Almost depolarizing channels with short Kraus decompositions

Guillaume AUBRUN

Université Lyon 1, France
Completely positive maps

\( \mathcal{M}(\mathbb{C}^d) : d \times d \) complex matrices --- \( \langle A, B \rangle = \text{Tr} \, AB^* \).

**Definition (equivalent to the usual one)**

A linear map \( \Phi : \mathcal{M}(\mathbb{C}^d) \rightarrow \mathcal{M}(\mathbb{C}^d) \) is **completely positive** (CP) if there is a random matrix \( V : (\Omega, \mathbf{P}) \rightarrow \mathcal{M}(\mathbb{C}^d) \) so that

\[
\Phi(X) = E \, V X V^*.
\]

- Depends only on the covariance matrix of \( V \) (\( \in \mathcal{M}_+(\mathcal{M}(\mathbb{C}^d)) \)).
- Therefore \( V \) can be chosen to be finitely supported.

**Kraus decomposition.** Any CP \( \Phi : \mathcal{M}(\mathbb{C}^d) \rightarrow \mathcal{M}(\mathbb{C}^d) \) can be decomposed as a sum of Kraus operators

\[
\Phi(X) = \sum_{i=1}^{N} V_i X V_i^* \quad \text{with} \quad N \leq d^2.
\]

The length \( N \) measures the complexity of \( \Phi \).
Quantum channels

Definition

A state $\rho \in \mathcal{M}(\mathbb{C}^d)$ is a positive self-adjoint matrix with trace 1. The state $\frac{\text{Id}}{d}$ (the maximally mixed state) plays a central role.

Definition

A CP map $\Phi : X \rightarrow \mathcal{E VXV}^*$ is a quantum channel if it preserves trace

$$\text{Tr} \Phi(X) = \text{Tr} X$$

- A quantum channel maps states to states.
- If $V$ is supported in the unitary group $\mathcal{U}(d)$, then $\Phi$ is a quantum channel — not all quantum channels are like this.
- **A canonical example.** Let $U$ be Haar-distributed on $\mathcal{U}(d)$. This leads to the « depolarizing » or « randomizing » channel $\Psi$.

$$\Psi(X) = \mathcal{E UXU}^* = \text{Tr} X \frac{\text{Id}}{d}.$$
Kraus decompositions of the depolarizing channel

Since the covariance matrix of $U$ (= a multiple of identity) has full rank, any Kraus decomposition of $\Psi$ has length at least $d^2$.

**Example**: if $\omega = \exp(2i\pi/d)$, let $A, B$ defined as

$$A(e_j) = \begin{bmatrix} \omega & 0 \\ \omega^2 & 0 \\ 0 & 1 \end{bmatrix}, \quad B(e_j) = \begin{bmatrix} 0 & 1 & 0 \\ \vdots & \ddots & \vdots \\ 1 & \vdots & 0 \end{bmatrix}.$$ 

The set $U = \{A^k B^l\}_{1 \leq k, l \leq d}$ is a orthogonal family of unitary matrices.

$$\Psi(X) = \text{Tr} X \frac{\text{Id}}{d} = \frac{1}{d^2} \sum_{k, l=1}^{d} A^k B^l X (A^k B^l)^*.$$
**$\varepsilon$-randomizing channels**

- Kraus decompositions of $\Psi \leftrightarrow$ Exact encryption protocols.
- Approximate decomposition of $\Psi \leftrightarrow$ Approximate encryption protocols.

**Definition (Hayden, Leung, Shor and Winter)**

A quantum channel $\Phi : \mathbb{C}^d \to \mathbb{C}^d$ is $\varepsilon$-randomizing if for any state $\rho$

$$\left\| \Phi(\rho) - \frac{\text{Id}}{d} \right\|_\infty \leq \frac{\varepsilon}{d}$$

i.e. the spectrum of $\Phi(\rho)$ belongs to $[\frac{1-\varepsilon}{d}, \frac{1+\varepsilon}{d}]$.

- The depolarizing channel $\Psi$ is 0-randomizing, but has Kraus decompositions of length $d^2$.
- Problem: find $\varepsilon$-randomizing channels with short Kraus decompositions (low-cost encryption).
Theorem (Hayden, Leung, Shor, Winter — A.)

Let $U_1, \ldots, U_N$ be i.i.d. Haar-distributed random $d \times d$ unitary matrices. Then for $N \geq Cd/\varepsilon^2$, the quantum channel

$$
\Phi : X \mapsto \frac{1}{N} \sum_{i=1}^{N} U_i X U_i^* 
$$

is $\varepsilon$-randomizing with exponentially large probability.

