Supplementary Material:

A Proof of Theorem 1 and Corollary 1

We now turn to the proof of Theorem 1. We will use the index $_\Omega$ to denote that we take a vectorized version of the elements of a matrix corresponding to the revealed entries. For example, A^R_Ω is thus a vector containing all the revealed entries, while A_Ω contains the real values of the entries that have been revealed. The 2-norm of such vectors is equivalent to their Frobenius norm; for example,

$$||\Delta||_F = \left|\left|A^R_\Omega - A_\Omega\right|\right|_F = \left|\left|A^R_\Omega - A_\Omega\right|\right|_2 .$$

We begin by deriving an expression for the logarithmic error $\log \hat{A}_{ij} - \log A_{ij}$, which we will need both as intermediate step towards our final bound. Before stating this equality, we recall our notation

$$D = \log A^R_\Omega - \log A_\Omega = \log(A_\Omega + \Delta) - \log A_\Omega.$$

Lemma A.1. *The error on the logarithm of the individual estimates satisfies*

$$\log \hat{A}_{ij} - \log A_{ij} = (e_i - e_j)^T L^\dagger_{W^R} B W^R D.$$

Proof. Eq. (2) may be rewritten as

$$B^T \log z = \log A_\Omega,$$

where, recall, B is the edge-vertex incidence matrix of the graph. Recalling Eq. (4) we obtain as a consequence that

$$L_{W^R} \log z = B W^R \log A^R_\Omega.$$

One solution of this system is $z = L^\dagger_{W^R} B W^R \log A_\Omega$ where $\dagger$ represents the Moore-Penrose pseudoinverse.

Recall that $\hat{z}$ is the solution constructed by our algorithm (see Eq. (5)), and we therefore have

$$\begin{aligned}\log \hat{z} - \log z &= L^\dagger_{W^R} B W^R (\log A^R_\Omega - \log A_\Omega)\\ &= L^\dagger_{W^R} B W^R D.\end{aligned}$$

Hence using again $\log A_{ij} = \log z_i - \log z_j$, we have that

$$\begin{aligned}\log \hat{A}_{ij} - \log A_{ij} &= (\log \hat{z}_i - \log z_i) - (\log \hat{z}_j - \log z_j)\\ &= (e_i - e_j)^T L^\dagger_{W^R} B W^R D.\end{aligned}$$

□

Our next step is to bound how much effect the perturbation Δ can have in terms of the resulting perturbation D in the "log space."

Lemma A.2. *If $\Delta_{ij} \leq (c-1)A_{ij}$ for every $(i,j) \in \Omega$ for some $c \geq 1$, then*

$$\left|A^R_{ij} D_{ij}\right| \leq c \left|\Delta_{ij}\right|,$$

for every $(i,j) \in \Omega$ and

$$\left|\left|(W^R)^{1/2} D\right|\right|_F \leq c\, ||\Delta||_F .$$

Proof. By concavity of the logarithm, we have

$$\begin{aligned}\left|\log A^R_{ij} - \log A_{ij}\right| &\leq \left|A^R_{ij} - A_{ij}\right| \max\left(\frac{1}{A^R_{ij}}, \frac{1}{A_{ij}}\right) \\ &= |\Delta_{ij}| \max\left(\frac{1}{A^R_{ij}}, \frac{1}{A_{ij}}\right)\end{aligned}$$

Moreover, the assumption of the Lemma implies $A^R_{ij} \leq cA_{ij}$ for $c \geq 1$. Hence we can bound

$$\begin{aligned}\left|A^R_{ij} D_{(i,j)}\right| &= A^R_{ij}\left|\log A^R_{ij} - \log A_{ij}\right| \\ &\leq |\Delta_{ij}| \max\left(\frac{A^R_{ij}}{A^R_{ij}}, \frac{A^R_{ij}}{A_{ij}}\right) \\ &\leq c\,|\Delta_{ij}|\,,\end{aligned}$$

proving the first claim of the Lemma. The second one follows from the definition of the $|\Omega| \times |\Omega|$ diagonal matrix W^R whose elements are the $(A^R_{ij})^2$. □

Our next lemma provides a first bound on the "logarithmic error." We also include an estimate on how big the (unrevealed) entries $\hat{A}_{ij}$ can get, which we will use in the sequel.

Proposition A.1. *If $\Delta_{ij} \leq (c-1)A_{ij}$ for every $(i,j) \in \Omega$ for some $c \geq 1$, then*

$$\left|\log \hat{A}_{ij} - \log A_{ij}\right| \leq c\sqrt{R_{W^R,ij}}\,||\Delta||_F\,, \tag{7}$$

where $R_{W^R,ij}$ is the resistance distance between i and j on the weighted graph G_{W^R}. As a consequence,

$$\hat{A}_{ij} \leq A_{ij}\exp\left(c\sqrt{R_{W^R,ij}}\,||\Delta||_F\right). \tag{8}$$

Proof. Let us introduce the notation

$$Q_{ij} = W^{R^{\frac{1}{2}}} B^T L^\dagger_{W^R}(e_i - e_j)(e_i - e_j)^T L^\dagger_{W^R} B W^{R^{\frac{1}{2}}}.$$

Then, using that $\log \hat{A}_{ij} - \log A_{ij}$ is a scalar and Lemma A.1, we have

$$\begin{aligned}(\log \hat{A}_{ij} - \log A_{ij})^2 &= (\log \hat{A}_{ij} - \log A_{ij})^T(\log \hat{A}_{ij} - \log A_{ij}) \\ &= D^T W^{R^{\frac{1}{2}}} Q_{ij} W^{R^{\frac{1}{2}}} D.\end{aligned} \tag{9}$$

where, recall, $L_{W^R} = BW^R B^T$. This implies that

$$(\log \hat{A}_{ij} - \log A_{ij})^2 \leq \left|\left|W^{R^{\frac{1}{2}}} D\right|\right|_F^2 \lambda^{ij}_{\max}, \tag{10}$$

where $\lambda^{ij}_{\max}$ is defined as

$$\begin{aligned}\lambda^{ij}_{\max} &:= \lambda_{\max}(Q_{ij}) \\ &= \lambda_{\max}\left(W^{R^{\frac{1}{2}}} B^T L^\dagger_{W^R}(e_i - e_j)(e_i - e_j)^T L^\dagger_{W^R} B W^{R^{\frac{1}{2}}}\right) \\ &= \lambda_{\max}\left((e_i - e_j)^T L^\dagger_{W^R} B W^{R^{\frac{1}{2}}} W^{R^{\frac{1}{2}}} B^T L^\dagger_{W^R}(e_i - e_j)\right) \\ &= (e_i - e_j)^T L^\dagger_{W^R} L_{W^R} L^\dagger_{W^R}(e_i - e_j) \\ &= (e_i - e_j)^T L^\dagger_{W^R}(e_i - e_j),\end{aligned} \tag{11}$$

This quantity equals the resistance $R_{W^R,ij}$, see Vishnoi [2013]. We now have that Eq. (7) follows then from (10) and the bound $||W^{R^{\frac{1}{2}}} D||^2_F \leq c\,||\Delta||^2_F$ of Lemma A.2. Finally, Eq. (7) immediately implies the bound of Eq. (8). □

Our next step is to define some additional notation. The quantity K_{W_0} will be the weighted Laplacian corresponding to the bipartite graph with weights A_{ij}^2, formally defined as

$$K_W := \sum_{i \in I_x, j \in I_y} (e_i - e_j)(A_{ij})^2 (e_i - e_j)^T.$$

Observe that if, in this definition, we replaced A_{ij}^2 by the squares of the revealed entries, and also replaced the sum with only the sum over the revealed entries, then the resulting quantity would be exactly L_{W^R}. We may thus intuitively view the quantity K_W as the Laplacian corresponding to the hypothetical scenario that all entries are revealed without perturbations.

