hexagon:2611.00292v1

Energy complexity and dissipation of noisy consensus

Zhenyu Sun*

November 2026

Abstract

We characterize the signal-energy complexity of distributed agreement over independent Gaussian point-to-point channels. For n processes deciding L binary coordinates under coordinatewise unanimity validity and a joint failure probability of at most ε, every finite causal protocol requires a worst-input expected signal energy of $\Omega(N_0 n[L+\ln(n/\varepsilon)])$. This lower bound combines two distinct mechanisms: an exact relative-entropy identity for locally interchanged reception histories yields the reliability term even in the presence of shared randomness and active noisy feedback, while a coordinatewise direct-sum argument establishes the payload term. In the fault-free setting, a deterministic fixed-leader protocol attains this optimal order for all $0 < \varepsilon \le 1/8$. In the presence of $f \le (1/3-\kappa)n$ static Byzantine faults, the same order is achieved in the polynomial-error regime $\varepsilon \ge n^{-a}$ via public committee sampling and robust joint decoding within explicit finite storage and computation budgets. Consequently, batching amortizes the reliability overhead, yielding an energy cost of $\Theta(N_0 n[1 + L^{-1}\ln(n/\varepsilon)])$ per coordinate. We then connect this communication optimum to physical dissipation through a stochastic-thermodynamic model of a thermal channel port. Under a non-recovering mode replacement interface, the communication converse translates directly into lower bounds on entropy production and, when the endpoint mean internal energies agree, cyclic work. When work recovery is permitted prior to mode reset, we derive the exact minimum reset entropy production within a bounded-displacement control class as a function of the information accessible to the recovery controller, thereby separating the fundamental communication cost from its control-dependent thermodynamic realization.

Keywords. consensus; Byzantine agreement; Gaussian channels; energy complexity; entropy production

1 Introduction

Reliable distributed agreement requires all honest participants to converge on a common decision from initially dispersed inputs. Over noisy physical channels, the fundamental signal-energy cost of this task decomposes into two parts: the energy needed to disseminate the decision payload across the network, and the energy required to suppress the joint probability of disagreement or validity violation. Crucially, deciding a block of coordinates simultaneously need not incur the same per-coordinate cost as deciding a single coordinate in isolation. Clarifying this distinction addresses the minimum-energy consensus question posed by Zhang

*School of Mathematics and Statistics, Northeastern University at Qinhuangdao, China.

[21] and exposes the structural boundary between abstract communication complexity and physical dissipation.

We study n processes communicating over independent memoryless Gaussian point-to-point channels of the form $Y = A + Z$, where $Z \sim N(0, N_0/2)$ and each transmission of amplitude A incurs signal energy A^2. The communication schedule is input-independent, and any feedback must traverse these same charged noisy channels. The processes must decide L binary coordinates subject to coordinatewise unanimity validity and a single joint error budget ε covering all honest participants and coordinates. We prove that both the expected and pathwise optimal signal energies scale as

$$N_0 n[L + \ln(n/\varepsilon)]. \tag{1.1}$$

The converse holds for arbitrary finite causal protocols and requires no faults. In the fault-free regime, a simple fixed-leader protocol matches this bound for all $0 < \varepsilon \leq 1/8$ without shared randomness. Against a static Byzantine adversary corrupting fewer than a third of the processes, the same scaling is achieved in the polynomial-error regime by combining a publicly sampled committee with robust joint reception.

Our converse rests on two complementary techniques. First, locally interchanging the complete reception histories of a single process between two coupled executions yields an exact relative-entropy identity governed solely by the expected energies of the two actual executions. Packing mutually exclusive single-process disagreement events into the global error budget ε then forces the $\Omega(N_0 n \ln(n/\varepsilon))$ reliability term. Second, a two-population mutual-information inequality, summed across coordinates and paired receivers via a direct-sum decomposition, forces the $\Omega(N_0 nL)$ payload term. Because the logarithmic reliability cost does not scale with L, encoding a batch of $L = \Omega(\ln(n/\varepsilon))$ coordinates amortizes the overhead to $\Theta(N_0 n)$ energy per coordinate.

To link signal energy to physical dissipation, we analyze a continuous thermal port in which a harmonic mode in equilibrium is displaced, sampled, and subsequently replaced by a fresh bath mode. When the displaced mode is discarded without work recovery, each reset generates entropy equal to the normalized signal energy plus the conditional mutual information between the mode and the retained measurement record. Combining this reset identity with a mixed input distribution and continuous-time Markovian register updates yields lower bounds on entropy production and on supplied work minus the internal-energy change over repeated fault-free consensus batches. The latter is a work bound when the endpoint mean internal energies agree. Conversely, if a bounded-speed recovery displacement is applied before the mode is discarded, the work recoverable by the permitted translations can depend strictly on the σ-field available to the controller. We solve this constrained variational problem in closed form and demonstrate a strict gap between amplitude-controlled and sample-controlled recovery.

