Quantum Gaussian Maximizers and Generalizations of Log-Sobolev Inequalities

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

****

BIID2022, Shenzhen, 26-30.09.2022

A long-standing problem in quantum Shannon theory is the classical capacity of bosonic Gaussian channels of various kinds. *Hypothesis of Gaussian Maximizers* (HGM) states that the full capacity of such channels is attained on Gaussian encodings. A breakthrough was made in [Giovannetti-AH-GarciaPatron'14], [Giovannetti-AH-Mari'15], where HGM was proved for important class of multimode gauge co- or contra-variant Gaussian channels. In [AH'16] the solution was extended to broader class of channels satisfying threshold condition, ensuring that the upper bound for the capacity as a difference between the maximum and the minimum output entropies is attainable.

At the same time, HGM remains open for large variety of Gaussian channels lying beyond the scope of the threshold condition. Here we outline a novel approach [AH'22] (Arxiv:2204.10626; Arxiv:2206.02133) to such problems based on principles of convex programming, and illustrate it on the characteristic case of approximate position measurement with the energy constraint, underlying noisy Gaussian homodyning in quantum optics. Remarkably, for this model, as well as for Gaussian heterodyning, the method reduces the solution of the optimization problem to new generalizations of the celebrated log-Sobolev inequality.

We believe that these two basic models reveal a path to a general solution of Hypothesis of Gaussian Maximizers.

Let $\mathfrak{S}$ be the convex set of all density operators in a separable Hilbert space $\mathcal{H}$ of the quantum system, $\mathcal{X}$ and $\mathcal{Y}$ are standard measurable spaces. *Measurement channel*:

$$M : \rho \to p_\rho(y) = \mathrm{Tr}\, \rho m(y), \quad \rho \in \mathfrak{S},$$

where $m(y)$ is a uniformly bounded positive-operator-valued function, such that $\int m(y)\mu(dy) = I$ (identity operator on $\mathcal{H}$, μ a measure). *Encoding* $\mathcal{E} = \{\pi(dx), \rho(x)\}$ is a probability measure $\pi(dx)$ with a measurable family of states $\rho(x), x \in \mathcal{X}$.

Let H be a Hamiltonian on $\mathcal{H}$, $E > 0$. The *energy-constrained classical capacity* of the measurement channel M is

$$C(M, H, E) = \sup_{\mathcal{E}:\mathrm{Tr}\bar{\rho}_{\mathcal{E}} H \leq E} I(\mathcal{E}, M), \tag{1}$$

where $I(\mathcal{E}, M)$ is the mutual information between x and y, and $\bar{\rho}_{\mathcal{E}} = \int \rho(x)\pi(dx)$ is the *average state* of the encoding.

Introducing the well defined *output differential entropy*

$$h_M(\rho) = -\int p_\rho(y) \ln p_\rho(y) \mu(dy),$$

we have

$$I(\mathcal{E}, M) = h_M(\bar{\rho}_{\mathcal{E}}) - \int h_M(\rho(x)) \pi(dx);$$

$$C(\mathcal{M}, H, E) = \sup_{\rho:\mathrm{Tr}\rho H \leq E} [h_{\mathcal{M}}(\rho) - e_{\mathcal{M}}(\rho)] \tag{2}$$

$$\leq \sup_{\rho:\mathrm{Tr}\rho H \leq E} h_M(\rho) - \inf_\rho h_M(\rho), \tag{3}$$

where we introduced an analog of the *convex closure of the output differential entropy* for quantum channel

$$e_M(\rho) = \inf_{\mathcal{E}:\bar{\rho}_{\mathcal{E}}=\rho} \int h_M(\rho(x)) \pi(dx). \tag{4}$$

There are important cases where (3) turns into equality thus giving the value of the capacity. This happens when the maximizer of the first term in (2) can be represented as a mixture of (pure) states minimizing $h_M(\rho)$. In particular, all the instances where the Hypothesis of Gaussian Maximizer was proved for Gaussian quantum channels so far refer to that case.

In the present talk we propose a method allowing to prove this hypothesis for the first time where this condition is violated, the inequality in (3) is strict and hence becomes useless. The method relies upon computation of the quantity $e_M(\rho)$ given by (4). Then we illustrate the method on the basic cases of Gaussian noisy homo- and hetero-dyning.

Our method of computation of the quantity $e_M(\rho)$ is based on similarity of the optimization problem in the right side of (4) and general quantum Bayes estimation problem.

