CONVEX CLOSURE OF THE OUTPUT ENTROPY AND SHARP ENTROPY INEQUALITIES

A.S. Holevo
Steklov Mathematical Institute,
Russian Academy of Sciences

CONTENTS

- Probability measures on the set of quantum states in convex optimization problems

- Computing accessible information of quantum ensemble

- Sharp entropy inequalities from "quantum pyramids"

A number of optimization problems in quantum information:
- (one-shot) constrained classical capacity of quantum channel;
- entanglement of formation of a composite state;
- classical capacity of a quantum observable;
- accessible information of an ensemble of quantum states

can be formulated as calculation of the convex closure of output entropy of certain quantum channel.

This is a convex programming problem for which the dual problem along with the necessary and sufficient optimality conditions can be established. These optimality conditions imply new sharp lower bounds for Shannon entropy which are related to generalizations of log-Sobolev inequality.

Ensembles as probability measures

Let $\mathfrak{S}(\mathcal{H})$ be the convex set of quantum states, i.e. density operators $\rho, \sigma, \ldots$ in a Hilbert space $\mathcal{H}$, $\mathfrak{P}(\mathcal{H})$ the subset of pure states $\boldsymbol{\psi} = |\psi\rangle\langle\psi|$, where $|\psi\rangle \in \mathcal{H}$ – unit vectors of the space $\mathcal{H}$.

A number of mathematical problems in quantum information theory require optimization over the set of ensembles of (pure) quantum states. The usual definition of an ensemble assumes a finite set $\mathcal{E} = \left\{p_j, \rho_j\right\}_{j=1,\ldots,m}$ where $\rho_j \in \mathfrak{S}(\mathcal{H})$, and $p = \left\{p_j\right\}$ – probability distribution. A more general definition is a *probability measure* $\pi(d\rho)$ on the set $\mathfrak{S}(\mathcal{H})$. Thus, the convex set structure is introduced in the set of ensembles, making optimization more amenable. Similarly, an ensemble of pure states is a probability measure $\pi(d\boldsymbol{\psi})$ on the set $\mathfrak{P}(\mathcal{H})$; the set of such measures is denoted by $\mathcal{P}(\mathfrak{P}(\mathcal{H}))$.

General formulation of the problem

Let Φ be a channel that maps the input quantum states into the output quantum or classical states, and let $\mathrm{H}(\cdot)$ be, respectively, quantum or classical entropy. We are considering the problem of calculating the *convex closure of the channel output entropy*

$$\begin{aligned} F(\pi) \equiv \int_{\mathfrak{P}(\mathcal{H})} \mathrm{H}(\Phi[\psi])\pi(d\psi) &\longrightarrow \min_{\pi} \qquad (1)\\ \pi &\in \mathcal{P}(\mathfrak{P}(\mathcal{H})) \\ \overline{\psi}_{\pi} \equiv \int_{\mathfrak{P}(\mathcal{H})} \psi\, \pi(d\psi) &= \overline{\rho}, \qquad (2) \end{aligned}$$

where $\overline{\rho}$ is a fixed state. This is a convex programming problem, the general solution of which is given below. First we present a number of tasks, the solution of which reduces to that general problem.

Cases from quantum information theory

1. *One-shot (constrained) classical capacity of the quantum channel* Φ

Let Φ be a quantum channel with input space $\mathcal{H}_1$ and output space $\mathcal{H}_2$. Let also H be a positive self-adjoint operator (Hamiltonian) and E be a positive number (energy limit), then this capacity is equal to

$$C_\chi(\Phi) = \sup_{\overline{\rho}:\mathrm{Tr}\,\overline{\rho}H \leq E} \left[\mathrm{H}\left(\Phi\left[\overline{\rho}\right]\right) - \inf_{\pi:\overline{\boldsymbol{\psi}}_\pi = \overline{\rho}} \int_{\mathfrak{P}(\mathcal{H})} \mathrm{H}(\Phi[\boldsymbol{\psi}])\pi(d\psi) \right],$$

where in square brackets stands the $\overline{\rho}-$*constrained capacity*. Here $\mathrm{H}(\rho) = -\mathrm{Tr}\,\rho \log \rho$ is the von Neumann entropy.

