Submitted to the Annals of Statistics

SUPPLEMENT TO "OPTIMIZATION VIA LOW-RANK APPROXIMATION, WITH APPLICATIONS TO COMMUNITY DETECTION IN NETWORKS"

By Can Le, Elizaveta Levina, and Roman Vershynin

University of Michigan

In this supplement we prove results of Sections 3.1, 3.2, 3.3, and 3.4. Let us first describe the projection of the cube under regular block models, which will be used to replace Assumption (2). See Figure 3.1 for an illustration.

Lemma 6. *Consider the regular block models and let $\mathcal{R} = U_{\mathbb{E}[A]}[-1,1]^n$. Then $\mathcal{R}$ is a parallelogram; the vertices of $\mathcal{R}$ are $\{\pm U_{\mathbb{E}[A]}(c), \pm U_{\mathbb{E}[A]}(\mathbf{1})\}$, where c is a true label vector. The angle between two adjacent sides of $\mathcal{R}$ does not depend on n.*

Proof of Lemma 6. Eigenvectors of $\mathbb{E}[A]$ are computed in Lemma 3. Let

$$x = \left(r_1\left(n(\pi_1 r_1^2+\pi_2)\right)^{-1/2}, r_2\left(n(\pi_1 r_2^2+\pi_2)\right)^{-1/2}\right)^T,$$

$$y = \left(\left(n(\pi_1 r_1^2+\pi_2)\right)^{-1/2}, \left(n(\pi_1 r_2^2+\pi_2)\right)^{-1/2}\right)^T.$$

Then $\mathcal{R} = \{(\epsilon_1+\cdots+\epsilon_{\bar{n}_1})x + (\epsilon_{\bar{n}_1+1}+\cdots+\epsilon_n)y, \epsilon_i \in [-1,1]\}$, and it is easy to see that $\mathcal{R}$ is a parallelogram. Vertices of $\mathcal{R}$ correspond to the cases when $\epsilon_1 = \cdots = \epsilon_{\bar{n}_1} = \pm 1$ and $\epsilon_{\bar{n}_1+1} = \cdots = \epsilon_n = \pm 1$. The angle between two adjacent sides of $\mathcal{R}$ equals the angle between $\sqrt{n}x$ and $\sqrt{n}y$, which does not depend on n. □

5.1. *Proof of results in Section 3.1.* Under degree-corrected block models, let us denote by $\bar{A}$ the conditional expectation of A given the degree parameters $\theta = (\theta_1, ..., \theta_n)^T$. Note that if $\theta_i \equiv 1$ then $\bar{A} = \mathbb{E}A$. Since $\bar{A}$ depends on θ, its eigenvalues and eigenvectors may not have a closed form. Nevertheless, we can approximate them using ρ_i and $\bar{u}_i$ from Lemma 3. To do so, we need the following lemma.

AMS 2000 subject classifications: Primary 62E10; secondary 62G05

Keywords and phrases: Community detection, Spectral clustering, Stochastic block model, Social networks

LEMMA 7. *Let $M = \rho_1 x_1 x_1^T + \rho_2 x_2 x_2^T$, where $x_1, x_2 \in \mathbb{R}^n$, $\|x_1\| = \|x_2\| = 1$, $\rho_1 \neq 0$, and $\rho_2 \neq 0$. If $c = \langle x_1, x_2\rangle$ then the eigenvalues z_i and corresponding eigenvectors y_i of M have the following form. For $i = 1, 2$,*

$$\begin{aligned} z_i &= \frac{1}{2}\left[(\rho_1+\rho_2) + (-1)^{i-1}\sqrt{(\rho_2-\rho_1)^2 + 4\rho_1\rho_2 c^2}\right], \\ y_i &= (c\rho_1)x_1 + (z_i - \rho_1)x_2. \end{aligned}$$

If ρ_1 and ρ_2 are fixed, $\rho_1 \geq \rho_2$, and $c = o(1)$ as $n \to \infty$ then eigenvalues and eigenvectors of M have the form

$$\begin{aligned} z_1 &= \rho_1 + O(c^2), \quad z_2 = \rho_2 + O(c^2), \\ y_1 &= x_1 + O(c)x_2, \quad y_2 = x_2 + O(c)x_1. \end{aligned}$$

PROOF OF LEMMA 7. It is easy to verify that $My_i = z_i y_i$ for $i = 1, 2$. The asymptotic formulas of z_i and y_i then follow directly from the forms of z_i and y_i. □

The next lemma shows the approximation of eigenvalues and eigenvectors of $\bar{A}$.

LEMMA 8. *Consider the degree-corrected block models (described in Section 3.1) and let $D_\theta = \mathrm{diag}(\theta)$. Denote by $\bar{A}$ the conditional expectation of A given θ. Then for any $\delta \in (0,1)$, with probability at least $1-\delta$, the nonzero eigenvalues ρ_i^θ and corresponding eigenvectors $\bar{u}_i^\theta$ of $\bar{A}$ have the following form. For $i = 1, 2$,*

$$\rho_i^\theta = \rho_i \|D_\theta \bar{u}_i\|^2 \left(1 + O(1/n)\right),$$

$$\bar{u}_1^\theta = \frac{\tilde{u}_1^\theta}{\|\tilde{u}_1^\theta\|}, \quad \textit{where} \quad \tilde{u}_1^\theta = \frac{D_\theta \bar{u}_1}{\|D_\theta \bar{u}_1\|} + O\left(n^{-1/2}\right)\frac{D_\theta \bar{u}_2}{\|D_\theta \bar{u}_2\|},$$

$$\bar{u}_2^\theta = \frac{\tilde{u}_2^\theta}{\|\tilde{u}_2^\theta\|}, \quad \textit{where} \quad \tilde{u}_2^\theta = \frac{D_\theta \bar{u}_2}{\|D_\theta \bar{u}_2\|} + O\left(n^{-1/2}\right)\frac{D_\theta \bar{u}_1}{\|D_\theta \bar{u}_1\|},$$

where ρ_i, $\bar{u}_i$, and r_i are defined in Lemma 3.

PROOF OF LEMMA 8. Let $M = \rho_1 \bar{u}_1 \bar{u}_1^T + \rho_2 \bar{u}_2 \bar{u}_2^T$ be the expectation of the adjacency matrix in the regular block model setting. In the degree-corrected block model setting, given θ, we have

$$\begin{aligned} \mathbb{E}[A] &= D_\theta M D_\theta = \rho_1 D_\theta \bar{u}_1 (D_\theta \bar{u}_1)^T + \rho_2 D_\theta \bar{u}_2 (D_\theta \bar{u}_2)^T \\ &= \rho_1 \|D_\theta \bar{u}_1\|^2 \frac{D_\theta \bar{u}_1}{\|D_\theta \bar{u}_1\|}\frac{(D_\theta \bar{u}_1)^T}{\|D_\theta \bar{u}_1\|} + \rho_2 \|D_\theta \bar{u}_2\|^2 \frac{D_\theta \bar{u}_2}{\|D_\theta \bar{u}_2\|}\frac{(D_\theta \bar{u}_2)^T}{\|D_\theta \bar{u}_2\|}. \end{aligned}$$

We are now in the setting of Lemma 7 with

$$\begin{aligned} c &= \left(\|D_\theta \bar{u}_1\|\|D_\theta \bar{u}_2\|\right)^{-1}\langle D_\theta \bar{u}_1, D_\theta \bar{u}_2\rangle \\ &= c_\theta \left[\pi_1\sqrt{(\pi_1 r_1^2+\pi_2)(\pi_1 r_2^2+\pi_2)}\|D_\theta \bar{u}_1\|\|D_\theta \bar{u}_2\|\right]^{-1}, \\ \text{where } \ c_\theta &= \frac{1}{n}\left[\pi_1(\theta_{\bar{n}_1+1}^2+\cdots+\theta_n^2)-\pi_2(\theta_1^2+\cdots+\theta_{\bar{n}_1}^2)\right]. \end{aligned}$$

Note that the two sums in the formula of c_θ have the same expectation. It remains to apply Hoeffding's inequality to each sum. □

Since we do not have closed-form formulas for eigenvectors of $\bar{A}$, we can not describe $U_{\bar{A}}[-1,1]^n$ explicitly. Lemma 9 provides an approximation of $U_{\bar{A}}[-1,1]^n$. It will be used to replace Assumption (2).

