Classical limits and Gaussian-average gaps for compound bosonic fading channels

Lluis Eriksson

26 October 2026

Abstract

We study unassisted classical communication through memoryless mixtures of noisy, phase-insensitive bosonic Gaussian channels. Both thermal noise and photon budget scale with an energy parameter. An explicit measure–prepare representation converts every quantum block code, including entangled codewords and collective receivers, into a classical code with identical error and an exactly quantified cost increment. The same simulation works for an unknown fixed mixing law. A Gaussian-smoothing estimate and a uniform finite-alphabet reduction give an explicit $O(E^{-1})$ comparison with a classical compound variational value. This value optimizes one input distribution before taking the worst mixing law. We also prove a strict separation between classical optimal modulation and the time-sharing hull of Gaussian coherent modulations for a matched-reset example. Matching bounds determine the high-energy limit of a regularized Holevo functional for ensembles with an exactly Gaussian average quantum state, permitting squeezed and entangled averages. Its gap from the unknown-fixed-phase operational capacity remains strictly positive in that limit. The classical coding tools and Gaussian measure–prepare representations are established ingredients. The Gaussian-average restriction is an ensemble condition, not a converse for every finite codebook described as Gaussian.

1 Model and operational statement

All logarithms are natural, and h_2 denotes binary Shannon entropy.

Write $\phi_v(z) = (\pi v)^{-1} e^{-|z|^2/v}$, so that $\mathbb{E}|G_v|^2 = v$. A thermal state γ_t has mean photon number t. Let $\Phi_{a,t}$ be the phase-insensitive channel sending a coherent input $|z\rangle$ to $D(az)\gamma_t D(az)^*$. Its characteristic function is

$$\chi_{\Phi_{a,t}(\rho)}(\xi) = \chi_\rho(\bar{a}\xi) e^{-[t+(1-|a|^2)/2]|\xi|^2}.$$

Physicality requires $t \geq \max(0, |a|^2 - 1)$.

Fix $K < \infty$, $0 < v_* \leq v^* < \infty$, and the compact parameter space

$$\Theta = \{(a, v) \in \mathbb{C} \times \mathbb{R} : |a|^2 \leq K, \quad v_* \leq v \leq v^*\}.$$

Let Γ be a nonempty weakly closed set of probability laws on Θ. For $\nu \in \Gamma$ define

$$N_{E,\nu} = \int_\Theta \Phi_{a,vE}\, d\nu(a, v), \qquad W_\nu(y|x) = \int_\Theta \phi_v(y - ax)\, d\nu(a, v).$$

The unknown law ν is fixed throughout each code; conditional on it, branches are sampled independently in different uses. Neither party receives the branch label or the unknown law. This is a compound memoryless model, not an arbitrarily varying channel.

A quantum code with M equally likely messages has m-mode states ρ_j, a single decoding POVM $\{D_j\}$, and expected photon cost

$$M^{-1}\sum_j \operatorname{Tr}\rho_j \sum_{i=1}^{m} n_i \le mE.$$

The average message error must tend to zero uniformly over $\nu \in \Gamma$. Write $C_{\mathrm{opt}}(E,\Gamma)$ for its supremum of rates. It is an unassisted *classical* capacity of a quantum channel, not a capacity for transmitting qubits.

For a law μ on $\mathbb{C}$, write $I_\mu(W_\nu) = I(X;Y)$, and set

$$V_\Gamma(P) = \sup_{\mu:\int |x|^2 d\mu \le P} \inf_{\nu\in\Gamma} I_\mu(W_\nu).$$

It is essential that the input optimization precede the worst-law operation. This is generally different from the worst separately optimized capacity.

Theorem 1 (Uniform classical comparison)**.** *For $E > 0$, $E \ge 2K/v_*$, the preceding model obeys*

$$V_\Gamma(1) - \frac{2}{v_* E} \;\le\; C_{\mathrm{opt}}(E,\Gamma) \;\le V_\Gamma(1) + \frac{\log(1+K/v_*) + 4K/v_*}{E}. \tag{1}$$

For $K = 0$ both operational capacities are zero. The constants are independent of Γ; K, v_*, v^* stay fixed.