Inspired by the proof above, we will also define

$$Q = W^{R^{\frac{1}{2}}} B^T L_{W^R}^{\dagger} K_W L_{W^R}^{\dagger} B W^{R^{\frac{1}{2}}},$$

and, finally, we will use the shorthand

$$\lambda_{\max} := \lambda_{\max}(Q).$$

The symmetry and nonnegative definiteness of Q implies $\lambda_{\max}$ is real and nonnegative.

Observe that the existing bounds from Eq. (7) and Eq. (8) allow us to derive a bound on the error $\hat{A} - A$ via a few straightforward manipulations. This can then be turned into a bound on $||A - \hat{A}||_F$. However, this approach would be extremely conservative because the bounds of Eq. (7) and Eq. (8) are all worst-case, and a single Δ will not be worst for all (i, j). Our next proposition shows how to exploit this fact.

Proposition A.2. *Suppose $\Delta_{ij} \leq (c-1)A_{ij}$ for every $(i,j) \in \Omega$ for some $c \geq 1$ and $\hat{A}_{ij} \leq \gamma A_{ij}$ for every $i \in I_x, j \in I_y$ and some $\gamma \geq 1$. Then*

$$\left|\left|\hat{A} - A\right|\right|_F^2 \leq \gamma^2 c^2 \lambda_{\max}\left(K_W {L_{W^R}}^{\dagger}\right) ||\Delta||_F^2 .$$

Before proceeding to the proof, we remark that $K_W L_{W^R}^{\dagger}$ is the product of two symmetric matrices, so its eigenvalues are real; consequently, writing $\lambda_{\max}\left(K_W {L_{W^R}}^{\dagger}\right)$ makes sense.

Proof. Since $\left|e^b - e^a\right| \leq \max(e^a, e^b)\,|b - a|$ and $\max(\hat{A}_{ij}, A_{ij}) \leq \gamma A_{ij}$ by assumption, we have

$$\left|\hat{A}_{ij} - A_{ij}\right| \leq \gamma A_{ij} \left|\log \hat{A}_{ij} - \log A_{ij}\right|.$$

If follows then from Eq. (9) that

$$(\hat{A}_{ij} - A_{ij})^2 \leq \gamma^2 (A_{ij})^2 (W^{R^{\frac{1}{2}}} D)^T Q_{ij} (W^{R^{\frac{1}{2}}} D).$$

Summing over all pairs $(i, j) \in I_x \times I_y$ leads to

$$\left|\left|\hat{A} - A\right|\right|_F^2 \leq \gamma^2 \left|\left|W^{R^{\frac{1}{2}}} D\right|\right|_F^2 \lambda_{\max}, \tag{12}$$

Finally using $L_{W^R} = BW^R B^T$,

$$\begin{aligned}
\lambda_{\max} &= \lambda_{\max}(W^{R^{\frac{1}{2}}} B^T L_{W^R}^{\dagger} K_W L_{W^R}^{\dagger} B W^{R^{\frac{1}{2}}}) \\
&= \lambda_{\max}\left(K_W L_{W^R}^{\dagger} B W^{R^{\frac{1}{2}}} W^{R^{\frac{1}{2}}} B^T L_{W^R}^{\dagger}\right) \\
&= \lambda_{\max}\left(K_W L_{W^R}^{\dagger} L_{W^R} L_{W^R}^{\dagger}\right) \\
&= \lambda_{\max}\left(K_W L_{W^R}^{\dagger}\right).
\end{aligned} \tag{13}$$

The result now follows immediately from Eq. (12) and Lemma A.2. □

Theorem 1 is then obtained by using Proposition A.2 with the bound $\gamma = \exp\left(c\sqrt{R_{W^R,\max}}\,||\Delta||_F\right)$ guaranteed by (8) in Proposition A.1.

A.1 Corollary 1

In order to relate the bound of Theorem 1 to more usual characteristics of the graph, we now bound the eigenvalue $\lambda_{\max}(K_W L^\dagger_{W^R})$.

Proposition A.3. *Let $\bar{\alpha}^0, \underline{\alpha}^R$ be respectively an upper bound on the entries of A and a lower bound on the entries of A^R. We then have*

$$\lambda_{\max}(K_W L^\dagger_{W^R}) \leq \left(\frac{\bar{\alpha}^0}{\underline{\alpha}^R}\right)^2 \frac{m+n}{\lambda_2(L)},$$

where we recall that $A \in \mathbb{R}^{m\times n}$ and $\lambda_2(L)$ is the algebraic connectivity (i.e., second-smallest eigenvalue) of the unweighted *bipartite graph G.*

Proof. It follows from Lemma C.1 in Appendix C that

$$\lambda_{\max}\left(K_W L^\dagger_{W^R}\right) \leq \lambda_{\max}(K_W)\lambda_{\max}(L^\dagger_{W^R})$$

Since L_{W^R} is symmetric and has rank $n-1$, we have $\lambda_{\max}(L^\dagger_{W^R}) = \lambda_2(L_{W^R})^{-1}$. Because the absolute values of the off-diagonal elements of L_{W^R} (i.e. the weights) are all at least $(\underline{\alpha}^R)^2$, Lemma C.3 in Appendix C implies then

$$\lambda_2(L_{W^R}) \geq (\underline{\alpha}^R)^2 \lambda_2(L), \tag{14}$$

where we remind that L is the Laplacian of the unweighted bipartite graph G representing the mask Ω. A parallel argument shows that $\lambda_{\max}(K_W) \leq (\bar{\alpha}^0)^2 \lambda_{\max}(K) = (m+n)(\bar{\alpha}^0)^2$, where K is the Laplacian of the complete bipartite graph on $I_x \cup I_y$, whose maximal eigenvalue is $m+n$, from which the statement of this proposition follows. □

We note that the bound of Proposition A.3 could be conservative in terms of the interplay between the values in A, A^R and the graph, but is not very conservative in terms of the graph properties. Indeed, a slightly more complicated argument shows that

$$\lambda_{\max}\left(K_W L^\dagger_{W^R}\right) \geq \left(\frac{\bar{\alpha}^R}{\underline{\alpha}^0}\right)^2 \frac{\min(m,n)}{\lambda_2(L)},$$

where $\bar{\alpha}^R, \underline{\alpha}^0$ are respectively an upper bound on the entries of A^R and a lower bound on those of A.

Having established proposition A.3, we now have that Corollary 1 follows almost immediately. Indeed, since $(\underline{\alpha}^R)^2$ bounds all weight in G_{W^R} from below, the largest resistance $R_{W^R,\max}$ in that graph is at most $(\underline{\alpha}^R)^{-2} R_{\max}$, with $R_{\max}$ the largest resistance of the corresponding unweighted graph G. The first part of Corollary 1 follows from this observation, Proposition A.3 and Theorem 1. Let now $\mathcal{D} \leq m+n$ be the diameter of the graph G. The second part of Corollary 1 follows from the classical bound $R_{\max} \leq \mathcal{D} \leq m+n$ and from $\lambda_2(L) \geq \frac{4}{\mathcal{D}(m+n)} \geq \frac{4}{(m+n)^2}$, see Mohar [1991a].