1.1 Relation to previous work

Point-to-point Gaussian communication has established energy bounds with and without noiseless feedback [18], as well as error-exponent bounds with active noisy feedback [11]. In distributed consensus, however, the correctness criterion is a global joint event across n nodes,

and local history perturbations simultaneously affect both incoming receptions and outgoing transmissions. Related distributed models—such as noisy gossip protocols [7] and awake-round energy complexity [17]—operate under different channel access abstractions. While information-theoretic lower bounds and direct-sum methods are foundational in communication complexity [2], our converse directly couples causal Gaussian path measures rather than lifting a discrete noiseless message bound to continuous channels.

On the algorithmic side, our upper bounds synthesize finite linear codes, the classical Phase King protocol [4], and concentration inequalities for sampling without replacement [10]. By combining committee internal agreement with majority-tournament joint decoding at external nodes, the protocol achieves the interactive energy lower bound while operating under strict finite quantization, storage, and bit-operation budgets. As is standard in Byzantine fault tolerance, achievable guarantees depend on the timing of corruptions and setup assumptions [1], and shared randomness constitutes a distinct distributed resource [5]. Here, public randomness is used solely to sample a static committee in the Byzantine regime, whereas fault-free optimality is achieved deterministically.

In the physics of computation, Landauer's principle and reversible computing architectures establish that logical irreversibility, rather than information transport per se, dictates fundamental heat dissipation [3, 14, 15]. Our thermodynamic analysis builds on the relative-entropy formulation of non-equilibrium entropy production [9], work extraction under operational and informational constraints [12], and thermodynamic speed limits governed by dynamical activity [8, 19]. Yadav and Wolpert derive mismatch-cost bounds for communication channels and for encoding and decoding linear codes [20]. Our results concern distributed consensus and isolate the port-level reset cost and register transition dissipation dictated by the consensus converse under explicit cyclic boundary conditions.

The remainder of the paper is organized as follows. Section 2 formalizes the consensus task, channel model, and resource accounting. Sections 3 and 4 establish the matching signal-energy lower and upper bounds. Sections 5 and 6 analyze the thermal port interface, optimal displacement recovery, and cyclic thermodynamic bounds.

2 Task, channels, and resources

All logarithms are natural. Let $h(p) = -p \ln p - (1-p)\ln(1-p)$ denote the binary entropy function, and write the binary relative entropy as

$$d_{\mathrm{Ber}}(p\|q) = p \ln \frac{p}{q} + (1-p) \ln \frac{1-p}{1-q}. \tag{2.1}$$

We consider a network of $n \geq 2$ uniquely identified processes. Each process $i \in \{1, \ldots, n\}$ receives an input vector $X_i \in \{0,1\}^L$ and must produce an output vector $O_i \in \{0,1\}^L$ by a prescribed deadline τ. At most $f < n/3$ processes may be Byzantine faulty. Point-to-point channels authenticate their senders, so faulty processes cannot spoof honest identities or tamper with transmissions between honest pairs.

Definition 2.1 (Joint correctness)**.** A protocol is ε-correct if, for every input assignment and every admissible adversarial strategy, with probability at least $1-\varepsilon$ all honest processes output the same vector $O_i = O$ and, for each coordinate $\ell \in \{1, \ldots, L\}$, whenever all honest

processes share a common input bit b at coordinate ℓ, their common output satisfies $O_\ell = b$. Throughout, we assume $0 < \varepsilon \leq 1/8$.

Note that the failure budget ε controls the joint error event across all honest processes and all L coordinates simultaneously, and validity is enforced coordinatewise rather than only when the entire L-bit input vectors are unanimous.

