Electronic Companion to "Information Retrieval under Network Uncertainty: Robust Internet Ranking"

by Anna Timonina-Farkas and Ralf W. Seifert

A. Appendix

LEMMA 2. Conic optimization problem

$$f(x) = \max_{\substack{z\in\mathbb{R}^N \\ ||z||_2\leq\varepsilon \\ |z_j|\leq\varepsilon_j}} z^T x \tag{A1}$$

is equivalent to the following minimization problem:

$$f(x) = \min_{\lambda+\mu=x} \left\{ \varepsilon\|\lambda\|_2 + \sum_{j=1}^{N} \varepsilon_j|\mu_j| \right\}. \tag{A2}$$

Proof Let us dualize the optimization problem (A1). First of all, notice that

$$\max_{\substack{z\in\mathbb{R}^N \\ ||z||_2\leq\varepsilon \\ |z_j|\leq\varepsilon_j}} z^T x \iff \max_{\substack{z\in\mathbb{R}^N \\ \sqrt{\sum_{i=1}^N z_i^2}\leq\varepsilon \\ z_j\leq\varepsilon_j \\ -z_j\leq\varepsilon_j}} z^T x.$$

Therefore, the Lagrangian $\mathscr{L}$ can be written in the following form for dual variables α,β,γ, where $\alpha\in\mathbb{R}_{\{0,+\}}$ and $\beta,\gamma\in\mathbb{R}^N_{\{0,+\}}$ $\mathscr{L} = z^T x - \alpha\left(\sqrt{\sum_{i=1}^N z_i^2} - \varepsilon\right) - \sum_{j=1}^N \beta_j(z_j-\varepsilon_j) - \sum_{j=1}^N \gamma_j(-z_j-\varepsilon_j)$. By strong duality, the following holds

$$\max_{\substack{z\in\mathbb{R}^N \\ ||z||_2\leq\varepsilon \\ |z_j|\leq\varepsilon_j}} z^T x = \max_{z\in\mathbb{R}^N} \min_{\substack{\alpha\in\mathbb{R}_{\{0,+\}} \\ \beta,\gamma\in\mathbb{R}^N_{\{0,+\}}}} \mathscr{L} = \min_{\substack{\alpha\in\mathbb{R}_{\{0,+\}} \\ \beta,\gamma\in\mathbb{R}^N_{\{0,+\}}}} \max_{z\in\mathbb{R}^N} \mathscr{L},$$

where the following must be true at the point of maximum over z: $\frac{\partial L(z,\alpha,\beta,\gamma)}{\partial z_j} = x_j - \beta_j + \gamma_j - \alpha\frac{z_j}{\|z\|_2} = 0$, $\forall j = 1,\ldots,N$. Substituting this equation into the Lagrangian, we get

$$\mathscr{L}(z,\alpha,\beta,\gamma) = \alpha\varepsilon + \sum_{j=1}^{N}(\beta_j+\gamma_j)\varepsilon_j. \tag{A3}$$

Now, let us make the following change of variables:

$$\lambda_j = \alpha\frac{z_j}{\|z\|_2},\ \forall j = 1,\ldots,N,$$
$$\mu_j = \beta_j - \gamma_j,\ \forall j = 1,\ldots,N.$$

Note that $\alpha = \|\lambda\|_2$ and that at the point of minimum over α, β, γ the term $\beta_j+\gamma_j$ behaves as $|\mu_j|$. This happens because at optimality $\beta_j=\mu_j,\ \gamma_j=0$ if $\mu_j\geq 0$ and $\beta_j=0,\ \gamma_j=-\mu_j$ if $\mu_j\leq 0$.

Hence, equation (A3) implies the statement of Lemma 2 under the proposed change of variables.

LEMMA 3. (see Theorem 3.1 in [18]) For $a_i \in \mathbb{R}^{n_i},\ \forall i = 0,\dots,N,\ \xi_j \in \mathbb{R}^{n_0\times n_j},\ j = 1,\dots,N$ the following holds:

$$\max_{\substack{\|\xi_1\|_F \le \varepsilon^{(\xi_1)}\\ \|\xi_2\|_F \le \varepsilon^{(\xi_2)}\\ \dots\\ \|\xi_N\|_F \le \varepsilon^{(\xi_N)}}} \left\|a_0 + \sum_{i=1}^N \xi_i a_i\right\|_2 = \|a_0\|_2 + \sum_{i=1}^N \varepsilon^{(\xi_i)} \|a_i\|_2.$$

Proof

$$\begin{aligned}
&\|a_0 + \sum_{i=1}^N \xi_i a_i\|_2^2 = \left(a_0 + \sum_{i=1}^N \xi_i a_i\right)^T \left(a_0 + \sum_{i=1}^N \xi_i a_i\right) = \\
&\le \|a_0\|_2^2 + \sum_{i=1}^N \left(\varepsilon^{(\xi_i)}\right)^2 \|a_i\|_2^2 + 2\|a_0\|_2 \sum_{j=1}^N \varepsilon^{(\xi_j)} \|a_j\|_2 + \\
&+2\sum_{i=1}^N \sum_{j=i+1}^N \varepsilon^{(\xi_i)}\varepsilon^{(\xi_j)} \|a_i\|_2 \|a_j\|_2 = \left(\|a_0\|_2 + \sum_{i=1}^N \varepsilon^{(\xi_i)} \|a_i\|_2\right)^2.
\end{aligned}$$

Hence, $\|a_0 + \sum_{i=1}^N \xi_i a_i\|_2 \le \|a_0\|_2 + \sum_{i=1}^N \varepsilon^{(\xi_i)} \|a_i\|_2$. Equality holds if $\xi_i = \xi_i^* = \frac{\varepsilon^{(\xi_i)} a_0 a_i^T}{\|a_0\|_2 \|a_i\|_2}$ for $a_0 \ne 0$.

A.1. Proof of Theorem 1

Proof Let $u^T = \left(u_1^T,\ u_2^T\right)$, where u_1 is a vector of the length N and u_2 is a vector of the length M. The following equality holds due to the duality of the second norm:

$$||Qx - x||_2 = \left\| \begin{pmatrix} (P+\xi - I_N)x^{(1)} + \zeta x^{(2)} \\ \psi x^{(1)} + (\chi - I_M)x^{(2)} \end{pmatrix} \right\|_2 = \max_{\substack{u \in \mathbb{R}^{N+M}\\ ||u||_2 \le 1}} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}^T \begin{pmatrix} (P+\xi - I_N)x^{(1)} + \zeta x^{(2)} \\ \psi x^{(1)} + (\chi - I_M)x^{(2)} \end{pmatrix}, \tag{A4}$$

where I_N and I_M are identity matrices of sizes $N\times N$ and $M \times M$ correspondingly.