LEMMA 9. *Consider the setting of Lemma 8 and let $\mathcal{R}^\theta = U_{\bar{A}}[-1,1]^n$ and*

$$\hat{\mathcal{R}}^\theta = \operatorname{conv}\left\{\pm U_{\bar{A}}(c), \pm U_{\bar{A}}(\mathbf{1})\right\}. \tag{5.4}$$

Then $\hat{\mathcal{R}}^\theta$ is a parallelogram and the angle between two adjacent sides is bounded away from zero and π; $\mathcal{R}^\theta$ is well approximated by $\hat{\mathcal{R}}^\theta$ in the sense that

$$\operatorname{dist}\left(\mathcal{R}^\theta, \hat{\mathcal{R}}^\theta\right) = \sup_{x\in\mathcal{R}^\theta}\inf_{y\in\hat{\mathcal{R}}^\theta}\|x-y\| = O(1)$$

as $n\to\infty$.

PROOF OF LEMMA 9. Let $v_i = \|D_\theta \bar{u}_i\|^{-1}D_\theta \bar{u}_i$, $i=1,2$, $V=(v_1,v_2)^T$, and $\mathcal{R}_V = V[-1,1]^n$. Following the same argument in the proof of Lemma 6, it is easy to show that $\mathcal{R}_V$ is a parallelogram with vertices $\{\pm Vc, \pm V\mathbf{1}\}$. By Lemma 8, $\|v_i - \bar{u}_i^\theta\| = O(n^{-1/2})$, which in turn implies $\operatorname{dist}\left(\mathcal{R}^\theta, \mathcal{R}_V\right) = O(1)$. The distance between two parallelograms $\mathcal{R}_V$ and $\hat{\mathcal{R}}^\theta$ is bounded by the maximum of the distances between corresponding vertices, which is also of order $O(1)$ because $\|v_i - \bar{u}_i^\theta\| = O(n^{-1/2})$. Finally by triangle inequality

$$\operatorname{dist}\left(\hat{\mathcal{R}}^\theta, \mathcal{R}^\theta\right) \le \operatorname{dist}\left(\hat{\mathcal{R}}^\theta, \mathcal{R}_V\right) + \operatorname{dist}\left(\mathcal{R}_V, \mathcal{R}^\theta\right) = O(1).$$

The angle between two adjacent sides of $\mathcal{R}_V$ equals the angle between $\sqrt{n}x$ and $\sqrt{n}y$, where x and y are defined in the proof of Lemma 6, which does not depend on n. Since $\operatorname{dist}(\hat{\mathcal{R}}^\theta, \mathcal{R}_V) = O(1)$, the angle between two adjacent sides of $\hat{\mathcal{R}}^\theta$ is bounded from zero and π. □

Before showing properties of the profile log-likelihood, let us introduce some new notations. Let $\bar{O}_{11}$, $\bar{O}_{12}$, $\bar{O}_{22}$, and $\bar{Q}_{DC}$ be the population version of O_{11}, O_{12}, O_{22}, and Q_{DC}, when A is replaced with $\bar{A}$. We also use $\bar{Q}_{BM}$, $\bar{Q}_{NG}$, and $\bar{Q}_{EX}$ to denote the population version of Q_{BM}, Q_{NG}, and Q_{EX} respectively. The following discussion is about $\bar{Q}_{DC}$, but it can be carried out for $\bar{Q}_{BM}$, $\bar{Q}_{NG}$, and $\bar{Q}_{EX}$ with obvious modifications and the help of Lemma 11.

Note that $\bar{O}_{11}$, $\bar{O}_{12}$, and $\bar{O}_{22}$ are quadratic forms of e and $\bar{A}$, therefore $\bar{Q}_{DC}$ depends on e through $U_{\bar{A}}e$, where $U_{\bar{A}}$ is the $2\times n$ matrix whose rows are eigenvectors of $\bar{A}$. With a little abuse of notation, we also use $\bar{O}_{ij}$, $i,j=1,2$, and $\bar{Q}_{DC}$ to denote the induced functions on $U_{\bar{A}}[-1,1]^n$. Thus, for example if $x\in U_{\bar{A}}[-1,1]^n$ then $\bar{Q}_{DC}(x)=\bar{Q}_{DC}(U_{\bar{A}}e)$ for any $e\in[-1,1]^n$ such that $x=U_{\bar{A}}e$.

To simplify $\bar{Q}_{DC}$, let ρ_1^θ and ρ_2^θ be eigenvalues of $\bar{A}$ as in Lemma 8 and let

$$t=(t_1,t_2)^T=U_{\bar{A}}\mathbf{1},\ \ \mu=(\rho_1^\theta t_1,\rho_2^\theta t_2)^T.$$

We parameterize $x\in U_{\bar{A}}[-1,1]^n$ by $x=\alpha t+\beta v$, where $v=(v_1,v_2)^T$ is a unit vector perpendicular to μ. If we denote $a=\frac{1}{4}(\rho_1^\theta t_1^2+\rho_2^\theta t_2^2)$ and $b=\frac{1}{4}(\rho_1^\theta v_1^2+\rho_2^\theta v_2^2)$, then

$$\begin{aligned}\bar{O}_{11} &= (\alpha+1)^2a+\beta^2b,\ \bar{O}_{22}=(\alpha-1)^2a+\beta^2b,\ \bar{O}_{12}=(1-\alpha^2)a-\beta^2b,\\ \bar{O}_1 &= \bar{O}_{11}+\bar{O}_{12}=2(1+\alpha)a,\ \bar{O}_2=\bar{O}_{22}+\bar{O}_{12}=2(1-\alpha)a.\end{aligned}$$

Note that $\bar{O}_{11}\bar{O}_{22}-\bar{O}_{12}^2=4\beta^2ab>0$ since ρ_1^θ and ρ_2^θ are positive by Lemma 8. With a little abuse of notation, we also use $\bar{Q}_{DC}(\alpha,\beta)$ to denote the value of $\bar{Q}_{DC}$ in the (α,β) coordinates described above. We now show some properties of $\bar{Q}_{DC}$.