Communication proceeds over at most M scheduled scalar channel uses. The t-th channel use occurs between a designated sender s_t and a distinct receiver r_t according to the additive white Gaussian noise (AWGN) model

$$Y_t = A_t + Z_t, \qquad Z_t \sim N(0, N_0/2), \tag{2.2}$$

where $N_0 > 0$ and the noise variables $(Z_t)_{t=1}^M$ are mutually independent. All initial public and private randomness is jointly independent of the inputs and channel noises. Each output is a measurable function of the corresponding process's input, initial randomness, and local reception history. The transmission schedule $(s_t, r_t)_{t=1}^M$ is fixed prior to execution; it may depend on public randomness that is independent of the inputs and channel noises. At each step t, an honest sender s_t selects its real-valued amplitude A_t via a Borel-measurable causal function of its input X_{s_t}, its initial private or public randomness, and its prior local channel observations. Because amplitudes are finite real numbers on every finite observation history, the induced path measure of any finite protocol has strictly positive density with respect to the Lebesgue product measure. Silence corresponds to setting $A_t = 0$, in which case the receiver r_t still observes pure Gaussian noise Z_t. All input-dependent interaction must pass through the noisy channels (2.2); there are no uncharged side channels such as input-dependent scheduling, noiseless control flags, or shared memory.

Each honest process maintains a digital memory of at most B bits, which must accommodate its initial input, any preloaded code descriptions or committee lists, and all retained quantized channel observations. Our achievability schemes explicitly respect finite B, whereas our converses hold even when processes retain unquantized real-valued observation histories. Local digital computation is assumed error-free. If the network has an aggregate channel-use rate limit of ν, the communication phase alone requires time at least M/ν. The achievability deadline bounds additionally assume that scheduled uses can be served serially at rate ν. Their operation and time budgets start after the input-independent code matrices, quantizer parameters, and any public committee roster have been installed; matrix construction and public sampling are setup resources.

For a protocol Π, an input assignment $x = (x_1, \ldots, x_n)$, and an adversarial strategy $\mathcal{A}$, the expected honest signal energy and its minimax optimum are defined as

$$E_\Pi(x, \mathcal{A}) = \mathbb{E} \sum_{t: s_t \text{ honest}} A_t^2, \qquad \mathcal{E}^{\mathrm{av}}_{\star,L} = \inf_{\Pi} \sup_{x,\mathcal{A}} E_\Pi(x, \mathcal{A}), \tag{2.3}$$

where the infimum is taken over all ε-correct protocols satisfying the prescribed resource budgets (B, M, τ) and setup conditions (with $\mathcal{E}^{\mathrm{av}}_{\star,L} = +\infty$ if no feasible protocol exists). We write $\mathcal{E}^{\mathrm{hard}}_{\star,L}$ for the corresponding optimum when expected energy is replaced by a deterministic pathwise energy cap. All lower bounds in Section 3 are derived in the fault-free setting ($f = 0$). In the Byzantine achievability analysis of Section 4, we consider a static corruption model

in which the input assignment and the set of at most f faulty processes are fixed prior to the realization of the public committee sample, after which faulty processes may choose their transmitted amplitudes adaptively based on the public sample and their causal view of the execution.

2.1 Main signal-energy result

To streamline notation, we leave the dependence of $\mathcal{E}^{\mathrm{av}}_{\star,L}$ and $\mathcal{E}^{\mathrm{hard}}_{\star,L}$ on $(n, f, L, \varepsilon; B, M, \tau)$ and the setup assumptions implicit. The following theorem synthesizes our communication bounds from Sections 3 and 4.

Theorem 2.2 (Signal-energy complexity)**.** *Let $n \geq 2$, $L \geq 1$, and $0 < \varepsilon \leq 1/8$. Every ε-correct protocol satisfies*

$$\mathcal{E}^{\mathrm{av}}_{\star,L} = \Omega(N_0 n[L + \ln(n/\varepsilon)]),$$

with a universal implied constant that holds even when $f = 0$. Conversely:

1. *For $f = 0$, a deterministic fixed-leader protocol with no public randomness attains this energy order for all $0 < \varepsilon \leq 1/8$ under the finite resource budgets specified in Theorem 4.2.*
2. *For any fixed $\kappa \in (0, 1/3)$ and $a > 0$, when $f \leq (1/3 - \kappa)n$ and $n^{-a} \leq \varepsilon \leq 1/8$, a protocol using independent public committee sampling attains this order uniformly in $L \geq 1$, provided the fault set is fixed prior to committee selection and the resource budgets satisfy* (4.12)–(4.14)*.*

In both achievable regimes, the expected and pathwise optima satisfy $\mathcal{E}^{\mathrm{av}}_{\star,L} = \Theta(N_0 n[L + \ln(n/\varepsilon)])$ and $\mathcal{E}^{\mathrm{hard}}_{\star,L} = \Theta(N_0 n[L + \ln(n/\varepsilon)])$.