2. *Entanglement of Formation*

Let ρ_{12} be a state in the tensor product of two Hilbert spaces $\mathcal{H}_1 \otimes \mathcal{H}_2$, then

$$E_F(\rho_{12}) = \inf_{\pi : \overline{\psi}_\pi = \rho_{12}} \int_{\mathfrak{P}(\mathcal{H}_1 \otimes \mathcal{H}_2)} \mathsf{H}(\mathrm{Tr}_2 \psi)\, \pi(d\psi),$$

where minimization is performed according to all possible probability distributions $\pi(d\psi)$ on $\mathfrak{P}(\mathcal{H}_1 \otimes \mathcal{H}_2)$, satisfying the condition

$$\int_{\mathfrak{P}(\mathcal{H}_1 \otimes \mathcal{H}_2)} \psi\, \pi(d\psi) = \rho_{12}.$$

In this case, $\Phi[\sigma_{12}] = \mathrm{Tr}_2 \sigma_{12}$ is the channel of partial trace w.r.t. the second space.

3. *The classical capacity of the measurement (q-c) channel with continuous outcomes*

In this case $\Phi : \rho \to p_\rho(y) = \operatorname{Tr} \rho m(y)$, where $m(y)$ is a measurable uniformly bounded positive operator-valued function of $y \in Y$, such that $\int m(y)dy = I$. Here (Y, dy) is the space of measurement outcomes. Then $p_\rho(y)$ is the probability density of the outcomes of a quantum measurement described by the function $m(y)$. The constrained classical capacity of the measurement channel Φ coincides with the one-shot capacity and is given by

$$C(M) = C_\chi(M) = \sup_{\rho : \operatorname{Tr} \rho H \leq E} \left[h\left(p_\rho\right) - \inf_{\pi : \overline{\psi}_\pi = \rho} \int_{\mathfrak{P}(\mathcal{H})} h\left(p_\psi\right) \pi(d\psi) \right],$$

where $h\left(p\right) = -\int p(y) \log p(y) dy$ is the differential entropy of the probability density.

4. *Accessible information of quantum state ensemble*

For definiteness, consider the finite ensemble $\mathcal{E} = \left\{p_j, \rho_j\right\}_{j=1,\dots,m}$ in a d-dimensional Hilbert space, and denote $\overline{\rho} = \sum p_j \rho_j$ the average state of the ensemble. By using *ensemble-measurement duality*, which interchanges input and output, the accessible information of the ensemble $\mathcal{E}$ can be represented as

$$A(\mathcal{E}) = \mathsf{H}(p) - \min_{\pi : \overline{\psi}_\pi = \overline{\rho}} \int_{\mathfrak{P}(\mathcal{H})} \mathsf{H}(\Phi[\psi])\pi(d\psi)$$

where

$$\Phi[\psi] = [\left\langle \psi | M_j' | \psi \right\rangle]_{j=1,\dots,m}, \quad M_j' = \overline{\rho}^{-1/2} \left(p_j \rho_j\right) \overline{\rho}^{-1/2}$$

is a measurement channel, H is the Shannon entropy, and the minimization is performed by ensembles with barycenter $\overline{\rho}$. Similarly for a continuous ensemble $\mathcal{E} = \{p(x), \rho_x\}$.

The criterion for optimality

Let $\mathcal{H}$ be finite-dimensional and $\overline{\rho}$ nondegenerate.