A Proofs of the Lower Bounds

A.1 Small disturbances: proof of Theorem 2

Let us recall the basic setup. We are given a mask Ω and a rank 1 matrix $A = xy^T$ of which we will be revealed the entries corresponding to Ω (i.e. A^R_Ω). Our approach is to construct, for a given value of $||\Delta||_F$ two matrices A^a, A^b whose entries in Ω are both within $||\Delta||_F$ of the revealed matrix A^R. Lower bounding $||A^a - A^b||$ will then produce a lower bound on the error achievable by any algorithm which only takes into account the set of revealed entries.

We take a fixed vector $\zeta \in \mathbb{R}^{m+n}$ and a sufficiently small constant δ, both to be specified later. We let then $A^a = x^a(y^a)^T, A^a = x^b(y^b)^T$ with

$$
\begin{aligned}
x^a_i &= x_i(1+\delta\zeta_i) \qquad & x^b_i &= x_i(1-\delta\zeta_i) \qquad & \forall i \in I_x\\
y^a_j &= y_j(1-\delta\zeta_j) \qquad & y^b_j &= y_j(1+\delta\zeta_j) \qquad & \forall j \in I_y
\end{aligned}
$$

We first compute the norm of $\Delta^a := (A^a)_\Omega - A^R_\Omega = (A^a - A)$.

$$
\begin{aligned}
||\Delta^a||_F^2 &= ||(A^a - A)_\Omega||_F^2 \qquad (15)\\
&= \sum_{(i,j)\in\Omega} \left(x_iy_j(1+\delta\zeta_i)(1-\delta\zeta_j) - x_iy_j\right)^2\\
&= \sum_{(i,j)\in\Omega} x_i^2y_j^2\left(\delta(\zeta_i-\zeta_j) - \zeta_i\zeta_j\delta^2\right)^2\\
&= \delta^2 \sum_{(i,j)\in\Omega} A_{ij}^2(\zeta_i-\zeta_j)^2 + o(\delta^2)\\
&= \delta^2\zeta^T L_W \zeta + o(\delta^2), \qquad (16)
\end{aligned}
$$

where L_W is the Laplacian of the weighted bipartite graph on $I_x \cup I_y$ corresponding to Ω where the edge (i,j) has weight A_{ij}^2. Parallel arguments show that

$$\left|\left|\Delta^b\right|\right|_F^2 = \left|\left|(A^b - A)_\Omega\right|\right|_F^2 = \delta^2\zeta^T L_W\zeta + o(\delta^2) \qquad (17)$$

and

$$\left|\left|A^b - A^a\right|\right|_F^2 = 4\delta^2\zeta^T K_W \zeta + o(\delta^2), \qquad (18)$$

where K_W is the Laplacian of the weighted complete bipartite graph on $I_x \cup I_y$ with weight A_{ij}^2. To select ζ, we let u be the eigenvector of $K_W L_W^\dagger$ corresponding to $\lambda_{\max}(K_W L_W^\dagger)$ and $\zeta := L_W^\dagger u$. It follows from (16) and (17) that

$$\left|\left|\Delta^\ell\right|\right|_F^2 = \delta^2\zeta^T u + o(\delta^2), \qquad (19)$$

for $\ell = a, b$, and from (18) that

$$
\begin{aligned}
\left|\left|A^b - A^a\right|\right|_F^2 &= 4\delta^2\zeta^T K_W L_W^\dagger u + o(\delta^2)\\
&= 4\delta^2\lambda_{\max}(K_W L_W^\dagger)\zeta^T u + o(\delta^2). \qquad (20)
\end{aligned}
$$

Since u is an eigenvector corresponding to the largest eigenvalue of $K_W L_W^\dagger$, it is not a multiple of the all-ones vector (that would make it an eigenvector corresponding to the smallest eigenvalue of $K_W L_W^\dagger$, which has nonnegative eigenvalues since it is the product of two nonnegative definite matrices). Thus the definition $\zeta := L_W^\dagger u$ implies $u = L_W\zeta$, and $\zeta^T u = u^T L_W^\dagger u > 0$, since u is not proportional to the all-ones vector. Hence (18) and (20) imply that for $\ell = a, b$,

$$\left|\left|A^b - A^a\right|\right|_F^2 = 4\lambda_{\max}(K_W L_W^\dagger)\left|\left|\Delta^\ell\right|\right|_F^2 + o(\left|\left|\Delta^\ell\right|\right|_F^2). \qquad (21)$$

Suppose now that $A^R_{ij} = A_{ij}$ for every $(i,j) \in \Omega$, and that A is in the interior of the set of allowed matrices $\mathcal{A}$. For sufficiently small δ and thus $||\Delta||_F$, we will have $A^a, A^b \in \mathcal{A}$, $||\Delta^a||_F^2 \le (1+\epsilon)\left|\left|\Delta^b\right|\right|_F^2$ and $\left|\left|\Delta^b\right|\right|_F^2 \le (1+\epsilon)\,||\Delta^a||_F^2$, so that both A^a, A^b would be possible values of A even if the algorithm explicitly uses the set $\mathcal{A}$ and a bound $\bar{\Delta} \geq (1+\epsilon) \left|\left|\Delta\right|\right|_F^2$. It follows then from the triangular inequality and (21) that for any estimate $\hat{A}$ there would hold

$$\left|\left|\hat{A} - A\right|\right|^2 \geq \lambda_{\max}(K_W L_W^\dagger) \left|\left|\Delta^\ell\right|\right|_F^2 + o(\left|\left|\Delta^\ell\right|\right|_F^2). \tag{22}$$

for at least one choice among $A = A^a$ or $A = A^b$. To conclude the result, we need to relate $\lambda_{\max}(K_W L_W^\dagger)$ to $\lambda_{\max}\left(K_W L_{W^R}^\dagger\right)$.

Observe first that $L_{W^R} = L_W^\dagger$ because $A_{ij}^R = A_{ij}$. We define the function $t \to \tilde{K}_W(t) \in \mathbb{R}^{(n+m)\times(n+m)}$ by

$$\begin{aligned}
(\tilde{K}_W(t))_{ij} &= x_i(1+t\zeta_i)y_j(1-t\zeta_j) && \forall i \in I_x, j \in I_y\\
(\tilde{K}_W(t))_{ji} &= (\tilde{K}_W(t))_{ij} && \forall i \in I_x, j \in I_y\\
(\tilde{K}_W(t))_{ii} &= -\sum_{j\in I_y} (\tilde{K}_W(t)) && \forall i \in I_x,\\
(\tilde{K}_W(t))_{jj} &= -\sum_{i\in I_x} (\tilde{K}_W(t)) && \forall j \in I_y,
\end{aligned}$$

and the other entries being 0. Observe that $\tilde{K}_W$ is analytic, $K_W = \tilde{K}_W(0)$, $K_W = \tilde{K}_W(\delta)$ if $A = A^a$ and $\tilde{K}_W(-\delta)$ if $A = A^b$. Besides, $\delta = \Theta(\left|\left|\Delta\right|\right|_F)$. Lemma C.2 and $L_W^\dagger = L_{W^R}$ imply then

$$\lambda_{\max}\left(K_W L_W^\dagger\right) = \lambda_{\max}\left(K_W L_{W^R}^\dagger\right) + o(\left|\left|\Delta\right|\right|_F),$$

which implies the result of Theorem 2 together with (22).

