кандидатів на профіль.
- Випадкові насіння: Експерименти Монте-Карло використовують унікальні насіння (42, 43, 44, ...) для кожного запуску для забезпечення статистичної незалежності. Візуалізації та скрипти одноразового запуску використовують фіксоване насіння 42 для відтворюваності.

Для кожного профілю ми обчислюємо:

1. Позиції виборців: $\mathbf{X} \in \mathbb{R}^{n \times d}$
2. Позиції кандидатів: $\mathbf{Y} \in \mathbb{R}^{M \times d}$
3. Матриця відстаней: $D \in \mathbb{R}^{n \times M}$ використовуючи гетерогенні або однорідні метрики
4. Матриця корисності: $U \in \mathbb{R}^{n \times M}$ через лінійну нормалізацію корисності
5. Ранжування: $R \in \{0, \ldots, M-1\}^{n \times M}$ через argsort корисностей
6. Результати правил голосування: Переможці для методів відносної більшості, Борда, IRV та Кондорсе

2.2 Відносна нормалізація корисності: Масштабування найкраще-найгірше

Порівняння корисностей між виборцями представляє фундаментальний виклик у просторових моделях голосування, оскільки різні виборці можуть сприймати відстані по-різному залежно від призначеної їм метрики. Для вирішення цього ми використовуємо схему нормалізації на виборця, яка відображає корисності на інтервал $[0, 1]$ використовуючи лінійне перетворення корисності. Протягом цього дослідження ми послідовно використовуємо лінійні функції корисності, а не гаусові або інші нелінійні форми.

Визначення 2.1 (Лінійна корисність з масштабуванням найкраще-найгірше). Для виборця v в позиції $\mathbf{x}_v$ та кандидата c в позиції $\mathbf{y}_c$, ми спочатку обчислюємо сирі корисності з відстаней:

$$\tilde{u}_{v,c} = -d(\mathbf{x}_v, \mathbf{y}_c) \tag{7}$$

де d позначає метрику відстані, призначену виборцю v (яка може варіюватися між виборцями в гетерогенних метричних конфігураціях). Ці сирі корисності потім нормалізуються на виборця через масштабування найкраще-найгірше:

$$u_{v,c} = \frac{\tilde{u}_{v,c} - \min_{c'} \tilde{u}_{v,c'}}{\max_{c'} \tilde{u}_{v,c'} - \min_{c'} \tilde{u}_{v,c'}} \tag{8}$$

Це перетворення гарантує, що найбільш переважений кандидат виборця (найближчий у їх призначеній метриці) отримує корисність 1, їх найменш переважений кандидат (найдальший) отримує корисність 0, а всі інші кандидати отримують корисності лінійно інтерпольовані між цими межами.

Нормалізація масштабування найкраще-найгірше надає кілька ключових властивостей:

- Обмежені корисності: $u_{v,c} \in [0, 1]$ для всіх пар виборець-кандидат, дозволяючи значущу агрегацію між виборцями
- Нормалізація найкращого кандидата: Найближчий кандидат кожного виборця (за їх призначеною метрикою) завжди отримує корисність 1
- Нормалізація найгіршого кандидата: Найдальший кандидат кожного виборця завжди отримує корисність 0
- Збереження метрики: Корисності вірно кодують відносини відстаней, обчислені використовуючи специфічну метрику відстані кожного виборця
- Лінійний спад: Корисність зменшується лінійно з відстанню, на відміну від експоненційного спаду в моделях гаусової корисності
- Масштабування відносно профілю: Корисності нормалізовані відносно лише кандидатів, присутніх у кожному електоральному профілі, без абсолютних порогів відстані

2.2.1 Декомпозиція міри розбіжності

Для розуміння того, як гетерогенність метрик впливає на результати, ми декомпозуємо розбіжність на два компоненти:

