Statistica Sinica **28** (2018), 1371-1393
doi:https://doi.org/10.5705/ss.202016.0103

AN ITERATIVE HARD THRESHOLDING ESTIMATOR FOR LOW RANK MATRIX RECOVERY WITH EXPLICIT LIMITING DISTRIBUTION

Alexandra Carpentier and Arlene K. H. Kim

Universität Potsdam and Sungshin Women's University

Abstract: We consider the problem of low rank matrix recovery in a stochastically noisy high-dimensional setting. We propose a new estimator for the low rank matrix, based on the iterative hard thresholding method, that is computationally efficient and simple. We prove that our estimator is optimal in terms of the Frobenius risk and in terms of the entry-wise risk uniformly over any change of orthonormal basis, allowing us to provide the limiting distribution of the estimator. When the design is Gaussian, we prove that the entry-wise bias of the limiting distribution of the estimator is small, which is of interest for constructing tests and confidence sets for low-dimensional subsets of entries of the low rank matrix.

Key words and phrases: High dimensional statistical inference, inverse problem, limiting distribution, low rank matrix recovery, numerical methods, uncertainty quantification.

1. Introduction

High-dimensional data have generated a great challenge in different fields of statistics, computer science, and machine learning. To consider cases where the number of covariates is larger than the sample size, new methodologies, applicable for the model under some structural constraints, have been developed. For instance, there have been substantial works under the sparsity assumption, including sparse linear regression, sparse covariance matrices estimation, and sparse inverse covariance matrices estimation (see e.g. Meinshausen and Bühlmann (2006); Bickel, Ritov and Tsybakov (2009); Huang, Ma and Zhang (2008); Friedman, Hastie and Tibshirani (2008); Cai and Zhou (2012)). In this paper, we focus on the problem of *low rank matrix recovery and uncertainty quantification*.

There have been quite a few works on estimating low rank matrices in the matrix regression setting (the trace regression setting, the matrix-compressed sensing setting, or the quantum tomography setting when the parameter is a density matrix). Many authors (e.g. Candès and Recht (2009); Candès and Tao

(2010); Recht (2011); Gross (2011)) considered the exact recovery of a low-rank matrix based on a subset of uniformly sampled entries. Recht (2011), Candès and Plan (2011), Flammia et al. (2015), Gross et al. (2010), and Liu (2011) considered matrix recovery based on a small number of noisy linear measurements in the framework of the Restricted Isometry Property (RIP). Negahban and Wainwright (2011) proved non-asymptotic bounds on the Frobenius risk, and investigated matrix completion under a row/column weighted random sampling. Koltchinskii, Lounici and Tsybakov (2011) proposed a nuclear norm minimisation method and derived a general sharp oracle inequality under the condition of the restricted isometry property. Cai and Zhang (2015) considered a rank-one projection model and used constrained nuclear norm minimization method to estimate the matrix. Flammia et al. (2015) and Gross et al. (2010) considered a specific quantum tomography problem where the parameter is a density matrix (for more details, please see Subsection 4.2), and Liu (2011) proved that the quantum tomography design setting satisfies the RIP. Koltchinskii (2011) proposed an estimator based on an entropy minimisation for solving a quantum tomography problem.

Goldfarb and Ma (2011) and Tanner and Wei (2012) adapted the iterative hard thresholding method (first introduced in the sparse linear regression setting, see e.g. Needell and Tropp (2009); Blumensath and Davies (2009)) to the problem of low rank matrix recovery when the noise is non-stochastic and of small L_2 norm. This procedure has the advantage of being computationally efficient. In the same vein, but applied to the more challenging stochastically noisy setting, Agarwal, Negahban and Wainwright (2012) introduced a soft thresholding technique that provides efficient result in this setting in Frobenius norm, see also Bunea, She and Wegkamp (2011), Chen and Wainwright (2015), and Klopp (2015) for other thresholding methods in related settings that provide results in Frobenius norm.

An important problem is to understand the uncertainty associated to these statistical methodologies, by e.g. characterizing the limiting distribution of the efficient estimators. Results in this area for high dimensional models are still scarce, available mainly for the sparse (generalised) linear regression models (Zhang and Zhang (2014); Javanmard and Montanari (2014); van de Geer et al. (2014); Nickl and van de Geer (2014)). In the papers Zhang and Zhang (2014), Javanmard and Montanari (2014), and van de Geer et al. (2014), the authors focus first on constructing an estimator for the sparse parameter that has good properties in L_∞ risk, then use this result to exhibit the limiting distribution of their estimator. Knowing this enables the construction of tests and confidence sets for

low-dimensional subsets of parameters.

The construction of an estimator that has an explicit limiting distribution does not exist in the low rank matrix recovery setting. To the best of our knowledge, moreover, the theoretical results on the estimation of the parameter, in the noisy setting, are derived in Frobenius risk.

We consider the problem of constructing an estimator for low-rank matrix in a stochastically noisy high-dimensional setting, under the assumption that a RIP-type isometry condition is satisfied. We provide error bounds for our estimator in all p Schatten norms for $p > 0$. We prove in particular that this estimator has optimal Frobenius and operator norm risk by proving that it has optimal L_∞ risk performance uniformly over any change of orthonormal basis. A slight modification of our estimator has an explicit Gaussian limiting distribution with bounded bias in operator norm and, when the design consists of uncorrelated Gaussian entries, we prove that the bias in L_∞ entry-wise norm is bounded as well; the last is useful for testing hypotheses and constructing confidence intervals for each parameter of interest, similar to the ideas in Zhang and Zhang (2014), Javanmard and Montanari (2014), and van de Geer et al. (2014). Our estimator is computationally efficient with an explicit algorithm. This algorithm is inspired by the iterative hard thresholding that refines its estimation of the matrix by iteratively estimating the low rank sub-space where the matrix's image is defined. It requires only $O(\log n)$ iteration steps to converge approximately, and the computational complexity of the method is of order $O(nd^2 \log n)$ where d is the dimension of the matrix and n is the sample size.

We provide some simulations to illustrate the efficiency of our method and explain how it can be used to create a confidence interval for the entries of the low rank matrix. We apply our method to a specific *quantum tomography application*, multiple ion tomography (see, e.g. Guta, Kypraios and Dryden (2012); Gross et al. (2010); Butucea, Guţă and Kypraios (2015); Haeffner et al. (2005); Acharya, Kypraios and Guta (2015); Holevo (2001); Nielsen and Chuang (2000)), where the assumptions required by our method are naturally satisfied (see e.g. Liu (2011); Flammia et al. (2015)). We compare our method with other existing estimation methods for the trace regression setting (Candès and Tao (2010); Gross et al. (2010); Koltchinskii, Lounici and Tsybakov (2011); Flammia et al. (2015)) using the gradient descent implementation of Agarwal, Negahban and Wainwright (2012) and regularized maximum likelihood based procedures (Butucea, Guţă and Kypraios (2015); Acharya, Kypraios and Guta (2015)).

In the online supplementary material, we adapt our method to the setting of

sparse linear regressionm and provide an estimator that has an explicit limiting distribution (recovering the results of Zhang and Zhang (2014); Javanmard and Montanari (2014); van de Geer et al. (2014)).

2. Setting

2.1. Preliminary notation

For $T > 0$, $q \in \mathbb{N}$, and $u \in \mathbb{C}^q$, we write $\lfloor u \rfloor_T =: v$ for the hard thresholded version of u at level T, $v_i = u_i \mathbf{1}\{|u_i| \geq T\}$ for $i = 1, \ldots, q$. For $q > 0$ and $u \in \mathbb{R}^q$, we write $\|u\|_2 = \sqrt{\sum_{i \leq q} |u_i|^2}$, and $\|u\|_\infty = \sup_i |u_i|$.

For a $q \times q$ complex matrix A, we write A^T as the conjugate transpose of A, $\mathrm{tr}(A) = \sum_k A_{k,k}$ for the trace of A, and $\mathrm{diag}(A)$ for the matrix whose diagonal entries are the same as A while its non-diagonal entries are all zeros. We write $\|A\|_\infty = \max_{i,j} |A_{i,j}|$, and $\|A\|_2^2 = \sum_{i,j} A_{i,j}^2$. We write the operator norm of A as $\|A\|_S = \max_i \lambda_i$, where the λ_i are the singular values of A, and the Schatten p norm of A for $p > 1$ as $\|A\|_{S_p} = (\sum_i \lambda_i^p)^{1/p}$, noting that $\|A\|_{S_2} = \|A\|_2$.

For $T > 0$, we write $\lfloor A \rfloor_T$ for the hard thresholded version of A at level T for each entry, $V_{i,j} = A_{i,j} \mathbf{1}\{|A_{i,j}| \geq T\}$ for $i, j = 1, \ldots, q$.