Coding length tends to infinity at fixed E, before energy tends to infinity. No uniform blocklength, efficient optimizer, fixed-error strong converse or positivity for every family is asserted. In particular, a family containing a completely replacing channel has value zero. If $\Gamma = \Gamma_E$ changes between energy levels, (1) compares to $V_{\Gamma_E}(1)$; it does not assert that this sequence converges.

2 An exact simulation preserving the common decoder

Put $k = |a|^2$. For $t \ge k$, the heterodyne measure–prepare representation is

$$\Phi_{a,t}(\rho) = \int_{\mathbb{C}} Q_\rho(\alpha) D(a\alpha)\gamma_{t-k} D(a\alpha)^*\, d^2\alpha, \qquad Q_\rho(\alpha) = \pi^{-1}\langle\alpha|\rho|\alpha\rangle. \tag{2}$$

This is a specialization of the established Gaussian entanglement-breaking representation [1, 2]. Indeed, the characteristic function of the heterodyne outcome is $\chi_\rho(\eta)e^{-|\eta|^2/2}$; multiplying it by the preparation factor $e^{-(t-k+1/2)|\xi|^2}$, at $\eta = \bar{a}\xi$, gives the characteristic function in Section 1. Thus the noise subtraction in (2) is k, and

$$\int |\alpha|^2 Q_\rho(\alpha)\, d^2\alpha = \operatorname{Tr}\rho n + 1. \tag{3}$$

The identity extends to correlated m-mode Husimi outcomes, with total cost $\operatorname{Tr}\rho\sum_i n_i + m$.

For $t > k$, refine the prepared thermal state by its positive coherent-state representation. For every quantum code this gives the classical Markov chain

$$J \longrightarrow A^m \longrightarrow B^m \longrightarrow \widehat{J}, \quad P(\widehat{J} = j | B^m = \beta^m) = \langle\beta^m|D_j|\beta^m\rangle. \tag{4}$$

Given A^m, the middle transition is a product of mixtures with variance $vE - k$. Both the stochastic Husimi encoder and the final Born decision kernel are independent of ν. Equation (4) reproduces the error of the quantum code for *every* law, with its correlations intact. The

hidden B^m is a converse simulation variable, not an observation supplied to the physical receiver.

Normalize $X = A/\sqrt{E}$, $Y = B/\sqrt{E}$. Define

$$W_{E,-,\nu}(y|x) = \int \phi_{v-k/E}(y - ax)\, d\nu, \qquad W_{E,+,\nu}(y|x) = \int \phi_{v+1/E}(y - ax)\, d\nu.$$

The lower channel is realized by coherent input $|\sqrt{E}x\rangle$ and heterodyne reception. The upper simulation has mean cost at most $1 + 1/E$, by (3). Consequently the operational classical compound capacities satisfy the exact sandwich

$$C_{\mathrm{cmp}}(W_{E,+}, \Gamma, 1) \le C_{\mathrm{opt}}(E, \Gamma) \le C_{\mathrm{cmp}}(W_{E,-}, \Gamma, 1 + 1/E). \tag{5}$$

No additivity or Gaussian optimizer of the quantum mixture is used. For continuous ν, the state integrals are trace-norm Bochner integrals; probabilities and nonnegative moments also permit Tonelli.

3 Uniform information estimates

Lemma 2 (Entropy continuity)**.** *Suppose planar densities $f_n, f \le L < \infty$ converge in L^1 and their second moments are at most $M < \infty$. Then $h(f_n) \to h(f)$.*

Proof. On a fixed ball use the uniform continuity of $-u \log u$ on $[0, L]$. Outside radius R, let δ be the mass. The restricted, normalized density has second moment at most M/δ. Gaussian maximum entropy therefore gives

$$-\delta \log L \le -\int_{|y|>R} f \log f \le \delta \log(\pi e M/\delta^2).$$

Here $\delta \le M/R^2$; both bounds tend uniformly to zero as $R \to \infty$, with value zero when $\delta = 0$. This also proves finiteness of the entropies in the assertion. □

Lemma 3 (Noise sensitivity)**.** *Fix any input law of finite second moment and any mixture of bounded gains. If component variances follow $v_\theta = v_0 + \theta d \in [\ell, U]$, $\ell > 0$, with $|d| \le D$, then*

$$|I_\mu(W_{\theta_1}) - I_\mu(W_{\theta_0})| \le \frac{2D}{\ell} |\theta_1 - \theta_0|. \tag{6}$$

The constant is independent of the input peak, its cost and the mixing law.