Визначення 2.2 (Сильна розбіжність). Сильна розбіжність D_{strong} — це відсоток профілів, де гетерогенний переможець відрізняється від обидвох базових переможців метрики центру та метрики екстремуму:

$$D_{\text{strong}} = \frac{1}{N}\sum_{p=1}^{N} \mathbf{1}[w_{\text{het},p} \neq w_{\text{center},p} \wedge w_{\text{het},p} \neq w_{\text{extreme},p}] \tag{9}$$

де $w_{\text{het},p}$, $w_{\text{center},p}$ та $w_{\text{extreme},p}$ — переможці для профілю p за гетерогенних, однорідних метрики центру та однорідних метрики екстремуму умов відповідно.

Сильна розбіжність представляє випадки, де гетерогенність створює новий результат, який жодна базова лінія не дала б.

Визначення 2.3 (Екстремально-вирівняна розбіжність). Екстремально-вирівняна розбіжність $D_{\text{ext-align}}$ — це відсоток профілів, де гетерогенний переможець дорівнює базовому переможцю метрики екстремуму, але відрізняється від базового метрики центру:

$$D_{\text{ext-align}} = \frac{1}{N}\sum_{p=1}^{N} \mathbf{1}[w_{\text{het},p} = w_{\text{extreme},p} \wedge w_{\text{het},p} \neq w_{\text{center},p}] \tag{10}$$

Екстремально-вирівняна розбіжність представляє випадки, де гетерогенність посилює результат метрики екстремуму, роблячи його переможцем навіть коли базова лінія метрики центру вибрала б інакше.

Загальна розбіжність проти базової лінії центру: $D_{\text{total}} = D_{\text{strong}} + D_{\text{ext-align}}$.

2.3 Математична логіка правил агрегації

2.3.1 Метод Борда

Борда призначає бали на основі позиції рангу:

$$\text{score}_c = \sum_{v=1}^{n} (M - \text{rank}_v(c)) \tag{11}$$

де $\text{rank}_v(c) \in \{0, \ldots, M-1\}$ — ранг кандидата c у порядку переваг виборця v (0 = найбільш переважений).

Переможець: $\arg\max_c \text{score}_c$

2.3.2 Ранжовані пари (метод Кондорсе)

Ранжовані пари конструюють соціальне впорядкування через:

1. Обчислення попарних маржж: $m_{i,j} = |\{v : i \succ_v j\}| - |\{v : j \succ_v i\}|$

2. Сортування пар (i, j) за спаданням маржі $m_{i,j}$
3. Фіксація пар по порядку, пропускаючи пари, які створили б цикли
4. Переможець — найкращий кандидат у фінальному впорядкуванні

Цей метод задовольняє критерій Кондорсе: якщо існує переможець Кондорсе, Ранжовані пари вибирають його.

2.3.3 Голосування за оцінкою (кардинальне)

Голосування за оцінкою використовує корисності безпосередньо:

$$\text{score}_c = \sum_{v=1}^{n} u_{v,c} \tag{12}$$

Переможець: $\arg\max_c \text{score}_c$

Це еквівалентно максимізації утилітарного соціального добробуту за нашої нормалізації.

2.3.4 Відносна більшість та IRV

- Відносна більшість: Переможець — кандидат, ранжований першим найбільшою кількістю виборців.
- Миттєве голосування другого туру (IRV): Ітеративно усуває кандидата з найменшою кількістю голосів першого місця, перерозподіляючи голоси до тих пір, поки один кандидат не матиме більшості.

2.4 Збіжність вимірності

Наші емпіричні результати демонструють, що ефекти гетерогенності метрик зберігаються для всіх протестованих вимірів, зі значними рівнями розбіжностей, які не демонструють просту монотонну збіжність. Декомпозиція розбіжності на сильні та екстремально-вирівняні компоненти виявляє складні патерни взаємодії, які варіюються за виміром та правилом голосування.

