INSTITUT DE FRANCE
Académie des sciences

Comptes Rendus
Mathématique

Vishesh Jain, Huy Pham and Thuy-Duong Vuong

Dimension reduction for maximum matchings and the Fastest Mixing Markov Chain

Volume 361 (2023), p. 869-876

Published online: 18 July 2023

https://doi.org/10.5802/crmath.447

Les Comptes Rendus. Mathématique *sont membres du*
Centre Mersenne pour l'édition scientifique ouverte
www.centre-mersenne.org
e-ISSN : 1778-3569

Comptes Rendus
Mathématique
2023, Vol. 361, p. 869-876
https://doi.org/10.5802/crmath.447

Combinatorics, Probability theory / *Combinatoire, Probabilités*

Dimension reduction for maximum matchings and the Fastest Mixing Markov Chain

Vishesh Jain[*, a], Huy Pham[b] and Thuy-Duong Vuong[c]

[a] Department of Mathematics, Statistics, and Computer Science, University of Illinois Chicago

[b] Department of Mathematics, Stanford University

[c] Department of Computer Science, Stanford University

E-mails: visheshj@uic.edu, huypham@stanford.edu, tdvuong@stanford.edu

Abstract. Let $G = (V,E)$ be an undirected graph with maximum degree Δ and vertex conductance $\Psi^*(G)$. We show that there exists a symmetric, stochastic matrix P, with off-diagonal entries supported on E, whose spectral gap $\gamma^*(P)$ satisfies

$$\Psi^*(G)^2 / \log\Delta \lesssim \gamma^*(P) \lesssim \Psi^*(G).$$

Our bound is optimal under the Small Set Expansion Hypothesis, and answers a question of Olesker-Taylor and Zanetti, who obtained such a result with $\log\Delta$ replaced by $\log|V|$.

In order to obtain our result, we show how to embed a negative-type semi-metric d defined on V into a negative-type semi-metric d' supported in $\mathbb{R}^{O(\log\Delta)}$, such that the (fractional) matching number of the weighted graph (V,E,d) is approximately equal to that of (V,E,d').

Manuscript received 9 March 2022, revised 21 July 2022 and 26 August 2022, accepted 22 November 2022.

1. Introduction

Let $G = (V,E,w)$ be a simple, undirected graph with non-negative weights $w : E \to \mathbb{R}_{\geq 0}$. In this note, we consider weights w that can be written in the form $w(u,v) = \|f(u) - f(v)\|^q$ for some function $f : V \to \mathbb{R}^{|V|}$ and $q \geq 1$, where $\|\cdot\|$ denotes the usual Euclidean or ℓ_2 norm. That is, w can be obtained as the restriction to the edges of G of a function $\widetilde{w} : V \times V \to \mathbb{R}_{\geq 0}$, where $\widetilde{w}^{1/q}$ embeds isometrically into Euclidean space for some finite $q \geq 1$. The special case $q = 2$ is known as a metric of negative-type and has been extensively studied in connection with semi-definite relaxations of combinatorial optimization problems [6]; in particular, it is known that all finite ℓ_p metric spaces, for $1 \leq p \leq 2$, embed isometrically into a metric space of negative-type [5].

Recall that a fractional matching of a weighted graph $G = (V,E,w)$ is a function $h : E \to [0,1]$ such that for all $v \in V$, $\sum_{e : v \in e} h(e) \leq 1$. A matching corresponds to maps $h : E \to \{0,1\}$.

* Corresponding author.

Let M (respectively, M_{frac}) denote a maximum weight matching (respectively, fractional matching) of (G, w). Here, the weight of a (fractional) matching $h : E \to [0,1]$ is defined by

$$w(h) := \sum_{e \in E} h(e) w(e).$$

The weight of a maximum weight (fractional) matching of (G, w) is called the (fractional) matching number of (G, w).