2.2. Models

Let $d, n \in \mathbb{N}$. Let $\mathcal{M}$ be the set of $d \times d$ matrices, $\mathcal{M}(k)$ be the set of $d \times d$ complex matrices of rank less than or equal to k, and $\mathcal{M}_\Omega$ for the set of orthonormal matrices in $\mathcal{M}$.

For $X^i \in \mathcal{M}, \Theta \in \mathcal{M}$, we consider the matrix regression problem, for any $i \leq n$,

$$Y_i = \mathrm{tr}\big((X^i)^T \Theta\big) + \epsilon_i,$$

where $\epsilon \sim \mathcal{N}(0, I_n)$ (our results hold in the same way for any sub-Gaussian independent noise ϵ: see Remark 3), and $d \leq n$ but $d^2 \gg n$. We write $\mathbb{X}$ for the linear operator such that, for any $A \in \mathcal{M}$,

$$\mathbb{X}(A) = (\mathrm{tr}((X^i)^T A))_{i \leq n}.$$

The model can be rewritten as

$$Y = \mathbb{X}(\Theta) + \epsilon,$$

where $Y = (Y_i)_{i \leq n}$. This model is directly related to the quantum tomography model (Flammia et al. (2015); Gross et al. (2010); Liu (2011); Gross (2011); Koltchinskii (2011)), also to e.g. matrix completion (Negahban and Wainwright

(2011); Koltchinskii (2011)).

Assumption 1. *Let $K \leq d$. For any $k \leq 2K$, it holds that*

$$\sup_{A \in \mathcal{M}(k)} \left| \frac{1}{n} \|\mathbb{X}(A)\|_2^2 - \|A\|_2^2 \right| \leq \tilde{c}_n(k) \|A\|_2^2,$$

where $\tilde{c}_n(k) > 0$.

Remark 1. This assumption is related to the Restricted Isometry Property. Typically, for uncorrelated Gaussian design with mean 0 and variance 1 entries, it holds with probability larger than $1 - \delta$ for $\tilde{c}_n(k) \leq C\sqrt{kd \log(1/\delta)/n}$, where $C > 0$ is a universal constant. For the Pauli design used in quantum tomography, it holds with probability larger than $1 - \delta$ for $\tilde{c}_n(k) \leq C\sqrt{kd \log(d/\delta)/n}$, where $C > 0$ is a universal constant (Liu, 2011).

3. Main Results

As a generalization of sparsity constraints in linear regression models, we impose a rank $k \leq d$ constraint on a matrix $\Theta \in \mathbb{R}^{d \times d}$. This constraint arises in such applications as quantum tomography, matrix completion, and matrix compressed sensing (see e.g. Flammia et al. (2015); Gross et al. (2010); Liu (2011); Gross (2011); Negahban and Wainwright (2011); Koltchinskii, Lounici and Tsybakov (2011)).

3.1. Method

We take the parameters $B > 0, \delta > 0, K > 0$. Here δ is a small probability that calibrates the precision of the estimate. The parameter K is an upper bound on two times the actual low rank of the parameter Θ; our final results will not depend on it as long $\sqrt{K}\tilde{c}_n(K) \ll 1$. The parameter B is an upper bound on the Frobenius norm of the parameter Θ.

We set the initial values for the estimator $\hat{\Theta}^0$ and the threshold T_0 as

$$\hat{\Theta}^0 = 0 \in \mathbb{R}^{d \times d}, \quad T_0 = B \in \mathbb{R}^+.$$

We update the thresholds

$$T_r = 4\tilde{c}_n(2K)\sqrt{K}T_{r-1} + \upsilon_n := \rho T_{r-1} + \upsilon_n,$$

where $\upsilon_n = C\sqrt{d\log(1/\delta)/n}$, C is an universal constant (see Lemma 2 in online supplement) and $\rho := 4\tilde{c}_n(2K)\sqrt{K}$.

Set recursively, for $r \in \mathbb{N}$, $r \geq 1$,

$$\hat{\Psi}^r = \frac{1}{n} \sum_{i=1}^{n} (X^i)^T \left(Y_i - \operatorname{tr}(X^i \hat{\Theta}^{r-1}) \right) \in \mathbb{R}^{d \times d},$$

and let $U^r, V^r \in \mathcal{M}^2_\Omega$ be two orthonormal matrices that diagonalise $\hat{\Theta}^{r-1} + \hat{\Psi}^r$. Then set

$$\hat{\Theta}^r = U^r \lfloor (U^r)^T (\hat{\Theta}^{r-1} + \hat{\Psi}^r) V^r \rfloor_{T_r} (V^r)^T. \tag{3.1}$$

This procedure provides a sequence of estimates, that is, with high probability, close to the true Θ as soon as r is of order $\log(n)$.

Remark 2. For implementing our method we need the quantities ρ, υ_n, T_0, and the stopping time r. We describe in Subsection 3.3 how to choose ρ, υ_n, T_0. In particular, T_0 can be chosen in a data driven way.

This method is related to Iterative Hard Thresholding (IHT), a method that has been developed for the sparse regression setting (see e.g. Blumensath and Davies (2009); Needell and Tropp (2009)). It is less straightforward to see this in this setting. In the sparse regression setting, we adapt our method in Subsection S2 of online supplement. For a discussion of the relation between our method and IHT, see the Remark 2 in online supplement. The IHT algorithms have been proved to work in settings where the noise is small and non-stochastic (see e.g. Blumensath and Davies (2009); Needell and Tropp (2009); Goldfarb and Ma (2011); Tanner and Wei (2012)), but apparently there are no results on IHT in a stochastically noisy setting.

3.2. Results for the low rank matrix recovery

Main result for our thresholded estimator We now show that the estimate $\hat{\Theta}^r$ after $O(\log(n))$ iterations has at most rank k, and its entry-wise L_∞ risk and Frobenius risk are bounded with the optimal rates.

Theorem 1. *If Assumption* 1 *holds and* $\tilde{c}_n(2K)\sqrt{K} < 1/4$, *with* $r \approx O(\log(n))$, *we have, for a constant* $C_1 > 0$, *that with probability larger than* $1 - \delta$ *and for any* $k \leq K/2$,

$$\sup_{\Theta \in \mathcal{M}(k), \|\Theta\|_2 \leq B} \|\Theta - \hat{\Theta}^r\|_S \leq C_1 \sqrt{\frac{d \log(1/\delta)}{n}},$$

$$\sup_{\Theta \in \mathcal{M}(k), \|\Theta\|_2 \leq B} \operatorname{rank}(\hat{\Theta}^r) \leq k,$$

and, for any $p > 0$,

$$\sup_{\Theta \in \mathcal{M}(k), \|\Theta\|_2 \leq B} \|\Theta - \hat{\Theta}^r\|_{S_p} \leq C_1 k^{1/p} \sqrt{\frac{d \log(1/\delta)}{n}}.$$

Observe that our estimate attains the minimax optimal Schatten p risk; other estimates in the literature also attain this for e.g. $p = 2$. It is also minimax-

optimal in operator norm (or entry-wise matrix L_∞ risk), and the entry-wise error is not more than $\sqrt{d/n}$ with high probability *for any orthonormal change of basis of the matrix* Θ. This is useful for measuring the uncertainty of an estimate (in particular since it does not require the a priori knowledge of the rank of the matrix Θ).

Asymptotic normality results To prove asymptotic normality, we modify the estimator of Theorem 1. Consider the estimator $\hat{\Theta}^r$ and define

$$\hat{\Theta} = \hat{\Theta}^r + \frac{1}{n}\sum_{i=1}^{n}(X^i)^T[Y_i - \mathrm{tr}((X^i)^T\hat{\Theta}^r)].$$

Theorem 2. *If*

$$Z := \frac{1}{\sqrt{n}}\sum_{i\leq n}(X^i)^T\epsilon_i,$$

$$\Delta := \sqrt{n}(\hat{\Theta}^r - \Theta) - \frac{1}{\sqrt{n}}\sum_{i\leq n}(X^i)^T\mathrm{tr}\big((X^i)^T(\hat{\Theta}^r - \Theta)\big),$$

then

$$\sqrt{n}(\hat{\Theta} - \Theta) = \Delta + Z, \tag{3.2}$$

where $Z|\mathbb{X} \sim \mathcal{N}\Big(0, \big(1/n\sum_{i\leq n}(X^i_{j,j'}X^i_{l,l'})\big)_{j,j',l,l'}\Big)$.