2.4.1 Емпіричні докази

Таблиця 2 представляє декомпозовані рівні розбіжностей для пар метрик L2-Косинусна для вимірів 1–10:

Дані виявляють значні ефекти гетерогенності метрик для всіх вимірів, з трьома ключовими спостереженнями:

Табл. 2: Рівні розбіжностей за виміром (L2 центр, Косинусна екстремум, поріг=0.5). Всі значення показують Сильна / Екстремально-вирівняна / Загальна розбіжність (%).

Вимір	Відносна більшість	Борда	IRV
1	6.5 / 12.0 / 18.5	6.5 / 0.5 / 7.0	10.5 / 5.5 / 16.0
2	6.5 / 7.5 / 14.0	5.0 / 3.5 / 8.5	6.0 / 6.5 / 12.5
3	2.5 / 7.0 / 9.5	1.5 / 2.5 / 4.0	1.0 / 9.0 / 10.0
4	1.0 / 7.5 / 8.5	8.0 / 3.5 / 11.5	3.5 / 10.0 / 13.5
5	4.0 / 7.0 / 11.0	2.0 / 6.5 / 8.5	4.5 / 13.0 / 17.5
7	5.0 / 13.0 / 18.0	6.0 / 8.0 / 14.0	1.5 / 10.0 / 11.5
10	3.0 / 9.5 / 12.5	4.5 / 6.0 / 10.5	4.0 / 9.0 / 13.0

1. Загальна розбіжність: Коливається від 4–18.5% залежно від виміру та правила голосування. Борда показує найнижчу розбіжність (4–14%), тоді як відносна більшість та IRV показують вищі рівні (8.5–18.5%).
2. Патерни декомпозиції: Співвідношення сильних до екстремально-вирівняних розбіжностей варіюється драматично:
   - Борда вимір 1: 92.9% сильних розбіжностей (гетерогенність вводить нові результати)
   - IRV вимір 3: 10.0% сильних розбіжностей (гетерогенність переважно вирівнюється з базовою лінією метрики екстремуму)
3. Немає монотонної тенденції: Розбіжність не зменшується монотонно з виміром, що вказує на складні ефекти взаємодії, а не просту збіжність.

Таблиця 3 представляє загальні рівні розбіжностей для всіх пар метрик у вимірі 2:

Табл. 3: Загальні рівні розбіжностей пар метрик у вимірі 2 (нотація A → B).

Пара метрик	Відносна більшість	Борда	IRV
L2 → Косинусна	14.0%	8.5%	12.5%
Косинусна → L2	16.0%	60.0%	4.5%
L1 → Косинусна	16.5%	8.5%	11.5%
Косинусна → L1	14.0%	52.5%	9.5%
Косинусна → Чебишева	20.0%	64.5%	8.0%
Чебишева → Косинусна	16.0%	11.5%	12.5%
L1 → L2	14.0%	11.5%	9.0%
L2 → L1	9.5%	5.5%	6.5%
L1 → Чебишева	19.0%	18.0%	11.0%
Чебишева → L1	13.0%	11.0%	10.5%
L2 → Чебишева	7.0%	7.5%	6.5%
Чебишева → L2	7.0%	7.0%	7.5%

Це виявляє драматичну асиметрію: пари Косинусна-Мінковський показують значно різні розбіжності, ніж пари Мінковський-Косинусна, і цей ефект залежить від правил. Асиметрія найбільш виражена для методу Борда, де напрямкові ефекти перевищують 50 процентних пунктів:

- Борда: Показує екстремальну напрямкову чутливість. Косинусна→L2 має 60.0% розбіжності проти 8.5% для L2→Косинусна — асиметрія 51.5 п.п. Аналогічно, Косинусна→Чебишева показує 64.5% проти 11.5% для Чебишева→Косинусна (асиметрія 53.0 п.п.). Це вказує, що Борда високо чутлива до того, яка метрика призначається центр- або екстремальним виборцям, з косинусною відстанню, що створює драматично різні результати, коли призначається центристським виборцям.
- IRV: Показує помірну асиметрію зі зворотними патернами. L2→Косинусна має 12.5% розбіжності проти 4.5% для Косинусна→L2 (асиметрія 8.0 п.п. у зворотному напрямку від Борда).
- Відносна більшість: Відносно симетрична (14–20% розбіжності), з меншими асиметріями зазвичай менше 6 п.п.