A.2 Larger disturbances: proof of Theorem 3

We begin with the claim (a) about the exponential factor. For any given n, we take $A = ee^T$, and the mask $\Omega = \{(i,i), i = 1,\ldots,n\} \cup \{(i,i-1), i = 2,\ldots,n\}$, that is, the entries on the main diagonal and the first other diagonal. We then take the disturbances

$$\Delta_{ii} = 0 \qquad \Delta_{i(i-1)} = \delta,$$

for all i for which these are defined and for some $\delta > 0$. The revealed entries are then

$$A_{ii}^R = 1 \qquad A_{i(i-1)}^R = 1+\delta,$$

Clearly, $\left|\left|\Delta\right|\right|_F^2 = (n-1)\delta^2$ so $\delta = \left|\left|\Delta\right|\right|_F / \sqrt{n-1}$. We then define the rank-1 matrix A by $A_{ij} = (1+\delta)^i(1+\delta)^{-j}$, and observe that A^R is an exact subsample of A because $A_{ij}^R = A_{ij}$ for every $(i,j) \in \Omega$. Moreover, it is not an exact subsample of any other matrix because the graph corresponding to the Ω is connected. Hence any consistent algorithm returns by definition $\hat{A} = A$, so that $(\hat{A} - A)_{ij} = (1+\delta)^{i-j} - 1$. In particular, remembering $\delta = \frac{\left|\left|\Delta\right|\right|_F}{\sqrt{n-1}}$, we have

$$\begin{aligned}
\left|\left|\hat{A} - A\right|\right|_F^2 &\geq (\hat{A}_{1n} - A_{1n})^2\\
&= \left((1+\delta)^{n-1} - 1\right)^2\\
&= \left(\left(1 + \frac{\left|\left|\Delta\right|\right|_F}{\sqrt{n-1}}\right)^{n-1} - 1\right)^2 =: E_n.
\end{aligned}$$

When n grows for a fixed $\left|\left|\Delta\right|\right|_F$, we obtain

$$\lim_{n\to\infty} E_n = \left(e^{\left|\left|\Delta\right|\right|_F \sqrt{n-1}} - 1\right)^2.$$

We conclude part (a) of Theorem 3 by observing that the both A^R_{ij} and A_{ij} are uniformly bounded, and that the graph G_{W^R} corresponding to the mask Ω is a line graph on $2n$ nodes, with $n-1$ weights $1+\delta$ and n weights 1, so that

$$R_{W^R,\max} = n + (n-1)(1+\delta) = n + (n-1)\left(1 + \frac{||\Delta||_F}{\sqrt{n-1}}\right),$$

so that $\sqrt{n-1} = \sqrt{R_{W^R,\max}(\frac{1}{2} - O(n^{-1/2}))}$.

We now move to part (b). For any fixed even n, we let again $A = ee^T$, and we consider the mask $\Omega = \{(i,j) : i,j \leq \frac{n}{2}\} \cup \{(i,j) : i,j \geq \frac{n}{2}\} \cup \{(1,n)\}$, i.e. we reveal the upper left-hand side quarter of the matrix and the lower right-hand side one, and the most upper right-hand side entry. We take $\Delta_{i,j} = 0$ for every revealed entry except $\Delta_{1,n} = \frac{1}{f} - 1$ for $f > 3$, so that $A^R_{ij} = 1$ for all $(i,j) \in \Omega$ except $A^R_{1,n} = 1/f$. Clearly, all $||\Delta||^2_F \leq 1$, and $\max_{(i,j)\in\Omega} \frac{\Delta_{ij}}{A_{ij}}$ and $\max_{(i,j)} A_{ij}$ are bounded independently of n, f, while $\min A^R_{ij} = f^{-1}$. Observe now that A^R is an exact subsample of the rank-1 matrix

$$A_f = \begin{pmatrix} ee^T & f^{-1}ee^T \\ fee^T & ee^T \end{pmatrix},$$

where the vectors e are of dimension $n/2$, and of no other rank-1 matrix. Hence any consistent algorithm would return $\hat{A} = A$ on the data A^R. Focusing on the error on the lower left-hand side block, and using $f > 3$, we would get

$$\left|\left|\hat{A} - A\right|\right|^2_F \geq \frac{n^2}{4}(f-1)^2 \geq \frac{n^2}{9} f^2 = \frac{n^2}{9}(\min_{ij} A^R_{ij})^{-2}.$$

B Proof of Theorem 4 and Corollary 2

High-level idea of proof: We have already seen in Section 5.1 that $(x, (y^{-1})^T$ (where the inverse is taken element-wise) is an eigenvector of the unperturbed matrix M. We therefore need to to argue that when looking at the eigenvector of the perturbed matrix M^R, we can recover a good approximation to $(x, (y^{-1})^T)$.

In general, this is tricky: it might involve expressions depending on the eigenvectors of the matrices M or M^R, which would be difficult to bound explicitly. However, two things make it possible in our case. The first is that the matrices M and M^R correspond to reversible Markov chains, which allows us to make use of a number of bounds appearing in the literature. The second is that the projection step is key: the assumption that the entries of A lie in $[\underline{\alpha}, \overline{\alpha}]$ allows us to project the resulting stationary distributions, which will therefore never be too small or too big; this allows us to go from an error bound on y^{-1} to an error-bound on y. As we have discussed in the main text of the paper, this seemingly minor difference is crucial: without a-priori bounds on entries of A, exponential growth is unavoidable due to our lower bounds.

Proof. We first observe that that dividing A^R by a constant c, and dividing the lower and upper bounds by a the same constant c, while multiplying the output of Algorithm 1 by c does not affect the final estimate $\hat{A}$. Moreover, both sides of Theorem 4 scale linearly with c if A^R, A, Δ are multiplied by c. Hence we can assume without loss of generality that $\mu = \sqrt{\overline{\alpha}\underline{\alpha}} = 1$, so that $A_{ij} \in [\rho^{-1}, \rho]$ for every i, j.

Our first step is to upper bound the difference between the matrices M^R and M.

Lemma B.1.

$$\left|\left|M^R - M\right|\right|_\infty \leq 2\max\left(||\Delta||_\infty, ||\Delta||_1\right),$$

where the norms are the induced matrix norms, with $\Delta_{ij} = 0$ for all $(i,j) \notin \Omega$.