It follows from Linear Programming (LP) duality that the fractional matching number of (G, w) is equal to the value of the following linear program:

$$\begin{aligned} \min_{g: V \to \mathbb{R}_{\geq 0}} \quad & \sum_{u \in V} g(u) \\ \text{s.t.} \quad & g(u) + g(v) \geq w(u,v) \quad \forall \{u,v\} \in E \end{aligned} \tag{1.1}$$

Motivated by the problem of the Fastest Mixing Markov Chain (see Section 1.1), we consider the following natural dimension reduction question: what is the minimum dimension of a Euclidean space $\mathbb{R}^m$ for which there is a function $f' : V \to \mathbb{R}^m$ such that $(G, w') = (V, E, w')$, with $w'(u, v) = \|f'(u) - f'(v)\|^q$, has approximately the same (fractional) matching number as $(G, w) = (V, E, w)$?

A direct application of the classical Johnson-Lindenstrauss lemma shows that $m = O_q(\log n/\varepsilon^2)$ suffices to preserve the (fractional) matching number to within a multiplicative factor of $(1 \pm \varepsilon)$. Our main result completely removes the dependence on the number of vertices and provides an essentially optimal bound depending only on the maximum degree of the graph.

Theorem 1 (Informal, see Theorem 7 for a precise version). *With notation as above, for any $q \geq 1$ and any weighted graph $(G, w) = (V, E, w)$ with maximum degree at most Δ, there exists a function $f' : V \to \mathbb{R}^m$, with $m = O(\log(\Delta/\varepsilon q))/\varepsilon^2$, such that the (fractional) matching number of $(G, w') = (V, E, w')$, with $w'(u, v) = \|f'(u) - f'(v)\|^q$, is within a factor of $(1 \pm \varepsilon)$ from the (fractional) matching number of (G, w).*

Remark. If we also demand that $w_\pi(E) \geq e^{-\varepsilon q} w(E)$ (which holds with high probability in our setting; see Section 2) then the factor of $\log(\Delta)$ in our bound on d cannot be improved under the Small Set Expansion Hypothesis (SSE) of Raghavendra and Steurer [12], the reason being that an improved dependence on Δ in Theorem 7 translates to a corresponding improved dependence on Δ in Theorem 4, which contradicts SSE (see the remark after Theorem 4).

Theorem 1 is analogous to, and motivated by, a growing body of work (see, e.g., the recent work [4] and the references therein) showing that for various geometric optimization problems in Euclidean space such as k-means clustering or low-rank approximation, the optimum cost and the optimizer are approximately preserved, even under an embedding of the problem into much fewer dimensions than a naïve application of the Johnson-Lindenstrauss lemma would suggest. Our result shows that a similar phenomenon holds for the classical combinatorial optimization problem of maximum matching.

Although Theorem 1 suffices for the application in Theorem 4, it raises the intriguing question of whether a stronger statement is true, namely that an embedding into $O_{q,\varepsilon}(\log(\Delta))$ dimensions is sufficient to preserve *all* edge weights within a multiplicative factor of $(1 \pm \varepsilon)$. We do not know whether this is true (although see the related work on local dimension reduction [1] and the references therein); however, we remark that if this were true, then the embedding must necessarily be non-linear[1]. In contrast, by using combinatorial information about maximum

[1] this follows, for instance, by taking the underlying graph to be a matching, mapping the left end point of each edge to 0, and mapping the right end point of the edges to $x_i - x_j$, where $x_1, \ldots, x_n$ is a set of points for which the Johnson-Lindenstrauss lemma is tight up to a constant [8]. We thank Ashwin Sah for this example.

matchings, we are able to show that the embedding in Theorem 1 can be simply taken to be a (suitably rescaled) random matrix with i.i.d. sub-Gaussian entries of mean 0 and variance 1.

1.1. *Fastest Mixing Markov Chain (FMMC)*

In the Fastest Mixing Markov Chain (FMMC) problem, introduced by Boyd, Diaconis, and Xiao [3], we are given a finite, undirected graph $G = (V, E)$ and are asked to design a symmetric $|V| \times |V|$ matrix P such that

(i) $P_{ij} \geq 0$ for all i, j and $P_{ij} > 0$ only if $\{i, j\} \in E$ or $i = j$;
(ii) $\sum_j P_{ij} = 1$ for all i; and
(iii) denoting the eigenvalues of P in decreasing order by $\lambda_1(P) \geq \lambda_2(P) \geq \cdots \geq \lambda_n(P)$, the second largest eigenvalue modulus (SLEM), defined by
$$\mu(P) := \max_{i=2,\dots,n} |\lambda_i(P)|,$$
is as small as possible.