Асиметрія Борда-Косинусна особливо вражаюча: коли Косинусна є метрикою центру, результати відрізняються від базової лінії L2 60% часу, але коли L2 є метрикою центру, розбіжність падає до 8.5%. Ця 7-кратна різниця вказує, що призначення метрик до груп виборців має таке ж значення, як і вибір самих метрик, і що механізм підсумовування переваг методу Борда сильно взаємодіє з напрямковими властивостями косинусної відстані.

2.4.2 Аналіз декомпозиції: Сильна проти екстремально-вирівняної розбіжності

Дослідження декомпозиції розбіжності на сильні та екстремально-вирівняні компоненти виявляє механізм, через який гетерогенність впливає на результати:

Табл. 4: Декомпозиція розбіжності для L2→Косинусна (Вимір 2)

Правило	Сильна	Екстремально-вирівняна	Сильна%
Відносна більшість	6.5%	7.5%	46.4%
Борда	5.0%	3.5%	58.8%
IRV	6.0%	6.5%	48.0%

Ключові результати:

- Борда показує найвищу пропорцію сильних розбіжностей (58.8%), що означає, що гетерогенність вводить справді нові результати, а не лише зсуває до базової лінії метрики екстремуму.

- Відносна більшість та IRV показують більш збалансовану декомпозицію (46–48% сильних), вказуючи, що ефекти гетерогенності розділені між новими результатами та екстремальним вирівнюванням.

- Для Косинусна→L2 Борда сильна розбіжність становить лише 11.5% (19.2% від загальної), з 48.5% екстремально-вирівняною (80.8% від загальної). Це виявляє, що висока загальна розбіжність (60%) переважно обумовлена вирівнюванням з базовою лінією косинусної однорідності, а не справді новими результатами. Гетерогенне правило ефективно посилює перевагу косинусної метрики, а не створює повністю нові рівноваги.

- Навпаки, L2→Косинусна Борда показує 58.8% сильних розбіжностей (5.0% від 8.5% загальної), що означає, що більшість розбіжностей представляють справді нові результати, не пояснені жодною базовою лінією.

Ця декомпозиція уточнює, що що робить гетерогенність критично залежить як від правила голосування, так і від напрямку призначення метрики. Та сама пара метрик (L2-Косинусна) створює фундаментально різні механізми залежно від того, яка метрика призначається центристським виборцям: L2-центр створює нові результати, тоді як Косинусна-центр посилює існуючі косинусні переваги.

2.4.3 Візуалізація сильних розбіжностей

Для ілюстрації механізму сильних розбіжностей, Рисунок 3 показує один електоральний профіль за трьох різних метричних конфігурацій. Ця візуалізація демонструє, як та сама просторова розстановка виборців та кандидатів дає трьох різних переможців залежно від призначення метрики.

Ключове розуміння з цієї візуалізації полягає в тому, що Кандидат 3 лише виграє, коли виборці використовують змішані метрики. Жодна однорідна базова лінія не дає цього результату:

- Якби всі виборці використовували L2 (Евклідова відстань), Кандидат 2 виграв би

- Якби всі виборці використовували Косинусну (кутова відстань), Кандидат 0 виграв би

- Але зі змішаними метриками (L2 для центристів, Косинусна для екстремалів), Кандидат 3 виникає як переможець

Це просторовий підпис сильних розбіжностей: гетерогенний переможець не пояснений жодною однорідною базовою лінією. Взаємодія між центристськими виборцями (сприймають відстань евклідово) та екстремальними виборцями (сприймають відстань кутово) створює нову рівновагу, яка не існувала б, якби всі виборці поділяли однакове сприйняття відстані.

