



Butucea, Cristina and Guță, Mădălin and Kypraios, Theodore (2016) Corrigendum: Spectral thresholding quantum tomography for low rank states (2015 New J. Phys. 17 113050). New Journal of Physics, 18 (6). 069501/1-069501/4. ISSN 1367-2630

Access from the University of Nottingham repository:

http://eprints.nottingham.ac.uk/46913/1/Butucea_2016_New_J._Phys._18_069501.pdf

Copyright and reuse:

The Nottingham ePrints service makes this work by researchers of the University of Nottingham available open access under the following conditions.

This article is made available under the Creative Commons Attribution licence and may be reused according to the conditions of the licence. For more details see: <http://creativecommons.org/licenses/by/2.5/>

A note on versions:

The version presented here may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher's version. Please see the repository url above for details on accessing the published version and note that access may require a subscription.

For more information, please contact eprints@nottingham.ac.uk

CORRIGENDUM • OPEN ACCESS

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

To cite this article: Cristina Butucea *et al* 2016 *New J. Phys.* **18** 069501

View the [article online](#) for updates and enhancements.

Related content

- [Spectral thresholding quantum tomography for low rank states](#)
Cristina Butucea, Mdlin Gu and Theodore Kypraios
- [Corrigenda](#)
P.J. McCarthy, K.S. Riedel, O.J.W.F. Kardaun *et al.*
- [Corrigenda](#)
B.B. Kadomtsev

Recent citations

- [Statistical analysis of compressive low rank tomography with random measurements](#)
Anirudh Acharya and Mdlin Gu



CORRIGENDUM

Corrigendum: Spectral thresholding quantum tomography for low rank states (2015 *New J. Phys.* **17** 113050)Cristina Butucea¹, Mădălin Guță² and Theodore Kypraios²¹ Université Paris-Est Marne-la-Vallée, LAMA(UMR 8050), UPEMLV F-77454, Marne-la-Vallée, France² University of Nottingham, School of Mathematical Sciences, University Park, Nottingham NG7 2RD, UKPUBLISHED
3 June 2016Original content from this
work may be used under
the terms of the [Creative
Commons Attribution 3.0
licence](#).Any further distribution of
this work must maintain
attribution to the
author(s) and the title of
the work, journal citation
and DOI.

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 than 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 $r \cdot \nu_c(\epsilon)^2 = r3^k/N$ 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 $\hat{\rho}_n^{(ls)}$ be the linear estimator of ρ . Then, for any $\epsilon > 0$, the following operator norm inequality holds, for n large enough, with probability larger than $1 - \epsilon$ under \mathbb{P}_ρ

$$\|\hat{\rho}_n^{(ls)} - \rho\| \leq \nu_c(\epsilon),$$

where

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

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

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

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

where $W_{s,i}$ 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, \dots, Y_n be independent, centered, $m \times m$ Hermitian random matrices. Suppose that, for some constants $V, W > 0$ we have $\|Y_j\| \leq V$, for all j from 1 to n , and that $\|\sum_j \mathbb{E}(Y_j^2)\| \leq W$. Then, for all $t \geq 0$,*

$$\mathbb{P}(\|Y_1 + \dots + Y_n\| \geq t) \leq 2 m \exp\left(-\frac{t^2/2}{W + tV/3}\right).$$

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

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

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 $W_{s,i} = \sum_{\mathbf{o}} B(\mathbf{o}|\mathbf{s}) (I(X_{s,i} = \mathbf{o}) - p(\mathbf{o}|\mathbf{s}))$. Then

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

In the last inequality we used that

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

which implies the following inequality between $2^k \times 2^k$ matrices: $\text{Cov}(\mathbf{s}) \leq p(\mathbf{s})$ 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 $E_{\mathbf{b}} := \{i : b_i = I\} \subset \{1, \dots, k\}$. 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 $E_{\mathbf{b}} \cap E_{\mathbf{b}'} := (E_{\mathbf{b}} \setminus E_{\mathbf{b}'}) \cup (E_{\mathbf{b}'} \setminus E_{\mathbf{b}})$ and $E_{\mathbf{b}} \Delta E_{\mathbf{b}'}$ 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

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

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 $E_b \Delta E_{b'}$, and $\tilde{\mathbf{s}}^c$ is the restriction to $(E_b \Delta E_{b'})^c$. Note that $\mathbf{g}(\mathbf{s}, E_b \Delta E_{b'})$ depends on \mathbf{s} only through the component $\tilde{\mathbf{s}}$. Then

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

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 $E_{\mathbf{g}} = (E_b \Delta E_{b'})^c$.

Note that any pair (E, E') is uniquely determined by three disjoint subsets, $(E \setminus E', E' \setminus E, E \cap E')$, or equivalently by the symmetric difference $D := E \Delta E'$ together with $F := E \setminus E' \subset D$ 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:

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

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.

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

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 $\nu_p(\mathcal{T}) := \sup_{\tau} \text{Tr}(\mathcal{T}(\tau)^p)^{1/p}$ be its p -norm, where the supremum is taken over all input states τ ; in particular for $p \rightarrow \infty$ this becomes the ∞ -norm $\nu_{\infty}(\mathcal{T}) := \sup_{\tau} \|\mathcal{T}(\tau)\|$. 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 $\|\mathcal{D}^{\otimes k}(\rho)\| \leq (3/5)^k$. Together with (3) this gives upper bound

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

Putting together (1) and (4) we obtain

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

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

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

We choose $t > 0$ such that

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

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

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

□

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 identification 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 Bernstein type inequalities *IMS Collections* **9** 213–26