We denote the set of all symmetric $V \times V$ matrices, satisfying properties (i), (ii), (iii) by $\mathscr{M}(G)$.

In words, we are asked to design a discrete-time, time-homogeneous Markov chain with state space V, whose transitions are supported on the edges of E and which is reversible with respect to the uniform distribution on V, with the smallest possible SLEM among all such Markov chains. Since $\mu(P) \leq 1$, this is equivalent to asking for the largest possible spectral gap, defined by
$$\gamma(P) := 1 - \mu(P).$$

Definition 2 (Optimal spectral gap). *Let $G = (V, E)$ be a finite, undirected graph. With notation as above, we define the optimal spectral gap (among reversible chains supported on E) by*
$$\gamma^*(G) := \sup\{\gamma(P) : P \in \mathscr{M}(G)\}.$$

For an extensive discussion of the history of this problem, we refer the reader to [11, Section 1.4], limiting ourselves here to only the most relevant results. All of these results involve the vertex conductance of a graph, whose definition we now recall.

Definition 3 (Vertex conductance). *The vertex conductance $\Psi^*(G)$ of a graph $G = (V, E)$ is defined as*
$$\Psi^*(G) := \min_{\emptyset \neq S \subseteq V : |S| \leq |V|/2} \frac{|\{v \notin S : \exists\, u \in S \ s.t.\ \{u, v\} \in E\}|}{|S|}$$

Roch [13] showed that $\gamma^*(G) \lesssim \Psi^*(G)$, thereby identifying the vertex conductance of the host graph as a barrier to designing Markov chains with large spectral gaps. In the other direction, Cheeger's inequality shows that $\gamma^*(G) \gtrsim \Psi^*(G)^2/\Delta^2$, where Δ denotes the maximum degree of G. Put together, we have
$$\Psi^*(G)^2/\Delta^2 \lesssim \gamma^*(G) \lesssim \Psi^*(G) \tag{1.2}$$
Simple examples (see [11]) show that the quadratic dependence on $\Psi^*(G)$ in the lower bound and the linear dependence on $\Psi^*(G)$ in the upper bound cannot be improved in general, leaving open the question of resolving the dependence on Δ (or removing it altogether).

Recently, Olesker-Taylor and Zanetti [11] provided a new lower bound on $\gamma^*(G)$, where the quadratic dependence on Δ is replaced by logarithmic dependence on $|V|$:
$$\Psi^*(G)^2/\log|V| \lesssim \gamma^*(G) \lesssim \Psi^*(G), \tag{1.3}$$
thereby providing an exponential improvement over (1.2) for dense graphs. However, in general, the lower bounds in (1.2) and (1.3) are incomparable, and Olesker-Taylor and Zanetti asked if $\log|V|$ in (1.3) may be replaced by $\log\Delta$, which would provide a common refinement of (1.2)

and (1.3) and moreover, would be tight under the Small Set Expansion Hypothesis [12] (see the remark after Theorem 4). As an application of Theorem 7, we answer this question in the affirmative.

Theorem 4. *Let $G = (V, E)$ be a graph of maximum degree Δ. Then, $\gamma^*(G)$ satisfies*

$$\Psi^*(G)^2/\log\Delta \lesssim \gamma^*(G) \lesssim \Psi^*(G)$$

Remark. Work of Louis, Raghavendra, and Vempala [9] shows that under the Small Set Expansion Hypothesis [12], for all sufficiently small $\varepsilon > 0$, there is no polynomial-time algorithm to distinguish between $\Psi^*(G) \le \varepsilon$ and $\Psi^*(G) \gtrsim \sqrt{\varepsilon\log\Delta}$ for graphs with maximum degree Δ satisfying $\log\Delta \asymp 1/\varepsilon$. As noted in [11], a lower bound of the form $\Psi^*(G)^2/o(\log\Delta) \lesssim \gamma^*(G)$ would allow us to distinguish between the two scenarios (using $\gamma^*(G) \le \varepsilon$ versus $\gamma^*(G) = \omega(\varepsilon)$), which would contradict the Small Set Expansion Hypothesis since $\gamma^*(G)$ can be computed in time polynomial in the size of the graph.