Theorem 1. *The problem dual to (1) has the form*

$$G(\Lambda) \equiv \mathrm{Tr}\,\overline{\rho}\Lambda \longrightarrow \max_{\Lambda^*=\Lambda} \tag{3}$$

$$\langle\psi|\Lambda|\psi\rangle \leq \mathrm{H}(\Phi[\boldsymbol{\psi}]), \quad \boldsymbol{\psi} \in \mathfrak{P}(\mathcal{H}). \tag{4}$$

The following statements are equivalent:

(i.) π_0 *is a solution to the problem (1), (2);* Λ_0 *is the (unique) solution to the problem (3)*;

(**ii.a**) $\langle\psi|\Lambda_0|\psi\rangle \leq \mathrm{H}(\Phi[\boldsymbol{\psi}]), \quad \boldsymbol{\psi} \in \mathfrak{P}(\mathcal{H})$,
(**ii.b**) $\langle\psi|\Lambda_0|\psi\rangle = \mathrm{H}(\Phi[\boldsymbol{\psi}]) \quad \mathrm{mod}\, \pi_0(d\boldsymbol{\psi})$,
for a measure π_0 *with the barycenter* $\overline{\rho}$.

If Φ *is* completely noisy* *then this is equivalent to:*

(**iii.a**) $\Lambda_0 \leq K(\boldsymbol{\psi}), \quad \boldsymbol{\psi} \in \mathfrak{P}(\mathcal{H})$,
(**iii.b**) $[K(\boldsymbol{\psi}) - \Lambda_0]\,|\psi\rangle = 0 \quad \mathrm{mod} \pi_0(d\boldsymbol{\psi})$,
for a measure π_0 *with the baricenter* $\overline{\rho}$, *where*

$$K(\boldsymbol{\psi}) = -\Phi^*\left(\log \Phi(\boldsymbol{\psi})\right), \quad \boldsymbol{\psi} \in \mathfrak{P}(\mathcal{H}).$$

*$\Phi(\boldsymbol{\psi}) > 0$, $\boldsymbol{\psi} \in \mathfrak{P}(\mathcal{H})$. One can relax this condition by using forms rather than operators.

Remarks: 1. Λ is the operator-valued Lagrange multiplier. Multiplying the equation (**ii.b**) on the right by $\langle\psi|$ and integrating w.r.t. $\pi_0(d\psi)$, we get a useful relation between Λ_0 and π_0 :

$$\Lambda_0\overline{\rho} = \int_{\mathfrak{P}(\mathcal{H})} K(\psi)\psi\pi_0(d\psi).$$

Non-degeneracy of $\overline{\rho}$ implies the uniqueness of Λ_0.

2. In the case of an infinite-dimensional Hilbert space $\mathcal{H}$ the proof of sufficiency of conditions (**ii**) works with some amendments: the operator Λ_0 can be unbounded self-adjoint and the inequality (**ii.a**) holds only on a dense domain in $\mathcal{H}$.

Optimizers for Gaussian q-c channels

This situation occurs in the case of proving the *Hypothesis of quantum Gaussian optimizers* for Gaussian measurement (q-c) channels, where Λ_0 becomes an operator quadratic in position and momentum observables. Strategy in solving the hypothesis of Gaussian optimizers for the capacity of Gaussian measurement channels [AH'2023, AH-Filippov'2023] consisted in finding optimizers in the limited class of Gaussian procedures and then verifying that condition (**ii**) of Theorem 1 is fulfilled for them by using some generalizations of Gaussian log-Sobolev inequality.

Accessible information ($d < \infty$)

Take a finite quantum state ensemble $\mathcal{E} = \left\{p_j, \rho_j\right\}_{j=1,\dots,m}$ with the average state $\overline{\rho} = \sum_j p_j \rho_j$.

An *observable* is a collection of Hermitean operators $\mathcal{M} = \{M_k\}_{k=1,\dots,n}$ where $M_k \geq 0$, $\sum_{k=1}^n M_k = I$ and the joint p.d. of "input" j and "output" k is $p_{jk} = p_j \mathrm{Tr}\ \rho_j M_k$.

Accessible information of the ensemble $\mathcal{E}$ is

$$A(\mathcal{E}) = \sup_{\mathcal{M}} I(\mathcal{E}, \mathcal{M}),$$

where $I(\mathcal{E}, \mathcal{M}) = \sum_{j,k} p_{jk} \log \frac{p_{jk}}{p_j p_{\cdot k}}$ is the *Shannon mutual information* between j and k. It is known [Davies'1978] that supremum is attained on observable $\mathcal{M}_0$ with linearly independent components of rank one $M_k = |\varphi_k\rangle\langle\varphi_k|$, $k = 1, \dots, n \leq d^2$.