LEMMA 10. *Consider $\bar{Q}=\bar{Q}_{DC}$ on $\hat{\mathcal{R}}^\theta$ defined by (5.4). Then*

(*a*) *$\bar{Q}(\alpha,0)$ is a constant.*
(*b*) *$\frac{\partial^2\bar{Q}}{\partial\beta^2}\geq 0$, $\frac{\partial\bar{Q}}{\partial\beta}>0$ if $\beta>0$ and $\frac{\partial\bar{Q}}{\partial\beta}<0$ if $\beta<0$. Thus, $\bar{Q}$ achieves minimum when $\beta=0$ and maximum on the boundary of $\hat{\mathcal{R}}^\theta$.*
(*c*) *$\bar{Q}$ is convex on the boundary of $\hat{\mathcal{R}}^\theta$. Thus, $\bar{Q}$ achieves maximum at $\pm U_{\bar{A}}(c)$.*
(*d*) *For any $x\in U_{\bar{A}}[-1,1]^n$, if $\bar{Q}(U_{\bar{A}}(c))-\bar{Q}(x)\leq\epsilon$ then*

$$\|U_{\bar{A}}(c)-x\|\leq 4\epsilon\sqrt{n}\left(\bar{Q}(U_{\bar{A}}(c))-\min_{\hat{\mathcal{R}}^\theta}\bar{Q}\right)^{-1}.$$

(*e*) *For any* $\delta \in (0,1)$, $\max_{\hat{\mathcal{R}}^\theta} \bar{Q} - \min_{\hat{\mathcal{R}}^\theta} \bar{Q}$ *is of order* $n\lambda_n$ *with probability at leat* $1-\delta$.

Parts (a) and (b) are used to prove part (c), which together with Lemma 9 will be used to replace Assumption (2). Parts (d) verifies Assumption (4), and part (e) provides a way to simplify the upper bound in part (d).

Proof of Lemma 10. Note that because $\hat{\mathcal{R}}^\theta \subset \mathcal{R}^\theta$, $\bar{O}_{11}$, $\bar{O}_{12}$, and $\bar{O}_{22}$ are nonnegative on $\hat{\mathcal{R}}^\theta$. Also, if we multiply $\bar{O}_{11}$, $\bar{O}_{12}$, and $\bar{O}_{22}$ by a constant $\eta > 0$ then the resulting function has the form $\eta\bar{Q} + C$, where C is a constant not depending on (α, β), and therefore the behavior of $\bar{Q}$ that we are interested in does not change. In this proof we use $\eta = 1/a$. Since $\bar{Q}$ is symmetric with respect to β, after multiplying by $1/a$, we replace $\beta^2 b/a$ with β and only consider $\beta \geq 0$. Thus, we may assume that

$$
\begin{aligned}
\bar{O}_{11} &= (\alpha+1)^2 + \beta,\ \bar{O}_{22} = (\alpha-1)^2 + \beta,\ \bar{O}_{12} = (1-\alpha^2) - \beta, \\
\bar{O}_1 &= \bar{O}_{11} + \bar{O}_{12} = 2(1+\alpha),\ \bar{O}_2 = \bar{O}_{22} + \bar{O}_{12} = 2(1-\alpha).
\end{aligned} \tag{5.5}
$$

(*a*) With (5.5) and $\beta = 0$, it is straightforward to verify that $Q(\alpha, 0)$ does not depend on α.

(*b*) Simple calculation shows that

$$
\frac{\partial \bar{Q}}{\partial \beta} = \log \frac{\bar{O}_{11}\bar{O}_{22}}{\bar{O}_{12}^2} \geq 0, \quad \frac{\partial^2 \bar{Q}}{\partial \beta^2} = \frac{1}{\bar{O}_{11}} + \frac{1}{\bar{O}_{22}} + \frac{2}{\bar{O}_{12}} \geq 0.
$$

(*c*) We show that $\bar{Q}$ is convex on the boundary line connecting $U_{\bar{A}}(\mathbf{1})$ and $U_{\bar{A}}(c)$. Let $(\alpha_0, \beta_0)^T$ be the coordinates of $U_{\bar{A}}(c)$, where $\beta_0 > 0$ and $\alpha_0 \in (-1, 1)$. We parameterize the segment connecting $U_{\bar{A}}(c)$ and $U_{\bar{A}}(\mathbf{1})$ by

$$
\left\{ \left(\alpha, \frac{\beta_0(1-\alpha)}{1-\alpha_0} \right)^T,\ \alpha \in [\alpha_0, 1] \right\}. \tag{5.6}
$$

With this parametrization, $\bar{O}_{11}$, $\bar{O}_{12}$, and $\bar{O}_{22}$ have the forms

$$
\begin{aligned}
\bar{O}_{11} &= (\alpha+1)^2 + \rho(\alpha-1)^2,\ \bar{O}_{22} = (\alpha-1)^2 + \rho(\alpha-1)^2 \\
\bar{O}_{12} &= (1-\alpha^2) - \rho(\alpha-1)^2,\ \rho = \frac{\beta_0^2}{(1-\alpha_0)^2}.
\end{aligned}
$$

Simple calculation shows that

$$
\begin{aligned}
\frac{1}{2}\frac{d^2\bar{Q}}{d\alpha^2} &= (\rho+1)\log \frac{(\rho+1)\bar{O}_{11}}{[\alpha+1+\rho(\alpha-1)]^2} \\
&+ \frac{4\rho}{[\alpha+1][\alpha+1+\rho(\alpha-1)]} - \frac{8\rho}{\bar{O}_{11}}.
\end{aligned}
$$

Note that the value of the right-hand side at $\alpha = 1$ is $(\rho+1)\log(\rho+1)-\rho \geq 0$ for any $\rho \geq 0$. Therefore to show that $\frac{d^2\bar{Q}}{d\alpha^2} \geq 0$, it is enough to show that $\frac{d^2\bar{Q}}{d\alpha^2}$ is non-increasing. Simple calculation shows that

$$\begin{aligned}\frac{d^3\bar{Q}}{d\alpha^3} &= 16\rho^2\left[(\alpha-1)^2\rho+\alpha^2-2\alpha-3\right]\times\\ &\times \left[(3\alpha+1)(\alpha-1)\rho+3(\alpha+1)^2\right]D^{-1},\end{aligned}$$

where $D = \bar{O}_{11}^2(\alpha+1)^2\left[\alpha+1+\rho(\alpha-1)\right]^2$. Since $\rho(1-\alpha) \leq (1+\alpha)$ because $\bar{O}_{12} \geq 0$, it follows that

$$(\alpha-1)^2\rho+\alpha^2-2\alpha-3 \leq (1-\alpha)(1+\alpha)+\alpha^2-2\alpha-3 = -2(\alpha+1) \leq 0.$$

Note that if $(3\alpha+1)(\alpha-1) \geq 0$ then $(3\alpha+1)(\alpha-1)\rho+3(\alpha+1)^2 \geq 0$. Otherwise $3\alpha+1 \geq 0$ and since $\rho(\alpha-1) \geq -(1+\alpha)$, it follows that

$$(3\alpha+1)(\alpha-1)\rho+3(\alpha+1)^2 \geq -(3\alpha+1)(\alpha+1)+3(\alpha+1)^2 = 2(\alpha+1) \geq 0.$$

Thus $\frac{d^3\bar{Q}}{d\alpha^3} \leq 0$. We have shown that $\bar{Q}$ is convex on the segment connecting $U_{\bar{A}}(c)$ and $U_{\bar{A}}(\mathbf{1})$. The same argument applies for other sides of the boundary of $\hat{\mathcal{R}}^\theta$.