For $r \approx O(\log(n))$, if Assumption 1 holds for some $K > 0$, $\tilde{c}_n(2K)\sqrt{K} = o(1)$, the rank of Θ is smaller than $2K$, and its Frobenius norm is bounded by B, there is a constant $C_1 > 0$ such that with probability larger than $1 - \delta$

$$\frac{\|\Delta\|_S}{\sqrt{d}} \leq 4C_1\tilde{c}_n(2K)\sqrt{K}\log(\frac{1}{\delta}) = o_{\mathbb{P}}(1).$$

If the elements in the design matrices $X^i \in \mathcal{M}$ are i.i.d. Gaussian with mean 0 and variance 1, and $\max(K^2d, Kd\log(d)) = o(n)$, we have that $\|\Delta\|_\infty = o_{\mathbb{P}}(1)$. Note that this implies the previous result.

This result follows the works in the context of sparse linear regression of Zhang and Zhang (2014), Javanmard and Montanari (2014), and van de Geer et al. (2014), and implies that there exists an estimator of Θ that has a Gaussian limiting distribution, and whose rescaled bias Δ with respect to Θ can be bounded in operator norm under Assumption 1, and in L_∞ norm as well when the design is Gaussian.

Remark 3. Theorems 1 and 2 are proved for Gaussian noise ϵ, but generalise to independent, sub-Gaussian noise with a similar, but more technical proof based on Talagrand's inequality. Here Z, conditioned on the design $\mathbb{X}$, would not be

Gaussian, but would have a limiting Gaussian distribution using the Central Limit Theorem.

Stopping rule r Theorem 1 applies after $r = O(\log(n))$ iterations of our thresholding strategy. It is possible to propose a data-driven stopping rule that performs well: for a desired precision $e > 0$, stop the algorithm as soon (after having thresholded a last time) as

$$T_r \leq (1+e)\frac{1}{1-\rho}v_n. \tag{3.3}$$

Write $\hat{r}$ for the time the stopping rule stops.

Theorem 3. *If Assumption* 1 *holds and* $\tilde{c}_n(2K)\sqrt{K} < 1/8$, *for* $e \leq 0.1$ *in* (3.3), $\hat{\Theta}^{\hat{r}}$ *satisfies, with probability larger than* $1-\delta$ *and for any* $k \leq K/2$,

$$\sup_{\Theta\in\mathcal{M}(k),\|\Theta\|_2\leq B} \|\Theta - \hat{\Theta}^{\hat{r}}\|_S \leq \frac{1.1}{1-\rho}v_n = 2.2C\sqrt{\frac{d\log(1/\delta)}{n}},$$

$$\sup_{\Theta\in\mathcal{M}(k),\|\Theta\|_2\leq B} \operatorname{rank}(\hat{\Theta}^{\hat{r}}) \leq k,$$

and so for any $p > 0$,

$$\sup_{\Theta\in\mathcal{M}(k),\|\Theta\|_2\leq B} \|\Theta - \hat{\Theta}^{\hat{r}}\|_{S_p} \leq 2.2C(2k)^{1/p}\sqrt{\frac{d\log(1/\delta)}{n}},$$

$$\hat{r} \leq 1 + \frac{\log\left(10(1-\rho)T_0/(v_n)\right)}{\log(1/\rho)} \leq O(\log(n)).$$

This empirical stopping rule guarantees minimax optimal results in less than $\log(n)$ iterations, and Theorem 2 holds using this stopping rule; this last can be proved in the same way as Theorem 3 is proved.

3.3. Discussion

Comparison of our results with the literature Our Theorem 1 gives bounds for our estimators in all Schatten $p > 0$ norms (including the operator norm, and therefore uniform entry wise bounds in all rotation basis). These results are minimax optimal in both Frobenius and operator norm. The corresponding lower bound in Frobenius norm can be found in e.g. Candès and Plan (2011) under a same assumption or Theorem 5 of Koltchinskii, Lounici and Tsybakov (2011) under a related assumption. The corresponding lower bound in operator norm can be found in e.g. Carpentier et al. (2015). Koltchinskii and Xia (2015) contains further lower bounds results proving that the operator norm rate

$\sqrt{d/n}$ (and associated Schatten q norm $k^{1/q}\sqrt{d/n}$) is optimal also in the case of quantum tomography under the additional assumptions that the parameter is a density matrix and that the design is random Pauli. Our method is apparently the first iterative method that has such an optimality property in operator norm; Koltchinskii and Xia (2015) provides results for Schatten norms with $q \in [1, 2]$, but not for other Schatten norms. A slight modification of our estimator has an explicit Gaussian limiting distribution, and apparently the first iterative method for low rank matrix recovery with such a property. The computational complexity of our algorithm is as low as for any procedure based on iterative hard thresholding; see the papers (Goldfarb and Ma (2011) and Tanner and Wei (2012)). Our assumption 1 is a strong RIP condition. But it is satisfied in the interesting application of multiple ion tomography for the natural Pauli design when the number of settings is large enough, see Subsection 4.2.

Operator norm bounds provide an entrywise bound up to any change of orthonormal basis, and provide a bound on the eigen values; since these bounds do not depend on the true rank k, they can be used to implement conservative confidence sets. As highlighted in the papers Zhang and Zhang (2014), Javanmard and Montanari (2014), and van de Geer et al. (2014), having a bound on the entrywise risk, and then an estimator with explicit limiting distribution, leads to construct tests and confidence intervals for subsets of coordinates of the parameter Θ. The bound on the bias term Δ in L_∞ norm in Theorem 2 requires that the design be Gaussian, but the bound on it requires only the fact that Assumption 1 is satisfied.

Stopping rule r We have defined an empirical stopping rule, see (3.3) and Theorem 3, and use it for all the experiments in Section 4.

Calibration of the parameters of the proposed method There are three quantities that need to be calibrated: ρ and υ_n enter in the definition of the thresholds sequence $(T_r)_r$. ρ controls the rate at which we make our threshold decay, and $\upsilon_n/(1-\rho)$ is the quantity toward which it converges when r goes to infinity; T_0 is the initialisation of the threshold sequences. Here are some comments on how to choose them.

Rate of decay ρ: The parameter ρ can be taken between 1 and $4\sqrt{K}\tilde{c}_n(2K)$ where K is an upper bound on the rank of the parameter and $\tilde{c}_n(2K)$ is the constant associated to the design such that Assumption 1 in the Supplement is satisfied. While one may be unable to compute K or $\tilde{c}_n(2K)$ without more assumptions on the design, the random Pauli design for quantum tomography

has an upper bound on $\tilde{c}_n(2K)$ for all K that is of order $\sqrt{Kd\log(d)/n}$ with high probability. In this design if n is large enough, we know that taking $\rho = 1/2$ will work.

Smallest threshold calibration υ_n: The interpretation and theoretical value of υ_n is clear: it should be taken to be larger than the δ quantile of the LHS quantity defined in Equation (S1.5), divided by $\|A\|_2$. As we do not have access to this quantile, we calibrate it as an empirical estimator of this asymptotic quantile (using Theorem 2).

Initialisation threshold T_0: The constant T_0 needs to be taken as an upper bound on the Frobenius norm of Θ. Estimating an upper bound on $\|\Theta\|_2^2$ from the data is easy under Assumption 1 in the Supplement:

$$\frac{1}{n}\|Y\|_2^2(1+\kappa)$$

overestimates $\|\Theta\|_2^2$. In our simulations, we propose a slightly more refined heuristic upper bound and use the same T_0 and T_r in Section 4 and 4.2.

In specific cases, we know enough about the design and the noise level to know that these calibrations will work, provided that the target matrix is indeed low rank.

4. Experiments

In this section we present some experiments, first some simulation for the construction of confidence intervals, and then a formal comparison of our thresholded estimator with other methods for multiple ion tomography.

4.1. Simulation results for the construction of entry-wise confidence intervals

We performed experiments for low-rank matrix recovery, with matrix dimension d. We considered a Gaussian design where the $X^i_{j,j'} \sim \mathcal{N}(0,1)$ and independent. We also considered a Gaussian uncorrelated noise $\epsilon \sim \mathcal{N}(0, I_n)$. We took a parameter Θ of rank k stochastically generated in an isotropic way, as

$$\Theta = \sum_{l=1}^{k} N_l N_l^T, \quad \text{where,} \quad N_l \sim \mathcal{N}(0, I_d).$$

We implemented our method choosing a data-driven heuristic for the choice of our parameters. We first set $\hat{\Theta}^0 = 0$. We set, for any $r \geq 1$,

$$\hat{\sigma}_r^2 = \|Y - (\mathrm{tr}\big((X^i)^T\hat{\Theta}^{r-1}\big)_{i\leq n}\|_2^2 \frac{1}{n}, \tag{4.1}$$

$$\upsilon_n(r) = \hat{\sigma}_r \sqrt{\frac{d}{n}} q_{90\%}, \tag{4.2}$$

where $q_{90\%}$ is the 90% quantile of the $\mathcal{N}(0,1)$. Here, $\upsilon_n(r)$ replaces υ_n, and is a heuristic high probability bound on the error for each coordinate.