Example 5. A popular use of Markov Chain Monte Carlo methods is to sample from the uniform distribution on an exponentially sized subset V of a product space $\{1,\dots,r\}^n$ (where $r \asymp 1$ and n is large) using "local chains". In our notation, the transitions of the graph are supported on $G = (V, E)$, where E connects vertices within Hamming distance s, where $s \asymp 1$. Thus, $\Delta \asymp \log|V| \asymp n$. Therefore, (1.2) gives the lower bound $\gamma^*(G) \gtrsim \Psi^*(G)^2/n^2$, (1.3) gives the lower bound $\gamma^*(G) \gtrsim \Psi^*(G)^2/n$, whereas Theorem 4 gives the lower bound $\gamma^*(G) \gtrsim \Psi^*(G)^2/\log n$. In examples supporting "rapidly mixing Markov chains", $\Psi^*(G) = \Omega(1/\operatorname{poly}\log|V|) = \Omega(1/\operatorname{poly}(n))$, so that the improvement from Theorem 4, compared to (1.3), is of order polynomial in the vertex conductance.

Example 6. We now illustrate that, even in very simple examples, the improvement coming from Theorem 4 over Cheeger's inequality or (1.3) is substantial, and can possibly essentially capture the truth. Let $G = (V, E)$ denote the standard hypercube graph, with vertex set $V = \{0,1\}^n$ and edges E connecting vertices which differ in exactly one coordinate. It is well-known that $\Psi^*(G) \asymp \frac{1}{\sqrt{n}}$, whereas the edge conductance

$$\Phi^*(G) := \min_{\emptyset\neq S\subseteq V:|S|\le|V|/2} \frac{|\{\{u,v\}\in E : u\in S, v\notin S\}|}{n|S|}$$

satisfies $\Phi^*(G) \asymp \frac{1}{n}$. Moreover, the spectral gap $\gamma(G)$ of the lazy random walk on the hypercube satisfies $\gamma(G) \asymp \frac{1}{n} \asymp \gamma^*(G)$ [2]. Cheeger's inequality implies that $\frac{1}{n^2} \lesssim \gamma(G) \lesssim \frac{1}{n}$ whereas (1.3) implies that $\frac{1}{n^2} \lesssim \gamma^*(G) \lesssim \frac{1}{\sqrt{n}}$; in both cases, the lower bound on the spectral gap is a factor of $1/n$ worse than the truth. On the other hand, Theorem 4 implies that

$$\frac{1}{n\log n} \lesssim \gamma^*(G) \lesssim \frac{1}{\sqrt{n}},$$

so that the lower bound on the spectral gap is within only a logarithmic factor of the truth.

1.2. *Concurrent work*

Shortly after the appearance of our manuscript on the arXiv, we were informed of upcoming work of Kwok, Lau, and Tung [7], which proves Theorem 4 using a different proof technique (in particular, Theorem 7 does not appear in [7]), as well as a weighted version of Theorem 4, which is not considered in our work.

1.3. *Acknowledgements*

We thank Sidhanth Mohanty for bringing several typos to our attention as well as anonymous referees for several suggestions, which have improved the presentation.

2. Proof of Theorem 7

In this section, we prove Theorem 1 in the following precise form.

Theorem 7. *For every $K > 0$, there exists a constant $C_7 = C_7(K) \ge 1$ for which the following holds. Let $G = (V, E, w)$ with $|V| = n$ and assume that w can be written in the form $w = \widetilde{w}|_E$, where $\widetilde{w}(u, v) = \|f(u) - f(v)\|^q$ for some $f : V \to \mathbb{R}^n$ and $q \ge 1$. Let Δ denote the maximum degree of G. Let M (respectively, M_{frac}) denote a maximum weight matching (respectively, fractional matching) of (G, w).*