Proof. First take a compactly supported input. Let $f_{a,v} = (a_{\#}\mu) * \phi_v$, $f = \int f_{a,v} d\nu$, and $J = \int d \nabla f_{a,v} d\nu$. Complex variance v means real-coordinate variance $v/2$, hence $\partial_v \phi_v = \Delta \phi_v / 4$. Differentiation and integration by parts yield

$$\partial_\theta h(f) = \tfrac{1}{4} \int J \cdot \nabla \log f.$$

The Gaussian score conditional-expectation identity and Jensen give $F(f_{a,v}) \le 4/v \le 4/\ell$, where $F(f) = \int |\nabla f|^2 / f$ is planar Fisher information. Weighted Cauchy–Schwarz gives

$$F(f) \le 4/\ell, \qquad \int |J|^2 / f \le 4D^2/\ell.$$

Thus $|\partial_\theta h(f)| \le D/\ell$. To justify the differentiations, bounded centers and variances give common Gaussian upper tails, a Gaussian lower density bound, and polynomial-times-Gaussian

bounds for derivatives; boundary terms vanish. These bounds hold uniformly over atoms or continuous mixtures.

For a general input, replace its tail outside radius R by zero. The coupled outputs converge in L^1, are bounded by $(\pi\ell)^{-1}$, and have second moment at most $K\mathbb{E}|X|^2 + U$. Lemma 2 passes the integrated entropy inequality to the limit. This does not require differentiability for the untruncated heavy-tailed input. For a fixed letter x, the same entropy bound is independent of x. Conditional entropies are integrable, bounded below by a constant and above by $\log[\pi e(K|x|^2+U)]$. Integrate the conditional bound and subtract it from the output bound to obtain (6). □

Revealing the independent branch only for an information upper bound gives

$$0 \le I_\mu(W_\nu) \le \int \log(1 + |a|^2 P/v)\, d\nu \le \log(1 + KP/v_*). \tag{7}$$

No branch is revealed to the actual decoder. The function V_Γ is finite, nondecreasing and concave, and $V_\Gamma(0) = 0$. Indeed input concavity holds simultaneously for every law and survives the infimum. In particular

$$V_\Gamma(1 + u) \le (1 + u)V_\Gamma(1). \tag{8}$$

4 One code for a continuous compound family

We supply the reduction to the established finite-alphabet compound coding mechanism [3, 4]; a finite-output theorem alone does not justify continuous observations or physical costs.

Proposition 4. *For the Gaussian-smoothed family above and $P > 0$,*

$$C_{\mathrm{cmp}}(W, \Gamma, P) = V_\Gamma(P). \tag{9}$$

The same holds for either family in (5) at fixed admissible E.

Proof. For the converse, Fano and data processing applied to a common stochastic encoder give, for each ν,

$$(1 - \epsilon)\log M - h_2(\epsilon) \le I(J; Y_\nu^m) \le \sum_i I_{\mu_i}(W_\nu) \le m I_{\bar\mu}(W_\nu), \qquad \bar\mu = m^{-1}\sum_i \mu_i.$$

The law $\bar\mu$ has cost at most P and is independent of ν. Taking the infimum over ν, then optimizing feasible $\bar\mu$, gives the required order. The block inequality can be proved without subtracting infinite differential entropies: the KL chain rule with product output marginals gives

$$\sum_i I(X_i; Y_i) = I(X^m; Y^m) + D(P_{Y^m} \| \textstyle\prod_i P_{Y_i}).$$

All terms are finite in our bounded-noise regime by (7).

For achievability choose a feasible input near $V_\Gamma(P)$. Mixing it with a small mass at zero makes its cost strictly below P; concavity loses arbitrarily little information uniformly in ν by (7). Truncate its tail to zero, then quantize a bounded input ball. The Gaussian kernel has the global input-gradient bound $\sqrt{2K}/(\pi\sqrt{e}\, v_*^{3/2})$. Thus their density difference is at most this constant times $\mathbb{E}|X_l - X|$. Under this coupling the output densities converge uniformly on every output ball, uniformly in ν, and have common density and second-moment bounds. The proof of Lemma 2, split into that ball and its complement, gives uniform output entropy convergence. The conditional entropy is jointly continuous on compact input and parameter

domains and is bounded by a constant plus $\log(1+|x|^2)$. Its tail integral under the fixed input vanishes. Thus one finite input alphabet with strict cost slack approximates all the information quantities simultaneously.