(d) Let (α_x, β_x) be the parameters of x, $\hat{x}$ be the point with parameters $(\alpha_x, 0)$, and x^* be the point on the boundary of $\hat{\mathcal{R}}_U$ with parameters (α_x, β_x^*). Without loss of generality we assume that x^* is on the line connecting $x_c = U_{\bar{A}}(c)$ and $x_{\mathbf{1}} = U_{\bar{A}}(\mathbf{1})$. Note that (a),(b), and (c) imply

$$\bar{Q}(x_c) \geq \bar{Q}(x^*) \geq \bar{Q}(x) \geq \bar{Q}(\hat{x}) = \bar{Q}(x_{\mathbf{1}}).$$

Let $\ell = \bar{Q}(x_c) - \min_{\hat{\mathcal{R}}^\theta}\bar{Q}$. Since $\bar{Q}(\alpha_x, \beta)$ is convex in β (by (b)), we have

$$\frac{\|x^*-x\|}{\|x^*-\hat{x}\|} \leq \frac{\bar{Q}(x^*)-\bar{Q}(x)}{\bar{Q}(x^*)-\bar{Q}(\hat{x})} \leq \frac{\bar{Q}(x_c)-\bar{Q}(x)}{\bar{Q}(x_c)-\bar{Q}(\hat{x})} \leq \frac{\epsilon}{\ell}.$$

Therefore $\|x^*-x\| \leq \epsilon\ell^{-1}\|x^*-\hat{x}\| \leq 2\epsilon\sqrt{n}\ell^{-1}$. Since $\bar{Q}$ is convex on the boundary of $\hat{\mathcal{R}}^\theta$, we have

$$\frac{\|x_c-x^*\|}{\|x_c-x_{\mathbf{1}}\|} \leq \frac{\bar{Q}(x_c)-\bar{Q}(x^*)}{\bar{Q}(x_c)-\bar{Q}(x_{\mathbf{1}})} \leq \frac{\bar{Q}(x_c)-\bar{Q}(x)}{\bar{Q}(x_c)-\bar{Q}(x_{\mathbf{1}})} \leq \frac{\epsilon}{\ell},$$

which in turn implies $\|x_c-x^*\| \leq \epsilon\ell^{-1}\|x_c-x_{\mathbf{1}}\| \leq 2\epsilon\sqrt{n}\ell^{-1}$. Finally by triangle inequality

$$\|x_c-x\| \leq \|x_c-x^*\|+\|x^*-x\| \leq 4\epsilon\sqrt{n}\ell^{-1}.$$

(*e*) Note that $\min_{\hat{\mathcal{R}}^\theta} \bar{Q} = \bar{Q}(\alpha_0, 0) = \bar{Q}(0,0)$. Also, to find $\bar{Q}(c) - \bar{Q}(0)$ we do not have to calculate $\bar{O}_1 \log \bar{O}_1 + \bar{O}_2 \log \bar{O}_2$ since along the line $\alpha = \alpha_0$, $\bar{O}_1$ and $\bar{O}_2$ do not change. Simple calculation with Hoeffding's inequality show that with probability at least $1-\delta$ the following hold

$$\bar{O}_{11}(0) = \bar{O}_{22}(0) = \bar{O}_{12}(0) = \frac{n\lambda_n}{4}\left(\pi_1^2 + \omega\pi_2^2 + 2\pi_1\pi_2 r\right) + O(\lambda_n\sqrt{n}),$$

$$\bar{O}_{11}(c) = n\lambda_n\pi_1^2 + \bar{O}(\lambda_n\sqrt{n}),\ \bar{O}_{22}(c) = n\lambda_n\omega\pi_2^2 + O(\lambda_n\sqrt{n}),$$
$$\bar{O}_{12}(c) = n\lambda_n\pi_1\pi_2 r + O(\lambda_n\sqrt{n}).$$

By the remark at the beginning of the proof of Lemma 10, we can take $\eta = n\lambda_n$, and therefore $\bar{Q}(U_{\bar{A}}(c)) - \min_{\hat{\mathcal{R}}^\theta} \bar{Q}$ is of order $n\lambda_n$. □

PROOF OF THEOREM 2. Note that $\bar{Q} = \bar{Q}_{DC}$ does not satisfy all Assumptions (1)–(4), therefore we can not apply Theorem 1 directly. Instead we will follow the idea of the proof of Lemma 1.

We first show that $\bar{Q}$ satisfies Assumption (1). For $\bar{Q}$, the functions g_j in (2.1) has the form $g(z) = z\log(z)$. We can assume that $z > 1$ because otherwise $g(z)$ is bounded by a constant. Since $g'(z) = 1 + \log(z)$, $g'(z)$ does not grow faster than $\log(z)$, and therefore assumption (1) holds.

Note that by Lemma 9, $\operatorname{dist}\left(\hat{\mathcal{R}}, \hat{\mathcal{R}}^\theta\right)$ is bounded by a constant; by Lemma 2, the Lipschitz constant of $\bar{Q}$ is of order $O\left(\sqrt{n}\log(n)\|\bar{A}\|\right)$. Therefore, to prove Lemma 1, and in turn Theorem 2, it is enough to consider $\bar{Q}$ on $\hat{\mathcal{R}}^\theta$.

Note also that $\bar{Q}$ may not be convex, therefore Assumption (2) may not hold. But we now show that the convexity of $\bar{Q}$ is not needed. In the proof of Lemma 1, the convexity of f_B is used only at one place to show that (A.4) implies (A.5), or more specifically, that $f_B(y) \le f_B(U_B(\hat{e}))$. Note that by A.3, $\|y - U_{\bar{A}}(c)\| \le 2\sqrt{n}\|U_A - U_{\bar{A}}\|$. By Lemma 10 part c, $\bar{Q}$ achieves maximum at $U_{\bar{A}}(c)$, a vertex of $\hat{\mathcal{R}}^\theta$; by Lemma 9, the angle between two adjacent sides of $\hat{\mathcal{R}}^\theta$ is bounded away from zero and π. Thus, there exists $s \in \mathcal{E}_A$ such that $\|y - U_{\bar{A}}(s)\| \le M\sqrt{n}\|U_A - U_{\bar{A}}\|$. By Lemma 2 we have

$$|\bar{Q}(y) - \bar{Q}(U_{\bar{A}}(s))| \le Mn\log(n)\|\bar{A}\| \cdot \|U_A - U_{\bar{A}}\|.$$

Therefore in (A.4) we can replace y with $U_{\bar{A}}(s)$, and (A.5) follows by definition of $\hat{e}$.

We now check assumptions (3) and (4). To check the assumption (3), we first assume that $U_{\bar{A}} = (D_\theta(\bar{u}_1, \bar{u}_2))^T$, where $\bar{u}_1$ and $\bar{u}_2$ are from Lemma 3, and $D_\theta = \operatorname{diag}(\theta)$. The first $\bar{n}_1 = n\pi_1$ column vectors of $(\bar{u}_1, \bar{u}_2)^T$ are equal and we denote by ξ_1. The last $\bar{n}_2 = n\pi_2$ column vectors of $(\bar{u}_1, \bar{u}_2)^T$ are also equal and we denote by ξ_2. Then

$$\begin{aligned} U_{\bar{A}}(c) - U_{\bar{A}}(e) &= \sum_{i=1}^{\bar{n}_1} \theta_i(1-e_i)\xi_1 + \sum_{i=\bar{n}_1+1}^{n} \theta_i(-1-e_i)\xi_2 \\ &= k_1 \sum_{i=1}^{\bar{n}_1} \theta_i \xi_1 - k_2 \sum_{i=\bar{n}_1+1}^{n} \theta_i \xi_2, \end{aligned}$$

where $k_1 = \sum_{i=1}^{\bar{n}_1}(1-e_i)$, $k_2 = \sum_{i=\bar{n}_1}^{n}(1+e_i)$, and $\|e-c\|^2 = k_1 + k_2$. By Lemma 3, entries of ξ_1, ξ_2 are of order $1/\sqrt{n}$ and the angle between ξ_1, ξ_2 does not depend on n, it follows that $\sqrt{n}\|U_{\bar{A}}(c) - U_{\bar{A}}(e)\|$ is of order $k_1 + k_2$. By Lemma 8, it is easy to see that the argument still holds for the actual $U_{\bar{A}}$.