Proof. For any scalars $x, y > 0$, we have the inequality

$$\begin{aligned}\left|\frac{x}{1+x} - \frac{y}{1+y}\right| &= \left|\tfrac{1}{1+x} - \tfrac{1}{1+y}\right| \\ &= \tfrac{|y-x|}{(1+x)(1+y)} \\ &\leq |y-x|.\end{aligned}$$

Hence we have for every $(i, j) \in \Omega$

$$\left|M^R_{ij} - M_{ij}\right| = \left|M^R_{ji} - M_{ji}\right| \le |\Delta_{ij}| \, . \tag{23}$$

Observe now that for a given matrix N whose rows sum to 0, we have

$$\begin{aligned} ||N||_\infty = \max_\ell \sum_k |N_{\ell k}| &= \max_\ell \left(|N_{\ell\ell}| + \sum_{k \ne \ell} |N_{\ell k}| \right) \\ &= \max_\ell \left(\left| -\sum_{k \ne \ell} N_{\ell k} \right| + \sum_{k \ne \ell} |N_{\ell k}| \right) \\ &\le 2 \max_\ell \sum_{k \ne \ell} |N_{\ell k}| \, . \end{aligned}$$

Since the rows of $M^R - M$ sum to zero, we have then

$$\left|\left|M^R - M\right|\right|_\infty \le 2 \max_{\ell \in I_x \cup I_y} \sum_{k \in I_x \cup I_y} \left|M^R_{\ell k} - M_{\ell k}\right| \, . \tag{24}$$

Consider first a $\ell = i \in I_x$. Then by the bipartite structure of M^R, M, the only off-diagonal nonzero $\left|M^R_{ik} - M_{ik}\right|$ are those for which $k \in I_y$. Hence, using (23), we have

$$\begin{aligned} \sum_{k \in I_x \cup I_y} \left|M^R_{ik} - M_{ik}\right| &= \sum_{j \in I_y} \left|M^R_{ij} - M_{ij}\right| \\ &\le \sum_{j \in I_y} |\Delta_{ij}| \\ &\le ||\Delta||_\infty \, . \end{aligned}$$

On the other hand, if $\ell = j \in I_y$, then

$$\begin{aligned} \sum_{k \in I_x \cup I_y} \left|M^R_{jk} - M_{jk}\right| &= \sum_{i \in I_x} \left|M^R_{ji} - M_{ji}\right| \\ &\le \sum_{i \in I_x} |\Delta_{ij}| \\ &\le ||\Delta||_1 \, . \end{aligned}$$

The result follows then from Eq. (24). □

The next part of the proof exploits results on the perturbations of stationary distributions of (discrete-time) Markov chains. Our first step is to introduce a reference stationary distribution associated with the true matrix. Recall that we have assumed all entries of A to be in $[\rho^{-1}, \rho]$ so that $\mu = 1$; Lemma C.4, proved in a subsequent appendix, implies that $A = xy^T$ for some vectors $x \in \mathbb{R}^m, y \in \mathbb{R}^n$ with all x_i and y_j^{-1} in $[\rho^{-1}, \rho]$.

Furthermore, we have seen in Section 5.1 that $(x^T, (y^{-1})^T)$ is a left-eigenvector of M corresponding to its eigenvalue 0. We next define normalized version π^0 for which $\left|\left|\pi^0\right|\right|_1 = 1$. Due to the bounds on the entries of $(x^T, (y^T)^{-1})$ discussed in the previous paragraph, we see that the elements of π^0 all lie in $[\frac{\rho^{-2}}{m+n}, \frac{\rho^2}{m+n}]$. Moreover, the values of $\hat{\pi}$ (see Algorithm 1 for a definition of $\hat{\pi}$) lie in the same interval (see Algorithm 1: those values of $\hat{\pi}$ that are outside this interval are projected onto it).

Proposition B.1.

$$\left|\left|\hat{\pi} - \pi^0\right|\right|_1 \le \frac{\log \rho \sqrt{m+n}}{2\lambda_2(M)} \left|\left|M^R - M\right|\right|_\infty$$

Proof. We will leverage results for perturbations of stationary distributions of discrete-time Markov-chains; to that end, we introduce the auxiliary matrices $P^R = I - \frac{1}{2d_{\max}} M^R$ and $P^0 = I - \frac{1}{2d_{\max}} M$, where $d_{\max}$ is the largest degree in G, i.e. the largest number of revealed elements on any row or column. Observe that the off-diagonal elements of M^R, M are non-negative and bounded by 1, and that each row or column contains at most $d_{\max}$ of them. Moreover, using e to denote the all-ones vector, $M^R e = Me = 0$, which implies P^R, P^0 are row-stochastic matrices with positive diagonals.

The left-eigenvectors π^R and π^0 of M^R and M corresponding to the eigenvalue 0 are also the principal left-eigenvectors of P^R, P^0, and thus the stationary distributions of the corresponding Markov chains since we have assumed them to be stochastic vectors. It follows then from Agarwal et al. [2018] (Theorems 2, 3 and the discussion immediately after the statement of Theorem 3 in the supplementary materials) that

$$\begin{aligned} \left|\left|\pi^R - \pi^0\right|\right|_1 &\leq \frac{1}{2}\left|\left|P^R - P^0\right|\right|_\infty \left(\frac{\log R}{-\log \lambda_2(P^0)} + \frac{1}{1-\lambda_2(P^0)}\right) \\ &\leq \frac{1}{2}\left|\left|P^R - P^0\right|\right|_\infty \frac{\log R + 1}{1-\lambda_2(P^0)}. \end{aligned} \tag{25}$$

where the second line was obtained via the standard inequality $\log x \leq x - 1$; here

$$R = \max_{\ell \in I_x \cup I_y} \sqrt{\frac{1-\pi_\ell^0}{4\pi_\ell^0}}.$$

Our first step is to bound R. Indeed, since $P^0 = I - \frac{1}{2d_{\max}} M$ we have

$$1 - \lambda_2(P^0) = \frac{1}{2d_{\max}} \lambda_2(M),$$

where we remark on the difference in the (standard) convention: while $\lambda_2(M)$ is the second-smallest eigenvalue of M, $\lambda_2(P^0)$ refers to the second-largest eigenvalue of the latter.

Since $\pi_\ell^0 \geq \frac{\rho^{-2}}{m+n}$ for every ℓ and $\frac{1-x}{4x}$ is decreasing, we have then

$$\begin{aligned} \max_{\ell \in I_x \cup I_y} \sqrt{\frac{1-\pi_\ell^0}{4\pi_\ell^0}} &\leq \sqrt{\frac{1-\rho^{-2}/(m+n)}{4\rho^{-2}/(m+n)}} \\ &= \sqrt{\frac{\rho^2(m+n)-1}{4}} \\ &\leq \frac{1}{2}\rho\sqrt{m+n}. \end{aligned}$$

Reintroducing this and the expression $1 - \lambda_2(P^0) = \frac{1}{2d_{\max}} \lambda_2(M)$ into Eq. (25) leads to

$$\begin{aligned} \left|\left|\pi^R - \pi^0\right|\right|_1 &\leq \left(\frac{1}{2}\right) \frac{1 + \log(\rho\sqrt{m+n}/2)}{\frac{1}{2d_{\max}}\lambda_2(M)} \left|\left|P^R - P^0\right|\right|_\infty \\ &\leq \frac{d_{\max} \log \rho\sqrt{m+n}}{\lambda_2(M)} \left|\left|P^R - P^0\right|\right|_\infty, \end{aligned}$$

and the result now follows from

$$\left|\left|P^R - P^0\right|\right|_\infty = \frac{1}{2d_{\max}} ||M^R - M||_\infty,$$

and from

$$||\hat{\pi} - \pi^0||_1 \leq ||\pi^R - \pi^0||_1,$$

which holds since each entry $\hat{\pi}_\ell$ is the projection of π_ℓ^R on an interval to which π_ℓ^0 belongs. □

The last ingredient in the proof is a relation between the error $||\hat{\pi} - \pi^0||_1$ on the stationary distribution and the error $||\hat{A} - A||_F$ on the matrix.

