New Journal of Physics
The open access journal at the forefront of physics
Deutsche Physikalische Gesellschaft DPG IOP Institute of Physics

CORRIGENDUM • OPEN ACCESS

Corrigendum: Spectral thresholding quantum tomography for low rank states (2015 *New J. Phys.* **17** 113050)

IOP Publishing New J. Phys. 18 (2016) 069501 doi:10.1088/1367-2630/18/6/069501

New Journal of Physics
The open access journal at the forefront of physics

Deutsche Physikalische Gesellschaft DPG
IOP Institute of Physics

Published in partnership with: Deutsche Physikalische Gesellschaft and the Institute of Physics

CrossMark

OPEN ACCESS

PUBLISHED
3 June 2016

CORRIGENDUM

Corrigendum: Spectral thresholding quantum tomography for low rank states (2015 *New J. Phys.* **17** 113050)

Cristina Butucea[1], **Mădălin Guţă**[2] **and Theodore Kypraios**[2]

[1] Université Paris-Est Marne-la-Vallée, LAMA(UMR 8050), UPEMLV F-77454, Marne-la-Vallée, France
[2] University of Nottingham, School of Mathematical Sciences, University Park, Nottingham NG7 2RD, UK

In this corrigendum to the paper Butucea *et al* (2015 *New J. Phys.* **17** 113050) we point out an error in one of the theoretical results describing the upper bound to the operator norm error of the least squares estimator. We provide a corrected version of the upper bound with a new convergence rate, and discuss the implications for other results which rely on the above upper bound.

Proposition 1 as stated in the paper is incorrect, in particular the dependence of the upper bound $\nu(\epsilon)^2$ on the number of atoms k is not valid. The error lies in the evaluation of the upper bound W of the variance term in the concentration bound. Below we provide a new version of proposition 1 with a corrected rate $\nu_c(\epsilon)^2$ replacing the rate $\nu(\epsilon)^2$ stated in the paper. Ignoring the logarithmic factors, the new upper bound scales as $3^k/N$ compared to erroneous rate $2^k/N$, where k is the number of atoms and $N = n3^k$ is the total number of measurements. We note that although the corrected bound is weaker that the one claimed in the paper, it is still an improvement compared to the previously known bound [2] which scaled as $4^k/N$.

We will now discuss the implication of the correction to subsequent results in the paper. Proposition 2, theorem 1, corollary 1, and theorem 2 establish error rates for estimators obtained by normalising, penalising or thresholding the least square estimator. The proofs of these results use the operator norm error rate $\nu(\epsilon)^2$ as a generic expression, and are therefore not affected by its concrete dependence on the number of atoms k. Therefore proposition 2, theorem 1, corollary 1, and theorem 2 hold true when the operator norm rate is taken to have expression $\nu_c(\epsilon)^2$ in proposition 1 below. In particular, the upper bounds on the Frobenius square norm error in corollary 1 and theorem 2, will scale as $r \cdot \nu_c(\varepsilon)^2 = r3^k/N$ rather than $rd/N = r2^k/N$. The remaining results including the lower bound in theorem 3 and the simulation results are independent of proposition 1 and do not require any correction.

Proposition 1. *Let $\widehat{\rho}_n^{(ls)}$ be the linear estimator of ρ. Then, for any $\varepsilon > 0$, the following operator norm inequality holds, for n large enough, with probability larger than $1 - \varepsilon$ under $\mathbb{P}_\rho$*

$$\|\widehat{\rho}_n^{(ls)} - \rho\| \leqslant \nu_c(\varepsilon),$$

where

$$\nu_c(\varepsilon)^2 = \frac{4 \cdot 3^k}{N} \log\left(\frac{2^{k+1}}{\varepsilon}\right)$$

with $N := n \cdot 3^k$ the total number of measurements. The same bound holds when $k = k(n)$ as long as $\nu(\varepsilon) \to 0$.

