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

View the article online for updates and enhancements.
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 bound.

Proposition 1 as stated in the paper is incorrect, in particular the dependence of the upper bound \( \nu(\varepsilon)^2 \) 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 \( \nu(\varepsilon)^2 \) 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 hold true when the operator norm rate is taken 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 \( \nu(\varepsilon)^2 \) in proposition 1 below. In particular, the upper bounds on the Frobenius norm error in corollary 1 and theorem 2, will scale as \( r \cdot \nu(\varepsilon)^2 = r3^k / N \) 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 \( \hat{\rho}^{(b)}_n \) be the linear estimator of \( \rho \). Then, for any \( \varepsilon > 0 \), the following operator norm inequality holds, for \( n \) large enough, with probability larger than \( 1 - \varepsilon \) under \( \mathbb{P}_\rho \),

\[
\| \hat{\rho}_n^{(b)} - \rho \| \leq \nu(\varepsilon),
\]

where

\[
\nu(\varepsilon)^2 = 4 \cdot \frac{3^k \log(2^{k+1})}{N}
\]

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 \( f(o|s) = \frac{1}{n} \sum_{j=1}^n I(X_{s,i} = o) \), where the random variables \( X_{s,i} \) are independent for all settings \( s \) and all \( i \) from 1 to \( n \). To estimate the risk of the linear estimator we write

\[
\hat{\rho}_n^{(b)} - \rho = \sum_b \sum_o \sum_s (f(o|s) - p(o|s)) \frac{A_b(o|s)}{2^k 3^d} \sigma_b
\]

\[
= \sum_b \sum_o \sum_s \frac{1}{n} \sum_i (I(X_{s,i} = o) - p(o|s)) \frac{A_b(o|s)}{2^k 3^d} \sigma_b
\]

\[
= \sum_s \sum_i W_{s,i},
\]
where $W_i$ 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 $\|Y_j\| \leq V$, for all $j$ from 1 to $n$, and that $\|\sum_j \mathbb{E}(Y_j^2)\| \leq W$. Then, for all $t \geq 0$,

$$
P(\|Y_1 + \ldots + Y_n\| \geq t) \leq 2m\exp\left(-\frac{t^2/2}{W + tV/5}\right).
$$

In our setup, $W_i$ play the role of $Y_j$. We bound $\|W_i\| \leq V$ for all $i$ and $\|\sum_i \mathbb{E}(W_i\hat{W}_i)\| \leq W$, where $V, W$ are evaluated below. We have

$$
\|W_i\| \leq \frac{1}{n} \sum_{h} \sum_{a} A_h(a|s) \cdot |I(X_{si} = o) - p(o|s)| \cdot \|c_b\|
$$

$$
\leq \frac{1}{n} \sum_{h} \sum_{a} \frac{1}{2^{k^2 - 2d(b)}} \prod_{j \in E_h} \delta_{b_{ij}^h} \sum_{o} |I(X_{si} = o) - p(o|s)|
$$

$$
\leq \frac{2}{n^2 \delta} \sum_{i \in E_h} \sum_{a} \sum_{o \in \mathbb{D}} \mathbb{E} \left( \mathbb{E}(W_{j} W_{s}^{*}) \right) \Cov(I(X_{s} = o), I(X_{s} = o')) \cdot B(o'|s)
$$

Now, denote by $B(o|s) = \sum_{i} 2^{-k^2 - 2d(b)} A_h(a|s) \sigma_b$ so that $W_i = \sum_o B(o|s)(I(X_{si} = o) - p(o|s))$. Then

$$
\frac{1}{n} \sum_{s} \sum_{i} \mathbb{E} \left( \mathbb{E}(W_{j} W_{s}^{*}) \right) \Cov(I(X_{s} = o), I(X_{s} = o')) \cdot B(o'|s)
$$

In the last inequality we used that

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

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

By expressing $B(o|s)$ in terms of $\sigma_b$, as above, we get

$$
\frac{1}{n} \left\| T \right\| = \frac{1}{n} \left\| \sum_{s} \sum_{o} B(o|s) \sum_{h} \frac{A_h(a|s)}{2^{k^2 - 2d(b)}} \cdot \sigma_b \sigma_b' \right\|
$$