Proposition B.2.

$$\left|\left|\hat{A} - A\right|\right|_F \leq 3(m+n)^2\rho^4 \left|\left|\hat{\pi} - \pi\right|\right|_1 .$$

Proof. Recall that $A_{ij} = \pi_i^0/\pi_j^0$ and that by construction $\hat{A}_{ij} = \hat{\pi}_i/\hat{\pi}_j$ for all i, j. We can decompose the error on an individual entry as

$$\begin{aligned}\left|\hat{A}_{ij} - A_{ij}\right| &= \left|\frac{\hat{\pi}_i}{\hat{\pi}_j} - \frac{\pi_i^0}{\pi_j^0}\right| \\ &\leq \left|\frac{\hat{\pi}_i}{\hat{\pi}_j} - \frac{\hat{\pi}_i}{\pi_j^0}\right| + \left|\frac{\hat{\pi}_i}{\pi_j^0} - \frac{\pi_i^0}{\pi_j^0}\right| \\ &= \hat{\pi}_i \left|\frac{1}{\hat{\pi}_j} - \frac{1}{\pi_j^0}\right| + \frac{1}{\pi_j^0}\left|\hat{\pi}_i - \pi_i^0\right|,\end{aligned}$$

so that

$$\begin{aligned}\sum_{i,j}\left|\hat{A}_{ij} - A_{ij}\right| \leq& \sum_i \hat{\pi}_i \left|\left|(\hat{\pi})^{-1} - (\pi^0)^{-1}\right|\right|_1 \\ &+ \sum_j \frac{1}{\pi_j^0}\left|\left|\hat{\pi} - \pi^0\right|\right|_1 .\end{aligned} \tag{26}$$

We first bound $\sum_i \hat{\pi}_i$. Observe[1] that

$$\begin{aligned}\sum_i \hat{\pi}_i &= \sum_i \pi_i^R + \sum_i (\hat{\pi}_i - \pi_i^R) \\ &\leq 1 + \sum_i (\hat{\pi}_i - \pi_i^R).\end{aligned}$$

Moreover, by construction $(\hat{\pi}_i - \pi_i^R)$ is positive only when

$$\pi_i^R < \rho^{-2}/(m+n),$$

in which case

$$\hat{\pi}_i = \rho^{-2}/(m+n).$$

Hence

$$\sum_i \hat{\pi}_i \leq 1 + \sum_i \frac{\rho^{-2}}{m+n} \leq 1 + \rho^{-2} \leq 2. \tag{27}$$

Secondly, since

$$\pi_j^0 \geq \frac{\rho^{-2}}{m+n},$$

we have

$$\sum_j \frac{1}{\pi_j^0} \leq \rho^2 m(m+n).$$

Plugging this into Eq. (26), we obtain

$$\sum_{i,j} |\hat{A}_{ij} - A_{ij}| \leq 2||\hat{\pi}^{-1} - (\pi^0)^{-1}||_1 + \rho^2 m(m+n)||\hat{\pi} - \pi^0||_1. \tag{28}$$

Moreover, since

$$\left|(\hat{\pi}_i)^{-1} - (\pi_i^0)^{-1}\right| = \frac{\left|\hat{\pi}_i - \pi_i^0\right|}{\hat{\pi}_i \pi_i^0}. \tag{29}$$

[1] It might be tempting to say that $\sum_i \hat{\pi}_i \leq 1$, but this may not be the cause because $\hat{\pi}_i$ is the projection of the statationary distribution, that that projection could increase the 1-norm.

we have that

$$||\hat{\pi}^{-1} - (\pi^0)^{-1}||_1 \leq ||\hat{\pi} - \pi^0||_1 \rho^4 (m+n)^2.$$

Plugging this into Eq. (28) leads to

$$\begin{aligned}\sum_{i,j} \left|\hat{A}_{ij} - A_{ij}\right| &\leq 2(m+n)^2 \rho^4 \left|\left|\hat{\pi} - \pi\right|\right|_1 \\ &\quad + m(m+n)\rho^2 \left|\left|\hat{\pi} - \pi\right|\right|_1 \\ &\leq 3(m+n)^2 \rho^4 \left|\left|\hat{\pi} - \pi\right|\right|_1,\end{aligned}$$

and the result follows then from

$$\begin{aligned}\left|\left|\hat{A} - A\right|\right|_F &= \left|\left|\mathrm{vec}(\hat{\mathrm{A}} - \mathrm{A})\right|\right|_2 \\ &\leq \left|\left|\mathrm{vec}(\hat{\mathrm{A}} - \mathrm{A})\right|\right|_1 \\ &= \sum_{i,j} \left|\hat{A}_{ij} - A_{ij}\right|.\end{aligned}$$

□

Theorem 4 now immediately follows from the combination of Lemma B.1, Propositions B.1 and B.2, together with the bound

$$\begin{aligned}\max\left(\left|\left|\Delta\right|\right|_\infty, \left|\left|\Delta\right|\right|_1\right) &\leq \max\left(\sqrt{n}\left|\left|\Delta\right|\right|_2, \sqrt{m}\left|\left|\Delta\right|\right|_2\right) \\ &\leq \sqrt{\max(m,n)}\left|\left|\Delta\right|\right|_F.\end{aligned}$$

Finally, to prove Corollary 2, observe first that all off-diagonal entries in M^R are at least $\frac{\rho^{-1}}{1+\rho^{-1}} \geq \rho^{-1}/2$ in absolute values. Moreover, as we have discussed in Section 5.1, $M_{k\ell}\pi_k = M_{\ell k}\pi_\ell$ for every $k, \ell \in I_x \cup I_y$; another way to say this is that $\mathrm{diag}(\pi^0)M$ is symmetric. Lemma C.3 implies then

$$\lambda_2(M) \geq \frac{\min_k \pi_k^0}{\max_k \pi_k^0} \frac{\rho^{-1}}{2} \lambda_2(L).$$

Finally, recall that $(\pi^0)^T = K(x^T, (y^{-1})^T)$ for some constant K, and it follows from Lemma C.4 that x, y can be chosen so that $x_i, y_j \in [\rho^{-1}, \rho]$. As a consequence, $\frac{\min_k \pi_k^0}{\max_k \pi_k^0} \geq \rho^{-2}$, and thus $\lambda_2(M) \geq \frac{\rho^{-3}}{2}\lambda_2(L)$. Corollary 2 follows from the combination of this bound with Theorem 4. □