Any encoding $\mathcal{E} = \{\pi(dx), \rho(x)\}$ is equivalent to probability distribution $\pi(d\rho)$ on the set of quantum states $\mathfrak{S}$. Another equivalent description of $\mathcal{E}$ is given by the positive operator-valued measure $\Pi(d\rho) = \rho\pi(d\rho)$ with values in $\mathfrak{S}$. The average state is

$$\bar{\rho}_{\mathcal{E}} = \int_{\mathfrak{S}} \rho\pi(d\rho) = \Pi(\mathfrak{S}),$$

and the minimized functional

$$F(\mathcal{E}) = \int_{\mathfrak{S}} h\left(p_\rho\right)\pi(d\rho) = \int_{\mathfrak{S}} \mathrm{Tr}\ K(\rho)\Pi(d\rho),$$

where

$$K(\rho) = -\int m(y) \ln p_\rho(y)\mu(dy).$$

By fixing an average state $\overline{\rho}$, we arrive at optimization problem

$$\begin{aligned} \int_{\mathfrak{S}} \mathrm{Tr}\, K(\rho)\Pi(d\rho) &\longrightarrow \min \\ \Pi(d\rho) &\geq 0 \\ \Pi(\mathfrak{S}) &= \overline{\rho}, \end{aligned}$$

which is formally similar to one arising in the general quantum Bayes problem [AH'72]. The minimized functional is affine in $\mathcal{E} = \{\Pi(d\rho)\}$ and the constraints are convex, so it is a convex programming problem. Under certain regularity assumption the problem was investigated in [AH'76], where the following necessary and sufficient conditions for optimality of an ensemble $\mathcal{E}_0 = \{\Pi_0(d\rho)\}$ were given, which we reproduce here *formally*:

By introducing $K(\rho) = -\int m(y) \ln p_\rho(y) \mu(dy)$, the optimality condition for encoding $\mathcal{E} = \{\pi_0(dx), \rho_0(x)\}$ can be written as:

There exists selfadjoint operator Λ_0 *such that*

(i) $\Lambda_0 \leq K(\rho)$ *for* $\rho \in \mathfrak{S}$;

(ii) $[K(\rho_0(x)) - \Lambda_0]\, \rho_0(x) = 0 \pmod{\pi_0}$.

By integrating (ii), the equation for determination of Λ_0 is

$$\int_{\mathfrak{S}} K(\rho_0(x))\rho_0(x)\, \pi_0(dx) = \Lambda_0 \bar{\rho}. \tag{5}$$

Passing to bosonic Gaussian systems, we denote ρ_α centered Gaussian state of canonical commutation relations with the covariance matrix α, by $\mathfrak{S}(\alpha)$ the set of all states ρ with the fixed matrix of second moments α, and $C(M;\alpha) \equiv \sup_{\mathcal{E}:\bar{\rho}_{\mathcal{E}} \in \mathfrak{S}(\alpha)} I(\mathcal{E}, M)$. The following theorem was proved in [AH'21]:

Theorem. *Let M be a general Gaussian measurement channel. The optimizing density operator ρ in (3) is a (centered) Gaussian density operator ρ_α :*

$$C(M;\alpha) = h_M(\rho_\alpha) - e_M(\rho_\alpha),$$

and hence for a quadratic bosonic Hamiltonian $H = R\epsilon R^t$

$$C(M, H, E) = \max_{\alpha:\mathrm{Tr}\,\alpha\,\epsilon \leq E} C(M;\alpha) = \max_{\alpha:\mathrm{Tr}\,\alpha\,\epsilon \leq E} \left[h_M(\rho_\alpha) - e_M(\rho_\alpha)\right].$$

Here R is the collection of canonical variables q, p's, ϵ is real positive definite energy matrix.

The *approximate (unsharp) measurement of position* q in one mode $R = (q, p)$ (a mathematical model for noisy homodyning) is given by POVM $M(dy) = m(y)dy$, where

$$m(y) = \frac{1}{\sqrt{2\pi\beta}} \exp\left[-\frac{(q-y)^2}{2\beta}\right] \equiv g_\beta(q-y),$$

where $\beta > 0$ is the power of the Gaussian noise.

The problem is to compute $e_M(\rho_\alpha)$ and hence the classical capacity $C(M, H, E)$ for the oscillator Hamiltonian $H = \frac{1}{2}\left(q^2 + p^2\right)$. In this case one can restrict to Gaussian states ρ_α with the diagonal covariance matrix

$$\alpha = \left[\begin{array}{cc} \alpha_q & 0 \\ 0 & \alpha_p \end{array}\right].$$