Theorem 2. *Observable* $\mathcal{M}_0 = \{|\varphi_k\rangle\langle\varphi_k|\}$ *maximizes accessible information iff there exists a Hermitean operator* Λ_0 *such that the entropy inequality*

$$-\sum_j \left\langle\psi|M_j'|\psi\right\rangle \log\left\langle\psi|M_j'|\psi\right\rangle \ \geq \langle\psi|\Lambda_0|\psi\rangle\,, \quad M_j' = \overline{\rho}^{-1/2}\left(p_j\rho_j\right)\overline{\rho}^{-1/2}, \tag{5}$$

holds for all unit vectors $\psi \in \mathcal{H}$*, and the unit vectors* $|\phi_k\rangle = \overline{\rho}^{1/2}|\varphi_k\rangle/\left\|\overline{\rho}^{1/2}\varphi_k\right\|$ *turn this into equality:*

$$-\sum_j \left\langle\phi_k|M_j'|\phi_k\right\rangle \log\left\langle\phi_k|M_j'|\phi_k\right\rangle = \langle\phi_k|\Lambda_0|\phi_k\rangle\,, \quad k = 1,\ldots.n.$$

The value of the accessible information is $A(\mathcal{E}) = H(\pi) - \mathrm{Tr}\,\overline{\rho}\Lambda_0$.

Of special interest is the case of *ensemble of pure states* $\rho_j = |\psi_j\rangle\langle\psi_j|$, where $|\psi_j\rangle; j = 1, \dots, d$, are *linearly independent* vectors. In this case $M'_j = |e_j\rangle\langle e_j|$, where the vectors

$$|e_j\rangle = \sqrt{p_j}\bar{\rho}^{-1/2}|\psi_j\rangle; \quad j = 1, \dots, d,$$

by necessity form an orthonormal basis. Then the entropy inequality (5) goes into a sort of *discrete log-Sobolev inequality*

$$-\sum_{j=1}^{d} \left|z_j\right|^2 \log \left|z_j\right|^2 \geq \sum_{j=1}^{d} \lambda_{jk} \overline{z_j} z_k,$$

where $\lambda_{jk} = \langle e_j|\Lambda_0|e_k\rangle$, and $z_j = \langle e_j|\psi\rangle$ are (complex) variables subject to the only constraint $\sum_{j=1}^{d} \left|z_j\right|^2 = 1$. The left part is simply the Shannon entropy of the probability distribution $\left\{\left|z_j\right|^2; j = 1, \dots, d\right\}$, and therefore it is possible to approach such an inequality with the tools of classical analysis.

Entropy inequalities from "quantum pyramids"

Consider the ensemble $\mathcal{E}$ of $m = d$ equiprobable equiangular pure states on $d-$dimensional Hilbert space – the *quantum pyramid* [Englert and Řeháček'2010]. Let $|e_j\rangle$, $j = 1, \ldots, d$, be an orthonormal basis, with the "height" ort

$$|e_0\rangle = \frac{1}{\sqrt{d}} \sum_{j=1}^{d} |e_j\rangle, \quad \langle e_0 | e_j \rangle \equiv \frac{1}{\sqrt{d}}.$$

A quantum pyramid is the set of unit vectors (edges of the pyramid)

$$|\psi_j\rangle = \sqrt{d r_1} |e_j\rangle + \left(\sqrt{r_0} - \sqrt{r_1}\right) |e_0\rangle, \quad j = 1, \ldots, d,$$

with parameters

$$r_0, r_1 \geq 0, \quad r_0 + (d-1) r_1 = 1.$$

The cosine between any two edges of pyramid is

$$\xi \equiv \langle \psi_j | \psi_k \rangle = r_0 - r_1 \, , \, j \neq k.$$

Acute pyramid: $0 < \xi < 1, (r_0 > r_1)$;

Orthogonal pyramid: $\xi = 0, (r_0 = r_1)$;

Obtuse pyramid: $-\frac{1}{d-1} < \xi < 0, (r_0 < r_1)$;

Extreme values: *the ort* $|\psi_j\rangle \equiv |e_0\rangle$, $\xi = 1, (r_1 = 0)$;

Flat pyramid: $\xi = -\frac{1}{d-1}, (r_0 = 0)$ in $(d-1)$-dimensional hyperplane orthogonal to $|e_0\rangle$.