Proof. The lower bound is established in Corollary 3.6, fault-free achievability in Theorem 4.2, and static Byzantine achievability in Theorem 4.3. □

3 Signal-energy lower bounds

3.1 Interchanging local reception histories

To compare protocol behavior across two fixed fault-free input assignments x^0 and x^1, we couple two executions by sharing the complete initial random seed R (encompassing both public and private randomness) while drawing independent channel noise sequences for the two runs. The resulting joint path measure takes the form

$$P(dR, dy^0, dy^1) = \mu_R(dR) p_0(y^0 \mid R) p_1(y^1 \mid R)\, dy^0 dy^1, \tag{3.1}$$

where conditioning on R renders the transmission schedule and all local encoding rules deterministic functions of the observed channel outputs. Each marginal $\int p_b(y^b \mid R)\mu_R(dR)$ faithfully represents an actual execution under input x^b.

Suppose process i holds the same local input in both assignments, $x^0_i = x^1_i$. Define the transformation T_i on the coupled path space that swaps all channel outputs received by

process i between the two executions ($y_t^0 \leftrightarrow y_t^1$ for all t with $r_t = i$) while leaving R and all receptions at other processes $j \neq i$ unchanged. Clearly, T_i is a measure-preserving involution with respect to the reference measure $\mu_R \otimes \mathrm{Leb}^{2M}$. Because $x_i^0 = x_i^1$, applying T_i exchanges the entire local observation trajectory—and hence the final decision O_i—of process i between the two runs, while leaving the local trajectories and decisions of all other processes $j \neq i$ invariant. Let $P_i = T_{i*}P$ denote the pushforward law, and let $\delta(i) = \{t : s_t = i \text{ or } r_t = i\}$ denote the set of channel uses incident to process i.

Lemma 3.1 (Gaussian interchange identity)**.** *For any protocol with finite M and finite expected energy,*

$$D(P_i\|P) = D(P\|P_i) = \frac{2}{N_0}\mathbb{E}\sum_{t\in\delta(i)}(A_t^0 - A_t^1)^2 \leq \frac{4}{N_0}(E_{0,i} + E_{1,i}), \tag{3.2}$$

where $E_{b,i} = \mathbb{E}\sum_{t\in\delta(i)}(A_t^b)^2$ is the expected energy incident to process i in execution $b \in \{0,1\}$.

Proof. Condition on R. By causality, the joint density of the two observation sequences factors chronologically over $t = 1, \ldots, M$ as a product of conditional Gaussian densities. For any channel use $t \notin \delta(i)$, neither the sender's history nor the received observation is affected by T_i, so the corresponding density factors in P and P_i are identical and cancel in the log-likelihood ratio. If $r_t = i$, T_i exchanges the received observations Y_t^0 and Y_t^1 while the sender's amplitudes A_t^0 and A_t^1 remain unchanged. Conversely, if $s_t = i$, T_i exchanges the prior local histories of process i between the two runs, which swaps the transmitted amplitudes A_t^0 and A_t^1 while leaving the observations Y_t^0 and Y_t^1 at receiver $r_t \neq i$ unchanged. In either case, each incident channel use $t \in \delta(i)$ contributes to $\ln(dP/dP_i)$ the term

$$\begin{aligned}&\frac{(Y_t^1 - A_t^0)^2 + (Y_t^0 - A_t^1)^2 - (Y_t^0 - A_t^0)^2 - (Y_t^1 - A_t^1)^2}{N_0}\\ &\qquad = \frac{2}{N_0}(Y_t^0 - Y_t^1)(A_t^0 - A_t^1).\end{aligned} \tag{3.3}$$

Conditioned on R and the histories prior to step t, the amplitudes A_t^0, A_t^1 are fixed and the fresh noises Z_t^0, Z_t^1 have mean zero, yielding a conditional expectation of $\frac{2}{N_0}(A_t^0 - A_t^1)^2$. Finite second moments and the Cauchy–Schwarz inequality ensure integrability of the log-likelihood ratio. Summing over $t \in \delta(i)$ and taking expectations establishes the identity for $D(P\|P_i)$. Because T_i is a reference-measure-preserving involution and P, P_i are mutually absolutely continuous, a change of variables gives $D(P_i\|P) = D(P\|P_i)$. The final inequality follows from $(a-b)^2 \leq 2(a^2+b^2)$. □

A key structural feature of Lemma 3.1 is that the right-hand side of (3.2) depends solely on the expected energies $E_{0,i}$ and $E_{1,i}$ under the true execution laws. Unlike change-of-measure arguments that evaluate a single execution against a counterfactual encoder driven by an altered observation path, the coupled involution naturally accommodates adaptive protocols whose amplitudes may take rare, large excursions.