Theorem. *The maximum communication rate*

$$C(M;\alpha) = \frac{1}{2}\ln\frac{\alpha_q+\beta}{\frac{1}{4\alpha_p}+\beta}.$$

is attained on the Gaussian encoding $\{\pi_0(dx), \rho_0(x)\}$ *where* $\rho_0(x) = |x\rangle_\delta \langle x|$ *are the squeezed states,* $|x\rangle_\delta = \mathrm{e}^{-ipx}|0\rangle_\delta$, *with* $\delta = 1/(4\alpha_p)$,

$$_\delta\langle 0|\, q^2 \,|0\rangle_\delta = \delta, \quad \mathrm{Re}\ {}_\delta\langle 0|\, qp \,|0\rangle_\delta = 0, \quad {}_\delta\langle 0|\, p^2 \,|0\rangle_\delta = \frac{1}{4\delta}.$$

The distribution $\pi_0(dx)$ *is centered Gaussian* $g_\gamma(x)dx$ *with the variance* $\gamma = \alpha_q - \frac{1}{4\alpha_p}$.

Thus the information is encoded solely into the displacement x of the position q leaving the momentum p ignored.

The constrained classical capacity is

$$C(M, H, E) = \max_{\alpha_q + \alpha_q \leq 2E} \frac{1}{2} \left[\ln \left(\alpha_q + \beta \right) - \ln \left(1/\left(4\alpha_p\right) + \beta \right) \right] \tag{6}$$

whence

$$C(M, H, E) = \ln \left(\frac{\sqrt{1 + 8E\beta + 4\beta^2} - 1}{2\beta} \right).$$

Sketch of proof of Theorem:

Integrating over $\pi_0(dx)$, and taking into account that

$$\int |x\rangle_\delta \langle x| \, \pi_0(dx) = \rho_\alpha,$$

after a lengthly computation we get

$$\int K(\rho_0(x))\rho_0(x)\pi_0(dx) = \left[\ln\sqrt{2\pi(\beta+\delta)} + \frac{\beta+2\delta}{2(\beta+\delta)} - \frac{2\delta^2 p^2}{(\beta+\delta)}\right]\rho_\alpha.$$

Comparing with (5), we obtain

$$\Lambda_0 = \left[\ln\sqrt{2\pi(\beta+\delta)} + \frac{\beta+2\delta}{2(\beta+\delta)}\right] I - \frac{2\delta^2}{(\beta+\delta)}p^2.$$

This is Hermitian operator satisfying the condition (ii).

Checking the condition

$$(i) \quad \langle\psi|\Lambda_0|\psi\rangle \leq \langle\psi|K(\rho)|\psi\rangle$$

requires a generalization of the logarithmic Sobolev inequality. Let $f(x) = |\psi(x)|^2$ be a smooth probability density on $\mathbb{R}$. Then the inequality reduces to:

$$\int [g_\beta(y) * f(y)] \ln[g_\beta(y) * f(y)] dy + \ln\sqrt{2\pi e\,(\beta+\delta)} + \frac{\delta}{2\,(\beta+\delta)}$$

$$\leq \frac{2\delta^2}{(\beta+\delta)} \int |\psi\prime(x)|^2\, dx \qquad (7)$$

for $\beta, \delta \geq 0$. For $\beta = 0$ this is equivalent to the version of the *log-Sobolev inequality* in [Lieb-Loss'01]. That can be used as starting point for the proof of (7), see [AH'22] Arxiv:2204.10626.

Noisy heterodyning

Next we summarize results from [AH'21], [AH-Kuznetsova'21], [AH-Filippov'22] (Arxiv:2206.02133) concerning the *unsharp joint position-momentum measurement* (with the noisy optical heterodyning as the physical prototype). This is described by the POVM

$$M(dxdy) = D(x,y)\rho_\beta D(x,y)^* \frac{dxdy}{2\pi},$$

where $D(x,y) = \exp i\,(yq - xp)\,, \quad x, y \in \mathbb{R}$, are unitary position-momentum displacement operators, and ρ_β is centered Gaussian density operator with the covariance matrix

$$\beta = \begin{bmatrix} \beta_q & 0 \\ 0 & \beta_p \end{bmatrix}; \quad \beta_q \beta_p \geq \frac{1}{4}.$$

Here β_q (β_p) are the noise power in position (momentum) quadratures. We denote

$$m(x,y) = \frac{1}{2\pi} D(x,y) \rho_\beta D(x,y)^*.$$

Let ρ_α be a centered Gaussian density operator with the diagonal covariance matrix.

The problem is to compute $e_M(\rho_\alpha)$ and hence the capacitles $C(M,\alpha)$, $C(M,H,E)$ for the oscillator Hamiltonian $H = \frac{1}{2}\left(q^2 + p^2\right)$.