Consider *quantum pyramid ensemble* $\mathcal{E} = \left\{p_j, \rho_j\right\}_{j=1,\dots,d}$ where $p_j = \frac{1}{d}$, $\rho_j = |\psi_j\rangle\langle\psi_j|$ with the average state

$$\bar{\rho} \equiv \frac{1}{d}\sum_{j=1}^{d} |\psi_j\rangle\langle\psi_j| = r_1 I + (r_0 - r_1)\,|e_0\rangle\langle e_0|.$$

[Englert and Řeháček'2010] conjectured and numerically tested certain hypotheses regarding optimal observables in the problem of accessible information for quantum pyramid ensemble. In [AH-Utkin'2025] we applied our Theorem 1 to verify the fulfillment of the optimality conditions, solved the dual problem, and derived corresponding sharp entropy inequalities.

It will be convenient to enter the parameter

$$p = |\langle e_j|\psi_j\rangle|^2 = \frac{1}{d}\left[(d-1)\sqrt{r_1} + \sqrt{r_0}\right]^2.$$

Acute pyramids ($\xi > 0$, $\frac{1}{d} \leq p \leq 1$)

The conjectured optimal observable belongs to the family that depends on the parameter $0 \leq t \leq 1$, and has the form $\mathcal{M}_0 = \{|\varphi_k\rangle\langle\varphi_k|; k = 0, 1, \ldots, d\}$, with

$$|\varphi_k\rangle = |e_k\rangle + \frac{t-1}{\sqrt{d}}|e_0\rangle; \quad k = 1, \ldots, d, \quad |\varphi_0\rangle = \sqrt{1-t^2}|e_0\rangle. \quad (6)$$

The optimal value of the parameter t

$$t(p,d) = \begin{cases} 1, & \frac{d-1}{d} \leq p \leq 1, \text{ moderately acute} \\ \frac{2(d-1)}{d-2}\sqrt{\frac{r_1}{r_0}}, & \frac{1}{d} \leq p \leq \frac{d-1}{d} \text{ strongly acute} \end{cases}$$

The case $t = 1$ corresponds to the natural assumption that the optimal measurement is given by the orthonormal basis $|e_j\rangle$, $j = 1, \ldots, d$. However, this is so only for moderately acute pyramids close to the base. Strongly acute pyramids require an adjustment for which t is chosen in the optimal way.

Guided by the symmetry of the problem (rotations around ort $|e_0\rangle$ leaving the pyramid invariant), we look for Λ_0 of the form

$$\Lambda_0 = \lambda_1(p)I + \lambda_0(p)|e_0\rangle\langle e_0|.$$

The optimality equation (**iii.b**), applied to the hypothesis (6), after a series of calculations results in

$$\lambda_0(p) = \frac{d\sqrt{\frac{p(1-p)}{d-1}}\left(\log p - \log\frac{1-p}{d-1}\right)}{\left(\sqrt{p} + (d-1)\sqrt{\frac{1-p}{d-1}}\right)\left(\sqrt{p} - \sqrt{\frac{1-p}{d-1}}\right)},$$

$$\lambda_1(p) = -\frac{\sqrt{p}\log p - \sqrt{\frac{1-p}{d-1}}\log\frac{1-p}{d-1}}{\left(\sqrt{p} - \sqrt{\frac{1-p}{d-1}}\right)}$$

for *moderately acute pyramid* ($\frac{d-1}{d} < p \leq 1$).