C Technical Lemmas

Lemma C.1. *Let A, B be two PSD matrices. Then every eigenvalue of AB is real and non-negative, and*

$$\lambda_{\max}(AB) \leq \lambda_{\max}(A)\lambda_{\max}(B).$$

Proof. Since A is PSD, its singular value decomposition is of the form $A = U\Sigma U^T$. The diagonal matrix Σ only contains non-negative values, so $\Sigma^{\frac{1}{2}}$ is well defined. Hence the eigenvalues of $AB = U\Sigma^{\frac{1}{2}}\Sigma^{\frac{1}{2}}U^T B$ are exactly the eigenvalues of $M := \Sigma^{\frac{1}{2}}U^T BU\Sigma^{\frac{1}{2}}$, and are thus real since M is symmetric. Moreover, M is positive semi-definite because for any x,

$$x^T Mx = x^T \Sigma^{\frac{1}{2}} U^T BU\Sigma^{\frac{1}{2}} x = (U\Sigma^{\frac{1}{2}} x)^T B(U\Sigma^{\frac{1}{2}} x) \geq 0,$$

due to B being positive semi-definite. Since the spectral radius is lower bounded-by the induced 2-norm, with equality for symmetric matrices, there holds

$$\lambda_{\max}(AB) \leq \left|\left|AB\right|\right|_2 \leq \left|\left|A\right|\right|_2 \left|\left|B\right|\right|_2 = \lambda_{\max}(A)\lambda_{\max}(B).$$

□

Lemma C.2. *Let $A : \delta \in I \to A(\delta)$ be an analytical function of the real variable δ for some interval I and whose values are symmetric PSD matrices, and B a PSD matrix. Then $\lambda_{max}(A(\delta)B)$ is Lipschitz continuous with respect to δ on I.*

Proof. Because $A(\delta)$ is analytic and symmetric, we can rewrite it as $A(\delta) = U(\delta)\Sigma(\delta)U(\delta)^T$, where Σ is diagonal and contain the non-negative eigenvalues of $A(\delta)$, U is orthonormal, and both Σ and U are analytic functions of δ, see Kato [2013]. As in Lemma C.1, we see that $\lambda_{\max}(A(\delta)B) = \lambda M(\delta)$, with

$$M(\delta) := \Sigma(\delta)^{\frac{1}{2}}U(\delta)^T BU(\delta)\Sigma(\delta)^{\frac{1}{2}}$$

an analytical function of δ that is always positive semi-definite, so its eigenvalues are real. It follows then from Theorem 6.8 in that the eigenvalues of M can be expressed as analytical functions of δ, and hence that $\max_i \lambda_i(M(\delta))$ is a Lipschitz-continuous function of δ on the interval I. □

Lemma C.3. *Let L be a directed Laplacian: this means that L is a matrix whose rows sum to zero and whose off-diagonal elements are positive (but L may not be symmetric). Let A_{ij} denote these offdiagonal weights, let $a_{\min}$ be a lowr bound on the smallest positive A_{ij}. Finally, let $\bar{L}$ the corresponding Laplacian when all positive weights ar replced by one.*

We make the assumption that $\bar{L}$ is symmetric. If further L is symmetric, then

$$\lambda_2(L) \geq a_{\min}\lambda_2(\bar{L})$$

If instead DL is symmetric for some positive diagonal D whose smallest and largest diagonal entries are $d_{\min}$ and $d_{\max}$, then $\lambda_2(L)$ is real and

$$\lambda_2(L) \geq \frac{d_{\min}}{d_{\max}} a_{\min}\lambda_2(\bar{L})$$

Proof. We assume first that L is symmetric. In that case its eigenvectors are orthogonal, and since the vector e corresponds to the its eigenvalue 0, we have

$$\lambda_2(L) = \min_{e^T x=0} \frac{x^T L x}{x^T x}$$

Using the classical expression of $x^T L x$ for symmetric Laplacian, we see that for any x, we have

$$\begin{aligned} x^T L x &= \sum_{i<j} A_{ij}(x_i - x_j)^2 \\ &\geq \sum_{i<j, A_{ij}>0} a_{\min}(x_i - x_j)^2 \\ &= a_{\min} \sum_{i<j, \bar{L}_{ij}\neq 0} (x_i - x_j)^2 \\ &= a_{\min} x^T \bar{L} x. \end{aligned}$$

Hence we have

$$\lambda_2(L) \geq a_{\min} \min_{e^T x=0} \frac{x^T \bar{L} x}{x^T x} = a_{\min}\lambda_2(L).$$

We now move to the second claim. Observe that

$$L = D^{-1/2}D^{-1/2}DL$$

which implies that L and $D^{-1/2}DLD^{-1/2}$ are similar. IF DL is symmetric, the latter matrix is also symmetric, and we obtain that all the eigenvalues of L are real. Thus it makes sense to talk about

$$\lambda_2(L) = \lambda_2(D^{-1/2}(DL)D^{-1/2}),$$

which is the second-smallest eigenvalue of L after the smallest eigenvalue of zero.

Observe that that $D^{1/2}e$ is an eigenvector of $D^{-1/2}(DL)D^{-1/2}$ with eigenvalue 0. Hence

$$\begin{aligned}\lambda_2(L) &= \lambda_2(D^{-1/2}(DL)D^{-1/2}) \\ &= \min_{x:e^T D^{1/2}x=0} \frac{x^T D^{-1/2}(DL)D^{-1/2}x^T}{x^T x} \\ &= \min_{y:e^T y=0} \frac{y^T DLy}{y^T D^{-1}y} \\ &\geq \frac{1}{d_{\max}} \min_{y:e^T y=0} \frac{y^T DLy}{y^T y} = \frac{\lambda_2(DL)}{d_{\max}}.\end{aligned}$$

Observe now that all nonzero off-diagonal elements of DL have an absolute value at least $d_{\min}a_{\min}$. The first claim of this lemma implies then $\lambda_2(DL) \geq d_{\min}a_{\min}$, from which the second claim follows. □

Lemma C.4. *Let $A \in \mathbb{R}^{m\times n}$ be a positive rank-1 matrix such that $A_{ij} \in [\rho^{-1}, \rho]$. Then A can be written as $A = xy^T$ for vectors $x \in \mathbb{R}^m, y \in \mathbb{R}^n$ such that $x_i, y_j \in [\rho^{-1}, \rho]$ for every $i \in I_x, j \in I_y$.*

Proof. Since A is rank-1 and positive it can be written as $\hat{x}\hat{y}^T$ for positive vectors $\hat{x}, \hat{y}$. We use the indices $\min, \max$ to denote the indices of the smallest and largest values of the vectors. Observe first that for an arbitrary index $j \in I_y$, we have

$$\frac{\hat{x}_{\max}}{\hat{x}_{\min}} = \frac{\hat{y}_j \hat{x}_{\max}}{\hat{y}_j \hat{x}_{\min}} = \frac{\max_{i\in I_x} A_{ij}}{\min_{i\in I_x} A_{ij}} \leq \rho^2.$$

The same argument shows $\frac{\hat{y}_{\max}}{\hat{y}_{\min}} \leq \rho^2$. We define

$$x = \frac{\rho}{\hat{x}_{\max}}\hat{x}, \qquad y = \frac{\hat{x}_{\max}}{\rho}\hat{y}.$$

There holds again $A = xy^T$. For the vector x we have just constructed we have that

$$x_{\max} = \frac{\rho}{\hat{x}_{\max}}\hat{x}_{\max} = \rho.$$

This implies $x_{\min} \geq \rho^{-1}$ by the same argument as above.