To compute the lower bound for the norm $||Qx - x||_2$, we choose such feasible $u_1 = u_1^*$, that $\|(P+\xi)x^{(1)} - x^{(1)}\|_2 = (u_1^*)^T(P+\xi - I_N)x^{(1)},\ u_1^* \in \mathbb{R}^N,\ \|u_1^*\|_2 \le 1$ (which exists due to the duality of the norm) and we fix $u_2 = \mathbb{0}_M$ (i.e., zero-vector of the length M). In this case, we can write the following:

$$\max_{Q \in \Xi^{(l_2)}} ||Qx - x||_2 \ge \|(P+\xi)x^{(1)} - x^{(1)}\|_2 + (u_1^*)^T \zeta x^{(2)},\ \forall\ \xi \text{ and } \zeta \in \Xi^{(l_2)}. \tag{A5}$$

Given $\varepsilon_j^{(\chi)} \ge 1\ \forall j$ and $\varepsilon^{(\chi)} \ge 1$, we can set $\zeta = 0$, as the bounds hold for all ξ, ζ in the perturbation set. Thus, the lower bound (6) holds. Note that conditions $\varepsilon_j^{(\chi)} = 1\ \forall j$ and $\varepsilon^{(\chi)} = 1$ are necessary and sufficient to include a column-stochastic matrix χ (and, thus, $\zeta = 0$) to the feasibility set, while the matrix $\chi = \left(\frac{1}{M}\right)_{i,j=1,\dots,M}$ among column-stochastic matrices provides the minimal Frobenius norm. Too low values of uncertainty parameters $\varepsilon_j^{(\chi)}$ and $\varepsilon^{(\chi)}$ would result in a violation of the column-stochasticity constraints on the matrix Q, e.g., column-wise sums $\|[\zeta]_j\|_1 + \|[\chi]_j\|_1 = 1$ could become infeasible.

Note also that the equality holds in (6) if $x^{(2)} = \mathbb{0}_M$ and $\psi = 0$ in line with (A4). Differently, if $x^{(1)} \ne \mathbb{0}_N$, $x^{(2)} \ne \mathbb{0}_M$ and $\psi \ne 0$ and if we relax the non-negativity constraint $\zeta \ge 0$ similar to the article [25], the inequality (A5) becomes strict. Indeed, $\zeta = -r\left(\frac{1}{N}\right)_{i=1,\dots,N;\ j=1,\dots,M}$, where $r \le 1$ is a constant, guarantees that $(u_1^*)^T \zeta x^{(2)} > 0$. This ζ is feasible as r can be chosen arbitrarily small to satisfy $\varepsilon_j^{(\zeta)} > 0\ \forall j$ and $\varepsilon^{(\zeta)} > 0$.

A.2. Proof of Theorem 2

Proof Let $u^T = \left(u_1^T,\, u_2^T\right)$, where u_1 is a vector of the length N and u_2 is a vector of the length M. Note that the following equality holds due to the duality of the second norm:

$$||Qx-x||_2 = \left\| \begin{pmatrix} P+\xi & \zeta \\ \psi & \chi \end{pmatrix} \begin{pmatrix} x^{(1)} \\ x^{(2)} \end{pmatrix} - \begin{pmatrix} x^{(1)} \\ x^{(2)} \end{pmatrix} \right\|_2 = \left\| \begin{pmatrix} (P+\xi-I_N)x^{(1)}+\zeta x^{(2)} \\ \psi x^{(1)}+(\chi-I_M)x^{(2)} \end{pmatrix} \right\|_2 =$$
$$= \max_{\substack{u\in\mathbb{R}^{N+M} \\ ||u||_2\le 1}} \left(u_1^T(P-I_N)x^{(1)} + \left(u_1^T\xi + u_2^T\psi\right)x^{(1)} + \left(u_1^T\zeta + u_2^T(\chi - I_M)\right)x^{(2)} \right), \tag{A6}$$

where I_N and I_M are identity matrices of sizes $N\times N$ and $M\times M$ correspondingly.

Furthermore, one can bound the norm $||Qx-x||_2$ from above based on the triangle inequality and norm duality, i.e., $||Qx-x||_2 \le \|Px^{(1)}-x^{(1)}\|_2 + \max_{\substack{u\in\mathbb{R}^{N+M} \\ ||u||_2\le 1}} \left(u_1^T\xi + u_2^T\psi\right)x^{(1)} + \max_{\substack{u\in\mathbb{R}^{N+M} \\ ||u||_2\le 1}} \left(u_1^T\zeta + u_2^T(\chi-I_M)\right)x^{(2)}$, which would lead to the upper bound for the value $v = \max_{Q\in\Xi^{(l_2)}} \|Qx-x\|_2$.

$$v \le \|Px^{(1)}-x^{(1)}\|_2 + \max_{\substack{\|[\xi]_j\|_1\le\varepsilon_j^{(\xi)} \\ \|\xi\|_F\le\varepsilon^{(\xi)} \\ \|[\psi]_j\|_1\le\varepsilon_j^{(\psi)} \\ \|\psi\|_F\le\varepsilon^{(\psi)}}} \max_{\substack{u\in\mathbb{R}^{N+M} \\ ||u||_2\le 1}} \left(u_1^T\xi + u_2^T\psi\right)x^{(1)} + \max_{\substack{\|[\zeta]_j\|_1\le\varepsilon_j^{(\zeta)} \\ \|\zeta\|_F\le\varepsilon^{(\zeta)} \\ \|[\chi]_j\|_1\le\varepsilon_j^{(\chi)} \\ \|\chi\|_F\le\varepsilon^{(\chi)}}} \max_{\substack{u\in\mathbb{R}^{N+M} \\ ||u||_2\le 1}} \left(u_1^T\zeta + u_2^T(\chi-I_M)\right)x^{(2)}. \tag{A7}$$

Considering the subproblem $g_1(x) = \max_{\substack{\|[\xi]_j\|_1\le\varepsilon_j^{(\xi)},\, \|\xi\|_F\le\varepsilon^{(\xi)} \\ \|[\psi]_j\|_1\le\varepsilon_j^{(\psi)},\, \|\psi\|_F\le\varepsilon^{(\psi)}}} \max_{\substack{u\in\mathbb{R}^{N+M} \\ ||u||_2\le 1}} \left(u_1^T\xi + u_2^T\psi\right)x^{(1)}$ and denoting $z = \xi^T u_1 + \psi^T u_2$, we reformulate it based on the conditions of the problem (5):

$$|z_j| = \Big|\sum_{i=1}^N \xi_{ij}u_{1i} + \sum_{i=1}^M \psi_{ij}u_{2i}\Big| \le \sum_{i=1}^N \left|\xi_{ij}\right| + \sum_{i=1}^M \left|\psi_{ij}\right| \le \varepsilon_j^{(\xi)} + \varepsilon_j^{(\psi)},$$
$$\|z\|_2 \le \|\xi\|_F\|u_1\|_2 + \|\psi\|_F\|u_2\|_2 \le \varepsilon^{(\xi)} + \varepsilon^{(\psi)},$$

where $u_{1i},\ \forall i = 1,\ldots,N$ and $u_{2i},\ \forall i = 1,\ldots,M$ are i-th coordinates of vectors u_1 and u_2 correspondingly.