Before giving the upper bound we introduce some combinatorial notations which will be used below. Let $b = \{x, y, z, I\}^2$ and recall that $E_b = \{i: b_i = 1\} \subseteq [1, \ldots, k]$. We say that $b$ agrees with a setting $s$ if $b_i = s_i$ for all $j \in E_b$. In this case $b$ is completely determined by the set $E_b$, for a fixed $s$. This fact will be used to replace the sums over $b$ and $b'$ with those over $E_b$ and $E_b'$ in $T$. Indeed since $A_h(o|s)$ is proportional to $\prod_{j \in E_b} \delta_{b_{ij}}$ the sums over $b$ and $b'$ in $T$ are restricted to sequences which agree with $s$. We denote by $E_b \cap E_b' = (E_b \setminus E_b') \cup (E_b' \setminus E_b)$ and $E_b \cap E_b'$ the symmetric difference and respectively the intersection of $E_b$ and $E_b'$. With these notations we have

$$
A_h(o|s)A_h'(o|s)\sigma_b \sigma_b' = \prod_{j \in E_b \Delta E_b'} o_j \cdot \sigma_g
$$

where $g = g(s, E_b \Delta E_b')$ is the sequence with $E_g = (E_b \Delta E_b')'$, and it agrees with $s$. With these notations we have
\[
T = \frac{1}{2^{kL}} \sum_s \sum_o p(o|s) \prod_{E_n \epsilon E} o_i \gamma_{s, E_n \Delta E_n} \\
= \frac{1}{2^{kL}} \sum_s \sum_o p(o|s) \prod_{E_n \epsilon E} o_i \gamma_{s, E_n \Delta E_n} \\
= \frac{1}{2^{kL}} \sum_s \sum_o p(o|s) \prod_{E_n \epsilon E} o_i \gamma_{s, E_n \Delta E_n} \\
(2)
\]

In the last expression we rewrite the sum over settings \(s\) as a double sum over \(s\) and \(s'\) where \(s\) is the restriction of \(s\) to \(E_n \Delta E_n\) and \(s'\) is the restriction to \((E_n \Delta E_n)'\). Note that \(g(s, E_n \Delta E_n)\) depends on \(s\) only through the component \(s'.\) Then

\[
\sum_s \sum_o p(o|s) \prod_{E_n \epsilon E} o_i \gamma_{s, E_n \Delta E_n} = \sum_s \left( \sum_o p(o|s) \prod_{E_n \epsilon E} o_i \gamma_{s, E_n \Delta E_n} \right) \gamma_{s', E_n \Delta E_n} \\
= \frac{1}{2^{kL}} \sum_{E_n \epsilon E} \sum_{o\epsilon E} \text{Tr}(\rho \gamma_{s', E_n \Delta E_n}) \gamma_{s', E_n \Delta E_n} \\
= \frac{1}{2^{kL}} \sum_{E_n \epsilon E} \sum_{o\epsilon E} \text{Tr}(\rho \gamma_{s', E_n \Delta E_n}) \gamma_{s', E_n \Delta E_n} \\
\]

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 \(\rho\). In the third equality we replaced the sum over \(s\) with an equivalent sum over sequences \(g\) such that \(E_n = (E_n \Delta E_n)'\).

Note that any pair \((E, E')\) is uniquely determined by three disjoint subsets, \((E \setminus E', E' \setminus E, E \cup E')\), 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 \(E, E'\) is (2) is computed by summing over all triples \(D, F, M\) satisfying the above conditions:

\[
T = \frac{1}{2^{kL}} \sum_{E, E'} \sum_{o\epsilon E} \sum_{g\epsilon E} \frac{3^{k-|E|}}{3^{k+|E|}} \gamma_{s', D} \gamma_{s', D'} \\
= \frac{1}{2^{kL}} \sum_{D \subseteq E} \sum_{M \subseteq D} \sum_{M \subseteq D} \frac{3^{k-|D|}}{3^{k+|D|}} \sum_{g\epsilon E} \sum_{g\epsilon E} \text{Tr}(\rho \gamma_{s', D}) \gamma_{s', D'} \\
= \frac{1}{2^{kL}} \sum_{D \subseteq E} \sum_{M \subseteq D} \sum_{M \subseteq D} \frac{3^{k-|D|}}{3^{k+|D|}} \sum_{g\epsilon E} \sum_{g\epsilon E} \text{Tr}(\rho \gamma_{s', E}) \gamma_{s', E'} \\
\]