We set

$$T_1 = B = \hat{\sigma}_1 + \upsilon_n(1), \tag{4.3}$$

which is by construction larger than the Frobenius norm of Θ with high probability, and

$$T_r = \rho T_{r-1} + \upsilon_n(r), \tag{4.4}$$

where we took $\rho = 1/2$ (1/2 so that the decay is not too fast, but $1/(1-\rho)$ is not too large).

We used the heuristic stopping rule described in Equation (3.3), iterating until

$$T_r \leq (1+e) \times \frac{1}{1-\rho} \upsilon_n(r) = 2.2 \upsilon_n(r),$$

for $e = 0.1$.

We constructed, using the limiting distribution results provided in Theorem 2, a confidence set for the all the entries of Θ so that, for any entry (m, m'),

$$C_n^{m,m'} = [\hat{\theta}_{m,m'} - c_{m,m'}, \hat{\theta}_{m,m'} + c_{m,m'}],$$

where

$$c_{m,m'} = \hat{\sigma}_r \hat{\Sigma}_{m,m'} \frac{q_{95\%}}{\sqrt{n}},$$

and $\hat{\Sigma}^2_{m,m'} = 1/n \times \sum_{i \leq n} (X^i_{m,m'})^2$.

We provide several results, depending on the values of (n, p, k), averaged over 100 iterations of simulations. For these simulations, we present three kinds of results: Figure 1 presents, for different values of p, k, and increasing n, the logarithm of the rescaled Frobenius risk of the estimate $\hat{\Theta}$,

$$\log\left(\frac{\|\hat{\Theta} - \Theta\|_2}{\|\Theta\|_2}\right).$$

Figure 2 presents, for different values of p, k, and increasing n, the logarithm of the averaged diameter of the confidence intervals $C_n^{m,m'}$,

$$\log\left(\frac{1}{d^2} \sum_{m,m'} c_{m,m'}\right).$$

Figure 3 gives, for different values of p, k, and increasing n, the averaged coverage probability of the confidence intervals $C_n^{m,m'}$,

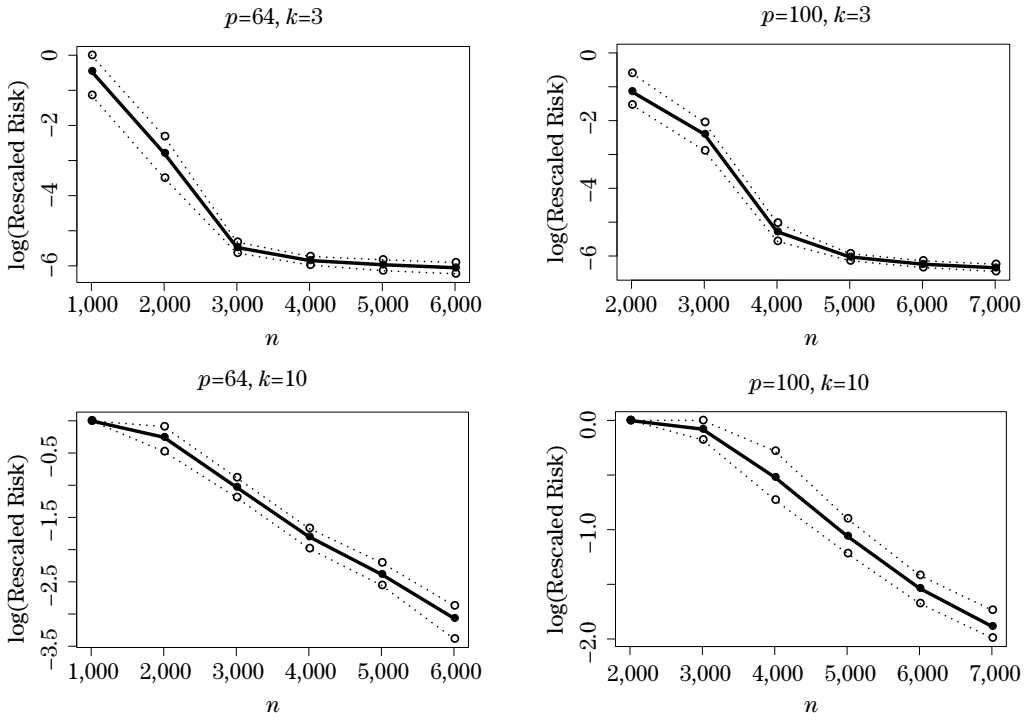

Figure 1. Logarithm of the rescaled Frobenius risk of the estimate in function of n, for different values of p, k. The solid line is the average over 100 iterations, the dotted lines form 95% confidence intervals.

$$\frac{1}{d^2}\sum_{m,m'} \mathbf{1}\{\theta_{m,m'} \in C_n^{m,m'}\}.$$

These graphs exhibit 95% confidence intervals (upper and lower 2.5% quantile values from 100 iterations) around their means (dotted lines in the graphs, the solid line being the mean).

These figures exhibit different behaviours depending on the difficulty of the problems (increasing with p, and more importantly with k). The graphs in Figure 1 for $k = 3$ and $p \in \{64, 100\}$ exhibit first a fast decay of the risk, until some critical threshold $n = ckd$, with c seemingly between 10 and 20. At this point, one can see that the method recovers in most case the true rank k of the matrix, whereas before it recovered only a smaller rank approximate of Θ—with a too small n, it could not distinguish all the signal from the noise, and the fact that it gradually does for larger n explains the fast decay of the logarithm of the rescaled risk. Subsequently, the curve has a kink and the decay becomes slower,

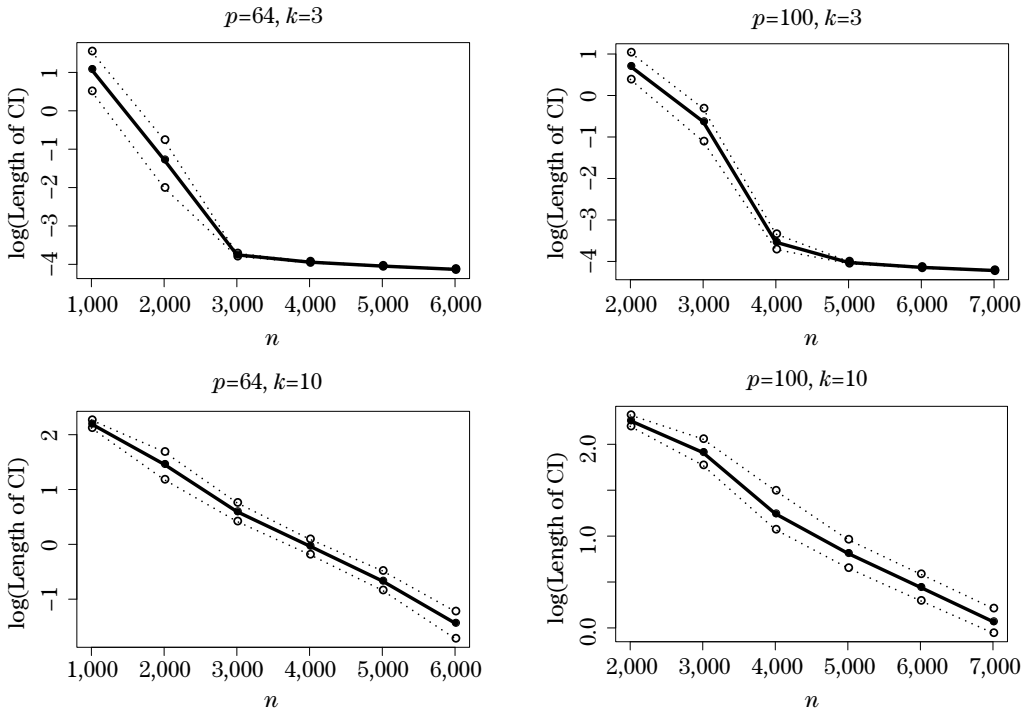

Figure 2. Logarithm of the averaged rescaled length of the confidence intervals of the in function of n, for different values of p, k. The solid line is the average over 100 iterations, the dotted lines form 95% confidence intervals.

after which all the k "rank directions" have been identified, and the logarithm of the rescaled risk starts decreasing slower, according to the theoretical rate of $-\log(n)$. The graphs in Figure 1 for $k = 10$ and $p \in \{64, 100\}$ exhibit mainly the first regime, since k is larger and the second regimes comes for larger values of n—empirically, we can observe that the method recovers most of the time all k "directions" as soon as $n = 4{,}000$ for $p = 64$, or as soon as $n = 6{,}000$ for $p = 100$.