Therefore, $g_1(x) \le \max_{\substack{z\in\mathbb{R}^N \\ ||z||_2\le\varepsilon^{(\xi)}+\varepsilon^{(\psi)} \\ |z_j|\le\varepsilon_j^{(\xi)}+\varepsilon_j^{(\psi)}}} z^T x^{(1)} = \left(\varepsilon^{(\xi)}+\varepsilon^{(\psi)}\right) \max_{\substack{z\in\mathbb{R}^N \\ ||z||_2\le 1 \\ |z_j|\le\frac{\varepsilon_j^{(\xi)}+\varepsilon_j^{(\psi)}}{\varepsilon^{(\xi)}+\varepsilon^{(\psi)}}}} z^T x^{(1)}$, which is a strictly feasible conic optimization problem. Dualizing the constraints (see Lemma 2 in the Appendix for details) we obtain

$$g_1(x) \le \left(\varepsilon^{(\xi)}+\varepsilon^{(\psi)}\right) \min_{\lambda+\mu=x^{(1)}} \left\{ \|\lambda\|_2 + \sum_{j=1}^N \frac{\varepsilon_j^{(\xi)}+\varepsilon_j^{(\psi)}}{\varepsilon^{(\xi)}+\varepsilon^{(\psi)}} |\mu_j| \right\}. \tag{A8}$$

Note that $g_1(x)$ upper bound is a norm.

Analogically, we bound the problem $g_2(x) = \max_{\substack{\|[\zeta]_j\|_1\le\varepsilon_j^{(\zeta)},\, \|\zeta\|_F\le\varepsilon^{(\zeta)} \\ \|[\chi]_j\|_1\le\varepsilon_j^{(\chi)},\, \|\chi\|_F\le\varepsilon^{(\chi)}}} \max_{\substack{u\in\mathbb{R}^{N+M} \\ ||u||_2\le 1}} \left(u_1^T\zeta + u_2^T(\chi-I_M)\right)x^{(2)}$ by its dualization. Let $z = \zeta^T u_1 + (\chi^T - I_M)u_2$. Based on the conditions of the problem (5), we can write

$$|z_j| = \left|\sum_{i=1}^N \zeta_{ij}u_{1i} + \sum_{i=1}^M \chi_{ij}u_{2i} - u_{2j}\right| \le \sum_{i=1}^N \left|\zeta_{ij}\right| + \sum_{i=1}^M \left|\chi_{ij}\right| + 1 \le \varepsilon_j^{(\zeta)} + \varepsilon_j^{(\chi)} + 1,$$
$$\|z\|_2 \le \|\zeta\|_F\|u_1\|_2 + \|\chi\|_F\|u_2\|_2 + \|u_2\|_2 \le \varepsilon^{(\zeta)} + \varepsilon^{(\chi)} + 1.$$

Therefore, $g_2(x) \le \max\limits_{\substack{z\in\mathbb{R}^M \\ ||z||_2 \le \varepsilon^{(\zeta)}+\varepsilon^{(\chi)}+1 \\ |z_j| \le \varepsilon_j^{(\zeta)}+\varepsilon_j^{(\chi)}+1}} z^T x^{(2)} = \left(\varepsilon^{(\zeta)}+\varepsilon^{(\chi)}+1\right) \max\limits_{\substack{z\in\mathbb{R}^M \\ ||z||_2\le 1 \\ |z_j| \le \frac{\varepsilon_j^{(\zeta)}+\varepsilon_j^{(\chi)}+1}{\varepsilon^{(\zeta)}+\varepsilon^{(\chi)}+1}}} z^T x^{(2)}$, which is a strictly feasible conic optimization problem. Dualizing the constraints (see Lemma 2 in the Appendix for details) we obtain:

$$g_2(x) \le \left(\varepsilon^{(\zeta)}+\varepsilon^{(\chi)}+1\right) \min_{\lambda+\mu=x^{(2)}} \left\{ \|\lambda\|_2 + \sum_{j=1}^{M} \frac{\varepsilon_j^{(\zeta)}+\varepsilon_j^{(\chi)}+1}{\varepsilon^{(\zeta)}+\varepsilon^{(\chi)}+1} |\mu_j| \right\}. \tag{A9}$$

Equations (A8) and (A9) imply the statement of Proposition 2.

A.3. Proof of Lemma 1

Proof Consider the problem (8). Based on the fact that $\|x^{(1)}\|_1 + \|x^{(2)}\|_1 = 1$, $x^{(1)} \ge 0$, $x^{(2)} \ge 0$, let us denote the total importance of current pages by $s = ||x^{(1)}||_1$ with $s \in [0,\ 1]$ and change variables as:

$$\begin{aligned} x^{(1)} &= s y^{(1)}, \\ x^{(2)} &= (1-s) y^{(2)}, \end{aligned} \tag{A10}$$

where $\|y^{(1)}\|_1 = 1$, $\|y^{(2)}\|_1 = 1$ and $y^{(1)} \ge 0$, $y^{(2)} \ge 0$. As before, $\|x^{(1)}\|_1 = s$ and $\|x^{(2)}\|_1 = 1-s$, $\forall s \in [0,\ 1]$. Hence, simplex constraints on $x^T = \left(x^{(1)^T}\ x^{(2)^T}\right)$ are satisfied.