3.2 Reliability without faults

Throughout our converse analysis, we set

$$s = \lfloor n/2 \rfloor, \qquad q_\varepsilon = \frac{1-3\varepsilon}{2}, \qquad c_\varepsilon = \frac{(1-3\varepsilon)^2}{4}. \tag{3.4}$$

Theorem 3.2 (Reliability converse)**.** *Every ε-correct protocol (Definition 2.1), even for $f = 0$, satisfies*

$$\mathcal{E}^{\mathrm{av}}_{\star,L} \geq \frac{N_0 s}{16} d_{\mathrm{Ber}}(q_\varepsilon \| \varepsilon/s) = \Omega(N_0 n \ln(n/\varepsilon)), \tag{3.5}$$

where the implied asymptotic constant is uniform over all $n \geq 2$ and $0 < \varepsilon \leq 1/8$.

Proof. Consider first a single coordinate ($L = 1$). Fix a balanced input assignment $z \in \{0,1\}^n$ containing s zeros and $n - s$ ones. By ε-correctness under z, all nodes agree with probability at least $1 - \varepsilon$, so there exists a bit $b \in \{0,1\}$ such that the event $\{O_1 = \cdots = O_n = b\}$ occurs with probability at least $(1-\varepsilon)/2$. Now consider the unanimous input $(1-b)^n$, under which validity requires $\{O_1 = \cdots = O_n = 1-b\}$ to occur with probability at least $1 - \varepsilon$. Coupling the executions under $x^0 = z$ and $x^1 = (1-b)^n$ as in Lemma 3.1, let G denote the joint event that execution 0 outputs all-b and execution 1 outputs all-$(1-b)$. By the union bound, $P(G) \geq (1-\varepsilon)/2 - \varepsilon = q_\varepsilon$.

Because z contains s zeros and $n - s \geq s$ ones, z and $(1-b)^n$ agree on at least s processes. Fix a subset I of exactly s such processes. For each $i \in I$, set $F_i = T_i(G)$ and $p_i = P(F_i)$. By the invariance properties of T_i, on the event F_i, execution 1 (under input $(1-b)^n$) produces $O_i^1 = b$ at process i and $O_j^1 = 1 - b$ at every other process $j \neq i$. Consequently, the events $(F_i)_{i \in I}$ correspond to distinct single-node disagreement patterns in execution 1; they are mutually disjoint and all lie within the disagreement failure event of execution 1. Thus,

$$\sum_{i \in I} p_i \leq \varepsilon, \qquad P_i(F_i) = P(G) \geq q_\varepsilon. \tag{3.6}$$

Apply the data-processing inequality for relative entropy to the indicator of F_i. Since $p_i \leq \varepsilon < q_\varepsilon \leq P_i(F_i)$, monotonicity in the first argument gives $d_{\mathrm{Ber}}(P_i(F_i)\|p_i) \geq d_{\mathrm{Ber}}(q_\varepsilon\|p_i)$. Convexity in the second argument, followed by its monotone decrease for $p < q_\varepsilon$ and the bound $s^{-1}\sum_{i \in I} p_i \leq \varepsilon/s$, then gives

$$\sum_{i \in I} D(P_i\|P) \geq \sum_{i \in I} d_{\mathrm{Ber}}(P_i(F_i)\|P(F_i)) \geq \sum_{i \in I} d_{\mathrm{Ber}}(q_\varepsilon\|p_i) \geq s d_{\mathrm{Ber}}(q_\varepsilon\|\varepsilon/s), \tag{3.7}$$

where the inequality holds trivially if any $p_i = 0$ since the left-hand side is infinite. Because each channel use has one sender and one receiver, it is incident to at most two processes in I, so $\sum_{i \in I}(E_{0,i} + E_{1,i}) \leq 2(E_z + E_{(1-b)^n})$. Combining (3.2) and (3.7) yields $\frac{8}{N_0}(E_z + E_{(1-b)^n}) \geq s d_{\mathrm{Ber}}(q_\varepsilon\|\varepsilon/s)$, so at least one of these two input energies is at least $\frac{N_0 s}{16} d_{\mathrm{Ber}}(q_\varepsilon\|\varepsilon/s)$. The worst-case energy of this protocol is therefore at least that value. Taking the infimum over all feasible protocols gives (3.5).