Assumption (4) follows directly from part (e) of Lemma 10.

Combining Assumptions (3), (4), and Lemma 1, we see that Theorem 1 holds. Note that the conclusion of Lemma 4 still holds if we replace $\mathbb{E}[A]$ with $\bar{A}$, except the constant M now also depends on ξ, that is $M = M(r, \omega, \pi, \delta) > 0$. The upper bound in Theorem 1 is simplified by Lemma 4 and part d of Lemma 10. The bound in Theorem 1 is simplified by (B.1) of Lemma 4 and part e of Lemma 10:

$$\|e^* - c\|^2 \leq Mn\log n \left(\lambda_n^{-1/2} + \|U_A - U_{\mathbb{E}[A]}\|\right).$$

If U_A is formed by eigenvectors of A then using (B.2) of Lemma 4, we obtain

$$\|e^* - c\|^2 \leq \frac{Mn\log n}{\sqrt{\lambda_n}}.$$

The proof is complete. □

5.2. *Proof of Results in Section 3.2.* We follow the notation introduced in the discussion before Lemma 10. Lemma 11 provides the form of n_1 and n_2 as functions defined on the projection of the cube.

LEMMA 11. *Consider the block models and let $\mathcal{R} = U_{\mathbb{E}[A]}[-1,1]^n$. In the coordinate system $x_e = U_{\mathbb{E}[A]}(e)$, the functions n_1 and n_2 defined by (3.4) admit the forms*

$$n_1 = \sqrt{n}(\sqrt{n} + \vartheta^T x)/2, \quad n_2 = \sqrt{n}(\sqrt{n} - \vartheta^T x)/2,$$

where ϑ is a vector with $\|\vartheta\| < M$ for some $M > 0$ not depending on n. In the coordinate system (α, β), n_1 and n_2 admit the forms

$$n_1 = \frac{\sqrt{n}}{2}\big[(1+\alpha) + s\beta\big], \quad n_2 = \frac{\sqrt{n}}{2}\big[(1-\alpha) - s\beta\big],$$

where s is a constant.

PROOF OF LEMMA 11. Let $U^* = (U_{\mathbb{E}[A]}^T, \frac{1}{\sqrt{n}}\mathbf{1})^T$ and $\mathcal{R}_{U^*} = U^*[-1,1]^n$. For each $e \in [-1,1]^n$, let $z = \frac{1}{\sqrt{n}}\mathbf{1}^T e$, so that $U^* e = \binom{x}{z}$. Then

$$n_1 = \sqrt{n}(\sqrt{n} + z)/2, \quad n_2 = \sqrt{n}(\sqrt{n} - z)/2.$$

By Lemma 6, the first $\bar{n}_1$ row vectors of $U_{\mathbb{E}[A]}$ are equal, and the last $\bar{n}_2$ row vectors of $U_{\mathbb{E}[A]}$ are also equal. Therefore U^* has rank two, and $\mathcal{R}_{U^*}$ is contained in a hyperplane. It follows that z is a linear function of x, and in turn, a linear function of (α, β).

In the coordinate system x, $n_1(0) = n/2$ implies $z(0) = 0$; $n_1(\mathbf{1}) = n$ implies $z(x_{\mathbf{1}}) = \sqrt{n}$; $n_1(c) = \bar{n}_1 = n\pi_1$ implies $z(x_c) = (2\pi_1 - 1)\sqrt{n}$. Since $\|x_{\mathbf{1}}\|$ and $\|x_c\|$ are of order $\sqrt{n}$ by Lemma 3 and Lemma 8, there exists a constant $M > 0$ such that $z = \vartheta^T x$ for some vector ϑ with $\|\vartheta\| < M$.

In the coordinate system (α, β), $n_1(0) = n_2(0) = n/2$ implies $z(0) = 0$; $n_1(\mathbf{1}) = n$ implies $z(1,0) = \sqrt{n}$; $n_1(-\mathbf{1}) = 0$ implies $z(-1,0) = -\sqrt{n}$. Therefore along the line $\beta = 0$, $z(\alpha, 0) = \sqrt{n}\alpha$. For any fixed α, z is a linear function of β with the same coefficient, so $z(\alpha, \beta) = \sqrt{n}\alpha + s\sqrt{n}\beta$ for some constant s. ☐

Lemma 12 show some properties of $\bar{Q}_{BM}$. Parts (b) gives a weaker version of convexity of $\bar{Q}_{BM}$. Part (c) together with Lemma 6 will be used to replace Assumption (2). Part (d) verifies Assumption (4), and part (e) simplifies the upper bound in part (d).

LEMMA 12. *Consider $\bar{Q} = \bar{Q}_{BM}$ on $\mathcal{R} = U_{\mathbb{E}[A]}[-1,1]^n$. Then*

(*a*) *$\bar{Q}(\alpha, 0)$ is a constant.*
(*b*) *$\frac{\partial^2 \bar{Q}}{\partial \beta^2} \geq 0$, $\frac{\partial \bar{Q}}{\partial \beta} > 0$ if $\beta > 0$ and $\frac{\partial \bar{Q}}{\partial \beta} < 0$ if $\beta < 0$. Thus, $\bar{Q}$ achieves minimum when $\beta = 0$ and maximum on the boundary of $\mathcal{R}$.*
(*c*) *$\bar{Q}$ is convex on the boundary of $\mathcal{R}$. Thus, $\bar{Q}$ archive maximum at $\pm U_{\mathbb{E}[A]}c$.*
(*d*) *If $\bar{Q}(U_{\mathbb{E}[A]}c) - \bar{Q}(x) \leq \epsilon$ then*

$$\|U_{\mathbb{E}[A]}c - x\| \leq 4\epsilon\sqrt{n}\left(\bar{Q}(U_{\mathbb{E}[A]}c) - \min_{\mathcal{R}} \bar{Q}\right)^{-1}.$$

(e) *$\bar{Q}(U_{\mathbb{E}[A]}(c)) - \min_{\mathcal{R}} \bar{Q}$ is of order $n\lambda_n$.*

Proof of Lemma 12. Let $G = \bar{O}_1 \log \frac{\bar{O}_1}{n_1} + \bar{O}_2 \log \frac{\bar{O}_2}{n_2}$, then $\bar{Q}_{BM} = \bar{Q}_{DCBM} + 2G$. By Lemma 10, to show (a), (b), and (c), it is enough to show that G satisfies those properties. Parts (d) and (e) follow from (a), (b), and (c) by the same argument used to prove Lemma 10. Note that if we multiply $\bar{O}_1$ and $\bar{O}_2$ by a positive constant, or multiply n_1 and n_2 by a positive constant, then the behavior of G does not change, since $\bar{O}_1 + \bar{O}_2$ is a constant. Therefore by Lemma 11 we may assume that

$$\begin{aligned} \bar{O}_1 &= 2(1+\alpha), \quad \bar{O}_2 = 2(1-\alpha), \\ n_1 &= (1+\alpha) + s\beta, \quad n_2 = (1-\alpha) - s\beta. \end{aligned}$$

(a) It is easy to see that $G(\alpha, 0)$ is a constant.
(b) Simple calculation shows that

$$\frac{\partial G}{\partial \beta} = \frac{4s^2\beta}{n_1 n_2}, \quad \frac{\partial^2 G}{\partial \beta^2} = \frac{4s^2}{(n_1 n_2)^2}\left(1 - \alpha^2 + s^2\beta^2\right),$$

and the statement follows.