Therefore, the optimization problem (8) can be rewritten as:

$$\tilde{y} \in \operatorname*{Argmin}_{y^{(1)}\in\Sigma_N,\, y^{(2)}\in\Sigma_M,\, s\in[0,\ 1]} \left\{ s\left(\|Py^{(1)} - y^{(1)}\|_2 + \varepsilon_1 \|y^{(1)}\|_{(a)} \right) + (1-s)\left(\varepsilon_2 \|y^{(2)}\|_{(b)} \right) \right\}, \tag{A11}$$

where $\tilde{y}^T = \left(\tilde{y}^{(1)^T}\ \tilde{y}^{(2)^T}\right)$ and s^* denote the optimal solution of the optimization problem (A11). Furthermore, $\tilde{y}^{(1)}$ and $\tilde{y}^{(2)}$ are independent of each other.

At optimality of the problem (A11), $s^* = 1$ if $\|P\tilde{y}^{(1)} - \tilde{y}^{(1)}\|_2 + \varepsilon_1 \|\tilde{y}^{(1)}\|_{(a)} < \varepsilon_2 \|\tilde{y}^{(2)}\|_{(b)}$ and $s^* = 0$ if $\|P\tilde{y}^{(1)} - \tilde{y}^{(1)}\|_2 + \varepsilon_1 \|\tilde{y}^{(1)}\|_{(a)} > \varepsilon_2 \|\tilde{y}^{(2)}\|_{(b)}$. In case $\|P\tilde{y}^{(1)} - \tilde{y}^{(1)}\|_2 + \varepsilon_1 \|\tilde{y}^{(1)}\|_{(a)} = \varepsilon_2 \|\tilde{y}^{(2)}\|_{(b)}$, s^* can take any value in the interval $[0,\ 1]$.

Hence, the statement of Lemma 1 follows. By comparing optimal values of problems (9) and (10), one can conclude the optimal solution s^* and, therefore, one can obtain the optimal solution via the system (A10).

A.4. Proof of Statement 1

Proof Consider the optimization problem

$$\min_{y^{(2)}\in\Sigma_M} \|y^{(2)}\|_{(b)} = \min_{\substack{y^{(2)}\in\Sigma_M \\ \lambda+\mu=y^{(2)}}} \left\{ \|\lambda\|_2 + \sum_{j=1}^{M} \frac{\varepsilon_j^{(\zeta)}+\varepsilon_j^{(\chi)}+1}{\varepsilon^{(\zeta)}+\varepsilon^{(\chi)}+1} |\mu_j| \right\},$$

which is equivalent to the following convex optimization problem:

$$\min_{y^{(2)}\in\Sigma_M} \|y^{(2)}\|_{(b)} = \min_{\substack{y^{(2)}\ge 0 \\ \sum_{j=1}^M y_j^{(2)}=1 \\ \lambda+\mu=y^{(2)} \\ \mu_j\le v_j,\, \forall j \\ -\mu_j \le v_j,\, \forall j}} \left\{ \|\lambda\|_2 + \sum_{j=1}^{M} \frac{\varepsilon_j^{(\zeta)}+\varepsilon_j^{(\chi)}+1}{\varepsilon^{(\zeta)}+\varepsilon^{(\chi)}+1} v_j \right\}. \tag{A12}$$

Let us denote $c_j = \frac{\varepsilon_j^{(\zeta)}+\varepsilon_j^{(\chi)}+1}{\varepsilon^{(\zeta)}+\varepsilon^{(\chi)}+1}$ and solve the optimization problem (A12) by its dualization.

The Lagrangian $\mathscr{L}$ for the problem (A12) is

$$\mathscr{L} = \|\lambda\|_2 + \sum_{j=1}^{M} c_j v_j + \alpha\left(\sum_{j=1}^{M} y_j^{(2)} - 1\right) + \sum_{j=1}^{M}\left(\beta_j(\lambda_j+\mu_j-y_j^{(2)}) + \gamma_j(\mu_j - v_j) + \eta_j(-\mu_j - v_j)\right),$$

where dual variables are $\alpha \in \mathbb{R}$, $\beta \in \mathbb{R}^M$, $\gamma, \eta \in \mathbb{R}^M_{\{0,+\}}$.

Differently, one can rewrite the Lagrangian in the following form:

$$\mathscr{L} = \left(\|\lambda\|_2 + \sum_{j=1}^{M}\beta_j\lambda_j\right) + \sum_{j=1}^{M} v_j\left(c_j - \gamma_j - \eta_j\right) + \sum_{j=1}^{M} y_j^{(2)}\left(\alpha - \beta_j\right) + \sum_{j=1}^{M}\mu_j\left(\beta_j + \gamma_j - \eta_j\right) - \alpha.$$

By strong duality (guaranteed by Slater's condition), the primal and the dual problems are equivalent:

$$\min_{y^{(2)}\in\Sigma_M} \|y^{(2)}\|_{(b)} = \min_{\substack{y^{(2)}\in\mathbb{R}^M_{\{0,+\}}\\ v\in\mathbb{R}^M_{\{0,+\}}\\ \lambda,\mu\in\mathbb{R}^M}} \max_{\substack{\alpha\in\mathbb{R}\\ \beta\in\mathbb{R}^M\\ \gamma,\eta\in\mathbb{R}^M_{\{0,+\}}}} \mathscr{L} = \max_{\substack{\alpha\in\mathbb{R}\\ \beta\in\mathbb{R}^M\\ \gamma,\eta\in\mathbb{R}^M_{\{0,+\}}}} \min_{\substack{y^{(2)}\in\mathbb{R}^M_{\{0,+\}}\\ v\in\mathbb{R}^M_{\{0,+\}}\\ \lambda,\mu\in\mathbb{R}^M}} \mathscr{L}.$$

For all j such that $\beta_j = 0$, we can conclude that $\lambda_j = 0$ at optimality. For all j such that $\beta_j \neq 0$, optimal λ_j must follow the equation: $\frac{\partial\mathscr{L}}{\partial\lambda_j} = \frac{\lambda_j}{\|\lambda\|_2} + \beta_j = 0,\ \forall j = 1,\ldots,N$. From these follows $\sum_{j:\,\beta_j\neq 0}\frac{\lambda_j^2}{\|\lambda\|_2^2} = \sum_{j:\,\beta_j\neq 0}\beta_j^2 \leq 1$, where the last inequality holds due to the fact that $\sum_{j=1}^{M}\frac{\lambda_j^2}{\|\lambda\|_2^2} = 1$ (i.e., the equality holds if $\beta_j \neq 0,\ \forall j$). Hence, the following conditions must be satisfied in order to guarantee feasibility of the optimization problem:

$$\begin{cases} \sum_{j=1}^{M}\beta_j^2 \leq 1, \\ \beta_j + \gamma_j - \eta_j = 0,\ \forall j = 1,\ldots,N, \\ c_j - \gamma_j - \eta_j \geq 0,\ \forall j = 1,\ldots,N, \\ \alpha - \beta_j \geq 0,\ \forall j = 1,\ldots,N, \end{cases}$$