For $L > 1$, we obtain the same bound by fixing the last $L - 1$ input coordinates to zero and projecting the outputs onto the first coordinate. Finally, since $q_\varepsilon \geq 5/16$ for $\varepsilon \leq 1/8$, the lower bound $d_{\mathrm{Ber}}(q_\varepsilon\|p) \geq q_\varepsilon \ln(1/p) - h(q_\varepsilon)$ establishes the $\Omega(N_0 n \ln(n/\varepsilon))$ scaling for

small ε/s, while strict positivity of $d_{\mathrm{Ber}}(q_\varepsilon\|\varepsilon/s)$ on the remaining compact parameter domain guarantees a uniform constant across all $n \geq 2$ and $0 < \varepsilon \leq 1/8$. □

Notice that the $\ln(n/\varepsilon)$ factor in Theorem 3.2 is not an artifact of union-bounding independent node errors of size ε/n; rather, it arises directly from packing s disjoint single-node isolation events F_i into the global failure budget ε. In particular, no protocol with finite M and finite expected energy can achieve $\varepsilon = 0$.

Remark 3.3 (General memoryless channels)**.** The interchange argument extends immediately to any sequence of memoryless point-to-point channels $W_t(\cdot \mid a)$ with a common output support and transmission cost $c_t(a)$ satisfying the symmetric divergence bound

$$D(W_t(\cdot \mid a)\|W_t(\cdot \mid b)) + D(W_t(\cdot \mid b)\|W_t(\cdot \mid a)) \leq \lambda[c_t(a) + c_t(b)].$$

Under this distinguishability condition, the prefactor $N_0 s/16$ in (3.5) is simply replaced by $s/(4\lambda)$.

3.3 Information required by parallel tasks

We next establish the payload lower bound by quantifying the mutual information that must cross a network bipartition.

Lemma 3.4 (Two-population information)**.** *Let U and V be independent fair bits, and let $A, B \in \{0,1\}$ be the outputs of two processes holding inputs U and V, respectively. If $\mathbb{P}(A \neq B \mid U = u, V = v) \leq \varepsilon$ for all $(u,v) \in \{0,1\}^2$ and $\mathbb{P}(A = b \mid U = b, V = b) \geq 1 - \varepsilon$ and $\mathbb{P}(B = b \mid U = b, V = b) \geq 1 - \varepsilon$ for each $b \in \{0,1\}$, then*

$$I(V; A \mid U) + I(U; B \mid V) \geq c_\varepsilon. \tag{3.8}$$

Proof. Let $q_{uv} = \mathbb{P}(A = 1 \mid U = u, V = v)$ and $r_{uv} = \mathbb{P}(B = 1 \mid U = u, V = v)$. Correctness implies

$$q_{00}, r_{00} \leq \varepsilon, \quad q_{11}, r_{11} \geq 1 - \varepsilon, \quad |q_{uv} - r_{uv}| \leq \varepsilon \quad \forall u, v \in \{0,1\}.$$

Applying Pinsker's inequality to the equal-weight Jensen–Shannon divergence (which equals $I(V; A \mid U = u)$ for a fair bit V) yields

$$\begin{aligned} I(V; A \mid U) &\geq \tfrac{1}{4}[(q_{01} - q_{00})^2 + (q_{11} - q_{10})^2], \\ I(U; B \mid V) &\geq \tfrac{1}{4}[(r_{10} - r_{00})^2 + (r_{11} - r_{01})^2]. \end{aligned}$$

Combining the boundary and agreement constraints gives

$$\begin{aligned} (q_{01} - q_{00}) + (r_{11} - r_{01}) &= (r_{11} - q_{00}) - (r_{01} - q_{01}) \geq 1 - 3\varepsilon, \\ (r_{10} - r_{00}) + (q_{11} - q_{10}) &= (q_{11} - r_{00}) - (q_{10} - r_{10}) \geq 1 - 3\varepsilon. \end{aligned}$$

Applying the elementary inequality $x^2 + y^2 \geq (x+y)^2/2$ to each pair and summing establishes (3.8). □

Theorem 3.5 (Payload and finite-use converse)**.** *Every ε-correct protocol requires $M > 0$ and satisfies*

$$\mathcal{E}^{\mathrm{av}}_{\star,L} \geq \frac{MN_0}{2}\left[\exp\left(\frac{2sc_\varepsilon L}{M}\right) - 1\right] \geq N_0 sc_\varepsilon L. \tag{3.9}$$