For $\varepsilon \in (0, 1/10)$ and $d \ge C_7 q \log(\Delta/\varepsilon q)/\varepsilon^2$, let $\Pi_{n,d}$ be a $d \times n$ random matrix whose entries are i.i.d. copies of a centered random variable with variance $1/d$ and sub-Gaussian norm at most $K/\sqrt{d}$. For $u, v \in V$, let $\widetilde{w}_{\Pi_{n,d}}(u, v) := \|\Pi_{n,d} f(u) - \Pi_{n,d} f(v)\|^q$. Let $M^{\Pi_{n,d}}$ (respectively, $M_{\text{frac}}^{\Pi_{n,d}}$) denote a maximum weight matching (respectively, fractional matching) of $(G, w_{\Pi_{n,d}})$.

Then, except with probability at most $\exp(-d\varepsilon^2/\sqrt{C_7})$ over the realisation $\pi \sim \Pi_{n,d}$, we have

$$e^{-\varepsilon q} w(M) \le w_\pi(M^\pi) \le e^{\varepsilon q} w(M),$$
$$e^{-\varepsilon q} w(M_{\text{frac}}) \le w_\pi(M^\pi_{\text{frac}}) \le e^{\varepsilon q} w(M_{\text{frac}}).$$

We begin by recalling some preliminary notions related to dimension reduction.

Definition 8 (cf. [10, Definition 2.1]). *Let $\Pi_{n,d}$ denote a distribution on linear maps from $\mathbb{R}^n \to \mathbb{R}^d$. For constants $q \ge 1$ and $c, C > 0$, we say that $\Pi_{n,d}$ is a (q, c, C)-good random dimension reduction if for all $1/10 > \varepsilon > \frac{C\sqrt{q}}{\sqrt{d}}$, there exist $0 < \delta, \rho < \exp(-c\varepsilon^2 d)$ satisfying the following properties.*

1. *For all $x, y \in \mathbb{R}^n$, $\mathbb{P}_{\pi\sim\Pi_{n,d}}[\|\pi(x) - \pi(y)\| \notin e^{\pm\varepsilon}\|x - y\|] \le \delta$.*
2. *For all $x, y \in \mathbb{R}^n$, letting $\mathscr{E}_{x,y}$ be the event that $\|\pi(x) - \pi(y)\| \ge e^{\varepsilon}\|x - y\|$, we have*

$$\mathbb{E}_{\pi\sim\Pi_{n,d}}\left[1_{\mathscr{E}_{x,y}}\left(\frac{\|\pi(x) - \pi(y)\|^q}{\|x - y\|^q} - e^{\varepsilon q}\right)\right] \le \rho.$$

Lemma 9 (see, e.g., [10, Lemma C.1]). *For every $K_9 > 0$, there exist constants $c_9, C_9 > 0$ for which the following holds. Let $\Pi_{n,d}$ denote a $d \times n$ random matrix whose entries are i.i.d. copies of a centered random variable with variance $1/d$ and sub-Gaussian norm at most $K_9/\sqrt{d}$. Then, $\Pi_{n,d}$ is a (q, c_9, C_9)-good random dimension reduction for all $q \ge 1$.*

We are now ready to prove Theorem 7. Below, we use the notation introduced in the statement of Theorem 7. For later use, we will prove a slightly stronger statement, which shows that in addition to the conclusion of Theorem 7, we also have that

$$\sum_{u,v\in V} \widetilde{w}_\pi(u, v) \ge e^{-\varepsilon q} \sum_{u,v\in V} w(u, v).$$

Proof of Theorem 7. We start by giving a high-level overview of the proof. It is quite immediate to show that the weight of any fixed matching (in particular, a maximum matching) is preserved with high probability under the passage from w to w_π. The challenge, then, is to show that there are no unexpectedly large matchings (in terms of additive error) corresponding to w_π (note that we cannot use simple union bound due to the extremely large number of matchings). To address this, we consider the (random) set of edges which have unexpectedly high weights under w_π and show that, with high probability, the total w_π-weight of this random set of edges is at most $\varepsilon/\operatorname{poly}(\Delta)$ times the total w-weight of all edges and hence, at most $\varepsilon/\operatorname{poly}(\Delta)$ times the w-matching number of the graph. Thus, while it is possible for the w_π-weight of a matching to be unexpectedly larger (in a multiplicative sense) than its w-weight, the magnitude of this increase is much smaller than the w-matching number of the graph and so, cannot significantly alter the w_π-matching number. We now proceed to the formal details.