Figure 2 is not surprising since length is supposed to reflect the risk. The averaged coverage of these intervals in Figure 3 is in average larger than 87% in all cases, and in more than 95% of the cases, it is higher than 74% in all cases.

4.2. Quantum tomography experiments

4.2.1. Description of the ion tomography setting

An important application to which our method can be applied is the estimation of quantum states.

We consider estimating the joint quantum state of m two-dimensional sys-

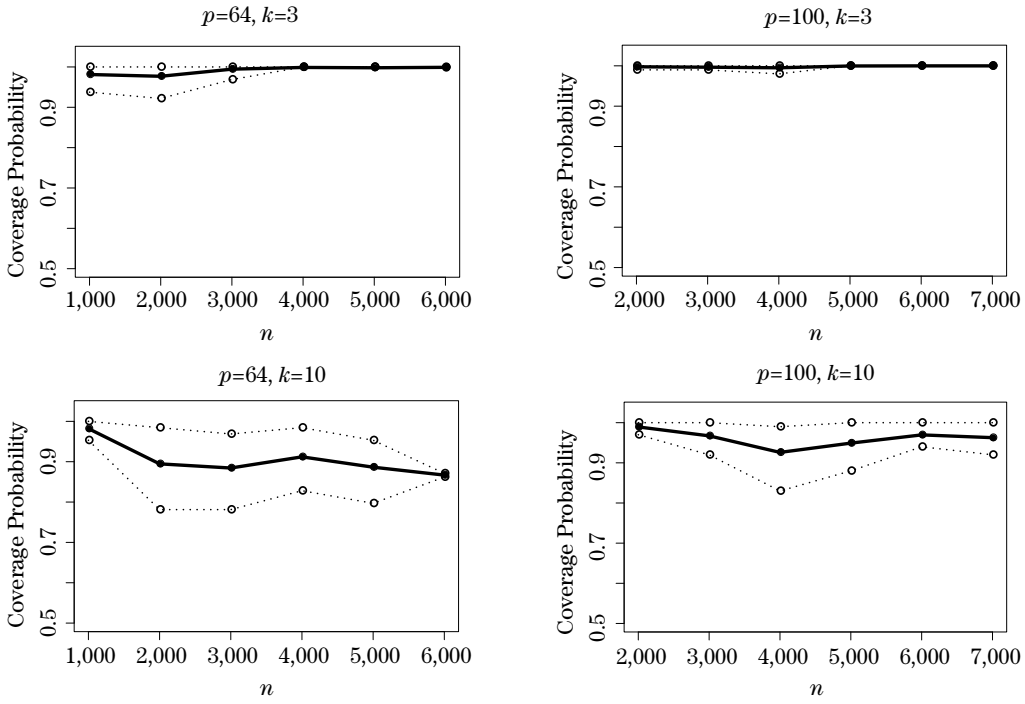

Figure 3. Averaged coverage of the confidence intervals of the in function of n, for different values of p, k. The solid line is the average over 100 iterations the dotted lines form 95% confidence intervals.

tems (qubits), as encountered in ion trap quantum tomography (see Guta, Kypraios and Dryden (2012); Gross et al. (2010); Butucea, Guţă and Kypraios (2015); Haeffner et al. (2005); Acharya, Kypraios and Guta (2015), or Holevo (2001); Nielsen and Chuang (2000) for textbooks on this problem). Such a system's *quantum state* can be represented by a positive semi-definite unit trace complex matrix Θ (a *density matrix*) of dimension $d := 2^m$.

For each individual qubit, the experimenter can measure one of the three Pauli observables described by the 2 by 2 *Pauli matrices* $\sigma_1, \sigma_2, \sigma_3$, where

$$\sigma^1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \ \sigma^2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \ \sigma^3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \tag{4.5}$$

and each of these measurements may yield one of two outcomes, denoted by $+1$ and -1 respectively.

An experiment that describes the measurement for each of the m qubits is then defined by a setting $S = (s_1, \ldots, s_m)$ where each $s_l \in \{\sigma_1, \sigma_2, \sigma_3\}$ for $l \leq m$, which specifies which of the 3 Pauli observables is measured for each

qubit. For each fixed setting S, the measurement produces random outcomes $O \in \{+1, -1\}^m$, with expected probability $p_{O,S} = \mathrm{tr}(P_{O,S}\Theta)$, where $P_{O,S} = \pi_{o_1,s_1} \otimes \ldots, \otimes \pi_{o_m,s_m}$, with π_{o_l,s_l} is the eigen projector of the the 2 by 2 Pauli matrix s_l associated to the eigen value o_l

Set

$$\sigma_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{4.6}$$

for the last 2×2 Pauli matrix such that $(\sigma_i)_{i\in\{0,\ldots,3\}}$ form an orthogonal basis of $\mathbb{C}^{2\times 2}$. Let O be the outcome of an experiment given a setting $S = (s_1, \ldots, s_m)$ (where each $s_l \in \sigma_1, \sigma_2, \sigma_3$). Write $\tilde{S}(E)$ for a setting where a subset $E \subset \{1, \ldots, m\}$ of the m matrices s_l of this setting have been replaced by σ_0, and $\tilde{O}(E)$ for the outcome where the same subset E of the m elements o_l have been replaced by 1. Since the only eigen value of σ_0 is 1, the outcome of the measurement of a qubit by σ_0 is always 1. This implies, in particular, that the distribution of $\tilde{O}(E)$ is the same as the distribution of the outcome of an experiment when the measurement setting is $\tilde{S}(E)$. For this reason, measuring according to setting S gives information about all settings $\tilde{S}(E)$ for any subset E of $\{1, \ldots, m\}$. Instead of measuring all settings which are tensor products of 2×2 Pauli matrices $\sigma_0, \ldots, \sigma_3$, it is enough to measure all settings which are tensor products of 2×2 Pauli matrices $\{\sigma_1, \sigma_2, \sigma_3\}$, as they provide information about corresponding settings that involve σ_0. If one measures all 3^m settings that correspond to the settings $S = (s_1, \ldots, s_m)$ where each $s_l \in \{\sigma_1, \sigma_2, \sigma_3\}$, we have observations about all measurement directions, and our measurement setting is complete.

We are interested in dealing with situations where one does not want to observe all 3^m settings, and where one has only a number of settings $N \leq 3^m$.

We consider a random measurement setting as in Flammia et al. (2015): each s_l in S is chosen uniformly at random among $\sigma_1, \ldots, \sigma_3$, with N the total number of measurement settings. For each chosen measurement setting S^i with $i \leq N$, we observe T independent outcomes O^{t,S^i} that are observations according to setting S^i.

4.2.2. Expression of the outcomes in the trace regression model

It is often convenient to express the information contained by a measurement (S, O) in a way that involves tensor products of 2×2 Pauli matrices, rather than their spectral projections, see Flammia et al. (2015). Indeed, the set of matrices that are created by m tensor products of 2×2 Pauli matrices $\sigma_0, \ldots, \sigma_3$ is exactly

the set of $2^m \times 2^m = d \times d$ Pauli matrices rescaled by $\sqrt{d}$ (introduced briefly in Remark 1), the $d \times d$ rescaled Pauli basis. Indeed if $f(O) = \prod_l o_l$, then

$$\begin{aligned}&\operatorname{tr}((s_1 \otimes \cdots \otimes s_m)\Theta)\\ &= \sum_{O \in \{1,-1\}^m} \Big(\prod_l o_l\Big) \operatorname{tr}\Big((\pi_{s_1,o_1} \otimes \cdots \otimes \pi_{s_m,o_m})\Theta\Big) = \mathbb{E}_{O|S}\big(f(O)\big),\end{aligned}$$

where $\mathbb{E}_{O|S}$ is the expectation according to the outcome O when measurement S is chosen. In this sense, the measurement described by the $d \times d$ rescaled Pauli matrix $P_S = s_1 \otimes \cdots \otimes s_m$ can be measured by the parity of the spins $f(O)$ that one gets when performing measurement S: $f(O)|S$ is a random variable that is 1 with probability $\big(\operatorname{tr}(P_S\Theta)+1\big)/2$, and -1 with probability $1 - \big(\operatorname{tr}(P_S\Theta)+1\big)/2$. Its expectation is $\operatorname{tr}(P_S\Theta)$ as noted. We write $\mathcal{R}(\operatorname{tr}(P_S\Theta))$ for this distribution.