For this finite input law choose nested finite output partitions: expanding dyadic boxes, their refined cells, and one exterior cell. They generate the Borel sigma algebra. Their mutual informations increase to the continuous-output information by posterior martingale convergence; the posterior entropy is bounded by the finite input entropy. Each partition information is continuous in ν: Gaussian kernels depend continuously on (a, v) in L^1, so every Borel cell probability is continuous, including the unbounded exterior cell. The limit is also continuous: bounded means and compact positive variances give common Gaussian tails and local density convergence. Dini's theorem on compact Γ now supplies one finite partition with uniformly small loss. This partition is fixed before code length.

For completeness, let A, B be these finite alphabets, and let $R < \inf_W I(p, W)$. Choose constant-composition types $p_n \to p$ on its positive support, retaining strict cost slack. Draw $M = \lceil e^{nR} \rceil$ codewords uniformly and independently from their type class. Decode a unique codeword whose empirical mutual information with the output exceeds $R+\eta$, with $\eta > 0$ below the rate slack. This decoder is independent of W. Hoeffding bounds for output counts within each input-symbol class, followed by finite-alphabet entropy continuity, make the transmitted word fail the threshold with probability e^{-cn}, uniformly in W. For any fixed output word a random competing codeword crosses the threshold with probability at most $\mathrm{poly}(n)e^{-n(R+\eta)}$: a conditional type class has size at most $e^{nH(X|Y)}$, whereas the input type class has size at least $(n+1)^{-|A|}e^{nH(p_n)}$. Sum over the polynomial number of joint types and over competitors.

To choose one deterministic codebook, cover the finite stochastic-matrix family in per-row total variation by a net of radius n^{-2}, of cardinality at most $(Cn^2)^{|A||B|}$. Choose its points from the family; all retain the rate slack. The expected *sum* of average errors on this net tends to zero, so a deterministic codebook has uniformly vanishing net error. Product-channel total variation changes every decoder error by at most nn^{-2}. Therefore its error vanishes over the whole family. There is no exchange of an expected supremum with a supremum of expectations.

Each word has exact additive cost $n\sum_x p_n(x)|x|^2 < nP$. Mapping x to $|\sqrt{E}x\rangle$ enforces physical expected photon cost $E\sum_i |x_i|^2$ for every message. Fix the slack and finite approximations first, let $n \to \infty$, then remove those approximations. Equivalently, for each rate strictly below the variational value, choose all finite approximations once before taking the coding limit. If that value is zero, the converse already settles the zero-rate case. The result asserts no uniform coding length or efficient search. □

Proof of Theorem 1. For the plus family use (6) with $D = 1, \ell = v_*$, over an interval of length $1/E$. For the minus family use $D = K, \ell = v_*/2$, valid when $E \geq 2K/v_*$. The bounds survive infima and suprema over the same laws and input constraints. Consequently

$$V_{E,+,\Gamma}(1) \geq V_\Gamma(1) - 2/(v_*E), \qquad V_{E,-,\Gamma}(P) \leq V_\Gamma(P) + 4K/(v_*E).$$

Apply (9) in (5), then (8) and (7). This gives (1). The optimized coherent-modulation heterodyne rate is within $[\log(1+K/v_*)+(4K+2)/v_*]/E$ of the unrestricted quantum rate. □

5 A matched-reset application: a strict Gaussian-hull gap

Fix $k, r > 0$, $0 < \lambda < 1$, $q = 1 - \lambda$, $s = k + r$, and

$$W_\theta(y|x) = \lambda\phi_r(y - \sqrt{k}e^{i\theta}x) + q\phi_s(y).$$

For $\mu_0 = \mathcal{CN}(0, 1)$, the average output is exactly ϕ_s. Let F denote the signal/reset flag and $H(x) = H(F|Y, X = x)$. It is radial, bounded, positive at finite x, and tends to zero at

infinity. The last assertion follows by separating the two widely separated Gaussian clouds and applying binary Fano. It follows that H is not affine in $|x|^2$.