Therefore,

$$\min_{y^{(2)}\in\Sigma_M}\|y^{(2)}\|_{(b)} = \max_{\substack{\tilde{a}\in\mathbb{R}\\ \tilde{b}\in\mathbb{R}^M\\ \gamma,\eta\in\mathbb{R}^M_{\{0,+\}}}} \Bigg\{\tilde{a},\ \text{subject to} \\ \sum_{j=1}^{M}\tilde{b}_j^2 \leq 1,\ c_j \geq \eta_j + \gamma_j,\ \forall j = 1,\ldots,N, \\ \tilde{a} \leq \tilde{b}_j,\ \forall j = 1,\ldots,N,\ \tilde{b}_j = \gamma_j - \eta_j,\ \forall j = 1,\ldots,N\Bigg\},$$

where we denoted $\tilde{a} = -\alpha$ and $\tilde{b}_j = -\beta_j$.

At optimality, $\eta_j = 0$. Hence, one can rewrite the optimization problem in the following equivalent form:

$$\min_{y^{(2)}\in\Sigma_M}\|y^{(2)}\|_{(b)} = \max_{\substack{\tilde{a}\in\mathbb{R}\\ \tilde{b}\in\mathbb{R}^M}}\left\{\tilde{a},\ \text{subject to}\ \sum_{j=1}^{M}\tilde{b}_j^2 \leq 1,\ \tilde{b}_j \leq c_j,\ \tilde{a} \leq \tilde{b}_j,\ \forall j = 1,\ldots,N\right\},$$

whose solution can be summarized as follows:

$$\min_{y^{(2)}\in\Sigma_M}\|y^{(2)}\|_{(b)} = \left[\begin{array}{l}\frac{1}{\sqrt{M}},\ \text{if}\ c_j \geq \frac{1}{\sqrt{M}},\ \forall j = 1,\ldots,M \\ \min_j\{c_j\},\ \text{if}\ \min_j\{c_j\} < \frac{1}{\sqrt{M}},\end{array}\right. \tag{A13}$$

which leads to the Statement (1).

A.5. Proof of Corollary 1

Proof Fix $c \in [0,1]$ and choose feasible $u_1 = (1-c)\frac{(P+\xi-I_N)x^{(1)}}{\|(P+\xi-I_N)x^{(1)}\|_2}$ and $u_2 = c\frac{\psi x^{(1)}}{\|\psi x^{(1)}\|_2}$ in equation (A4), implying $\max_{Q\in\Xi^{(l_2)}} ||Qx-x||_2 \geq u_1^T(P+\xi-I_N)x^{(1)} + u_1^T\zeta x^{(2)} + u_2^T\psi x^{(1)} + u_2^T(\chi - I_M)x^{(2)}$, which holds $\forall\, \xi,\, \zeta,\, \psi,\, \chi \in \Xi^{(l_2)}$.

As $\chi = I_M$ is feasible given $\varepsilon_j^{(\chi)} \geq 1$, $\forall j$ and $\varepsilon^{(\chi)} \geq \sqrt{M}$, we conclude that $\zeta = 0$ due to the column-stochasticity of matrix Q and obtain $\max_{Q\in\Xi^{(l_2)}} ||Qx-x||_2 \geq (1-c)\|(P+\xi)x^{(1)} - x^{(1)}\|_2 + c\|\psi x^{(1)}\|_2$. Further, minimizing lower and upper bounds of the optimization problem (5) under additional constraint $\mathbb{1}_N^T x^{(1)} = s$ and $\mathbb{1}_M^T x^{(2)} = 1-s$ with $s \in [0,1]$, we obtain

$$v_1(s) = \min_{\substack{x\geq 0\\ \mathbb{1}_N^T x^{(1)}=s\\ \mathbb{1}_M^T x^{(2)}=1-s}} \left\{\|Px^{(1)} - x^{(1)}\|_2 + \varepsilon_1\|x^{(1)}\|_{(a)} + \varepsilon_2||x^{(2)}||_{(b)}\right\} \geq \min_{\substack{x\geq 0\\ \mathbb{1}_N^T x^{(1)}=s\\ \mathbb{1}_M^T x^{(2)}=1-s}} \max_{Q\in\Xi^{(l_2)}} ||Qx-x||_2, \tag{A14}$$

$$v_2(s) = \min_{\substack{x\geq 0\\ \mathbb{1}_N^T x^{(1)}=s\\ \mathbb{1}_M^T x^{(2)}=1-s}} \left\{(1-c)\|(P+\xi)x^{(1)} - x^{(1)}\|_2 + c\|\psi x^{(1)}\|_2\right\} \leq \min_{\substack{x\geq 0\\ \mathbb{1}_N^T x^{(1)}=s\\ \mathbb{1}_M^T x^{(2)}=1-s}} \max_{Q\in\Xi^{(l_2)}} ||Qx-x||_2, \tag{A15}$$

where we denote optimal values depending on s by $v_1(s)$ and $v_2(s)$. Therefore, we compute the optimality gap as $\Delta(s) \leq v_1(s) - v_2(s)$.

Now, let us make the change of variables (A10) in equation (A14). In these notations, we write

$$\min_{\substack{x\geq 0\\ \mathbb{1}_N^T x^{(1)}=s\\ \mathbb{1}_M^T x^{(2)}=1-s}} \left\{\|Px^{(1)} - x^{(1)}\|_2 + \varepsilon_1\|x^{(1)}\|_{(a)} + \varepsilon_2||x^{(2)}||_{(b)}\right\} = s\tilde{\phi} + (1-s)\varepsilon_2 \min_{\substack{y^{(2)}\geq 0\\ \mathbb{1}_M^T y^{(2)}=1}} \|y^{(2)}\|_{(b)}. \tag{A16}$$

The equality holds because of the Lemma 1, equations (12) and (A11) given $\varepsilon_j^{(\chi)} \geq 1\ \forall j$, $\varepsilon^{(\chi)} \geq \sqrt{M}$, as

$$\min_{\substack{y^{(1)}\geq 0\\ \mathbb{1}_M^T y^{(1)}=1}} \left\{\|Py^{(1)} - y^{(1)}\|_2 + \varepsilon_1\|y^{(1)}\|_{(a)}\right\} = \min_{\substack{y^{(1)}\in\Sigma_N,\, y^{(2)}\in\Sigma_M\\ s\in[0,1]}} \left\{s\left(\|Py^{(1)} - y^{(1)}\|_2 + \varepsilon_1\|y^{(1)}\|_{(a)}\right) + (1-s)\varepsilon_2||y^{(2)}||_{(b)}\right\} = \tilde{\phi}.$$