We observe at each measurement S^i, for each replication $t \le T$ and for all $E \subset \{1, \dots m\}$,

$$y_{S^i,E} = f(\tilde{O}^{t,S_i}(E)) \sim \mathcal{R}(\operatorname{tr}(P_{\tilde{S}^i(E)}\Theta)).$$

In our trace regression model, we can average the observations $f(\tilde{O}^{t,S_i}(E))$ to obtain, for any $i \le s$ and for all $E \subset \{1, \dots m\}$,

$$\bar{Y}_{S^i,E} = \frac{1}{T}\sum_{t \le T} \bar{y}^t_{S^i,E} = \operatorname{tr}(P_{\tilde{S}^i(E)}\Theta)) + \bar{\epsilon}_{S^i,E},$$

where $\bar{y}^t_{S^i,E}$ is the t^{th} repetition (among T iterations) of the observation and where $\bar{\epsilon}_{S^i,E}$ is the averaged noise and is such that $\mathbb{E}_{(O^{t,S^i})|S^i}\bar{\epsilon}_{S^i,E} = 0$, and such that $\bar{\epsilon}_{S^i,E}$ is sub-Gaussian has a sum of bounded random variables satisfying $\mathbb{E}_{(O^{t,S^i}))|S^i} \exp(\lambda\bar{\epsilon}_{S^i,E}) \le \exp(\lambda^2/T)$ for any $\lambda \ge 0$.

For Assumption 1 to be satisfied for rank k matrices for a large enough number of settings N, we need to rescale our data. We set

$$Y_{S^i,E} = \sqrt{d}3^{-|E|/2}\Big(\frac{3}{4}\Big)^{m/2}\bar{Y}_{S^i,E} = \sqrt{d}3^{-|E|/2}\Big(\frac{3}{4}\Big)^{m/2}\operatorname{tr}(P_{\tilde{S}^i(E)}\Theta)) + \epsilon_{S^i,E},$$

where $|E|$ is the cardinality of E, and where $\epsilon_{S^i,E}$ is the rescaled noise such that $\mathbb{E}_{(O^{t,S^i})|S^i}\epsilon_{S^i,E} = 0$, and such that $\epsilon_{S^i,E}$ is sub-Gaussian and satisfies $\mathbb{E}_{(O^{t,S^i})|S^i}$ $\exp(\lambda\epsilon_{S^i,E}) \le \exp(\lambda^2 3^{-|E|}\left(\frac{3}{2}\right)^m/T)$ for any $\lambda \ge 0$. It is a direct consequence from results of Liu (2011) and our Remark 1 that if $N \ge O(k^2 d\log(d))$, then with high probability on the random draws of our settings we have Assumption 1 satisfied for rank k matrices. We can therefore apply our method and other low rank recovery methods such as trace regression methods to our rescaled data

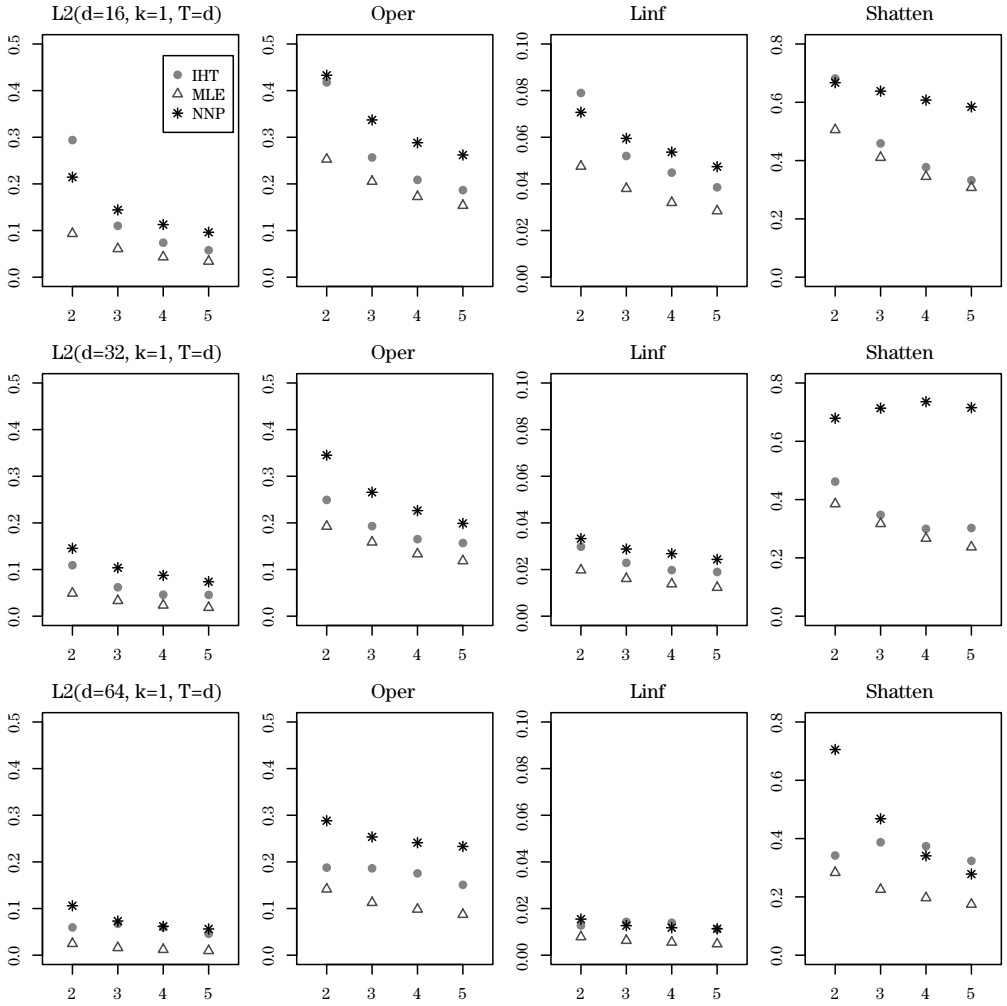

Figure 4. Squared Frobenius norm (L2), operator norm (Oper), entrywise L_∞ norm (Linf) and Shatten 1 norm (Shatten) of $\hat{\Theta} - \Theta$ in function of α (and therefore in function og the number of settings N), for different values of d for replication $T = d$ and $k = 1$ using the three methods described.

$$\left(Y_{S^i,E}, \sqrt{d} 3^{-|E|/2} \left(\frac{3}{4}\right)^{m/2} P_{\tilde{S}^i(E)} \right)_{i \leq N, E \subset \{1,\ldots,m\}} . \tag{4.7}$$

4.2.3. Experimental results

We let $m \in \{4, 5, 6\}$ so that $d \in \{16, 32, 64\}$, and took $k \in \{1, 2\}$ with $\alpha \in \{2, 3, 4, 5\}$. Consider $N = \alpha k d$ measurement settings, of which those with

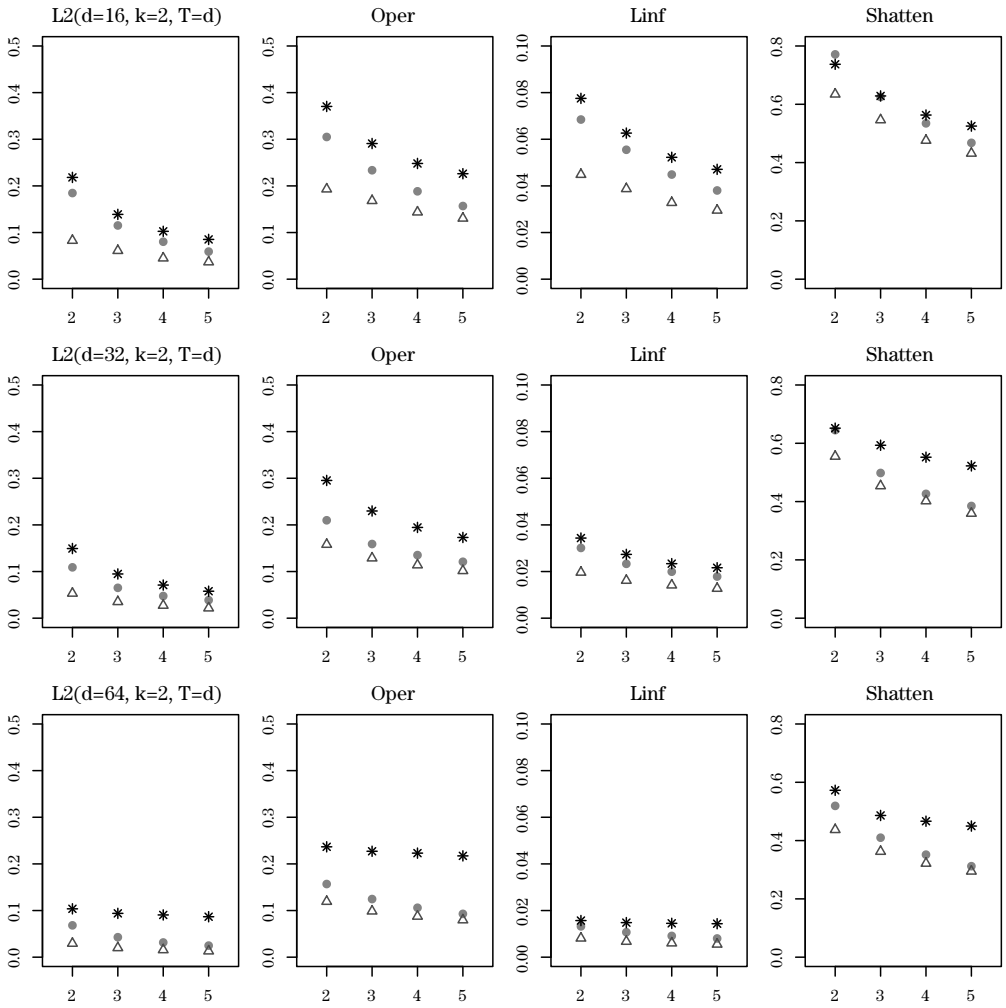

Figure 5. Squared Frobenius norm (L2), operator norm (Oper), entrywise L_∞ norm (Linf) and Shatten 1 norm (Shatten) of $\hat{\Theta} - \Theta$ in function of α, for different values of d for replication $T = d$ and $k = 2$ using three methods.