Proof of proposition 1. Note that the empirical frequencies can written as $f(\mathbf{o}|\mathbf{s}) = \frac{1}{n}\sum_i I(X_{\mathbf{s},i} = \mathbf{o})$, where the random variables $X_{\mathbf{s},i}$ are independent for all settings $\mathbf{s}$ and all i from 1 to n. To estimate the risk of the linear estimator we write

$$\begin{aligned}\widehat{\rho}_n^{(ls)} - \rho &= \sum_{\mathbf{b}}\sum_{\mathbf{o}}\sum_{\mathbf{s}}(f(\mathbf{o}|\mathbf{s})) - p(\mathbf{o}|\mathbf{s}))\frac{A_{\mathbf{b}}(\mathbf{o}|\mathbf{s})}{2^k 3^{d(\mathbf{b})}}\sigma_{\mathbf{b}}\\ &= \sum_{\mathbf{b}}\sum_{\mathbf{o}}\sum_{\mathbf{s}}\frac{1}{n}\sum_i(I(X_{\mathbf{s},i} = \mathbf{o}) - p(\mathbf{o}|\mathbf{s}))\frac{A_{\mathbf{b}}(\mathbf{o}|\mathbf{s})}{2^k 3^{d(\mathbf{b})}}\sigma_{\mathbf{b}}\\ &:= \sum_{\mathbf{s}}\sum_i W_{\mathbf{s},i}.\end{aligned}$$

where $W_{\mathbf{s},i}$ are independent and centered Hermitian random matrices. We will apply the following extension of the Bernstein matrix inequality [1] due to Tropp, see also [4, 6].

Proposition 2 (Bernstein inequality, Tropp). *Let $Y_1, \ldots, Y_n$ be independent, centered, $m \times m$ Hermitian random matrices. Suppose that, for some constants $V,\ W > 0$ we have $\|Y_j\| \leqslant V$, for all j from 1 to n, and that $\|\sum_j \mathbb{E}(Y_j^2)\| \leqslant W$. Then, for all $t \geqslant 0$,*

$$\mathbb{P}(\|Y_1 + \ldots + Y_n\| \geqslant t) \leqslant 2\, m \exp\left(-\frac{t^2/2}{W + tV/3}\right).$$

In our setup, $W_{\mathbf{s},i}$ play the role of Y_j. We bound $\|W_{\mathbf{s},i}\| \leqslant V$ for all $\mathbf{s}$ and i and $\|\sum_{\mathbf{s},i} \mathbb{E}(W_{\mathbf{s},i}^* W_{\mathbf{s},i})\| \leqslant W$, where $V,\ W$ are evaluated below. We have

$$\begin{aligned}
\|W_{\mathbf{s},i}\| &\leqslant \frac{1}{n}\sum_{\mathbf{b}}\sum_{\mathbf{o}}\left|\frac{A_{\mathbf{b}}(\mathbf{o}|\mathbf{s})}{2^k 3^{d(\mathbf{b})}}\right| \cdot |I(X_{\mathbf{s},i} = \mathbf{o}) - p(\mathbf{o}|\mathbf{s})| \cdot \|\sigma_{\mathbf{b}}\| \\
&\leqslant \frac{1}{n}\sum_{\mathbf{b}}\frac{1}{2^k 3^{d(\mathbf{b})}}\prod_{j\notin E_{\mathbf{b}}}\delta_{b_j,s_j}\sum_{\mathbf{o}}|I(X_{\mathbf{s},i} = \mathbf{o}) - p(\mathbf{o}|\mathbf{s})| \\
&\leqslant \frac{2}{n2^k}\sum_{\ell=0}^{k}\sum_{E:|E|=\ell}\frac{1}{3^\ell} = \frac{2}{n2^k}\sum_{\ell=0}^{k}\binom{k}{\ell}\frac{1}{3^\ell} = \frac{2}{n2^k}\left(1+\frac{1}{3}\right)^k = 2\frac{2^k}{n\cdot 3^k} \coloneqq V.
\end{aligned}$$

Now, denote by $B(\mathbf{o}|\mathbf{s}) = \sum_{\mathbf{b}} 2^{-k}3^{-d(\mathbf{b})}A_{\mathbf{b}}(\mathbf{o}|\mathbf{s})\sigma_{\mathbf{b}}$ so that $W_{\mathbf{s},i} = \sum_{\mathbf{o}} B(\mathbf{o}|\mathbf{s})(I(X_{\mathbf{s},i} = \mathbf{o}) - p(\mathbf{o}|\mathbf{s}))$. Then