Indeed, the sum over \(F \subseteq D\) gives a factor \(2^{kL}\) 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.

\[
T = \frac{1}{2^{kL}} \sum_{D \subseteq E} \sum_{M \subseteq D} \sum_{M \subseteq D} \frac{3^{k-|D|}}{3^{k+|D|}} \sum_{g\epsilon E} \sum_{g\epsilon E} \text{Tr}(\rho \gamma_{s', D}) \gamma_{s', D'} \\
= \left( \frac{3}{4} \right)^k \sum_{D \subseteq E} \left( \frac{2}{9} \right)^{|D|} \sum_{g\epsilon E} \sum_{g\epsilon E} \text{Tr}(\rho \gamma_{s', D}) \gamma_{s', D'} \\
= \left( \frac{5}{6} \right)^k \sum_{D \subseteq E} \left( \frac{1}{5} \right)^{|D|} \sum_{g\epsilon E} \sum_{g\epsilon E} \text{Tr}(\rho \gamma_{s', D}) \gamma_{s', D'} \\
= \left( \frac{5}{6} \right)^k 2^{kL} \gamma_{s', D'} (\rho) \\
\]

The final sum goes over subsets \(D\) and over sequences \(g\) such that \(E_g = D'\), and is similar to the Fourier decomposition of \(\rho\) except that each term 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 \(\rho\), where an individual depolarising channel is defined by

\[
D : \frac{1}{2} (I + r_z \sigma_z + r_y \sigma_y + r_x \sigma_x) = \frac{1}{2} \left( I + \frac{r_z}{5} \sigma_z + \frac{r_y}{5} \sigma_y + \frac{r_x}{5} \sigma_x \right) \\
\]

For an arbitrary quantum channel \(T\), let \(\nu_p (T) := \text{sup}_\tau \text{Tr}(T (\tau \rho)^p)/p\) be its \(p\)-norm, where the supremum is taken over all input states \(\tau\); in particular for \(p \rightarrow \infty\) this becomes the \(\infty\)-norm \(\nu_\infty(T) := \text{sup}_\|T \|\| T (\tau \rho) \|\). For the depolarising channel \(D\) defined above, the \(\infty\)-norm can be computed easily by applying the channel to an arbitrary pure state and is equal to \(\nu_\infty (D) = 3/5\). Moreover, it is known \([5]\) that the depolarising channel has multiplicative \(p\)-norm, i.e. \(\nu_p (D^\otimes k) = \nu_p (D)^k\), which implies that \(\|D^\otimes k (\rho)\| \leq (3/5)^k\). Together with (3) this gives upper bound
\[
\frac{1}{n} \| T \| \leq \frac{1}{n} \left( \frac{5}{6} \right)^k \cdot 2^k \cdot \left( \frac{3}{5} \right)^k = \frac{3^k}{n} \cdot 3^k = \frac{3^k}{N}.
\]

(4)

Putting together (1) and (4) we obtain

\[
\| \sum \sum (W_{s_i} W_{s_i}) \| \leq \frac{3^k}{n} \cdot 3^k = W.
\]

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

\[
P_\rho(\| \hat{\rho}_n - \rho \| \geq t) \leq 2^{k+1} \exp \left( -\frac{n \cdot t^2 / 2}{1 + t \cdot (2/3)^{k+1}} \right).
\]

We choose \( t > 0 \) such that

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

which leads, for \( n \) large enough, to \( t = \nu_\epsilon(\epsilon) \) such that

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

\[ \square \]

Acknowledgments

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

References