The entropy inequality (**ii.a**) for moderately acute pyramid is

$$-\sum_{j=1}^{d} \left|z_j\right|^2 \log \left|z_j\right|^2 \geq \frac{\lambda_0(p)}{d} \left|\sum_{j=1}^{d} z_j\right|^2 + \lambda_1(p).$$

The value of $p = \frac{d-1}{d}$ is the threshold: for all *strongly acute pyramids* ($\frac{1}{d} \leq p \leq \frac{d-1}{d}$) the optimality equation (**iii.b**) implies

$$\lambda_1(p) = \log d - \frac{d \log (d-1)}{d-2}, \quad \lambda_0(p) = \frac{d \log (d-1)}{d-2},$$

resulting in the entropy inequality

$$-\sum_{j=1}^{d} \left|z_j\right|^2 \log \left|z_j\right|^2 \geq \log d - \frac{d \log (d-1)}{d-2} \left[1 - \frac{1}{d} \left|\sum_{j=1}^{d} z_j\right|^2\right].$$

This is closely related to log-Sobolev inequality corresponding to the sharp LS-constant for simple random walk on Z_d (Pott's model) computed in [Diaconis and Saloff-Coste'1996].

Since $\lambda_0(p) \geq 0$, the inequality (**ii.a**) reduces to the following sharp lower bounds for the Shannon entropy of a discrete p.d. $P = (t_1, \ldots, t_d)$, where $t_j = \left|z_j\right|^2$

Theorem 3. *For integers $d \geq 2$ and $p > \frac{d-1}{d}$ (moderately acute pyramid), there is a sharp lower bound for the Shannon entropy of an arbitrary probability distribution $P = (t_1, \ldots, t_d)$:*

$$\mathsf{H}(P) \geq \lambda_0(p) B(P; P_U)^2 + \lambda_1(p), \tag{7}$$

where $B(P; P_U) = \sum_{j=1}^{d} \sqrt{t_j/d}$ – the Bhattacharya coefficient between probability distribution P and uniform distribution P_U.

This bound is exact: graphs of functions $t_1, \ldots, t_d$ in the left and right sides of the inequality (7) are tangent at points obtained by permutations $t_1 = p, t_k = \frac{1-p}{d-1}, k = 2, \ldots, d$.

For $d > 2$ and $\frac{1}{d} \leq p \leq \frac{d-1}{d}$ (strongly acute pyramid)

$$\mathsf{H}(P) \geq \log d - \frac{d \log(d-1)}{d-2}\left[1 - B(P; P_U)^2\right]. \qquad (8)$$

Graphs of functions $t_1, \ldots, t_d$ in the left and right parts (8) are tangents at points obtained by permutations from $t_1 = \frac{d-1}{d}, t_k = \frac{1}{d(d-1)}, k = 2, \ldots, d$, and at the point P_U.

The proof developed by A.V. Utkin as a response to the hypothesis formulated by AH, is based on a scrupulous analysis of the critical points of a function of many variables equal to the difference between the left and right sides of the inequality (7).

The case of flat pyramid ($\xi = -1/(d-1)$)

Theorem 4. *For all* z_j *satisfying* $\sum_{j=1}^d \left|z_j\right|^2 = 1, \quad \sum_{j=1}^d z_j = 0,$

$$-\sum_{j=1}^{d} \left|z_j\right|^2 \log \left|z_j\right|^2 \geq \begin{cases} 1, & d \leq 6; \\ \log d - \frac{d-2}{d} \log(d-1), & d \geq 7. \end{cases} \tag{9}$$

In the case $d \leq 6$ *the equality is attained for* $z_1 = -z_2 = 1/\sqrt{2}$, $z_j = 0$ *for* $j \geq 3$; *in the case* $d \geq 7$ – *for* $z_1 = \sqrt{\frac{d-1}{d}}$, $z_j = -\sqrt{\frac{1}{(d-1)d}}$, $j \geq 2$, *(and for all permutations of such* z_j*).*

The entropy inequality (9) was formulated and partially confirmed in [AH-Utkin'2025]. Recently (AI-assisted) proofs were suggested in [Zhang'2026], [Arulandu'2026].