$$\begin{aligned}
&\left\|\sum_{\mathbf{s}}\sum_{i}\mathbb{E}(W_{\mathbf{s},i}^* W_{\mathbf{s},i})\right\| \\
&= \frac{1}{n^2}\left\|\sum_{\mathbf{s}}\sum_{i}\sum_{\mathbf{o},\mathbf{o}'}B^*(\mathbf{o}|\mathbf{s})\cdot \mathrm{Cov}(I(X_{\mathbf{s},i} = \mathbf{o}), I(X_{\mathbf{s},i} = \mathbf{o}'))\cdot B(\mathbf{o}'|\mathbf{s})\right\| \\
&= \frac{1}{n}\left\|\sum_{\mathbf{s}}\sum_{\mathbf{o},\mathbf{o}'}B^*(\mathbf{o}|\mathbf{s})\cdot \mathrm{Cov}(I(X_{\mathbf{s},1} = \mathbf{o}), I(X_{\mathbf{s},1} = \mathbf{o}'))\cdot B(\mathbf{o}'|\mathbf{s})\right\| \\
&\leqslant \frac{1}{n}\left\|\sum_{\mathbf{s}}\sum_{\mathbf{o}}p(\mathbf{o}|\mathbf{s})B^*(\mathbf{o}|\mathbf{s})\cdot B(\mathbf{o}|\mathbf{s})\right\| \coloneqq \frac{1}{n}\|T\|. \qquad (1)
\end{aligned}$$

In the last inequality we used that

$$\mathrm{Cov}(\mathbf{s})_{\mathbf{o},\mathbf{o}'} \coloneqq \mathrm{Cov}(I(X_{\mathbf{s},1} = \mathbf{o}), I(X_{\mathbf{s},1} = \mathbf{o}')) = p(\mathbf{o}|\mathbf{s})\delta_{\mathbf{o},\mathbf{o}'} - p(\mathbf{o}|\mathbf{s})\cdot p(\mathbf{o}'|\mathbf{s})$$

which implies the following inequality between $2^k \times 2^k$ matrices: $\mathrm{Cov}(\mathbf{s}) \leqslant p(\mathbf{s})$ where $p(\mathbf{s})$ is the diagonal matrix with elements $p(\mathbf{s})_{\mathbf{o},\mathbf{o}'} = p(\mathbf{o}|\mathbf{s})\delta_{\mathbf{o},\mathbf{o}'}$.

By expressing $B(\mathbf{o}|\mathbf{s})$ in terms of $\sigma_{\mathbf{b}}$ as above, we get

$$\frac{1}{n}\|T\| = \frac{1}{n}\left\|\sum_{\mathbf{s}}\sum_{\mathbf{o}}p(\mathbf{o}|\mathbf{s})\sum_{\mathbf{b},\mathbf{b}'}\frac{A_{\mathbf{b}}(\mathbf{o}|\mathbf{s})}{2^k 3^{d(\mathbf{b})}}\frac{A_{\mathbf{b}'}(\mathbf{o}|\mathbf{s})}{2^k 3^{d(\mathbf{b}')}}\sigma_{\mathbf{b}}\sigma_{\mathbf{b}'}\right\|.$$

Before giving the upper bound we introduce some combinatorial notations which will be used below. Let $\mathbf{b} \in \{x, y, z, I\}^k$ and recall that $E_{\mathbf{b}} \coloneqq \{i : b_i = I\} \subset \{1, \ldots, k\}$. We say that $\mathbf{b}$ agrees with a setting $\mathbf{s}$ if $b_j = s_j$ for all $j \in E_{\mathbf{b}}^c$. In this case $\mathbf{b}$ is completely determined by the set $E_{\mathbf{b}}$, for a fixed $\mathbf{s}$. This fact will be used to replace the sums over $\mathbf{b}$ and $\mathbf{b}'$ with those over $E_{\mathbf{b}}$ and $E_{\mathbf{b}'}$ in T. Indeed since $A_{\mathbf{b}}(\mathbf{o}|\mathbf{s})$ is proportional to $\prod_{j\notin E_{\mathbf{b}}}\delta_{b_j,s_j}$, the sums over $\mathbf{b}$ and $\mathbf{b}'$ in T are restricted to sequences which agree with $\mathbf{s}$. We denote by $E_{\mathbf{b}} \cap E_{\mathbf{b}'} \coloneqq (E_{\mathbf{b}} \setminus E_{\mathbf{b}'}) \cup (E_{\mathbf{b}'} \setminus E_{\mathbf{b}})$ and $E_{\mathbf{b}} \cap E_{\mathbf{b}'}$ the symmetric difference and respectively the intersection of $E_{\mathbf{b}}$ and $E_{\mathbf{b}'}$. With these notations we have