Theorem. The optimal encoding is Gaussian with parameters:

Table 1			
range	L: $\frac{1}{2}\sqrt{\frac{\beta_q}{\beta_p}} < \frac{1}{4\alpha_p}$	C: $\frac{1}{4\alpha_p} \leq \frac{1}{2}\sqrt{\frac{\beta_q}{\beta_p}} \leq \alpha_q$	R: $\alpha_q < \frac{1}{2}\sqrt{\frac{\beta_q}{\beta_p}}$
δ_{opt}	$1/(4\alpha_p)$	$\frac{1}{2}\sqrt{\frac{\beta_q}{\beta_p}}$	α_q
$e_M(\rho_\alpha) - c$	$\frac{1}{2}\ln\left[\left(\frac{1}{4\alpha_p} + \beta_q\right) \times (\alpha_p + \beta_p)\right]$	$\ln\left(\sqrt{\beta_q\beta_p} + 1/2\right)$	$\frac{1}{2}\ln\left[\left(\frac{1}{4\alpha_q} + \beta_p\right) \times (\alpha_q + \beta_q)\right]$
$C(M;\alpha)$	$\frac{1}{2}\ln\frac{\alpha_q+\beta_q}{\frac{1}{4\alpha_p}+\beta_q}$	$\frac{1}{2}\ln\frac{(\alpha_q+\beta_q)(\alpha_p+\beta_p)}{\left(\sqrt{\beta_q\beta_p}+1/2\right)^2}$	$\frac{1}{2}\ln\frac{\alpha_p+\beta_p}{\frac{1}{4\alpha_q}+\beta_p}$

Here the column C corresponds to the case where the *threshold condition* holds, implying $e_M(\rho_\alpha) = \min_\rho h_M(\rho)$. Then the full validity of the HGM in the multimode situation was established in [AH-Kuznetsova'21]. In the cases of mutually symmetric columns L and R the problem was solved recently in [AH-Filippov'22].

Maximizing $C(M;\alpha)$ over α_q, α_p which satisfy the energy constraint $\alpha_q + \alpha_p = 2E$, we obtain $C(M, H, E)$ depending on the signal energy E and the measurement noise β_q, β_p :

Table 2: $C(M,H,E)$		
L: [AH-Filippov'22]	C: [AH-Kuznetsova'21]	R: [AH-Filippov'22]
$\beta_q \le \beta_p; E < E(\beta_p, \beta_q)$	$E \ge E(\beta_p, \beta_q) \vee E(\beta_q, \beta_p)$	$\beta_p \le \beta_q; E < E(\beta_q, \beta_p$
$\ln\left(\frac{\sqrt{1+8E\beta_q+4\beta_q^2}-1}{2\beta_q}\right)$	$\ln\left(\frac{E+(\beta_q+\beta_p)/2}{\sqrt{\beta_q\beta_p}+1/2}\right)$	$\ln\left(\frac{\sqrt{1+8E\beta_p+4\beta_p^2}-1}{2\beta_p}\right.$

where we introduced the energy threshold function

$$E(\beta_1, \beta_2) = \frac{1}{2}\left(\beta_1 - \beta_2 + \sqrt{\frac{\beta_1}{\beta_2}}\right).$$

The new log-Sobolev type inequality appears as condition (i):

$$\int \langle\psi| m(x,y) |\psi\rangle \ln \langle\psi| m(x,y) |\psi\rangle \, dxdy$$

$$+\ln 2\pi\sqrt{(\beta_q+\delta)(\beta_p+1/4\delta)}+\frac{\beta_q+2\delta}{2(\beta_q+\delta)}+\frac{\beta_p}{2(\beta_p+1/4\delta)}$$

$$\leq \frac{4\delta^2\beta_p-\beta_q}{2(\beta_q+\delta)(\beta_p+1/4\delta)}\int |\psi\prime(x)|^2\,dx. \tag{8}$$

A proof was found by Sergey Filippov [AH-Filippov'22]. For $\delta=\frac{1}{2}\sqrt{\frac{\beta_q}{\beta_p}}$ the right-hand side vanishes and the inequality turns into a generalization of the Wehrl inequality [AH'16]

$$\min_{\|\psi\|=1} h_M(|\psi\rangle\langle\psi|)=\ln 2\pi e\left(\sqrt{\beta_q\beta_p}+1/2\right).$$

Notably, both the proof of the original Wehrl inequality and the log-Sobolev inequality given by Lieb and Loss rely upon the sharp version of Young's inequality for convolution (with different functions). When $\delta > \frac{1}{2}\sqrt{\frac{\beta_q}{\beta_p}}$ (case **L**), the inequality (8) represents a new type of logarithmic Sobolev inequality which relates the generalized Wehrl entropy $h_M(|\psi\rangle\langle\psi|)$ with the wavefunction gradient.

$* * * * *$