By Lemma 9, $\Pi_{n,d}$ is a (q, c', C')-good random dimension reduction, where $c', C' > 0$ depend only on K. Let $\pi \sim \Pi_{n,d}$, and consider the following random subsets of $\binom{V}{2}$, corresponding to heavy edges, light edges, and light pairs.

$$\begin{aligned}
\mathcal{H} &= \{\{u,v\} \in E : w_\pi(u,v) \ge e^{\varepsilon q} w(u,v)\},\\
\mathcal{L}_1 &= \left\{\{u,v\} \in E : w_\pi(u,v) \le e^{-\varepsilon q} w(u,v)\right\},\\
\mathcal{L}_2 &= \left\{\{u,v\} \in \binom{V}{2} : w_\pi(u,v) \le e^{-\varepsilon q} w(u,v)\right\}.
\end{aligned}$$

These random subsets naturally give rise to the following quantities corresponding to the excess cost of heavy edges, the total original weight of light edges, and the total original weight of light pairs:

$$\begin{aligned}
\mathrm{Diff}(\mathcal{H}) &:= \sum_{\{u,v\}\in\mathcal{H}} \left(w_\pi(u,v) - e^{\varepsilon q} w(u,v)\right);\\
\mathrm{Cost}(\mathcal{L}_1) &:= \sum_{\{u,v\}\in\mathcal{L}_1} w(u,v);\\
\mathrm{Cost}(\mathcal{L}_2) &:= \sum_{\{u,v\}\in\mathcal{L}_2} \widetilde{w}(u,v).
\end{aligned}$$

We show that, in expectation, all of these quantities are sufficiently small. First, since any graph of maximum degree Δ can be written as a disjoint union of at most $(\Delta+1)$ matchings, it follows that $w(M) \ge \frac{1}{\Delta+1}\sum_{\{u,v\}\in E} w(u,v)$. Then, from Lemma 9 and Definition 8, it immediately follows that

$$\begin{aligned}
\mathbb{E}_\pi[\mathrm{Diff}(\mathcal{H})] &\le \rho \sum_{\{u,v\}\in E} w(u,v) \le \rho(\Delta+1)\, w(M),\\
\mathbb{E}_\pi[\mathrm{Cost}(\mathcal{L}_1)] &\le \delta \sum_{\{u,v\}\in E} w(u,v) \le \delta(\Delta+1)\, w(M),\\
\mathbb{E}_\pi[\mathrm{Cost}(\mathcal{L}_2)] &\le \delta \sum_{\{u,v\}\in\binom{V}{2}} \widetilde{w}(u,v) = \delta\, \widetilde{w}\left(\binom{V}{2}\right).
\end{aligned}$$

Therefore, by Markov's inequality, the union bound, and using $\max\{\delta,\rho\} \le \exp(-c'\varepsilon^2 d)$, we see that except with probability at most $3\exp(-c'\varepsilon^2 d/2)$, the following event holds:

$$\mathcal{G} = \left\{\mathrm{Diff}(\mathcal{H}) \le \sqrt{\rho}(\Delta+1)\, w(M), \mathrm{Cost}(\mathcal{L}_1) \le \sqrt{\delta}(\Delta+1)\, w(M), \mathrm{Cost}(\mathcal{L}_2) \le \sqrt{\delta}\, \widetilde{w}\binom{V}{2}\right\}.$$

Let π be a realisation of $\Pi_{n,d}$ for which $\mathcal{G}$ holds. Then, for any (fractional) matching M', we have

$$\begin{aligned}
w_\pi(M') = \sum_e M'(e)\, w_\pi(e) &= \sum_{e\notin\mathcal{H}} M'(e)\, w_\pi(e) + \sum_{e\in\mathcal{H}} M'(e) e^{\varepsilon q} w(e) + \sum_{e\in\mathcal{H}} M'(e)(w_\pi(e) - e^{\varepsilon q} w(e))\\
&\le e^{\varepsilon q} w(M') + \mathrm{Diff}(\mathcal{H}) \le e^{\varepsilon q} w(M') + \sqrt{\rho}(\Delta+1)\, w(M)\\
&\le (e^{\varepsilon q} + \sqrt{\rho}(\Delta+1))\, w(M).
\end{aligned}$$