(c) We show that G is convex on the segment connecting $U_{\mathbb{E}[A]}c$ and $U_{\mathbb{E}[A]}\mathbf{1}$. With the parametrization (5.6), n_1 and n_2 have the form

$$n_1 = (1+\alpha) + s(1-\alpha), \quad n_2 = (1-\alpha) - s(1-\alpha),$$

for some constant s. Simple calculation shows that

$$\frac{d^2 G}{d\alpha^2} = \frac{4}{\bar{O}_1} - \frac{2(1-s)}{n_1} - \frac{4s(1-s)}{n_1^2}.$$

Note that when $\alpha = 1$, the right hand side equals $s^2 \geq 0$. Therefore, to show that G is convex, it is enough to show that the second derivative of G is non-increasing. The third derivative of G has the form

$$\frac{d^3 G}{d\alpha^3} = \frac{8s^2}{n_1^3(1+\alpha)^2}\big[(3\alpha+1)s - 3\alpha - 3\big].$$

Note that $n_1 \geq 0$ implies $s \geq -\frac{1+\alpha}{1-\alpha}$; $n_2 \geq 0$ implies $s \leq 1$. Consider function $h(s) = (3\alpha+1)s - 3\alpha - 3$ on $\left[\frac{1+\alpha}{1-\alpha}, 1\right]$. Since

$$h\left(\frac{1+\alpha}{1-\alpha}\right) = \frac{-4(1+\alpha)}{1-\alpha} \leq 0, \quad h(1) = -2 < 0,$$

$h(s) \leq 0$ and G is convex. ∎

Note that $\bar{Q}_{BM}$ does not have the exact form of (2.1). A small modification shows that Lemma 1 still holds for $\bar{Q}_{BM}$.

LEMMA 13. *Let $Q = Q_{BM}$, $\bar{Q} = \bar{Q}_{BM}$, and U_A be an approximation of $U_{\mathbb{E}[A]}$. Under the assumptions of Theorem 3, there exists a constant $M = M(r, w, \pi, \delta) > 0$ such that with probability at least $1 - n^{-\delta}$, we have*

$$\bar{Q}(x_c) - \bar{Q}(x_{e^*}) \leq Mn \log n \left(\sqrt{\lambda_n} + \lambda_n \|U_A - U_{\mathbb{E}[A]}\| \right).$$

In particular, if U_A is the matrix whose row vectors are leading eigenvectors of A, then

$$\bar{Q}(x_c) - \bar{Q}(x_{e^*}) \leq Mn \log n \sqrt{\lambda_n}.$$

PROOF OF LEMMA 13. Let $G_i = O_i \log n_i$ and $\bar{G}_i = \bar{O}_i \log n_i$ for $i = 1, 2$. Also, let $G = Q_{DCBM}$ and $\bar{G} = \bar{Q}_{DCBM}$. Then

$$Q = G + G_1 + G_2, \quad \bar{Q} = \bar{G} + \bar{G}_1 + \bar{G}_2.$$

In the proof of Theorem 2 we have shown that G satisfies Assumption (1). Therefore inequality (A.1) in the proof of Lemma 1 also holds for G:

$$|G(e) - \bar{G}(e)| \leq Mn \log n \|A - \mathbb{E}A\|. \tag{5.7}$$

The same type of inequality holds for G_i as well. Indeed, since $\|\mathbf{1} + e\|^2 = 2(\mathbf{1} + e)^T \mathbf{1} = 4n_1$, we have

$$\begin{aligned} |G_i(e) - \bar{G}_i(e)| &= |\log n_1||(1+e)^T (A - \mathbb{E}[A])\mathbf{1}| \\ &\leq 2n \log(n) \|A - \mathbb{E}[A]\|. \end{aligned} \tag{5.8}$$

From (5.7) and (5.8) we obtain

$$|Q(e) - \bar{Q}(e)| \leq Mn \log n \|A - \mathbb{E}A\|. \tag{5.9}$$

Let $\hat{e} = \arg\max\{\bar{Q}(e), e \in \mathcal{E}_A\}$. Using (5.9) and definition of e^*, we have

$$\begin{aligned} \bar{Q}(\hat{e}) - \bar{Q}(e^*) &\leq \bar{Q}(\hat{e}) - Q(\hat{e}) + Q(e^*) - \bar{G}(e^*) \\ &\leq Mn \log(n) \|A - \mathbb{E}[A]\|. \end{aligned} \tag{5.10}$$

Let $y \in \mathrm{conv}(U_{\mathbb{E}[A]}\mathcal{E}_A)$ such that $\|U_{\mathbb{E}[A]}(c) - y\| = \mathrm{dist}\big(U_{\mathbb{E}[A]}(c), \mathrm{conv}(U_{\mathbb{E}[A]}\mathcal{E}_A)\big)$. Using the same argument as in the proof of Lemma 1, we obtain

$$\|U_{\mathbb{E}[A]}(c) - y\| \leq 2\sqrt{n}\ \|U_A - U_{\mathbb{E}[A]}\|, \tag{5.11}$$

and there exists a constant $M > 0$ such that

$$\begin{aligned} \left|\bar{O}_1(y) - \bar{O}_1(U_{\mathbb{E}[A]}(c))\right| &\leq Mn\|\mathbb{E}[A]\|.\|U_A - U_{\mathbb{E}[A]}\| \\ &\leq Mn\lambda_n\|U_A - U_{\mathbb{E}[A]}\|. \end{aligned}$$

By Lemma 6, the angle between two adjacent sides of $\mathcal{R}$ does not depend on n. Therefore (5.11) implies that there exists $s \in \mathcal{E}_A$ such that

$$\|U_{\mathbb{E}[A]}(c) - U_{\mathbb{E}[A]}(s)\| \leq M\sqrt{n}\|U_A - U_{\mathbb{E}[A]}\|. \tag{5.12}$$

Denote $x_e = U_{\mathbb{E}[A]}(e)$ for $e \in [-1,1]^n$. By Lemma 2 the Lipchitz constant of $\bar{G}$ on $U_{\mathbb{E}}[A][-1,1]^n$ is of order $\sqrt{n}\|\mathbb{E}[A]\|\log n \leq \sqrt{n}\lambda_n \log n$. Therefore from (5.12) we have

$$\bar{G}(x_c) - \bar{G}(x_s) \leq Mn\lambda_n \log n\|U_A - U_{\mathbb{E}[A]}\|. \tag{5.13}$$

We will show that the same inequality holds for $\bar{G}_i$, and thus also for $\bar{Q}$. By triangle inequality we have

$$\bar{G}_i(x_c) - \bar{G}_i(x_s) \leq |\bar{O}_i(x_s) - \bar{O}_i(x_c)||\log n_i(x_c)| + \bar{O}_i(x_s)\left|\log \frac{n_i(x_s)}{n_i(x_c)}\right|. \tag{5.14}$$