$$A_{\mathbf{b}}(\mathbf{o}|\mathbf{s})A_{\mathbf{b}'}(\mathbf{o}|\mathbf{s})\sigma_{\mathbf{b}}\sigma_{\mathbf{b}'} = \prod_{j\in E_{\mathbf{b}}\Delta E_{\mathbf{b}'}} o_j \cdot \sigma_{\mathbf{g}}$$

where $\mathbf{g} = \mathbf{g}(\mathbf{s}, E_{\mathbf{b}}\Delta E_{\mathbf{b}'})$ is the sequence with $E_{\mathbf{g}} = (E_{\mathbf{b}}\Delta E_{\mathbf{b}'})^c$, and it agrees with $\mathbf{s}$. With these notations we have

$$
\begin{aligned}
T &= \frac{1}{2^{2k}} \sum_{\mathbf{s}} \sum_{\mathbf{o}} p(\mathbf{o}|\mathbf{s}) \sum_{E_{\mathbf{b}}, E_{\mathbf{b}'}} \frac{\prod_{j \in E_{\mathbf{b}} \Delta E_{\mathbf{b}'}} o_j}{3^{|E_{\mathbf{b}}| + |E_{\mathbf{b}'}|}} \sigma_{\mathbf{g}(\mathbf{s}, E_{\mathbf{b}} \Delta E_{\mathbf{b}'})} \\
&= \frac{1}{2^{2k}} \sum_{E_{\mathbf{b}}, E_{\mathbf{b}'}} \sum_{\mathbf{s}} \sum_{\mathbf{o}} p(\mathbf{o}|\mathbf{s}) \frac{\prod_{j \in E_{\mathbf{b}} \Delta E_{\mathbf{b}'}} o_j}{3^{|E_{\mathbf{b}}| + |E_{\mathbf{b}'}|}} \sigma_{\mathbf{g}(\mathbf{s}, E_{\mathbf{b}} \Delta E_{\mathbf{b}'})} \\
&= \frac{1}{2^{k}} \sum_{E_{\mathbf{b}}, E_{\mathbf{b}'}} \frac{3^{k - |E_{\mathbf{b}} \Delta E_{\mathbf{b}'}|}}{3^{|E_{\mathbf{b}}| + |E_{\mathbf{b}'}|}} \sum_{\mathbf{s}} \sum_{\mathbf{o}} p(\mathbf{o}|\mathbf{s}) \frac{\prod_{j \in E_{\mathbf{b}} \Delta E_{\mathbf{b}'}} o_j}{2^k \cdot 3^{k - |E_{\mathbf{b}} \Delta E_{\mathbf{b}'}|}} \sigma_{\mathbf{g}(\mathbf{s}, E_{\mathbf{b}} \Delta E_{\mathbf{b}'})}
\end{aligned} \tag{2}
$$

In the last expression we rewrite the sum over settings $\mathbf{s}$ as a double sum over $\tilde{\mathbf{s}}$ and $\tilde{\mathbf{s}}^c$ where $\tilde{\mathbf{s}}$ is the restriction of $\mathbf{s}$ to $E_{\mathbf{b}} \Delta E_{\mathbf{b}'}$ and $\tilde{\mathbf{s}}^c$ is the restriction to $(E_{\mathbf{b}} \Delta E_{\mathbf{b}'})^c$. Note that $\mathbf{g}(\mathbf{s}, E_{\mathbf{b}} \Delta E_{\mathbf{b}'})$ depends on $\mathbf{s}$ only through the component $\tilde{\mathbf{s}}$. Then