Based on the equation (A16) and the first statement of Lemma 1, we can conclude that the upper bound (A14) is decreasing in s because $\tilde{\phi} \leq \varepsilon_2 \min_{\substack{y^{(2)}\geq 0\\ \mathbb{1}_M^T y^{(2)}=1}} \|y^{(2)}\|_{(b)}$. Conversely, the lower bound (A15) is increasing in s, as it depends only on the vector of components $x^{(1)}$. Thus, if the optimal sum $\mathbb{1}_N^T x^{(1)}$ denoted by s^* in the optimization problem (5) is greater than the estimated value of s (i.e., if $s < s^*$), we can conclude that $v_1(s) > v_1(s^*)$ and $v_2(s) < v_2(s^*)$. From this follows that $\Delta(s^*) \leq v_1(s^*) - v_2(s^*) < v_1(s) - v_2(s)$. Hence, the optimality gap between problems (5) and (8) is bounded if the value of s is underestimated while the first statement of Lemma 1 still holds. Clearly, if $s = s^*$, the bounds for the optimality gaps coincide.

Differently, the relationship $s > s^*$ implies that $v_1(s) < v_1(s^*)$ and $v_2(s) > v_2(s^*)$, from which follows $\Delta(s) \leq v_1(s) - v_2(s) < v_1(s^*) - v_2(s^*)$, and the valid bounds for the gap $\Delta(s^*)$ are $\tilde{\phi}$ and $v_1(s^*) - v_2(s^*)$. Overall, this may seem to imply that the decision-maker is better off choosing the minimal possible value of the sum s.

Nevertheless, it is not always correct, as we can bound the difference $\delta_v = \left[v_1(s^*) - v_1(s)\right] - \left[v_2(s^*) - v_2(s)\right]$ in terms of $\delta_s = s - s^*$ if $s > s^*$ based on the fact that $v_2(s) = sv_2(1)$ and equation (A16):

$$0 < \delta_v = \delta_s\left(v_2(1) - \tilde{\phi} + \varepsilon_2 \min_{\substack{y^{(2)}\geq 0\\ \mathbb{1}_M^T y^{(2)}=1}} \|y^{(2)}\|_{(b)}\right) =$$
$$= \delta_s\left(\varepsilon_2 \min\left\{\frac{1}{\sqrt{M}},\ \frac{\varepsilon_j^{(\zeta)} + \varepsilon_j^{(\chi)} + 1}{\varepsilon^{(\zeta)} + \varepsilon^{(\chi)} + 1}\ \forall j\right\} - \left[v_1(1) - v_2(1)\right]\right). \quad \text{(A17)}$$

This leads to the upper bound for the optimality gap $\Delta(s^*)$ between optimization problems (5) and (8), which holds $\forall s^* \in [0,1]$, where $s = 1$ is, clearly, an overestimation:

$$\Delta(s^*) \leq (1-\delta_1)(v_1(1) - v_2(1)) + \delta_1 \min\left\{\frac{\varepsilon_2}{\sqrt{M}},\ \varepsilon_j^{(\zeta)} + \varepsilon_j^{(\chi)} + 1\ \forall j\right\},\ \text{where } \delta_1 = 1 - s^*. \quad \text{(A18)}$$

The bound (A18) arises from the inequality $\Delta(s^*) \leq (v_1(s) - v_2(s)) + \delta_v\ \forall s > s^*$ taken with $s = 1$ and attains its maximal value at the point $\delta_1 = 1$ due to Lemma 1 (point 1). Importantly, due to the use of Propositions 1 and 2, the bound (A18) can provide a better guarantee for the gap $\Delta(s^*)$ than $v_1(s^*) - v_2(s^*)$ (Figure 17(c)). Next, note that the bound (A17) drops when s approaches s^*. Furthermore, if $\varepsilon_2 = \varepsilon^{(\chi)} + \varepsilon^{(\zeta)} + 1$ would be independent of the number of new pages M, the value δ_v would decrease in M, i.e., $\delta_v \to 0$ as $M \to \infty$. Also, the optimal values $v_1(1)$ and $v_2(1)$ would decrease in this case. This is due to the fact that $v_2(1) < \tilde{\phi} < \varepsilon_2 \min_{\substack{y^{(2)}\geq 0\\ \mathbb{1}_M^T y^{(2)}=1}} \|y^{(2)}\|_{(b)}$. Importantly, the convergence of δ_v to zero would imply $\Delta(s^*) \to 0$ if $M \to \infty$ in line with the bound (A18). Nevertheless, the bounds (A17) and (A18) do not necessarily converge to zero as ε_2 depends on M in the statement of Corollary 1. In Section 8 we analyze the case with $\varepsilon^{(\chi)} = \sqrt{M}$, $\varepsilon^{(\zeta)} = 0$.

A.6. Proof of Corollary 2

Proof In line with lower and upper bounds (6), (7) and (A5), we can write

$$\|(P + \xi - I_N)\tilde{x}^{(1)}\|_2 \leq \max_{Q \in \Xi^{(l_2)}} ||Q\tilde{x} - \tilde{x}||_2 \leq \tilde{\phi} = ||P\tilde{x}^{(1)} - \tilde{x}^{(1)}||_2 + \varepsilon_1 \|\tilde{x}^{(1)}\|_{(a)} + \varepsilon_2 ||\tilde{x}^{(2)}||_{(b)},$$

where we denote the optimal value of the upper bound $\tilde{\phi}(\tilde{x})$ by $\tilde{\phi}$.
In line with equation (A6), we write $\forall\ \xi,\ \zeta,\ \psi,\ \chi \in \Xi^{(l_2)}$

$$\min_{x \in \Sigma_{N+M}} \max_{Q \in \Xi^{(l_2)}} ||Qx - x||_2 \geq u_1^T(P - I_N)\hat{x}^{(l_2),(1)} + \left(u_1^T \xi + u_2^T \psi\right)\hat{x}^{(l_2),(1)} + \left(u_1^T \zeta + u_2^T(\chi - I_M)\right)\hat{x}^{(l_2),(2)} \quad \text{(A19)}$$