There exists a compactly supported bounded radial h, normalized by $\|h\|_\infty \le 1$, with

$$\int h\,d\mu_0 = \int |x|^2 h\,d\mu_0 = 0, \qquad a_0 = \int hH\,d\mu_0 > 0. \tag{10}$$

Indeed, if all such moment-annihilating tests annihilated H, two fixed radial moment correctors would imply that H equals an affine function of $|x|^2$ on every bounded radial domain, hence globally almost everywhere. Boundedness and decay would make it zero, contradicting positivity. Reverse the sign and normalize a nonzero test. The moment-corrector matrix is invertible by taking two disjoint annuli with distinct conditional squared radii.

For $\mu_\epsilon = (1+\epsilon h)\mu_0$, (10) preserves mass and cost. Since the output second moment stays s, Gaussian cross entropy and the binary flag identity give

$$I_{\mu_\epsilon}(W_0) - I_{\mu_0}(W_0) = \epsilon a_0 - D(p_{Y,\epsilon}\|\phi_s) \ge \epsilon a_0 - \epsilon^2 \int h^2\,d\mu_0. \tag{11}$$

The last inequality is KL data processing and $\log(1+u) \le u$. A fixed sufficiently small $\epsilon > 0$ gives a positive, non-numerical improvement.

Let G_ϑ range over all real two-dimensional Gaussian laws on $\mathbb{C}$, including displaced, anisotropic and singular laws. Write $\vartheta = (m, V)$, $p(\vartheta) = |m|^2 + \operatorname{Tr} V$, $J(\vartheta) = I_{G_\vartheta}(W_0)$, and define

$$H_G = \sup_{\eta:\int p\,d\eta \le 1} \int J\,d\eta. \tag{12}$$

Here the time-sharing setting is known to both parties and does not itself carry a message.

Proposition 5 (Strict Gaussian scheduling gap)**.** *For this matched-reset channel, $H_G < C(1)$, where $C(P) = \sup_{\mu:\mathbb{E}|X|^2 \le P} I_\mu(W_0)$.*

Proof. The symmetry, injectivity and strict-growth strategy has nearby antecedents in fading problems [8, 7]; we check it directly for the present channel and Gaussian scheduling class. The output-law map is injective: subtract $q\phi_s$ and Fourier transform the remaining convolution, whose Gaussian factor is nowhere zero. Thus output entropy, and hence mutual information, is strictly concave in the input law.

If a Gaussian law attained $C(1)$, rotation covariance and strict concavity would force it to be circular $\mathcal{CN}(0, P)$. The finite concave nondecreasing capacity $C(P)$ is unbounded, since revealing the independent flag costs at most $h_2(\lambda)$ and gives

$$C(P) \ge \lambda \log(1 + kP/r) - h_2(\lambda).$$

Such a function is strictly increasing: a flat interval would, by concavity and monotonicity, force all subsequent slopes to vanish. Therefore $P = 1$, contradicted by (11). No Gaussian law attains $C(1)$.

The parameter sublevel sets of p are compact. Laws η with $\int p\,d\eta \le 1$ are tight and weakly closed, hence compact. Gaussian smoothing and Lemma 2 show that J is continuous, including at singular covariance matrices. Furthermore

$$0 \le J(\vartheta) \le \lambda \log(1 + kp(\vartheta)/r).$$

The upper bound is sublinear in p, so its tails over $p > R$ are uniformly integrable under the mean-cost constraint. Compactly supported continuous truncation then makes $\eta \mapsto \int J\,d\eta$ continuous. The supremum (12) is attained.

For an attaining η_*, the barycenter $\bar{\mu} = \int G_\vartheta \, d\eta_*$ has cost at most one. To justify averaging entropies, every output density is bounded below by $q\phi_s$ and above by a common constant. Thus $|\log p_{Y,\bar{\mu}}(y)| \le C + C'|y|^2$, and the average output second moment is finite. Individual output entropies have integrable bounds by constants and by $C+\log(1+p(\vartheta))$; the conditional entropy of $W(.|x)$ is bounded through its binary flag identity. Fubini therefore gives the exact identity

$$I_{\bar{\mu}}(W_0) - H_G = \int D(p_{Y,\vartheta} \| p_{Y,\bar{\mu}}) \, d\eta_* \ge 0.$$

If $H_G = C(1)$, equality forces almost all output laws to coincide. Injectivity makes almost all Gaussian input laws identical to $\bar{\mu}$, which would be a Gaussian maximizer. This contradiction proves the strict gap. Its size and its optimizer are not computed. □