$$
\begin{aligned}
\sum_{\mathbf{s}} \sum_{\mathbf{o}} p(\mathbf{o}|\mathbf{s}) \frac{\prod_{j \in E_{\mathbf{b}} \Delta E_{\mathbf{b}'}} o_j}{2^k \cdot 3^{k - |E_{\mathbf{b}} \Delta E_{\mathbf{b}'}|}} \sigma_{\mathbf{g}(\mathbf{s}, E_{\mathbf{b}} \Delta E_{\mathbf{b}'})} &= \sum_{\tilde{\mathbf{s}}} \left(\sum_{\tilde{\mathbf{s}}^c} \sum_{\mathbf{o}} p(\mathbf{o}|\mathbf{s}) \frac{\prod_{j \in E_{\mathbf{b}} \Delta E_{\mathbf{b}'}} o_j}{2^k \cdot 3^{k - |E_{\mathbf{b}} \Delta E_{\mathbf{b}'}|}} \right) \sigma_{\mathbf{g}(\tilde{\mathbf{s}}, E_{\mathbf{b}} \Delta E_{\mathbf{b}'})} \\
= \sum_{\tilde{\mathbf{s}}} \rho_{\mathbf{g}(\tilde{\mathbf{s}}, E_{\mathbf{b}} \Delta E_{\mathbf{b}'})} \sigma_{\mathbf{g}(\tilde{\mathbf{s}}, E_{\mathbf{b}} \Delta E_{\mathbf{b}'})} &= \frac{1}{2^k} \sum_{\mathbf{g}: E_{\mathbf{g}} = (E_{\mathbf{b}} \Delta E_{\mathbf{b}'})^c} \mathrm{Tr}(\rho \sigma_{\mathbf{g}}) \sigma_{\mathbf{g}}
\end{aligned}
$$

In the second equality we have used formulas (2.3) and (2.6) in the paper [3], to evaluate the interior sum as a Fourier coefficient of ρ. In the third equality we replaced the sum over $\tilde{\mathbf{s}}$ with an equivalent sum over sequences $\mathbf{g}$ such that $E_{\mathbf{g}} = (E_{\mathbf{b}} \Delta E_{\mathbf{b}'})^c$.

Note that any pair (E, E') is uniquely determined by three disjoint subsets, $(E \backslash E', E' \backslash E, E \cap E')$, or equivalently by the symmetric difference $D := E \Delta E'$ together with $F := E \backslash E' \subset D$ and $M := E \cap E'$. The sum over E, E' in (2) is computed by summing over all triples D, F, M satisfying the above conditions:

$$
\begin{aligned}
T &= \frac{1}{2^{2k}} \sum_{E, E'} \frac{3^{k - |E \Delta E'|}}{3^{|E| + |E'|}} \sum_{\mathbf{g}: E_{\mathbf{g}} = (E \Delta E')^c} \mathrm{Tr}(\rho \sigma_{\mathbf{g}}) \sigma_{\mathbf{g}} \\
&= \frac{1}{2^{2k}} \sum_{D} \sum_{F \subset D} \sum_{M: M \cap D = \varnothing} \frac{3^{k - |D|}}{3^{|D| + 2|M|}} \sum_{\mathbf{g}: E_{\mathbf{g}} = D^c} \mathrm{Tr}(\rho \sigma_{\mathbf{g}}) \sigma_{\mathbf{g}} \\
&= \frac{1}{2^{2k}} \sum_{D} 2^{|D|} \sum_{M: M \cap D = \varnothing} \frac{3^{k - |D|}}{3^{|D| + 2|M|}} \sum_{\mathbf{g}: E_{\mathbf{g}} = D^c} \mathrm{Tr}(\rho \sigma_{\mathbf{g}}) \sigma_{\mathbf{g}}
\end{aligned}
$$

Indeed, the sum over $F \subset D$ gives a factor $2^{|D|}$ since the summands do not depend on F. Next, the sum over M is performed by summing over the size $m = |M|$ and the binomial coefficient represents the number of sets M of a given size.

$$
\begin{aligned}
T &= \frac{1}{2^{2k}} \sum_{D} 2^{|D|} \sum_{m=0}^{k - |D|} \binom{k - |D|}{m} \frac{3^{k - |D|}}{3^{|D| + 2m}} \sum_{\mathbf{g}: E_{\mathbf{g}} = D^c} \mathrm{Tr}(\rho \sigma_{\mathbf{g}}) \sigma_{\mathbf{g}} \\
&= \left(\frac{3}{4}\right)^k \sum_{D} \left(\frac{2}{9}\right)^{|D|} \left(\frac{10}{9}\right)^{k - |D|} \sum_{\mathbf{g}: E_{\mathbf{g}} = D^c} \mathrm{Tr}(\rho \sigma_{\mathbf{g}}) \sigma_{\mathbf{g}} \\
&= \left(\frac{5}{6}\right)^k \sum_{D} \left(\frac{1}{5}\right)^{|D|} \sum_{\mathbf{g}: E_{\mathbf{g}} = D^c} \mathrm{Tr}(\rho \sigma_{\mathbf{g}}) \sigma_{\mathbf{g}} \\
&= \left(\frac{5}{6}\right)^k 2^k \mathcal{D}^{\otimes k}(\rho).
\end{aligned} \tag{3}
$$