Moreover, we see that for $M' = M$ or $M' = M_{\mathrm{frac}}$, we have

$$\begin{aligned}
w_\pi(M') &= \sum_{e\notin\mathcal{L}_1} M'(e)\, w_\pi(e) + \sum_{e\in\mathcal{L}_1} M'(e)\, w_e + \sum_{e\in\mathcal{L}_1} M'(e)(w_\pi(e) - w(e))\\
&\ge e^{-\varepsilon q} w(M') - \mathrm{Cost}(\mathcal{L}_1)\\
&\ge (e^{-\varepsilon q} - \sqrt{\delta}(\Delta+1))\, w(M').
\end{aligned}$$

A similar computation shows that

$$\widetilde{w}_\pi\left(\binom{V}{2}\right) \geq (e^{-\varepsilon q} - \sqrt{\delta})\,\widetilde{w}\left(\binom{V}{2}\right).$$

Finally, by taking C to be sufficiently large depending on c', C', we can ensure that

$$\max\{\sqrt{\delta}, \sqrt{\rho}\}(\Delta+1) \leq \varepsilon q/10,$$

so that the desired conclusion follows by rescaling ε. □

3. Application to the Fastest Mixing Markov Chain: Proof of Theorem 4

We will need the following proposition due to Olesker-Taylor and Zanetti.

Proposition 10 ([11]). *Let $G = (V, E)$ be a simple, undirected graph with $|V| = n$, vertex conductance $\Psi^*(G)$, and optimal spectral gap $\gamma^*(G)$. For $1 \leq m \leq n$, let λ_m^* denote the optimum value of the following problem:*

$$\min_{f:V\to\mathbb{R}^m, g:V\to\mathbb{R}_{\geq 0}} \frac{\sum_{u\in V} g(u)}{\sum_{u\in V}\|f(u)\|^2}$$
$$\text{s.t.} \quad \sum_{v\in V} f(v) = 0, \quad \text{and} \quad g(u)+g(v) \geq \|f(u)-f(v)\|^2 \quad \forall\,\{u,v\}\in E$$

Let $m^* = \min\{1 \leq m \leq n : \lambda_n^* \leq \lambda_m^* \leq 2\lambda_n^*\}$. Then,

$$\Psi^*(G)^2/m^* \lesssim \gamma^*(G) \lesssim \Psi^*(G).$$

Given the previous proposition, Theorem 4 follows from Theorem 7 and LP duality (1.1).

Proof of Theorem 4. By Proposition 10, and since λ_m^* decreases as m increases, it suffices to show that for $d = O(\log\Delta)$, $\lambda_d^* \leq 2\lambda_n^*$. To this end, let $F : V \to \mathbb{R}^n$ be such that $\sum_{v\in V} F(v) = 0$, and that the value of the program

$$\min_{g:V\to\mathbb{R}_{\geq 0}} \frac{\sum_{u\in V} g(u)}{\sum_{u\in V}\|F(u)\|^2} \qquad \text{s.t.} \quad g(u)+g(v) \geq \|F(u)-F(v)\|^2 \quad \forall\,\{u,v\}\in E$$

is λ_n^*.

By LP duality (1.1), λ_n^* is the fractional matching number of the graph $G = (V, E, w/S_F)$, where $w(u,v) := \|F(u)-F(v)\|^2$ and $S_F := \sum_{u,v\in V}\|F(u)-F(v)\|^2$. Here, we have used that $\sum_{u,v\in V}\|F(u)-F(v)\|^2 = \sum_{u\in V}\|F(u)\|^2$, which holds since $\sum_{v\in V} F(v) = 0$.