The case of obtuse pyramid

When $\xi < 0$ the pyramid is *obtuse*, with extreme *flat* case when all the state vectors $|\psi_j\rangle$ lie in the hyperplane orthogonal to $|e_0\rangle$. The hypothetical optimal observable is then drastically different. To describe it let us start with the flat pyramid in the case $d = 3$ where $|\psi_j\rangle; j = 1, 2, 3$, are the "trine" of equiangular unit vectors in two-dimensional real plane, cf. [AH'1973]. Then the optimal observable is $\mathcal{M}^0 = \{M_k; k = 1, 2, 3\}$, where $M_k = \frac{2}{3}|\psi_k^\perp\rangle\langle\psi_k^\perp|$, and $|\psi_k^\perp\rangle$ are the unit vectors in the plane such that $\langle\psi_k^\perp|\psi_k\rangle = 0; k = 1, 2, 3$. In other words, $\mathcal{M}$ is "unambiguous discrimination" measurement (UDM) for the states $\rho_j = |\psi_j\rangle\langle\psi_j|$, $j = 1, 2, 3$. [Sasaki et al.'1999]

The analysis based on our optimality criterion shows that the conditions (**ii**) of the theorem 1 are fulfilled with $\Lambda^0 = I$, the condition (**ii.a**) is equivalent to the sharp entropy inequality (saturated for $\alpha = \frac{\pi}{2} + \pi k$):

$$-\sum_{j=1}^{3} t_j \log t_j \geq 1; \quad t_j = \frac{2}{3} \cos^2 \left(\alpha + \frac{2\pi j}{3} \right), \quad .$$

Most interesting is the case of "almost flat" ($-0.5 < \xi < -0.408$) *lifted trine* [Shor'2000]. Then the conjectured optimal observable has 6 outcomes and is a statistical mixture of the UDM described above and the sharp observable $M_k = |e_k\rangle\langle e_k|$, $k = 1, 2, 3$. The last becomes optimal for more lifted trines, up to the rectangular one.

The generalization of this picture to the case $d \geq 3$ based on numerical studies was elaborated in [Englert and Řeháček'2010]. Let us apply our optimality criterion to derive the corresponding entropy inequalities and to obtain additional insight.

Moderately obtuse pyramid: $\xi < 0$ and the parameter p varies in the range $[\tilde{p}(d), 1]$, where $\tilde{p}(d)$ is described below. The conjectured optimal observable is $\mathcal{M}_0 = \{|e_k\rangle\langle e_k|,\, k = 1, \ldots, d\}$.

The optimality condition (**ii.a**) amounts to the inequality:

$$-\sum_{j=1}^{d}\left|z_j\right|^2\log\left|z_j\right|^2 \geq \frac{\tilde{\lambda}_0(p)}{d}\left|\sum_{j=1}^{d} z_j\right|^2+\tilde{\lambda}_1(p),$$

$$\tilde{\lambda}_0(p)=\frac{d\sqrt{\frac{p(1-p)}{d-1}}\left(\log p-\log\frac{1-p}{d-1}\right)}{\left((d-1)\sqrt{\frac{1-p}{d-1}}-\sqrt{p}\right)\left(\sqrt{p}+\sqrt{\frac{1-p}{d-1}}\right)},$$

$$\tilde{\lambda}_1(p)=-\frac{\sqrt{p}\log p+\sqrt{\frac{1-p}{d-1}}\log\frac{1-p}{d-1}}{\left(\sqrt{p}+\sqrt{\frac{1-p}{d-1}}\right)}.$$

The equality is attained for $z_1=\sqrt{p}$, $z_j=-\sqrt{\frac{1-p}{d-1}}$, $j=2,\ldots,d$, and all the permutations of these z_j.