For any correlated multivariate Gaussian modulation on product coherent states with total cost mE, positive thermal preparation gives exactly

$$\chi(X^m : N_{E,\theta}^{\otimes m}) \le I(X^m : Y^m) \le \sum_i I_{\mu_i}(W_\theta) \le m H_G. \tag{13}$$

The first inequality uses the preparation chain $X^m \to Y^m \to |\sqrt{E} Y^m\rangle$, valid for coherent letters without the input heterodyne cost increment. Here $N_{E,\theta} = \lambda \Phi_{\sqrt{k} e^{i\theta}, rE} + q\gamma_{sE}$ Tr; we may take $E \ge 2k/r$, which also ensures physicality and the hypotheses of Theorem 1. The last uses Gaussian marginals and their average cost, allowing individual costs above one. Known, non-message-bearing time sharing of such blocks is also covered. This does *not* bound squeezed or entangled Gaussian quantum signal states, message-bearing covariance selection, or every protocol described as Gaussian.

Choose a finite input law of strict cost below one whose information exceeds H_G by a fixed amount, using Proposition 4's approximation argument. Heterodyne reception loses at most $2/(rE)$ by (6), so it eventually beats the entire benchmark (13). For an unknown fixed phase, $W_\theta(y|x) = W_0(e^{-i\theta} y|x)$ implies $I_\mu(W_\theta) = I_\mu(W_0)$ for every input law. The common code, rather than this equality alone, is supplied by (9). No phase oracle is assumed.

5.1 Ensembles with Gaussian average quantum states

Consider any ensemble $\{p_j, \rho_j\}$ on m modes with average state $\bar{\rho} = \sum_j p_j \rho_j$ Gaussian and $\operatorname{Tr} \bar{\rho} \sum_i n_i \le mE$. The average state may be displaced, squeezed or entangled across uses; the conditional states need not be Gaussian. Continuous ensembles use the relative-entropy definition of Holevo information. Then

$$\frac{1}{m} \chi\big(\{p_j, N_{E,\theta}^{\otimes m}(\rho_j)\}\big) \le H_G + \frac{\lambda \log(1 + k/r) + 4k/r}{E}. \tag{14}$$

The simulation in (4) implies

$$\chi(J : \text{output}) \le I(J; B^m) \le I(A^m; B^m) \le \sum_i I_{\mu_i}(W_{E,-,\theta}).$$

For continuous labels, the first inequality is $D(\sigma_j \| \bar{\sigma}) \le D(P_{B|j} \| P_B)$, integrated in j, without subtracting an entropy of the label. To justify it, approximate the density ratios by conditional expectations on increasing finite partitions of B, preparing the average coherent state in each cell. Finite-partition data processing, L^1 convergence of the ratios, trace-norm convergence of the prepared states, and lower semicontinuity of relative entropy pass to the limit. Tonelli averages the inequalities. The remaining steps are classical data processing and the block KL chain rule. Linearity of the Husimi function makes its average law $Q_{\bar{\rho}}$, which is multivariate

Gaussian for any Gaussian quantum state: heterodyne is a Gaussian measurement [9]. Its normalized marginals μ_i are real planar Gaussian laws, with average cost at most $1 + 1/E$. Correlations in A^m are allowed in the displayed block inequality. The noise-sensitivity bound, uniform in each marginal cost, gives

$$I_{\mu_i}(W_{E,-,\theta}) \le I_{\mu_i}(W_0) + \frac{4k}{rE}.$$

Writing $H_G(P)$ for (12) at cost P, the sum is therefore at most $mH_G(1+1/E)+4mk/(rE)$. The hull is finite, nondecreasing and concave in cost, with $H_G(0) = 0$; mix its parameter probability laws to see concavity. Thus $H_G(1+1/E) \le H_G+H_G/E$. Revealing the independent signal flag and applying Jensen in cost also yields $H_G \le \lambda \log(1+k/r)$, proving (14). All output entropies are finite under the photon constraint and the uniform preparation-noise floor. Known time sharing among such blocks is covered by summing their marginal costs before applying $H_G(P)$.

By Proposition 5, a fixed finite strict-cost law can have $I_\mu(W_0) \ge H_G + \delta$ for some $\delta > 0$. Its coherent-modulation heterodyne rate exceeds the supremum of (14), over all block lengths, by at least

$$\delta - \frac{\lambda \log(1 + k/r) + (4k + 2)/r}{E}, \tag{15}$$

which is positive for sufficiently large E. The common unknown-phase decoder is again supplied by (9).