To bound the first term on the right-hand side of (5.14), we note that by Lemma 2, the Lipchitz constant of $\bar{O}_i$ is of order $\sqrt{n}\|\mathbb{E}[A]\| \leq \lambda_n\sqrt{n}$. Using (5.12) we obtain

$$\begin{aligned} |\bar{O}_i(x_s) - \bar{O}_i(x_c)||\log n_i(x_c)| &\leq |\bar{O}_i(x_s) - \bar{O}_i(x_c)|\log n \\ &\leq Mn\lambda_n \log n\|U_A - U_{\mathbb{E}[A]}\|. \end{aligned} \tag{5.15}$$

We now bound the second term on the right-hand side of (5.14). By Lemma 11, there exist $M' > 0$ not depending on n and a vector ϑ such that $\|\vartheta\| \leq M'$ and

$$\begin{aligned} |n_i(x_c) - n_i(x_s)| &= |\vartheta^T(x_c - x_s)|/2 \leq M'\|x_c - x_s\| \\ &\leq M'\sqrt{n}\|U_A - U_{\mathbb{E}[A]}\|. \end{aligned} \tag{5.16}$$

Note that $n_i(x_c) = \bar{n}_i = n\pi_1$ and $|n_i(x_c) - n_i(x_s)| = o(n)$ by (5.16). Using (5.16) and the inequality $\log(1+t) \leq 2|t|$ for $|t| \leq 1/2$, we have

$$\begin{aligned} \left|\log \frac{n_i(x_s)}{n_i(x_c)}\right| &= \left|\log\left(1 + \frac{n_i(x_s) - n_i(x_c)}{n_i(x_c)}\right)\right| \\ &\leq \frac{2M'\sqrt{n}\|U_A - U_{\mathbb{E}[A]}\|}{n_i(x_c)}. \end{aligned} \tag{5.17}$$

By definition, $\bar{O}_i(x_s)$ is at most $O(n\lambda_n)$. Therefore from (5.17) we obtain

$$|\bar{O}_i(x_s)| \cdot \left|\log \frac{n_i(x_s)}{n_i(x_c)}\right| \leq M\lambda_n\sqrt{n}\|U_A - U_{\mathbb{E}[A]}\|. \tag{5.18}$$

Using (5.13), (5.14), (5.15), (5.18), and the fact that $\bar{Q}(x_s) \leq \bar{Q}(x_{\hat{e}})$, we get

$$\bar{Q}(x_c) - \bar{Q}(x_{\hat{e}}) \leq \bar{Q}(x_c) - \bar{Q}(x_s) \leq Mn\lambda_n \log n\|U_A - U_{\mathbb{E}[A]}\|. \tag{5.19}$$

Finally, from (5.10), inequality (B.1) of Lemma 4, and (5.19), we obtain

$$\bar{Q}(x_c) - \bar{Q}(x_{e^*}) \leq Mn \log n \left(\sqrt{\lambda_n} + \lambda_n\|U_A - U_{\mathbb{E}[A]}\|\right).$$

If U_A is formed by eigenvectors of A then it remains to use inequality (B.2) of Lemma 4. The proof is complete. $\square$

Proof of Theorem 3. The proof is similar to that of Theorem 2, with the help of Lemma 12 and Lemma 13. $\square$

5.3. *Proof of Results in Section 3.3.* We follow the notation introduced in the discussion before Lemma 10.

Proof of Theorem 4. Note that $\bar{Q} = \bar{Q}_{NG}$ does not have the exact form of (2.1). We first show that $\bar{Q}$ is Lipschitz with respect to $\bar{O}_1$, $\bar{O}_2$, and $\bar{O}_{12}$, which is stronger than assumption (1) and ensures that the argument in the proof of Lemma 1 is still valid.

To see that $\bar{Q}$ is Lipschitz, consider the function $h(x, y) = \frac{xy}{x+y}$, $x \geq 0, y \geq 0$. The gradient of h has the form $\nabla h(x, y) = \left(\frac{y^2}{(x+y)^2}, \frac{x^2}{(x+y)^2}\right)$. It is easy to see that $\nabla h(x, y)$ is bounded by $\sqrt{2}$. Therefore h is Lipschitz, and so is $\bar{Q}$.

Simple calculation shows that $\bar{Q} = 2b\beta^2$. Therefore $\bar{Q}$ is convex, and by Lemma 6, it achieves maximum at the projection of the true label vector. Thus, assumption (2) holds. Assumption (3) follows from Lemma 3 by the same argument used in the proof of Theorem 2. Assumption (4) follows from the convexity of $\bar{Q}$ and the argument used in the proof of part (e) of Lemma 10. Note that $\bar{Q}(0) = 0$ and $\bar{Q}(c)$ is of order $n\lambda_n$, therefore Theorem 4 follows from Theorem 1. $\square$

5.4. *Proof of Results in Section 3.4.* We follow the notation introduced in the discussion before Lemma 10. We first show some properties of $\bar{Q}_{EX}$. Parts (b) and (c) verify Assumption (2), and part (d) verifies Assumption (4).

LEMMA 14. *Let $\bar{Q} = \bar{Q}_{EX}$. Then*

(a) $\bar{Q}(\alpha, 0) = 0$.

(b) $\bar{Q}$ *is convex.*

(c) *If $\pi_1^2 > r\pi_2^2$ then the maximum value of $\bar{Q}$ is $n\lambda_n\pi_1\pi_2(1-r)$ and it is achieved at $x_c = U_{\mathbb{E}[A]}(c)$; if $\pi_1^2 \le r\pi_2^2$ then the maximum value of $\bar{Q}$ is $n\lambda_n\pi_1\pi_2 r(\frac{\pi_2^2}{\pi_1^2} - 1)$ and it is achieved at $x_{-c} = -U_{\mathbb{E}[A]}(c)$.*

(d) *Let $x_{\max}$ be the maximizer of $\bar{Q}$. If $\bar{Q}(x_{\max}) - \bar{Q}(x) \le \epsilon = o(n\lambda_n)$ then $\|x_{\max} - x\| \le 2\epsilon\sqrt{n}(\bar{Q}(x_{\max}))^{-1}$.*

PROOF OF LEMMA 14. Note that multiplying $\bar{O}_{11}, \bar{O}_{12}$ by a positive constant, or multiplying n_1 and n_2 by a constant does not change the behavior of $\bar{Q}$. Therefore by Lemma 11 we may assume that

$$\begin{aligned} \bar{O}_{11} &= (1+\alpha)^2 + b\beta^2, \quad \bar{O}_{12} = (1-\alpha^2) - b\beta^2, \\ n_1 &= 1 + \alpha + s\beta, \quad n_2 = 1 - \alpha - s\beta. \end{aligned}$$

(a) It is straightforward that $\bar{Q}(\alpha, 0) = 0$.

(b) Let $z = s\beta$, $r = s^2/b > 0$, and $h(\alpha, z) = \frac{z^2 - r(1+\alpha)z}{z+1+\alpha}$, then $\bar{Q} = \frac{2}{r}h(\alpha, z)$. Simple calculation shows that the Hessian of h has the form

$$\nabla h = \frac{2(r+1)}{(z+1+\alpha)^3}\begin{pmatrix} (1+\alpha)^2 & -z(1+\alpha) \\ -z(1+\alpha) & z^2 \end{pmatrix},$$

which implies that h and $\bar{Q}$ are convex.

(c) Since $\mathcal{R} = U_{\mathbb{E}[A]}[-1,1]^n$ is a parallelogram by Lemma 6 and $\bar{Q}$ is convex by part (b), it reaches maximum at one of the vertices of $\mathcal{R}$. The claim then follows from a simple calculation.