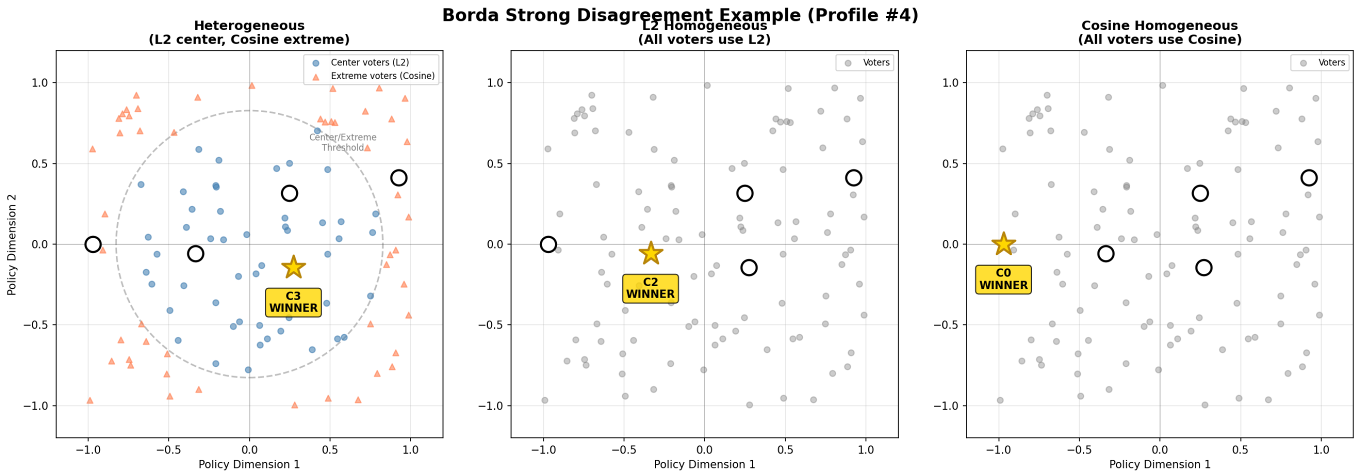

Рис. 3: Приклад сильних розбіжностей, що показує той самий профіль за трьох метричних конфігурацій. Зліва: Гетерогенна (L2 центр, Косинусна екстремум) обирає Кандидата 3. По центру: L2 однорідна обирає Кандидата 2. Справа: Косинусна однорідна обирає Кандидата 0. Сині кола представляють центристських виборців (використовують L2), червоні трикутники представляють екстремальних виборців (використовують Косинусну), а пунктирне коло показує поріг центр-екстремум. Кандидат 3 виграє лише за гетерогенності, демонструючи, що змішані метрики створюють справді емерджентні результати.

2.4.4 Теоретичне пояснення

Збереження ефектів гетерогенності через виміри, незважаючи на концентрацію міри, вказує, що вибір метрики залишається значущим навіть у високовимірних просторах. Хоча концентрація міри спричиняє, що більшість виборців кластеризуються біля ефективного радіуса (≈ 0.577), напрямкові властивості косинусної відстані створюють стійкі різниці в ранжуванні переваг, які не усуваються вимірним масштабуванням. Немонотонні патерни в рівнях розбіжностей (наприклад, вимір 4 Борда показує 11.5% розбіжності проти 4.0% для виміру 3) вказують на складні взаємодії між вимірністю, властивостями метрик та механізмами агрегації правил голосування, які не можуть бути пояснені простими аргументами збіжності.

2.5 Вплив метрик на парадокси Кондорсе

2.5.1 Взаємодія Борда-Косинусна

Наші симуляції виявляють помітний феномен: коли косинусна відстань використовується як метрика центру (з методом Борда), рівні розбіжностей різко зростають порівняно з конфігураціями L2-центру.