where $u = (u_1,\ u_2)^T \in \mathbb{R}^{N+M}$, $\|u\|_2 \leq 1$ and $\hat{x}^{(l_2)} = (\hat{x}^{(l_2),(1)},\ \hat{x}^{(l_2),(2)})^T$ is the optimal solution of the optimization problem (5). Equation (A19) holds for $\zeta = 0$ as it belongs to the uncertainty set given $\varepsilon_j^{(\chi)} \geq 1$ and $\varepsilon^{(\chi)} \geq 1$. Note that we require $\varepsilon_j^{(\chi)} \geq 1$ and $\varepsilon^{(\chi)} \geq 1$, as it suffices for the feasibility of column-stochasticity constraints of the transition matrix Q, e.g., with matrix $\chi = \left(\frac{1}{M}\right)_{i,j=1,\ldots,M}$.
Now we consider the optimal solution $\tilde{x} = (\tilde{x}^{(1)}, \tilde{x}^{(2)})$ of the optimization problem (8) and the vector $u_1 = \frac{(P+\xi-I_N)\tilde{x}^{(1)}}{\|(P+\xi-I_N)\tilde{x}^{(1)}\|_2}$, $u_2 = \mathbb{0}_M$ which is feasible in equation (A19). Thus, we conclude

$$\min_{x \in \Sigma_{N+M}} \max_{Q \in \Xi^{(l_2)}} ||Qx - x||_2 \geq \frac{(\tilde{x}^{(1)})^T(P + \xi - I_N)^T(P + \xi - I_N)\hat{x}^{(l_2),(1)}}{\|(P + \xi - I_N)\tilde{x}^{(1)}\|_2} \geq \frac{(\tilde{x}^{(1)})^T(P + \xi - I_N)^T(P + \xi - I_N)\hat{x}^{(l_2),(1)}}{\tilde{\phi}},$$

which results in an implicit equation $(\tilde{x}^{(1)})^T(P+\xi-I_N)^T(P+\xi-I_N)\hat{x}^{(l_2),(1)} \leq \tilde{\phi}^2$. Note that this equation holds $\forall \xi \in \Xi^{(l_2)}$ and we do not require $\psi = 0$ ($\psi = 0$ would imply column-stochasticity of $P+\xi$).
Analogically, we obtain the bound $\tilde{\phi} \geq \|(\chi - I_M)\hat{x}^{(l_2),(2)}\|_2 \; \forall \, \chi \in \Xi^{(l_2)}$ by using $u_1 = \mathbb{0}_N$, $u_2 = \frac{(\chi - I_M)\hat{x}^{(l_2),(2)}}{\|(\chi - I_M)\hat{x}^{(l_2),(2)}\|_2}$ and $\psi = 0$. As the uncertainty set includes matrices χ with columns summing up to values less than one, the left-hand side is non-zero in the worst case. Further, one can state the bound $\tilde{\phi}^2 \geq (\tilde{x}^{(1)})^T\psi^T\left[\psi\hat{x}^{(l_2),(1)} + (\chi - I_M)\hat{x}^{(l_2),(2)}\right] \; \forall \, \psi, \chi \in \Xi^{(l_2)}$, if the uncertainty set is large enough, i.e., if $\varepsilon^{(\chi)} \geq \sqrt{M}$. This statement combines old and new components and is obtained via the use of vectors $u_1 = \mathbb{0}_N$, $u_2 = \frac{\psi\tilde{x}^{(1)}}{\|\psi\tilde{x}^{(1)}\|_2}$ in the inequality (A19).

A.7. Proof of Corollary 3

Proof Consider the following optimization problem with $\hat{x}^{(F)} = (\hat{x}^{(1,F)}, \hat{x}^{(2,F)})$

$$\hat{x}^{(F)} \in \underset{x \in \Sigma_{N+M}}{\text{Argmin}} \left\{ \max_{Q} ||Qx - x||_2 : \right.$$
$$\left. \mathbb{1}_{N+M}^T [Q]_j = 1 \; \forall j, \; ||\xi||_F \leq \varepsilon^{(\xi)}, \; ||\psi||_F \leq \varepsilon^{(\psi)}, \; ||\zeta||_F \leq \varepsilon^{(\zeta)}, \; ||\chi||_F \leq \varepsilon^{(\chi)} \right\}.$$

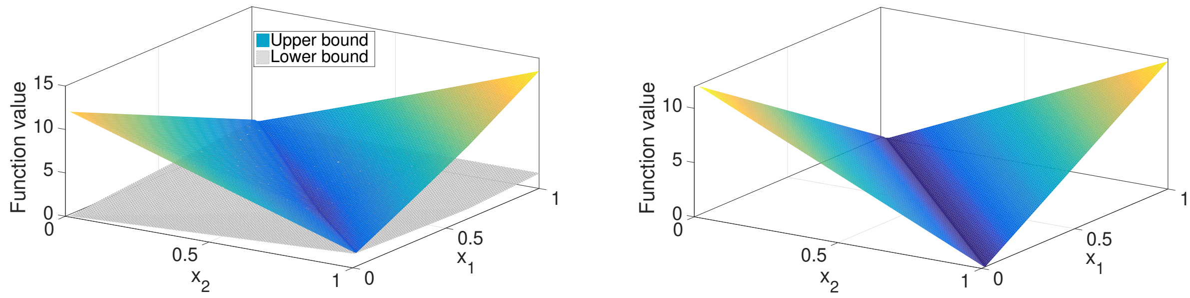
The first statement of Corollary 3 directly follows from equation (A6) and Lemma 3, implying

$$\max_{Q \in \Xi^{(F)}} ||Qx - x||_2 \leq \max_{\substack{||\xi||_F \leq \varepsilon^{(\xi)} \\ ||\zeta||_F \leq \varepsilon^{(\zeta)}}} \|(P+\xi-I_N)x^{(1)} + \zeta x^{(2)}\|_2 + \max_{\substack{||\psi||_F \leq \varepsilon^{(\psi)} \\ ||\chi||_F \leq \varepsilon^{(\chi)}}} \|\psi x^{(1)} + (\chi - I_M)x^{(2)}\|_2 =$$
$$= \|Px^{(1)} - x^{(1)}\|_2 + (\varepsilon^{(\xi)} + \varepsilon^{(\psi)})\|x^{(1)}\|_2 + (\varepsilon^{(\zeta)} + \varepsilon^{(\chi)} + 1)\|x^{(2)}\|_2.$$