(d) Note that $|\bar{Q}(x_c) - \bar{Q}(x_{-c})| = |\frac{\pi_2}{\pi_1}n\lambda_n(\pi_1^2 - r\pi_2^2)|$ is of order $n\lambda_n$, therefore if $\bar{Q}(x_{\max}) - \bar{Q}(x) \le \epsilon = o(n\lambda_n)$ then $x_{\max}$ and x belong to the same part of $\mathcal{R}$ divided by the line $\beta = 0$. In other words, if $\hat{x}$ is the intersection of the line going through x and $x_{\max}$ and the line $\beta = 0$, then x belongs to the segment connecting $x_{\max}$ and $\hat{x}$. By convexity of $\bar{Q}$ and the fact that $\bar{Q}(\hat{x}) = 0$ from part (a) and part (b), we get

$$\frac{\|x_{\max} - x\|}{\|x_{\max} - \hat{x}\|} \le \frac{\bar{Q}(x_{\max}) - \bar{Q}(x)}{\bar{Q}(x_{\max}) - \bar{Q}(\hat{x})} \le \frac{\epsilon}{\bar{Q}(x_{\max})}.$$

It remains to bound $\|x_{\max} - \hat{x}\|$ by $2\sqrt{n}$. □

Note that $\bar{Q}_{EX}$ does not have the exact form of (2.1). The following Lemma shows that the argument used in the proof of Lemma 1 holds for $\bar{Q}_{EX}$.

LEMMA 15. *Let $\bar{Q} = \bar{Q}_{EX}$ and assume that the assumption of Theorem 5 holds. Let U_A be an approximation of $U_{\mathbb{E}[A]}$. Then there exists a constant $M = M(r, \pi, \delta) > 0$ such that with probability at least $1 - n^{-\delta}$, we have*

$$\bar{Q}(c) - \bar{Q}(e^*) \le Mn\lambda_n \left(\lambda_n^{-1/2} + \|U_A - U_{\mathbb{E}[A]}\| \right). \tag{5.20}$$

In particular, if U_A is a matrix whose row vectors are eigenvectors of A, then

$$\bar{Q}(c) - \bar{Q}(e^*) \le Mn\sqrt{\lambda_n}.$$

PROOF OF LEMMA 15. Note that $\|\mathbf{1} + e\|^2 = 2(\mathbf{1} + e)^T \mathbf{1} = 4n_1$. Using inequality (B.1) of Lemma 4, we have

$$\begin{aligned}
\left| \frac{n_2}{n_1} O_{11} - \frac{n_2}{n_1} \bar{O}_{11} \right| &= \left| \frac{n_2}{n_1} (\mathbf{1} + e)^T (A - \mathbb{E}[A])(\mathbf{1} + e) \right| \\
&\le \frac{n_2}{n_1} \|\mathbf{1} + e\|^2 \|A - \mathbb{E}[A]\| \\
&\le Mn_2\sqrt{\lambda_n} \le Mn\sqrt{\lambda_n}, \\
|O_{12} - \bar{O}_{12}| &\le Mn\sqrt{\lambda_n}.
\end{aligned}$$

Therefore

$$|Q(e) - \bar{Q}(e)| \le Mn\sqrt{\lambda_n}.$$

Let $\hat{e} = \arg\max\{\bar{Q}(e), e \in \mathcal{E}_A\}$. Then $Q(e^*) \ge Q(\hat{e})$ and hence

$$\begin{aligned}
\bar{Q}(\hat{e}) - \bar{Q}(e^*) &\le \bar{Q}(\hat{e}) - Q(\hat{e}) + Q(e^*) - \bar{Q}(e^*) \\
&\le Mn\sqrt{\lambda_n}.
\end{aligned} \tag{5.21}$$

Let $y \in \mathrm{conv}(U_{\mathbb{E}[A]}\mathcal{E}_A)$ such that $\|U_{\mathbb{E}[A]}(c) - y\| = \mathrm{dist}\big(U_{\mathbb{E}[A]}(c), \mathrm{conv}(U_{\mathbb{E}[A]}\mathcal{E}_A)\big)$. By the same argument as in the proof of Lemma 1, we have

$$\|U_{\mathbb{E}[A]}(c) - y\| \le 2\sqrt{n}\ \|U_A - U_{\mathbb{E}[A]}\|. \tag{5.22}$$

From Lemma 2, the Lipchitz constant of $\bar{O}_i$ is of order $\sqrt{n}\|\mathbb{E}[A]\| \le \sqrt{n}\lambda_n$. Using (5.22), we get

$$\left| \bar{O}_{1i}(y) - \bar{O}_{1i}(U_{\mathbb{E}[A]}(c)) \right| \le Mn\lambda_n \|U_A - U_{\mathbb{E}[A]}\|. \tag{5.23}$$

Denote $x_e = U_{\mathbb{E}[A]}(e)$ for $e \in [-1, 1]^n$. By Lemma 11, there exist $M' > 0$ not depending on n and a vector ϑ such that $\|\vartheta\| \le M'$ and for $i = 1, 2$,

$$\begin{aligned}
|n_i(x_c) - n_i(y)| &= |\vartheta^T (x_c - y)|/2 \le M' \|x_c - y\| \\
&\le M'\sqrt{n}\|U_A - U_{\mathbb{E}[A]}\|, \quad \text{by (5.22)}.
\end{aligned} \tag{5.24}$$

Note that $n_i(x_c) = \bar{n}_i = \pi_i n$ and $|n_i(x_c) - n_i(y)| = o(n)$ by (5.24). Therefore from (5.24) we obtain

$$\left|\frac{\bar{n}_2}{\bar{n}_1} - \frac{n_2(y)}{n_1(y)}\right| \le Mn^{-1/2}\|U_A - U_{\mathbb{E}[A]}\|.$$

Together with (5.23) and the fact that $\bar{O}_{11}(y) \le n\lambda_n$, we get

$$\begin{aligned} |\bar{Q}(x_c) - \bar{Q}(y)| &\le \frac{\bar{n}_2}{\bar{n}_1}|\bar{O}_{11}(x_c) - \bar{O}_{11}(y)| + \left|\frac{\bar{n}_2}{\bar{n}_1} - \frac{n_2(y)}{n_1(y)}\right| \bar{O}_{11}(y) \\ &+ |\bar{O}_{12}(y) - \bar{O}_{12}(x_c)| \\ &\le Mn\lambda_n\|U_A - U_{\mathbb{E}[A]}\|. \end{aligned}$$

The convexity of $\bar{Q}$ by Lemma 14 then imply

$$\bar{Q}(x_c) - \bar{Q}(x_{\hat{e}}) \le Mn\lambda_n\|U_A - U_{\mathbb{E}[A]}\|. \tag{5.25}$$

Finally, adding (5.21) and (5.25) we get (5.20). If U_A is formed by eigenvectors of A, then it remains to use inequality (B.2) of Lemma 4. The proof is complete. □

PROOF OF THEOREM 5. The proof is similar to that of Theorem 2, with the help of Lemma 3, Lemma 14, and Lemma 15. □

439 West Hall, 1085 S. University Ave.
Ann Arbor, MI 48109-1107
E-mail: canle@umich.edu

439 West Hall, 1085 S. University Ave.
Ann Arbor, MI 48109-1107
E-mail: elevina@umich.edu

2074 East Hall, 530 Church St.
Ann Arbor, MI 48109-1043
E-mail: romanv@umich.edu