Moreover, $y_{\max} \leq 1$, for otherwise we would have $\max_{i,j} A_{ij} = x_{\max}y_{\max} > \rho$. So, if $y_{\min} \geq \rho^{-1}$, then we are done. Otherwise, we have $\frac{\rho^{-1}}{y_{\min}} > 1$, and we can define

$$x' = \frac{y_{\min}}{\rho^{-1}}x, \qquad y' = \frac{\rho^{-1}}{y_{\min}}y,$$

satisfying again $x'(y')^T = A$. By construction $y'_{\min} = \rho^{-1}$, so that $y'_{\max} \leq \rho$. Moreover, since $\frac{\rho^{-1}}{y_{\min}} > 1$, we have that

$$x'_{\max} \leq x_{\max} \leq \rho.$$

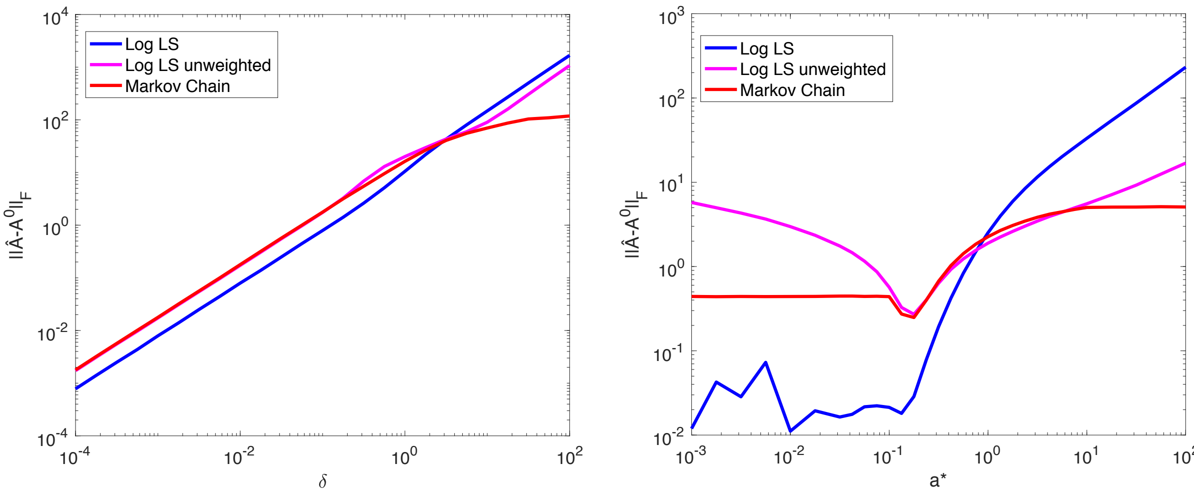
Finally,

$$x'_{\min} = \frac{x_{\min}y_{\min}}{\rho^{-1}} \geq \frac{\rho^{-1}}{\rho^{-1}} = 1,$$

so that x', y' satisfy the conditions we need. □

D Additional Experiments

Ridge regression: Our implementation of the ridge-regression (Eq. 1) uses gradient descent, projecting x and y at each step on the set $[\mu\rho^{-1}, \mu\rho]$ to which we know the real values belong. Different values of λ were tried. The gradient iterations were interrupted when $||x(k+1) - x(k)||_1 + ||y(k+1) - y(k)||_1 \leq 10^{-12}$ or after 200000 steps. Each problem was solved using 10 different initial $x(0), y(0)$, with values randomly selected in $[\mu\rho^{-1}, \mu\rho]$, and the best final iterate (in term of the objective function) was kept. Examples of results are presented in Figure 4, for the same experimental conditions as in Figure 3(a), except that results are averaged over 5 tests for each data point. We further note that large errors for small values of δ were consistently obtained on every single one of the realizations.



(a) (b)

Figure 3: Evolution of the average error $||\hat{A} - A||_F$ for the two proposed algorithms (Markov Chain, Log-LS) and for an unweighted version of the algorithm of Section 3 in (a) a scenario where all revealed entries are perturbed by a random noise of magnitude $\delta/2$ (50×50 matrices with on average 20% of revealed entries), and (b) a targeted scenario where the smallest revealed entry is replaced by $a*$ (10×10 matrices with on average 50% of revealed entries). Initial matrices have entries between 10^{-1} and 10.

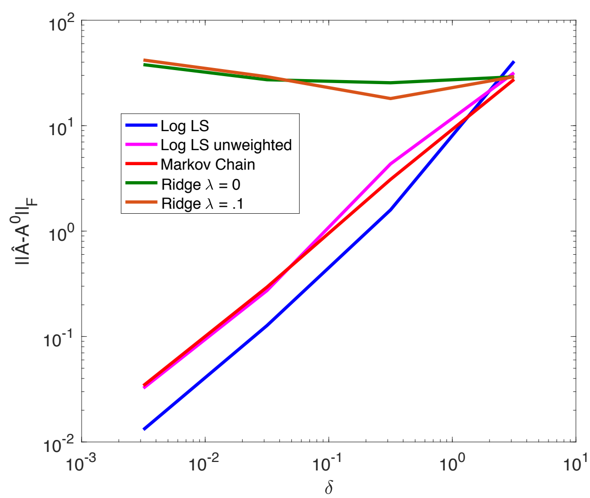
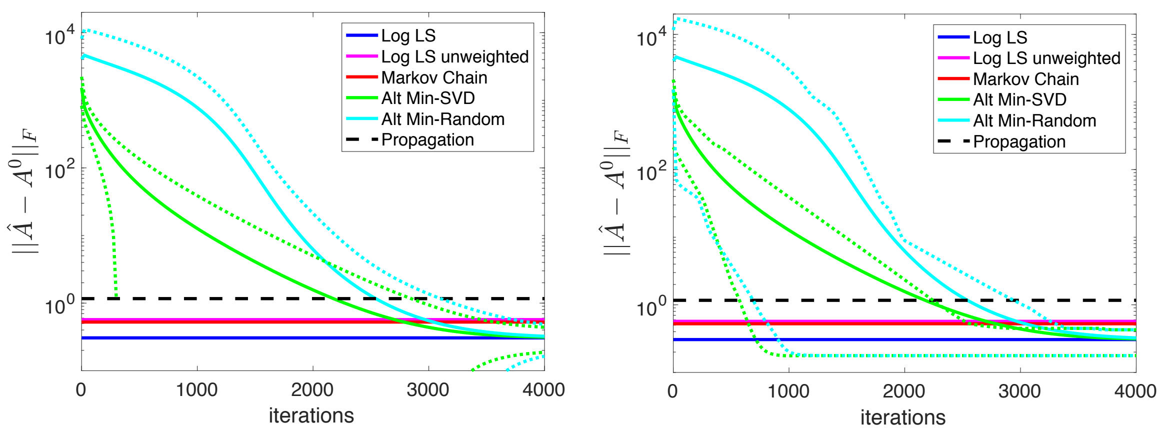


Figure 4: Evolution with δ of the average error $||\hat{A} - A||_F$ for our two algorithms, an unweighted version of the algorithm of Section 3, and our implementation of the ridge regression with $\lambda = 0$ (no regularization) and $\lambda = .1$, in a scenario where all revealed entries are perturbed by a random noise of magnitude $\delta/2$ (50×50 matrices with on average 50% of revealed entries). Initial matrices have entries between 10^{-1} and 10. Large errors are observed for the ridge regression methods, even for very small values of δ.



(a) (b)

Figure 5: Evolution with the number of iterations of the average error (with the variance in (a) and 80% Error Margins (b)) on "Star-Graph" Sampling (3 columns and 3 rows), for the Alt-Min-SVD and Alt-Min-Random methods, compared with the performance of non-iterative methods.