The average-state condition is essential to the statement. Its distinction from a Gaussian ensemble is standard; see [12, Remark 2], where a thermal average of Fock states gives an explicit example. That work concerns Gaussian measurement channels, not the present non-Gaussian signal/reset mixture. An ensemble of individually Gaussian states with message-dependent covariances can have a non-Gaussian average and is not covered. A finite deterministic codebook generally does not have an exactly Gaussian average either. Equation (14) is an exact Holevo-ensemble benchmark, not a converse for every codebook selected from a Gaussian random-coding distribution or every protocol called Gaussian.

5.2 A matching limit for the ensemble benchmark

Define the regularized Gaussian-average Holevo functional

$$R_G(E) = \sup_{m\ge1} \frac{1}{m} \sup_{\substack{\{p_j,\rho_j\}:\ \bar\rho \text{ Gaussian} \\ \operatorname{Tr}\bar\rho\sum_i n_i \le mE}} \chi\big(\{p_j, N_{E,0}^{\otimes m}(\rho_j)\}\big). \tag{16}$$

Corollary 6 (Gaussian-average Holevo limit)**.** *For $E \ge 2k/r$, the functional (16) satisfies*

$$H_G - \frac{2}{rE} \le R_G(E) \le H_G + \frac{\lambda \log(1 + k/r) + 4k/r}{E}. \tag{17}$$

In particular $R_G(E) \to H_G$.

Proof. The upper bound is (14). For the lower, take a parameter probability law nearly optimal in (12). Mixing it with a small mass at the zero-cost point Gaussian creates strict cost slack with arbitrarily small loss of its average J. Replace its high-cost tail by that point. The information loss tends to zero since $J(\vartheta) \le \lambda \log(1 + kp(\vartheta)/r) = o(p(\vartheta))$. On a compact parameter sublevel, continuity of p, J permits finite quantization; rational approximation of the resulting weights retains the cost slack. Thus rational finite time-sharing laws approximate H_G from below with strict mean cost below one.

For weights with denominator m, use m independent coherent Gaussian modulations in their respective parameter settings. Their average quantum input is a product Gaussian state,

of photon cost $E\sum_i p(\vartheta_i) < mE$. Heterodyne information adds over these independent modes, and Holevo information is at least the information of that physical measurement. Lemma 3 gives at least $\sum_i J(\vartheta_i) - 2m/(rE)$. Divide by m and remove the finite approximations at fixed E. This proves the lower bound in (17) without claiming a finite-message code with exactly Gaussian average. □

The quantum channels satisfy $N_{E,\theta} = \mathcal{U}_\theta \circ N_{E,0}$, where $\mathcal{U}_\theta$ rotates the output and leaves the thermal reset invariant. Consequently (16) is phase independent even for arbitrary conditional input states. For the compact family of unknown fixed phases, the classical compound value is $C(1)$, by the output-rotation identity. Equations (1), (17) and Proposition 5 give the exact limiting separation

$$\lim_{E\to\infty} \left[C_{\mathrm{opt}}(E, \text{all fixed phases}) - R_G(E)\right] = C(1) - H_G > 0. \tag{18}$$

The common decoder in the operational term comes from (9), whereas R_G remains the explicitly defined Holevo functional. Its block supremum is not silently replaced by an operational constrained-code capacity. The positive constant in (18) is not numerically evaluated.

6 Prior work, limits and verification context

Gaussian measure–prepare structure, the single Gaussian classical limit, data processing, Fano, heat-flow score methods, and type coding are prior tools [1, 2, 3, 4]. Recent general Gaussian-encoding optimality results for single-mode Gaussian channels are given by Adesso, Lami and Giovannetti [11]. Their Gaussian-channel hypothesis excludes the generic mixtures studied here; a mixture of Gaussian components is not in general a Gaussian channel. Their fixed-error energy-rescaling formula is not imported into our compound average-error problem. Uniform quantization using compactness and Dini-type arguments is also explicit in earlier compound-channel work [5, 6]. The Mitran–Devroye–Tarokh operational model assumes that the channel parameter is known to the receiver; its theorem is therefore not a substitute for our decoder without that side information. Our use of its quantization mechanism is separately justified above. Classical on/off fading already motivates non-Gaussian input optimization [7]. Non-Gaussian advantages in entanglement-assisted quantum fading are also prior research [8]. Near-capacity coherent modulation for non-Gaussian attenuator environments and additive classical noise is also established [10]. Those bounds depend on the environment and are independent of input energy at fixed channel. Our joint scaling and unknown-law comparison must be distinguished from that earlier setting; coherent near-optimality itself is not new. Our example therefore does not discover the broad non-Gaussian phenomenon. The contribution here is the uniform simulation and the persistent Gaussian-average Holevo gap in the declared joint-scaling model, with its ensemble condition and its unknown-fixed-law coding quantifiers.