Для Косинусна→L2 з Борда ми спостерігаємо:

- Загальна розбіжність: 60.0% (проти 8.5% для L2→Косинусна) — асиметрія 51.5 п.п.

- Сильна розбіжність: 11.5% (19.2% від загальної) — справді нові результати
- Екстремально-вирівняна розбіжність: 48.5% (80.8% від загальної) — посилення косинусних переваг
- Це вказує, що більшість розбіжностей представляють вирівнювання з базовою лінією косинусної однорідності, а не справді нові результати

Ця "Взаємодія Борда-Косинусна" виявляє фундаментальну асиметрію: коли Косинусна призначається центристським виборцям, гетерогенне правило дає результати, які сильно вирівнюються з тим, що обирала б косинусно-однорідна електоральна група (48.5% профілів), вказуючи, що косинусна відстань чинить потужний вплив на агрегацію Борда, коли застосовується до більшості виборців. Напрямкова природа косинусної відстані — фокусування на ідеологічному вирівнюванні, а не політичній відстані — взаємодіє з механізмом підсумовування переваг Борда способами, що посилюють вплив косинусної метрики, ефективно роблячи гетерогенну електоральну групу поводитися більш як косинусно-однорідна.

Зворотний напрямок (L2→Косинусна) показує протилежний патерн: лише 3.5% екстремально-вирівняної розбіжності, з 5.0% сильних розбіжностей (58.8% від загальної). Це вказує, що відстань L2, коли призначається центристським виборцям, створює справді нові результати, а не лише посилює переваги метрики екстремуму.

2.5.2 Взаємодії метрик Мінковського

Метрики L1 та L2 показують помірну чутливість до гетерогенності з відносно симетричними напрямковими ефектами. Для пар L1-L2 у вимірі 2:

- L1→L2: 14.0% (Відносна більшість), 11.5% (Борда), 9.0% (IRV)
- L2→L1: 9.5% (Відносна більшість), 5.5% (Борда), 6.5% (IRV)
- Асиметрія: 4.5 п.п. (Відносна більшість), 6.0 п.п. (Борда), 2.5 п.п. (IRV)

Декомпозиція виявляє, що пари L1-L2 дають переважно екстремально-вирівняну розбіжність (наприклад, L1→L2 Борда: 1.0% сильних, 10.5% екстремально-вирівняних), вказуючи, що гетерогенність зсуває результати до базової лінії метрики екстремуму, а не створює нові рівноваги. Це контрастує з косинусними взаємодіями, де напрямок має драматичне значення для Борда.

Пари Чебишева-L2 показують майже ідеальну симетрію (7.0–7.5% розбіжності в обох напрямках), вказуючи, що метрики Чебишева та L2 взаємодіють збалансовано, відносно нечутливо до напрямку призначення.

2.6 Перевірки ефективності задоволеності виборців (VSE)

Ефективність задоволеності виборців вимірює, наскільки добре правило голосування вибирає соціально оптимального кандидата:

Визначення 2.4 (VSE). Для корисностей $U \in \mathbb{R}^{n\times M}$ та переможця w:

$$\text{VSE} = \frac{\bar{u}_w - \bar{u}_{\min}}{\bar{u}_{\max} - \bar{u}_{\min}} \tag{13}$$

де $\bar{u}_c = \frac{1}{n}\sum_{v=1}^{n} u_{v,c}$ — середня корисність для кандидата c.

VSE коливається від 0 (переможець — найгірший кандидат) до 1 (переможець — оптимальний кандидат).

2.6.1 Порівняльна продуктивність за правилами

Таблиця 5 показує значення VSE для різних правил голосування за вимірами:

Табл. 5: VSE за виміром (L2-Косинусна, гетерогенна).