We now appeal to the slight extension of Theorem 7 (with $K = 1$) proved in the previous section. Let $\varepsilon = 1/100$, $d = 2C_7(1)\log(\Delta/2\varepsilon)/\varepsilon^2 = O(\log\Delta)$, and $\pi \in \mathscr{G}$. Here $\mathscr{G}$ is the event defined in the previous subsection, where it was also shown to be nonempty. Let $f : V \to \mathbb{R}^d$ be defined by $f(v) = \pi(F(v))$. Then $\sum_{v\in V} f(v) = 0$ and hence, λ_d^* is at most

$$\min_{g:V\to\mathbb{R}_{\geq 0}} \frac{\sum_{u\in V} g(u)}{\sum_{u,v\in V}\|f(u)-f(v)\|^2} \qquad \text{s.t.} \quad g(u)+g(v) \geq \|f(u)-f(v)\|^2 \quad \forall\,\{u,v\}\in E.$$

Once again, by LP duality (1.1), λ_d^* is at most the fractional matching number of the graph $G = (V, E, w'/S'_f)$, where $w'(u,v) := \|f(u)-f(v)\|^2$ and $S'_f := \sum_{u,v\in V}\|f(u)-f(v)\|^2$.

In the previous section, we showed that for $\pi \in \mathscr{G}$, $S'_f \geq e^{-2\varepsilon} S_F$ and the fractional matching number of (V, E, w') is at most $e^{2\varepsilon}$ times the fractional matching number of (V, E, w). By the choice of ε, this shows that $\lambda_d^* \leq 2\lambda_n^*$, as desired. □

References

[1] Y. Bartal, B. Recht, L. J. Schulman, "Dimensionality reduction: beyond the Johnson-Lindenstrauss bound", in *Proceedings of the twenty-second annual ACM-SIAM symposium on Discrete Algorithms*, SIAM, 2011, p. 868-887.

[2] S. Boyd, P. Diaconis, P. Parrilo, L. Xiao, "Fastest mixing Markov chain on graphs with symmetries", *SIAM J. Optim.* **20** (2009), no. 2, p. 792-819.

[3] S. Boyd, P. Diaconis, L. Xiao, "Fastest mixing Markov chain on a graph", *SIAM Rev.* **46** (2004), no. 4, p. 667-689.

[4] M. Charikar, E. Waingarten, "The Johnson–Lindenstrauss Lemma for Clustering and Subspace Approximation: From Coresets to Dimension Reduction", https://arxiv.org/abs/2205.00371, 2022.

[5] M. M. Deza, M. Laurent, *Geometry of cuts and metrics*, Algorithms and Combinatorics, vol. 15, Springer, 1997.

[6] M. X. Goemans, "Semidefinite programming in combinatorial optimization", *Math. Program.* **79** (1997), no. 1, p. 143-161.

[7] T. C. Kwok, L. C. Lau, K. C. Tung, "Cheeger Inequalities for Vertex Expansion and Reweighted Eigenvalues", https://arxiv.org/abs/2203.06168, 2022.

[8] K. G. Larsen, J. Nelson, "Optimality of the Johnson-Lindenstrauss lemma", in *2017 IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS)*, IEEE, 2017, p. 633-638.

[9] A. Louis, P. Raghavendra, S. Vempala, "The complexity of approximating vertex expansion", in *2013 IEEE 54th Annual Symposium on Foundations of Computer Science*, IEEE, 2013, p. 360-369.

[10] K. Makarychev, Y. Makarychev, I. Razenshteyn, "Performance of Johnson-Lindenstrauss transform for k-means and k-medians clustering", in *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, 2019, p. 1027-1038.

[11] S. Olesker-Taylor, L. Zanetti, "Geometric Bounds on the Fastest Mixing Markov Chain", in *13th Innovations in Theoretical Computer Science Conference (ITCS 2022)*, Schloss Dagstuhl-Leibniz-Zentrum für Informatik, 2022.

[12] P. Raghavendra, D. Steurer, "Graph expansion and the unique games conjecture", in *Proceedings of the forty-second ACM symposium on Theory of computing*, 2010, p. 755-764.

[13] S. Roch, "Bounding fastest mixing", *Electron. Commun. Probab.* **10** (2005), p. 282-296.