The results do not cover within-block changes of the unknown law, channel memory, unbounded gains, a vanishing normalized noise floor, arbitrary nonclassical reset states, or entanglement assistance. Expected photon cost is not a deterministic photon-number bound. A directed primary-source search is not an exhaustive literature audit. The proof does not identify the general classical fading optimizer.

AI and responsibility disclosure. The named author directed substantial OpenAI Codex assistance in derivation and writing. Technical audits requested GPT-6 Astra xhigh, with shared filesystem and reused context declared. Runtime identity, effort, isolation and statistical independence are not independently attested. No independent human review, Lean formalization, certified worldwide originality or predicted assessment score is claimed. Assessment reports must be read with their exact artifact hashes, declared participation and limitations. Earlier manuscript scores do not apply to this manuscript.

References

[1] A. S. Holevo, Entanglement-breaking channels in infinite dimensions, arXiv:0802.0235v1 (2008).

[2] A. S. Holevo and V. Giovannetti, Quantum channels and their entropic characteristics, arXiv:1202.6480v1 (2012), especially Sections 3.4–3.5 and 4.3.

[3] I. Csiszár, The method of types, IEEE Transactions on Information Theory 44, 2505–2523 (1998), doi:10.1109/18.720546.

[4] I. Bjelaković and H. Boche, Classical capacities of averaged and compound quantum channels, arXiv:0710.3027v2 (2009), Section II for the classical finite-alphabet formulation.

[5] P. Mitran, N. Devroye and V. Tarokh, On compound channels with side-information at the transmitter, author draft, 20 December 2005, Lemmas 3–5. Published in IEEE Transactions on Information Theory 52, 1745–1755 (2006), doi:10.1109/TIT.2006.871044. The cited draft is not identified as the final published byte version.

[6] R. F. Schaefer and S. Loyka, The secrecy capacity of compound Gaussian MIMO wiretap channels, arXiv:1509.02182v1 (2015), Section III for common quantization; its secrecy model and Gaussian optimizer are not imported here.

[7] Tu, Li and Dhillon, An information-theoretic framework for vision-aided communication over memoryless on-off fading channels, npj Wireless Technology 2, 80 (2026), doi:10.1038/s44459-026-00074-7.

[8] G. Catalano, M. Fanizza, F. A. Mele, G. De Palma and V. Giovannetti, Convex combinations of bosonic pure-loss channels, arXiv:2604.26874v2 (2026), selected discussion of entanglement-assisted fading and non-Gaussian inputs. Version 2, revised 30 September 2026.

[9] C. Weedbrook, S. Pirandola, R. García-Patrón, N. J. Cerf, T. C. Ralph, J. H. Shapiro and S. Lloyd, Gaussian quantum information, arXiv:1110.3234v1 (2011), Sections II.5 and II.5.2 for Gaussian measurement outcomes and heterodyne.

[10] S. Huber and R. König, Coherent state coding approaches the capacity of non-Gaussian bosonic channels, arXiv:1710.06734v2 (2018), Theorem 1. Journal of Physics A 51, 184001, doi:10.1088/1751-8121/aab7ff.

[11] G. Adesso, L. Lami and V. Giovannetti, Classical capacity of single-mode bosonic Gaussian channels, arXiv:2610.22308v1 (15 October 2026), Section IV and Theorem 2 for the Gaussian-channel scope.

[12] A. Holevo, On the classical capacity of general quantum Gaussian measurement, Entropy 23, 377 (2021), doi:10.3390/e23030377, Theorem 1 and Remark 2 for the distinction between Gaussian average and Gaussian ensemble; its Gaussian-measurement optimization is not imported here.