The final sum goes over subsets D and over sequences $\mathbf{g}$ such that $E_{\mathbf{g}} = D^c$, and is similar to the Fourier decomposition of ρ except that each terms is weighted by the factor $5^{-|D|}$. In fact, a closer inspection shows that the weighted sum is nothing but the output state of a product of depolarising channels acting in parallel in the state ρ, where an individual depolarising channel is defined by

$$
\mathcal{D}: \frac{1}{2}(I + r_x \sigma_x + r_y \sigma_y + r_z \sigma_z) : \mapsto \frac{1}{2}\left(I + \frac{r_x}{5}\sigma_x + \frac{r_y}{5}\sigma_y + \frac{r_z}{5}\sigma_z\right).
$$

For an arbitrary quantum channel $\mathcal{T}$, let $\nu_p(T) := \sup_{\tau} \mathrm{Tr}(T(\tau)^p)^{1/p}$ be its p-norm, where the supremum is taken over all input states τ; in particular for $p \to \infty$ this becomes the ∞-norm $\nu_\infty(T) := \sup_\tau \|T(\tau)\|$. For the depolarising channel $\mathcal{D}$ defined above, the ∞-norm can be computed easily by applying the channel to an arbitrary pure state and is equal to $\nu_\infty(\mathcal{D}) = 3/5$. Moreover, it is known [5] that the depolarising channel has multiplicative p-norm, i.e. $\nu_p(\mathcal{D}^{\otimes k}) = \nu_p(\mathcal{D})^k$, which implies that $\|\mathcal{D}^{\otimes k}(\rho)\| \leqslant (3/5)^k$. Together with (3) this gives upper bound

$$\frac{1}{n}\|T\| \leqslant \frac{1}{n}\left(\frac{5}{6}\right)^k \cdot 2^k \cdot \left(\frac{3}{5}\right)^k = \frac{3^k}{n \cdot 3^k} = \frac{3^k}{N}. \tag{4}$$

Putting together (1) and (4) we obtain

$$\|\sum_{\mathbf{s}} \sum_i \mathbb{E}(W_{\mathbf{s},i}^* W_{\mathbf{s},i})\| \leqslant \frac{3^k}{n \cdot 3^k} =: W.$$

We apply now the matrix Bernstein inequality in proposition 2 to get, for any $t > 0$:

$$\mathbb{P}_\rho(\|\widehat{\rho}_n^{(ls)} - \rho\| \geqslant t) \leqslant 2^{k+1} \exp\left(-\frac{n \cdot t^2/2}{1 + t \cdot (2/3)^{k+1}}\right).$$

We choose $t > 0$ such that

$$2^{k+1} \exp\left(-\frac{n \cdot t^2/2}{1 + t \cdot (2/3)^{k+1}}\right) \leqslant \varepsilon,$$

which leads, for n large enough, to $t = \nu_c(\varepsilon)$ such that

$$\nu_c(\varepsilon)^2 = \frac{4 \cdot 3^k}{N} \log\left(\frac{2^{k+1}}{\varepsilon}\right).$$

□

Acknowledgments

MG's work was supported by the EPSRC Grant No. EP/J009776/1.

References

[1] Ahlswede R and Winter A 2002 Strong converse for indentification via quantum channels *IEEE Trans. Inf. Theory* **48** 569–79
[2] Alquier P, Butucea C, Hebiri M, Meziani K and Morimae T 2013 Rank penalized estimation of a quantum system *Phys. Rev.* A **88** 032113
[3] Butucea C, Guta M and Kypraios T 2015 Spectral thresholding quantum tomography for low rank states *New J. Phys.* **17** 113050
[4] Gross D 2011 Recovering low-rank matrices from few coefficients in any basis *IEEE Transactions on Information Theory* **57** 1548–66
[5] King C 2002 A remark on low rank matrix recovery and noncommutative Bernstein type inequalities *J. Math. Phys.* **43** 4641
[6] Koltchinskii V 2013 A remark on low rank matrix recovery and noncommutative bersntein type inequalities *IMS Collections* **9** 213–26