- HLSW had the weaker estimate $N \geq Cd \log d/\varepsilon^2$.
- **Idea of the proof**: For unit vectors $x, y \in \mathbb{C}^d$, the random variable $\langle x, Uy \rangle$ is subgaussian. Therefore a net argument coupled with Bernstein inequalities will work.
- Optimal dependence in $d$. Can we achieve better dependence in $\varepsilon$ with another (non-random) construction?
Derandomization

The Haar measure is hard to generate in real-life situations. We show (answering a question of HLSW) that we can replace « reduce the amount of randomness » and replace it by any measure.

**Theorem**

Let \( U : (\Omega, P) \to \mathcal{U}(d) \) be a random unitary matrix so that

\[
\mathbb{E} UXU^* = \psi(X) = \text{Tr} \, X \cdot \frac{\text{Id}}{d}.
\]

Let \((U_i)\) be i.i.d. copies of \(U\). For \(N \geq Cd(\log d)^6/\varepsilon^2\), the quantum channel

\[
\Phi : X \mapsto \frac{1}{N} \sum_{i=1}^{N} U_i X U_i^*
\]

is \(\varepsilon\)-randomizing with probability \(\geq 1/2\).
**Isotropic measures**

**Definition**

*Say that a \( U(d) \)-valued random vector \( U \) is **isotropic** if*

\[
\forall X \in \mathcal{M}(\mathbb{C}^d), \quad \mathbb{E} UXU^* = \text{Tr} \ X \cdot \frac{\text{Id}}{d},
\]

\[\iff \forall X \in \mathcal{M}(\mathbb{C}^d), \quad \mathbb{E} |\text{Tr} \ UX|^2 = \frac{1}{d} \|X\|_\text{HS}^2.\]

1. Haar measure.
2. Uniform measure on \( \mathcal{U} \) (orthogonal basis of unitary matrices).
3. On \( \mathbb{C}^2 \) : uniform measure on the 4 Pauli matrices
   \[
   \sigma_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.
   \]
4. On \( (\mathbb{C}^2)^\otimes k \) : \( k \)-wise tensor product of the previous example.

Examples 2-4 are not subgaussian \(\longrightarrow\) net arguments cannot work.
Proof (1)

We need to estimate, for $U_i \in \mathcal{U}(d)$ i.i.d. isotropic

$$M := \mathbb{E} \sup_{\rho \text{ state}} \left\| \frac{1}{N} \sum_{i=1}^{N} U_i \rho U_i^* - \frac{\text{Id}}{d} \right\|_{\infty}$$

$$= \mathbb{E} \sup_{|x|=1} \left\| \frac{1}{N} \sum_{i=1}^{N} |U_i x\rangle\langle U_i x| - \frac{\text{Id}}{d} \right\|_{\infty}$$

$$= \mathbb{E} \sup_{|x|=|y|=1} \left| \frac{1}{N} \sum_{i=1}^{N} \langle U_i x, y \rangle^2 - \frac{1}{d} \right|$$

$$= \mathbb{E} \sup_{|x|=|y|=1} \left| \frac{1}{N} \sum_{i=1}^{N} \text{Tr} U_i |x\rangle\langle y| - \frac{1}{d} \right|$$

$$= \mathbb{E} \sup_{A \in B(S_1^d)} \left| \frac{1}{N} \sum_{i=1}^{N} \left| \text{Tr} U_i A \right|^2 - \mathbb{E} \text{Tr} |UA|^2 \right|$$

This is an empirical process in the Schatten space $S_1^d = (\mathcal{M}(\mathbb{C}^d), \| \cdot \|_1)$. 
Proof (2)

We can use results by Rudelson and Guédon, Mendelson, Pajor, Tomczak-Jaegermann about empirical processes in a Banach space with a good modulus of convexity (such as Hilbert space, \( \ell_1^d, S_1^d \)).

**Proof** (following [R],[GMPT])

- Symmetrization argument à la Giné–Zinn

\[
M \leq 2 \mathbb{E}_U \mathbb{E}_\varepsilon \sup_{A \in B(S_1^d)} \mathbb{E} \left| \frac{1}{N} \sum_{i=1}^{N} \varepsilon_i \left| \text{Tr} \, U_i A \right|^2 \right|
\]

- The theorem follows from the next lemma

**Lemma**

Let \( U_1, \ldots, U_N \in \mathcal{U}(d) \) be deterministic, \( N \geq d \). Then,

\[
\mathbb{E}_\varepsilon \sup_{A \in B(S_1^d)} \left| \sum_{i=1}^{N} \varepsilon_i \left| \text{Tr} \, U_i A \right|^2 \right| \leq C \log^3 N \sup_{A \in B(S_1^d)} \sum_{i=1}^{N} \left| \text{Tr} \, U_i A \right|^2.
\]
Proof of the lemma