Вимір	Відносна більшість	Борда	IRV
1	0.606	0.966	0.581
2	0.821	0.989	0.584
3	0.834	0.994	0.528
4	0.848	0.990	0.536
5	0.883	0.992	0.530
7	0.887	0.995	0.581
10	0.909	0.993	0.633

Ключові спостереження:

1. Борда послідовно перевершує: VSE ≥ 0.99 для всіх вимірів, демонструючи виняткову продуктивність соціального добробуту незалежно від вимірності.
2. Відносна більшість показує вимірне покращення: VSE зростає від 0.606 (вимір 1) до 0.909 (вимір 10), вказуючи, що вищовимірні простори дозволяють відносній більшості краще ідентифікувати соціально оптимальних кандидатів.
3. Продуктивність IRV залежить від виміру: VSE коливається від 0.528 (вимір 3) до 0.633 (вимір 10), показуючи слабшу продуктивність, ніж Борда, але покращення з виміром.
4. Вимірне покращення: Всі правила показують зростаючу або стабільну VSE зі збільшенням виміру, вказуючи, що високовимірні простори природно сприяють кращим соціальним виборам, хоча величина покращення варіюється за правилом.

2.6.2 Ефекти гетерогенності метрик на VSE

Для пар L2-Косинусна різниці VSE між гетерогенними та однорідними (базова лінія метрики центру) умовами малі для всіх вимірів та правил, зазвичай в межах

± 0.01. Це вказує, що хоча гетерогенність метрик може драматично впливати на який кандидат виграє (вимірюється рівнями розбіжностей), вона має мінімальний вплив на якість переможця, виміряну соціальним добробутом.

Однак аналіз декомпозиції виявляє важливий нюанс: коли Косинусна є метрикою центру, гетерогенне правило дає результати, які сильно вирівнюються з косинусно-однорідними перевагами (48.5% екстремально-вирівняної розбіжності для Косинусна→L2 Борда). Це вирівнювання може пояснити, чому різниці VSE залишаються малими — гетерогенне правило ефективно вибирає переможців, які оптимізують косинусно-базований соціальний добробут, який може бути подібним до L2-базованого соціального добробуту в багатьох профілях.

3 Висновки

3.1 Резюме: Гетерогенність метрик та чутливість правил голосування

Наш математичний та обчислювальний аналіз виявляє, що ефекти гетерогенності метрик є залежними від правил та напрямковими, зі значними впливами навіть у високовимірних просторах:

1. Стійкі ефекти гетерогенності: Рівні розбіжностей 4–18.5% зберігаються для всіх протестованих вимірів (1–10), суперечивши гіпотезі повної вимірної збіжності для сценаріїв гетерогенних метрик.

2. Драматична взаємодія Борда-Косинусна: Коли Косинусна є метрикою центру за методу Борда, розбіжність досягає 60% (проти 8.5% для L2-центру), представляючи асиметрію 51.5 п.п. Це найбільший напрямковий ефект, що спостерігається.

3. Декомпозиція виявляє механізми: Сильна розбіжність (гетерогенна відрізняється від обох базових ліній) коливається від 10–93% загальної розбіжності, залежно від правила та конфігурації. Пропорція варіюється драматично за напрямком: L2→Косинусна Борда показує 58.8% сильних розбіжностей (нові результати), тоді як Косинусна→L2 Борда показує лише 19.2% сильних розбіжностей (переважно посилення). Ця напрямкова асиметрія в механізмі така ж важлива, як і асиметрія за величиною.

4. Специфічні для правил чутливості: Борда високо чутлива до напрямку призначення метрики; IRV показує зворотні патерни асиметрії; відносна більшість залишається відносно симетричною.

Ці результати мають глибокі імплікації для демократичної теорії. Замість розгляду гетерогенності метрик як спотворення, ми повинні визнати її як фундаментальний аспект різноманітності виборців, який складним чином взаємодіє з механізмами агрегації.