small k and α were chosen since we are more interested in the truncated measurement setting (such that $N \leq 3^m$). For replications, we took $T \in \{d, 10d\}$. Using the data (4.7), we estimated Θ by our proposed method (IHT), the truncated maximum likelihood estimator (MLE) for the high dimensional multiple ion tomography model as described in Acharya, Kypraios and Guta (2015), and the nuclear norm penalization (NNP) method computed using a gradient descent method (e.g. Agarwal, Negahban and Wainwright (2012)).

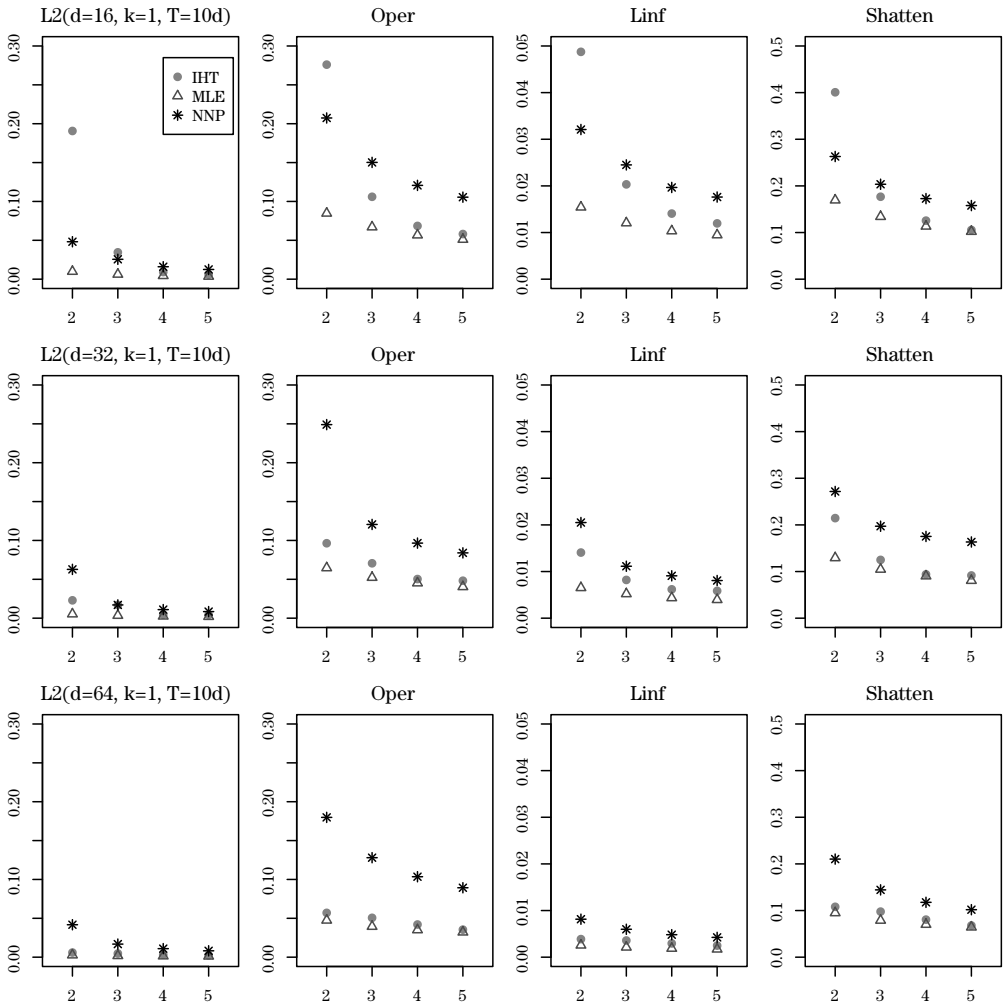

Figure 6. Squared Frobenius norm (L2), operator norm (Oper), entrywise L_∞ norm (Linf) and Shatten 1 norm (Shatten) of $\hat{\Theta} - \Theta$ in function of α, for different values of d for replication $T = 10d$ and $k = 1$ using three methods.

We used the same tuning parameters as in (4.1), (4.2), (4.3), and (4.4) and selected $\rho = 1/2$ and the same stopping rule as described in Section 4. For the stepsize of the gradient descent method used to compute the NNP estimator, we followed the recommendation in Subsection 3.1 in Agarwal, Negahban and Wainwright (2012).

Figure 4 and 5 present the results when the number T of replications is d, when the true rank is 1 or 2, respectively, and for four values of d. We

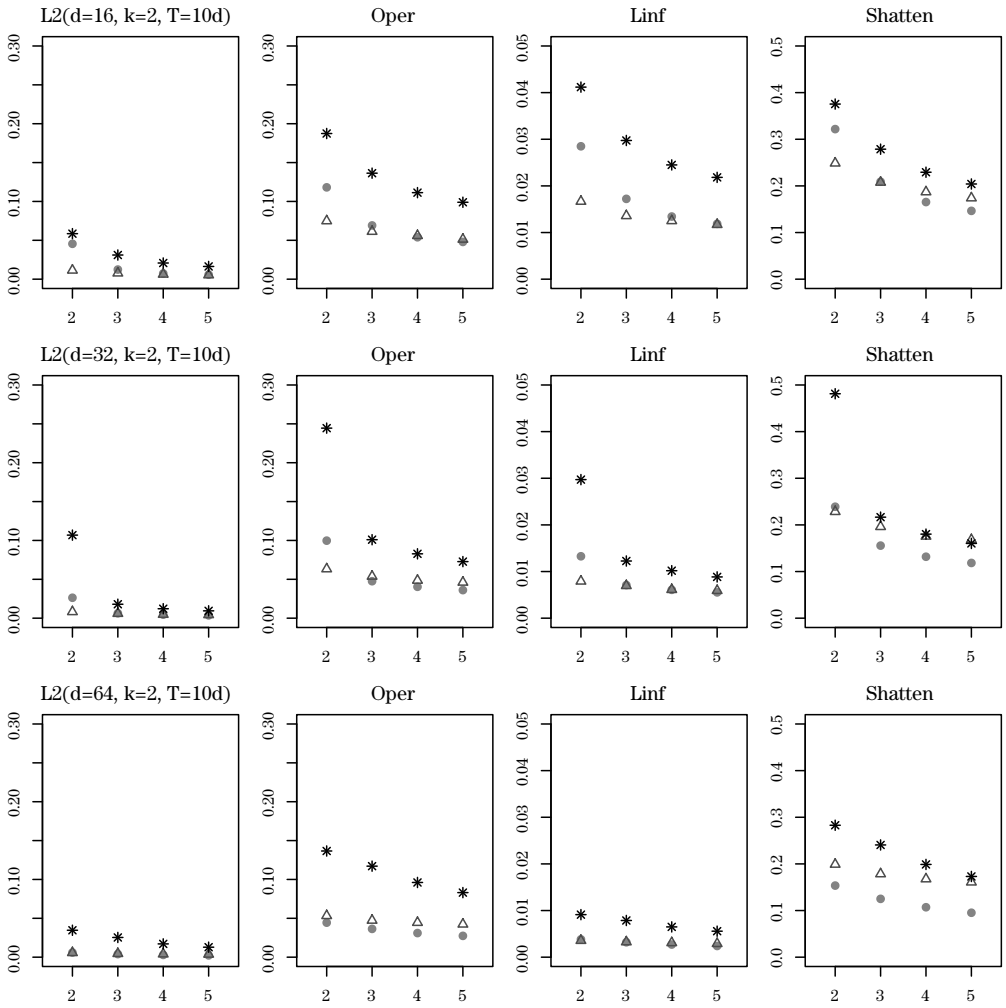

Figure 7. Squared Frobenius norm (L2), operator norm (Oper), entrywise L_∞ norm (Linf) and Shatten 1 norm (Shatten) of $\hat{\Theta} - \Theta$ in function of α, for different values of d for replication $T = 10d$ when $k = 2$ using three methods.