Proof. Partition the n processes into two subsets of sizes s and $n - s$. Assign a uniform random vector $U \in \{0,1\}^L$ to all processes in the first subset and an independent uniform vector $V \in \{0,1\}^L$ to all processes in the second. Form s disjoint pairs of processes across the two subsets. For any process i in the first subset, the independence of the coordinates of V and the subadditivity of conditional entropy give the direct-sum inequality

$$\begin{aligned} I(V; O_i \mid U) &= H(V \mid U) - H(V \mid O_i, U) \\ &\geq \sum_{\ell=1}^{L} \left[H(V_\ell \mid U_\ell) - H(V_\ell \mid O_{i\ell}, U_\ell)\right] \\ &= \sum_{\ell=1}^{L} I(V_\ell; O_{i\ell} \mid U_\ell), \end{aligned} \tag{3.10}$$

and symmetrically $I(U; O_j \mid V) \geq \sum_{\ell=1}^{L} I(U_\ell; O_{j\ell} \mid V_\ell)$ for its partner j in the second subset. Because coordinatewise validity and agreement hold on every realization of the remaining $L - 1$ coordinates, averaging over those coordinates preserves the hypotheses of Lemma 3.4. Applying that lemma to the resulting marginal output law gives $I(V_\ell; O_{i\ell} \mid U_\ell) + I(U_\ell; O_{j\ell} \mid V_\ell) \geq c_\varepsilon$ for each ℓ. Summing over $\ell = 1, \ldots, L$ and over all s disjoint pairs yields at least $sc_\varepsilon L$ nats of total conditional mutual information.

To upper-bound this information via channel capacity, let R be the complete input-independent initial random seed. By the independence of R from (U, V) and the data-processing inequality,

$$I(V; O_i \mid U) \leq I(V; O_i \mid U, R) \leq I(V; \text{receptions of } i \mid U, R).$$

First condition on R, so that the schedule and hence each receiver's list of receptions are fixed. Expanding by the causal chain rule over this list, the fresh noise at each reception is independent of the inputs and prior history. Averaging over the latter history and applying the Gaussian capacity bound and Jensen's inequality bounds the increment at use t by $\frac{1}{2}\ln(1 + 2e_t(R)/N_0)$, where $e_t(R) = \mathbb{E}[A_t^2 \mid R]$. For each fixed R, every use has a single receiver and appears in at most one reception history across the s disjoint pairs. Summing first at fixed R, then averaging over R and applying Jensen's inequality again, gives

$$\begin{aligned} sc_\varepsilon L &\leq \frac{1}{2}\mathbb{E}_R \sum_{t=1}^{M} \ln\left(1 + \frac{2e_t(R)}{N_0}\right) \\ &\leq \frac{M}{2}\ln\left(1 + \frac{2\sum_{t=1}^{M} e_t}{MN_0}\right), \qquad e_t = \mathbb{E}A_t^2. \end{aligned}$$

Solving for $\sum_{t=1}^{M} e_t$ yields the first inequality in (3.9), and the linear lower bound $e^x - 1 \geq x$ gives the second. □

Corollary 3.6. *For all $L \geq 1$,*

$$\mathcal{E}^{\mathrm{av}}_{\star,L} \geq \max\left\{\frac{N_0 s}{16} d_{\mathrm{Ber}}(q_\varepsilon \| \varepsilon/s),\ \frac{MN_0}{2}\left(e^{2sc_\varepsilon L/M} - 1\right)\right\} = \Omega(N_0 n[L + \ln(n/\varepsilon)])\,. \tag{3.11}$$

Although Theorems 3.2 and 3.5 evaluate the protocol on different adversarial input distributions, the worst-case energy $\mathcal{E}^{\mathrm{av}}_{\star,L}$ must exceed both lower bounds simultaneously, and $\max\{a, b\} \geq (a+b)/2$ yields the additive scaling in (3.11).

4 Matching signal-energy upper bounds

4.1 A finite code and quantizer

We first construct a finite block code and a quantized decoder that operate entirely with finite-precision integer arithmetic.

Lemma 4.1 (Block transmission)**.** *For every integer $b \geq 1$, there exists a binary linear code of dimension b, blocklength $k = 8b$, and minimum relative Hamming distance at least $1/4$, whose generator matrix requires $8b^2$ bits of storage. Using this code with a finite uniform quantizer and exhaustive maximum-correlation decoding of the quantized observations over (2.2), a b-bit packet can be transmitted with decoding error probability at most $\delta \in (0, 1)$ and deterministic signal energy*

$$w \leq \frac{N_0}{c_*}[b \ln 2 + \ln(2/\delta)], \qquad c_* = 9/64. \tag{4.1}$$

Proof. Draw a $b \times 8b$ binary generator matrix uniformly at random. For any nonzero message $u \in \{0,1\}^b \setminus \{0\}$, the codeword weight is distributed as $\mathrm{Bin}(8b, 1/2)$. By Chernoff's bound and the union bound, the probability that some nonzero message has Hamming weight strictly less than $2b$ is at most

$$(2^b - 1)e^{-8bd_{\mathrm{Ber}}(1/4\|1/2)} < 1,$$

since $8d_{\mathrm{Ber}}(1/4\|1/2) > 1.04 > \ln 2$. Thus, there exists a deterministic full-rank generator matrix with minimum distance at least $2b$, which we fix as an input-independent protocol parameter.