3.2 Практичні рекомендації

На основі нашого аналізу ми рекомендуємо:

1. Для систем методу Борда:
   - Бути обізнаними про екстремальну чутливість до призначення метрики. Конфігурація косинус-центру може дати 60% розбіжності проти 8.5% для L2-центру.
   - Якщо використовувати гетерогенні метрики, ретельно розглянути, які групи виборців отримують які метрики.
   - Борда дає найвищу пропорцію сильних розбіжностей (справді нові результати).
2. Для низьковимірних просторів ($d \leq 3$):
   - Ефекти гетерогенності метрик найсильніші та найбільш залежні від правил.
   - Пари Косинусна-Мінковський показують напрямкові асиметрії 50+ п.п. для Борда.
   - Розглянути пари L1/L2 для симетричної поведінки (15–22% розбіжності, мінімальні напрямкові ефекти).
3. Для високовимірних просторів ($d \geq 5$):
   - Ефекти гетерогенності зберігаються (8–18% розбіжності), але без чітких монотонних тенденцій.
   - Вибір метрики все ще має значення; вимірна збіжність НЕ усуває ефекти гетерогенності.
   - Вибір правила залишається критичним навіть у високих вимірах.
4. Для аналізу з обліком декомпозиції:
   - Відстежувати сильні проти екстремально-вирівняних розбіжностей для розуміння як гетерогенність впливає на результати.
   - Висока загальна розбіжність з низькою сильною розбіжністю вказує на зсув вирівнювання, а не нові результати.
   - Використовувати декомпозицію для діагностики, чи вводить гетерогенність справді нових переможців, або лише сприяє різним базовим лініям.

3.3 Напрямки майбутніх досліджень

Залишається кілька відкритих питань:

1. Теоретичне розуміння взаємодії Борда-Косинусна: Чому косинус-центр Борда дає 7-кратну вищу розбіжність, ніж L2-центр, і чому 80.8% цієї розбіжності обумовлено екстремальним вирівнюванням, а не новими результатами? Чи можемо ми вивести аналітичні межі цієї асиметрії та пояснити механізм, через який косинусна відстань посилює свої власні переваги під агрегацією Борда?
2. Декомпозиція у вищих вимірах: Як еволюціонують пропорції сильних проти екстремально-вирівняних за вимір 10? Чи є кінцева збіжність?
3. Нерівномірні геометрії: Як поляризовані або кластеризовані розподіли виборців впливають на патерни декомпозиції та асиметрії?
4. Альтернативні функції корисності: Гаусові або квадратичні корисності можуть показувати різні патерни взаємодії між метриками та правилами голосування.
5. Стратегічне голосування: Як гетерогенність метрик взаємодіє зі стратегічною поведінкою? Чи можуть виборці експлуатувати структури декомпозиції?
6. Емпірична валідація: Чи можемо ми вивести гетерогенність метрик з реальних електоральних даних? Чи відповідають спостережені патерни результатів нашим передбаченням декомпозиції?
7. Інші правила голосування: Як схвальне голосування, STAR голосування та інші методи взаємодіють з гетерогенністю метрик та декомпозицією?

Подяки

Це дослідження було проведено з використанням обчислювальних симуляцій, реалізованих у Python. Весь код та дані доступні для перевірки відтворюваності.

Література

[1] Downs, A. (1957). An Economic Theory of Democracy. Harper & Row.

[2] Black, D. (1958). The Theory of Committees and Elections. Cambridge University Press.

[3] Arrow, K. J. (1963). Social Choice and Individual Values (2nd ed.). Yale University Press.

[4] Enelow, J. M., & Hinich, M. J. (1984). The Spatial Theory of Voting: An Introduction. Cambridge University Press.

[5] Merrill, S., & Grofman, B. (2005). A Unified Theory of Voting: Directional and Proximity Spatial Models. Cambridge University Press.

[6] Ledoux, M. (2001). The Concentration of Measure Phenomenon. American Mathematical Society.

[7] Talagrand, M. (1995). Concentration of measure and isoperimetric inequalities in product spaces. Publications Mathématiques de l'Institut des Hautes Études Scientifiques, 81(1), 73–205.