provide average values of squared Frobenius norm, operator norm, entrywise L_∞ norm, and Shatten 1 norm of $\hat{\Theta} - \Theta$ averaged over 100 iterations. Dots, blank triangles, and asterisks are average value of IHT, MLE, and NNP, respectively. Intuitively when α increases these risks should decrease. Our estimator shows almost comparable results to the MLE except when $d = 4$ and $\alpha = 2$; then IHT estimates Θ by 0 a few times and pretty well for most cases, so that on average the Frobenius norm is still large. Figures 6 and 7 present the results with $10d$

replications when the true rank is 1 or 2, respectively, and for four values of d. There are similar patterns as for the cases $T = d$, but IHT performs well, especially for $\alpha \in \{4, 5\}$. These pictures illustrate that IHT performs the best relatively to other methods for a large number of replication T, and for difficult problems with high d and k (see in particular the plots in Figure 7 for large d). Our method is computationally more efficienct than the other two methods in the sense that when $d = 64, \alpha = 5, k = 2$, IHT takes about 40 seconds while MLE (and even NNP) takes about 2.5 minutes for one iteration on a regular laptop.

Supplementary Materials

The online supplementary material contains proofs and results for the sparse linear regression model. These materials are presented in Sections S1 and S2, respectively.

Acknowledgment

We thank Richard Nickl, Richard Samworth and Rajen Shah for insightful comments and discussions. Part of this work was produced when AC and AKHK were in the StatsLab in the University of Cambridge. AC's work is supported since 2015 by the DFG' Emmy Noether grant MuSyAD (CA 1488/1-1).

References

Acharya, A., Kypraios, T. and Guta, M. (2015). Efficient quantum tomography with incomplete measurement settings. *arXiv preprint arXiv:1510.03229*.

Agarwal, A., Negahban, S. and Wainwright, M. J. (2012). Fast global convergence of gradient methods for high-dimensional statistical recovery. *Ann. Statist.* **40**, 2452–2482.

Bickel, P., Ritov, Y. and Tsybakov, A. (2009). Simultaneous analysis of lasso and dantzig selector. *Ann. Statist.* **37**, 1705–1732.

Blumensath, T. and Davies, M. E. (2009). Iterative hard thresholding for compressed sensing. *Appl. Computat. Har. Analysis* **27**, 265–274.

Bunea, F., She, Y. and Wegkamp, M. (2011). Optimal selection of reduced rank estimators of high-dimensional matrices. *Ann. Statist.* **39**, 1282–1309.

Butucea, C., Guţă, M. and Kypraios, T. (2015). Spectral thresholding quantum tomography for low rank states. *New Journal of Physics* **17**, 113050.

Cai, T. and Zhou, H. H. (2012). Optimal rates of convergence for sparse covariance matrix estimation. *Ann. Statist.* **40**, 2389–2420.

Cai, T. and Zhang, A. (2015). ROP: matrix recovery via rank-one projections. *Ann. Statist.* **43**, 102–138.

Candès, E. and Tao, T. (2010). The power of convex relazation: near-optimal matrix completion.

IEEE Trans. Inform. Theory **56**, 2053–2080.

Candès, E. J. and Plan, Y. (2011). Tight oracle bounds for low-rank matrix recovery from a minimal number of random measurements. *IEEE Trans. Inform. Theory* **57**, 2342–2359.

Candès, E. and Recht, B. (2009). Exact matrix completion via convex optimization. *Found. Comput. Math.* **9**, 717–772.

Carpentier, A., Eisert, J., Gross, D. and Nickl, R. (2015). Uncertainty quan-tification for matrix compressed sensing and quantum tomography problems. *arXiv preprint arXiv:1504.03234*.

Chen, Y. and Wainwright, M. (2015). Fast low-rank estimation by projected gradient descent: General statistical and algorithmic guarantees. *arXiv preprint arXiv:1509.03025*.

Flammia, S. T, Gross, D., Liu, Y.-K. and Eisert, J. (2015). Quantum tomography via compressed sensing: error bounds, sample complexity and efficient estimators. *New J. Phys.* **14**, 095022, 2012.

Friedman, J., Hastie, T. and Tibshirani, R. (2008). Sparse inverse covariance estimation with the graphical lasso. *Biostatistics* **9**, 432–441.

van de Geer, S., Bühlmann, P., Ritov, Y. and Dezeure, R. (2014). On asymptotically optimal confidence regions and tests for high-dimensional models. *Ann. Statist.* **42**, 1166–1202.

Goldfarb, D. and Ma, S. (2011). Convergence of fixed-point continuation algorithms for matrix rank minimization. *Found. Comput. Math.* **11**, 183–210.

Gross, D. (2011). Recovering low-rank matrices from few coefficients in any basis. *IEEE Trans. Inform. Theory* **57**, 1548–1566.

Gross, D., Liu, Y.-K., Flammia, S. T., Becker, S. and Eisert, J. (2010). Quantum state tomography via compressed sensing. *Physical Rev. Letters* **105**, 150401.

Guta, M., Kypraios, T. and Dryden, I. (2012). Rank-based model selection for multiple ions quantum tomography. *New J. Phys.* **14**, 105002.

Haeffner, H., Haensel, W., Roos, C. F., Benhelm, J., al Kar, D. C., Chwalla, M., Koerber, T., Rapol, U. D., Riebe, M., Schmidt, P. O., Becher, C., Gühne, O., Dur, W. and Blatt, R. (2005). Scalable multi-particle entanglement of trapped ions. *Nature* **438**, 643.

Holevo, A. S. (2001). *Statistical Structure of Quantum Theory*. Springer.

Huang, H., Ma, S. and Zhang, C.-H. (2008). Adaptive lasso for sparse high-dimensional regression models. *Stat. Sinica* **18**, 1603–1618.

Javanmard, A. and Montanari, A. (2014). Confidence intervals and hypothesis testing for high-dimensional regression. *J. Mach. Learn. Res.* **15**, 2869–2909.

Klopp, O. (2015). Matrix completion by singular value thresholding: sharp bounds. *Electron. J. Statist.* **9**, 2348–2369.

Koltchinskii, V. (2011). on Neumann entropy penalization and low-rank matrix estimation. *Ann. Statist.* **39**, 2936–2973.

Koltchinskii, V., Lounici, K. and Tsybakov, A. B. (2011). Nuclear-norm penalization and optimal rates for noisy low-rank matrix completion. *Ann. Statist.* **39**, 2302–2329.

Koltchinskii, V. and Xia, D. (2015). Optimal estimation of low rank density matrices. *arXiv preprint arXiv:1507.05131*.

Liu, Y. K. (2011). Universal low-rank matrix recovery from Pauli measurements. *Adv. Neural Inf. Process. Syst.* 1638–1646.

Meinshausen, N. and Bühlmann, P. (2006). High dimensional graphs and variable selection with the lasso. *Ann. Statist.* **34**, 1436–1462.

Needell, D. and Tropp, J. A. (2009). CosaMP: Iterative signal recovery from incomplete and inaccurate samples. *Appl. Comput. Harmon. Anal.* **26**, 301–321.

Negahban, S. and Wainwright, M. (2011). Estimation of (near) low-rank matrices with noise and high-dimensional scaling. *Ann. Statist.* **39**, 1069–1097.

Nickl, R. and van de Geer, S. (2014). Confidence sets in sparse regression. *Ann. Statist.* **41**, 2852–2876.

Nielsen, M. A. and Chuang, I. L. (2000). *Quantum Computation and Quantum Information.* Cambridge: Cambridge University Press.

Recht, B. (2011). A simpler approach to matrix completion. *J. Mach. Learn. Res.* **12**, 3413–3430.

Tanner, J. and Wei, K. (2012). Normalized iterative hard thresolding for matrix completion. *SIAM J. Sci. Comput.* **35**, S104–S125.

Zhang, C.-H. and Zhang, S.-S. (2014). Confidence intervals for low dimensional parameters in high dimensional linear models. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **76**, 217–242.

Institut für Mathematik, Universitãt Potsdam, Karl-Liebknecht-Strasse 24-2, 14476 Potsdam, Germany.

E-mail: carpentier@math.uni-potsdam.de

Department of Statistics, Sungshin Women's University, 34Da-Gil 2, Bomun-Ro, Seongbuk-Gu, Seoul 02844, South Korea.

E-mail: arlenekim@sungshin.ac.kr

(Received February 2015; accepted January 2017)