Further, the following lower bound can be stated based on the equation (A4):

$$\max_{Q \in \Xi^{(F)}} ||Qx - x||_2 \geq \max_{\substack{\mathbb{1}_N^T[\xi]_j = 0, \, \mathbb{1}_N^T[\zeta]_j = 0 \\ \|\xi\|_F \leq \varepsilon^{(\xi)}, \, \|\zeta\|_F \leq \varepsilon^{(\zeta)}}} \|(P - I_N + \xi)x^{(1)} + \zeta x^{(2)}\|_2 = \|Px^{(1)} - x^{(1)}\|_2 + \varepsilon^{(\xi)}\|x^{(1)}\|_2 + + \varepsilon^{(\zeta)}\|x^{(2)}\|_2.$$

Note that the equality holds due to Lemma 3 with worst-case perturbations $\xi^* = \varepsilon^{(\xi)} \frac{(Px^{(1)} - x^{(1)})(x^{(1)})^T}{\|Px^{(1)} - x^{(1)}\|_2 \|x^{(1)}\|_2}$ and $\zeta^* = \varepsilon^{(\zeta)} \frac{(Px^{(1)} - x^{(1)})(x^{(2)})^T}{\|Px^{(1)} - x^{(1)}\|_2 \|x^{(2)}\|_2}$, for which conditions $\mathbb{1}_N^T[\xi^*]_j = 0$ and $\mathbb{1}_N^T[\zeta^*]_j = 0$ are satisfied due to column-stochasticity of matrix P. Condition $\varepsilon^{(\chi)} \geq 1$ suffices to account for column-wise constraints $\mathbb{1}_{N+M}^T[Q]_j = 1$, which could be violated if $\zeta = 0$ and $\varepsilon^{(\chi)}$ is too low.
The second statement of Corollary 3 follows if $\varepsilon^{(\psi)} = 0$ and $\min_{x^{(1)} \in \Sigma_N}\left(\|Px^{(1)} - x^{(1)}\|_2 + \varepsilon^{(\xi)}\|x^{(1)}\|_2\right) \leq \frac{\varepsilon^{(\zeta)}}{\sqrt{M}}$ in line with Lemma 1 as the minimal upper and lower bounds coincide in this case.
Figure A1 demonstrates a small-dimensional example for Corollary 3 with $N = 2$ and $M = 1$. Sampling a random transition matrix and computing both upper and lower bounds of the problem (17), we observe that the bounds obviously coincide at the points with $x^{(2)} = 0$.



(a) Upper and lower bounds. (b) Difference between the bounds.

Figure A1 **Optimality gap example.**

A.8. Solution process clarifications

In the solution process of the optimization problem (20) we assume that one page (page N) is known to have the highest rank [47] and, therefore, we use the nonstandard normalization $x_N = 1$ instead of $\sum_{i=1}^{N} x_i = 1$ in order to develop an efficient algorithm which does not iterate over the page N and which, thus, functions similarly to algorithms for unconstrained optimization. By this, we introduce the following problem:

$$z \in \operatorname*{Argmin}_{x \in \mathbb{X}} f(x), \; f(x) = \|Px - x\|_2 + \varepsilon \|x\|_2, \text{ where } \mathbb{X} = \{x \in \mathbb{R}^N, \; x_N = 1, \; x \geq 0\} \tag{A20}$$

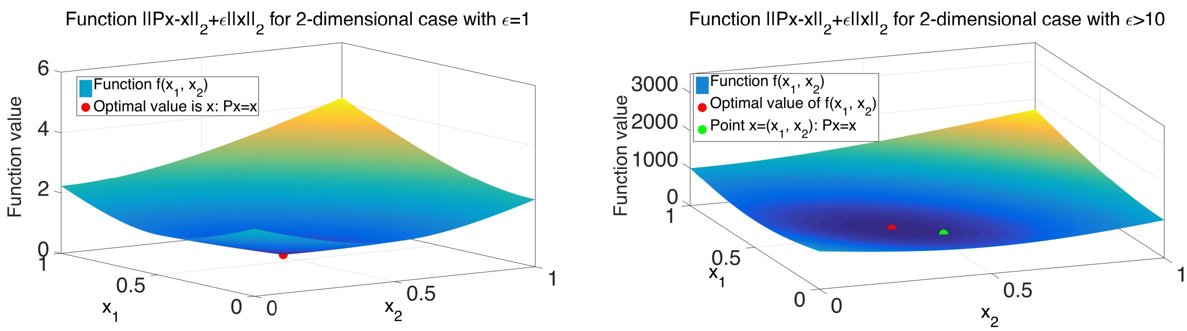
Note that the optimal value and the solution of the problem (20) are related to those of the problem (A20):

$$\begin{cases} f(\tilde{x}) = \tilde{x}_N f(z), \\ \tilde{x} = \tilde{x}_N z, \end{cases}$$

where $z = \left(\frac{\tilde{x}_1}{\tilde{x}_N}, \frac{\tilde{x}_2}{\tilde{x}_N}, \cdots, 1\right)^T$. This leads to the statement $\tilde{x}_N = \frac{1}{\sum_{i=1}^{N} z_i}$.

Next, we note that the optimization problem (A20) is non-smooth at least at one point $P\bar{x} = \bar{x}$, which is a possible optimal solution corresponding to the dominant (or principal) eigenvector of the non-perturbed matrix P. We check numerically if the optimal solution of the problem (A20) is always the point $\bar{x}$: $P\bar{x} = \bar{x}$. For this, we solve the optimization problem (A20) in small-dimensional cases with randomly chosen non-negative column-stochastic matrix P. In Figure A2, we observe that the following two cases are possible: (i) the optimal solution of the problem (A20) is the point $\bar{x}: \; P\bar{x} = \bar{x}$ (Figure A2 (a)), in which case the function $f(x)$ is non-smooth at optimality, and (ii) the optimal solution of the problem (A20) is a point $x: \; Px \neq x$ (Figure A2 (b)), in which case the function $f(x)$ is smooth at optimality.

As seen in Figure A2, the optimal solution of the problem (A20) is not necessarily the principal eigenvector of the matrix P. Therefore, our Algorithm 1 aims to find the solution for the problem (A20) and not directly the principle eigenvector of the matrix P (which would be possible via the use of algorithms proposed in [47]). Algorithm 1 reduces the value of the objective function (A20) at each iteration and converges to the point in some distance from the true optimal solution. In order to improve its performance, one could employ adaptive coordinate-wise algorithms, changing the matrix G at each iteration k, e.g., $G_k = \alpha_k P + (1 - \alpha_k) S$ with α_k adapted for the convergence similarly to the iterative regularization in [47].



(a) Non-smooth at optimality.

(b) Smooth at optimality.

Figure A2 Function $f(x)$ for randomly chosen P in small-dimensional cases.