Here $\tilde{\lambda}_0(p) \le 0$, which prevents us from passing to the probabilities $t_j = \left|z_j\right|^2$. The parameter p varies in the range $[\tilde{p}(d), 1]$, where $\tilde{p}(d) = \max\left\{\frac{d-1}{d}, p(d)\right\}$ and $p(d)$ is the root of the equation

$$\tilde{\lambda}_1(p(d)) = 1.$$

d	2	3	4	5	6	7
$\frac{d-1}{d}$	0,5	0,6(6)	0,75	0,8	0,8(3)	6/7≈ 0,857
$p(d)$	0,5	0,8725	0,8656	0,85699	0,84896	0,8417<6/7

For $d \geq 7$ one has $p(d) < (d-1)/d$! “Phase transition”.

Strongly obtuse pyramid $\left(\frac{d-1}{d} < p \leq p(d); d \leq 6\right)$.

The conjectured optimal observable $\mathcal{M}^0$ has $d + \frac{d(d-1)}{2} = \frac{d(d+1)}{2}$ components:

$$\begin{aligned} M_k &= t^{-2}|\varphi_k\rangle\langle\varphi_k|; \quad k = 1, \ldots, d; \\ M_{rs} &= \frac{2}{d}\left(1 - t^{-2}\right)|rs\rangle\langle rs|; \quad 1 \leq r < s \leq d, \end{aligned} \tag{10}$$

where the vectors $|\varphi_k\rangle$ are given by (6), and

$$|rs\rangle = \frac{1}{\sqrt{2dr_1}}\left(|\psi_r\rangle - |\psi_s\rangle\right) = \frac{1}{\sqrt{2}}\left(|e_r\rangle - |e_s\rangle\right).$$

The condition (**ii.a**) then amounts to the inequality

$$-\sum_{j=1}^{d} \left|z_j\right|^2 \log \left|z_j\right|^2 \geq \frac{\tilde{\lambda}_0(p(d))}{d} \left|\sum_{j=1}^{d} z_j\right|^2 + 1.$$

The equality is attained if $z_1 = -z_2 = 1/\sqrt{2}, z_j = 0$ for $j \geq 3$ and for all the permutations of these z_j ; for $z_1 = \sqrt{p(d)}$, $z_j = -\sqrt{\frac{1-p(d)}{d-1}}$, $j \geq 2$; and for all permutations and global change of sign of these z_j.

The entropy inequalities for obtuse pyramids were formulated and partially confirmed in [AH-Utkin'2025]. Recently (AI-assisted) proof was suggested in [Arulandu'2026].

Discussion

Here we give references to solutions of several problems akin to our inequalities. All cases one way or another concern minimization of the output entropy of certain quantum channel and rely upon the symmetry properties of the problem.

[Lieb'1978] gave a solution of the Wehrl conjecture which can be reformulated as a conjecture about the minimal output entropy of the measurement (quantum-classical) channel associated with the Glauber's coherent states with the underlying Heisenberg group, and generalized the conjecture to SU(2) group. The solution was based on the sharp versions of Young's and Hausdorff-Young inequalities in the classical harmonic analysis.

[Lieb and Solovej'2015] proved the Wehrl-type entropy conjecture for symmetric SU(N) coherent states and suggested a similar conjecture for larger class of Lie groups and their representation (for further progress in this direction see [Frank'2022] and the references therein). Lieb and Solovej used the "universal cloning channel" and established minimization for arbitrary concave function of the output distribution (the majorization).

Another relevant case is the solution of the Gaussian optimizers conjecture for the classical capacity of bosonic Gaussian channels by [Giovannetti, Holevo and Garcia-Patron'2015]. The result for the minimal Wehrl entropy problem can be obtained as a special limiting case of this.

For one mode Gaussian measurement channels the Gaussian optimizers conjecture was settled in [AH, AH-Filippov'2023]. In that case certain generalizations of the logarithmic Sobolev inequality were used.

We suggest that our sharp entropy inequalities can be regarded as a discrete relatives of the aforementioned problems, within the context of the symmetric group and its representations. Tentatively, the minimizers in the inequalities can play the role of a discrete analogue of coherent state vectors. However, what is unusual as compared to problems with continuous symmetry groups is the presence of two types of maximizers and a "phase transition" between them.

* * *