Lemma

Let $U_1, \ldots, U_N \in \mathcal{U}(d)$ be deterministic, $N \geq d$. Then,

$$E_{\varepsilon} \sup_{A \in B(S_1^d)} \left| \sum_{i=1}^{N} \varepsilon_i |\text{Tr} U_i A|^2 \right| \leq C \log^3 N \sqrt{\sup_{A \in B(S_1^d)} \sum_{i=1}^{N} |\text{Tr} U_i A|^2}.$$ 

Let $(g_i)$ be independent $N(0, 1)$

$$E_{\varepsilon} \sup_{A \in B(S_1^d)} \left| \sum_{i=1}^{N} \varepsilon_i |\text{Tr} U_i A|^2 \right| \leq \sqrt{\frac{\pi}{2}} E_{\varepsilon} \sup_{A \in B(S_1^d)} \left| \sum_{i=1}^{N} g_i |\text{Tr} U_i A|^2 \right| \leq C \int_{0}^{\infty} \sqrt{\log N(B(S_1^d), \delta, \varepsilon)} d\varepsilon$$

Here $\delta$ the distance induced by the Gaussian process and $N(K, \delta, \varepsilon)$ the number of balls of radius $\varepsilon$ in the metric $\delta$ needed to cover $K$. 
Proof of the lemma

The metric $\delta$ can be upper-bounded

$$
\delta(A, B)^2 = \sum_{i=1}^{N} \left| \text{Tr} U_i A \right|^2 - \left| \text{Tr} U_i B \right|^2 \leq \left( \sum_{i=1}^{N} \left| \text{Tr} U_i (A + B) \right|^2 \right) \left( \sup_{1 \leq i \leq N} \left| \text{Tr} U_i (A - B) \right|^2 \right).
$$

This leads to the bound

$$
E_{\varepsilon} \cdots \leq C \left( \sup_{A \in B(S_1^d)} \sum_{i=1}^{N} \left| \text{Tr} U_i A \right|^2 \right)^{1/2} \int_{0}^{\infty} \sqrt{\log N(B(S_1^d), ||| \cdot |||, \varepsilon)} d\varepsilon
$$

with $|||A||| = \sup_{1 \leq i \leq N} \left| \text{Tr} U_i A \right| \leq ||A||_1$.

The unit ball $L$ of $||| \cdot |||$ has $N$ « faces » and contains $B(S_1^d)$. 
Covering numbers

We need to estimate

\[ I = \int_{0}^{\infty} \sqrt{\log N(B(S^d_1), \| \cdot \|, \varepsilon)} \, d\varepsilon \leq C \log^3 N \]

Assume for the moment the duality property for covering numbers holds (it is still a conjecture)

\[ \log N(K, L, \varepsilon) \leq C \log N(L^\circ, K^\circ, c\varepsilon) \]

This leads to

\[ I \leq C \int_{0}^{\infty} \sqrt{\log N(L^\circ, B(S^d_\infty), \varepsilon)} \, d\varepsilon. \]

With \( L^\circ \) the unit ball for \( \| \cdot \| \cdot \|^* \) — a convex body with \( N \ll \) « vertices » contained in \( B(S^d_\infty) \).
Lemma (Maurey’s lemma)

If $K \subset L$ and $K$ has $N \ll$ vertices, then for all $\varepsilon > 0$,

$$\varepsilon \sqrt{\log N(K, L, \varepsilon)} \leq CT_2(L) \sqrt{\log N}$$

Here $T_2(L)$ is the type 2 constant of the norm associated to $L$.

1. In our case $T_2(S^d_{\infty}) \leq C \sqrt{\log d}$ (Tomczak-Jaegermann).

2. The duality conjecture holds up to a logarithmic factor. This follows from results by Bourgain, Pajor, Szarek and Tomczak–Jaegermann since $S^d_1$ has an equivalent norm which has a good modulus of convexity, namely the norm of $S^d_p$ for $p = 1 + 1/\log d$ (Tomczak-Jaegermann, Ball–Carlen–Lieb).

3. Collect all the logarithms.
Theorem

Let \((U_i) \in \mathcal{U}(d)\) be i.i.d. random matrices with isotropic law, and \(N \geq Cd \log^6 d/\varepsilon^2\). With probability \(\geq 1/2\),

\[
\sup_{\rho \geq 0, \text{Tr} \rho = 1} \left\| \frac{1}{N} \sum_{i=1}^{N} U_i \rho U_i^* - \frac{\text{Id}}{d} \right\|_{\infty} \leq \frac{\varepsilon}{d}.
\]

- The power of \(\log d\) can certainly be improved, e.g. using Talagrand’s majorizing measures instead of Dudley integral (however existing results in the litterature do not give better).
- You get \(d \log^4 d\) if you prove the duality conjecture.
- However, some power of \(\log d\) is needed, for example when \(U\) is distributed on a orthogonal set of unitary matrices.
- (Vague) question: is it possible to approximate any quantum channel (and not only \(\Psi\)) in a similar way?