Map each codeword for message u via antipodal modulation to $c(u) \in \{\pm\sqrt{a}\}^k$ with total block energy $w = ka$. Any two distinct messages $u \neq v$ satisfy $\|c(u) - c(v)\|^2 \geq 4 \cdot (2b)a = w$. At the receiver, set $\eta = 1/4$ and apply a uniform mid-tread quantizer with step size $\Delta = 2\eta\sqrt{a}$ and integer saturation threshold $J = \lceil R/\Delta \rceil$, where

$$R = \sqrt{a} + \sqrt{N_0 \ln(4k/\delta)}.$$

By the Gaussian tail bound $Q(x) \leq \frac{1}{2}e^{-x^2/2}$ and a union bound over the k symbols, the probability that any honest reception exceeds the saturation radius R is at most $\delta/2$. On the event that no sample saturates, each quantization error has magnitude at most $\eta\sqrt{a}$. Consequently, in the pairwise correlation comparison between the transmitted message u and a competitor $v \neq u$, quantization can shift the inner product $\langle \tilde{Y}, c(u) - c(v)\rangle$ by at most $\eta\|c(u) - c(v)\|_1\sqrt{a} = (\eta/2)\|c(u) - c(v)\|^2$. The unquantized correlation score has mean $\|c(u) - c(v)\|^2/2$ and noise variance $(N_0/2)\|c(u) - c(v)\|^2$. For v to defeat u, the Gaussian

noise projection must therefore be at most $-(1-\eta)\|c(u)-c(v)\|^2/2$. The Gaussian tail bound gives

$$\mathbb{P}(v \text{ defeats } u,\ \text{no saturation}) \le \exp\left(-\frac{(1-\eta)^2 w}{4N_0}\right) = e^{-c_* w/N_0}. \tag{4.2}$$

Taking a union bound over the $2^b - 1$ competing messages and setting the total error to $\delta/2 + \delta/2 = \delta$ yields (4.1). Because all codewords have equal energy and symbols in $\{\pm\sqrt{a}\}$, rescaling by $2\Delta\sqrt{a}$ reduces every pairwise correlation difference to an integer sum of stored quantizer indices with coefficients in $\{-1, 0, 1\}$. Exhaustive decoding therefore requires storing only the $8b^2$-bit generator matrix, the k quantized observations, a single candidate index, and finite integer accumulators. □

4.2 Fault-free attainability for every error level

In the absence of faults ($f = 0$), a direct coordinator broadcast matches our converse across the entire error range without shared randomness or interaction.

Theorem 4.2 (Fixed leader)**.** *For $f = 0$, $n \ge 2$, $L \ge 1$, and every $0 < \varepsilon \le 1/8$, there exists a deterministic protocol requiring no public randomness that achieves the pathwise energy bound*

$$E_{\text{lead}} \le \frac{N_0(n-1)}{c_*}\left[L \ln 2 + \ln \frac{2(n-1)}{\varepsilon}\right] \tag{4.3}$$

using $M_0 = 8L(n-1)$ scalar channel uses. With $\zeta_0 = O(\ln(2 + nL/\varepsilon))$, the protocol operates within per-process storage B_0 and total network bit operations G_0 satisfying

$$B_0 = O(L^2 + L\zeta_0 + \ln n), \qquad G_0 = O\Big(n 2^L (L+1)^2 \zeta_0\Big). \tag{4.4}$$

In particular, any deadline $\tau \ge M_0/\nu + G_0 t_{\text{bit}}$ suffices when each bit operation takes time at most t_{bit}.

Proof. Designate process 1 as the fixed leader. Process 1 outputs its own input $O_1 = X_1$ and transmits X_1 to each of the remaining $n-1$ processes via Lemma 4.1 with $b = L$ and per-packet error budget $\delta = \varepsilon/(n-1)$. By the union bound, all $n-1$ receivers decode X_1 correctly with probability at least $1-\varepsilon$. Whenever all honest inputs agree at coordinate ℓ, $X_{1\ell}$ equals that unanimous value, ensuring both agreement and coordinatewise validity. Summing the $n-1$ packet energies from (4.1) yields (4.3). Each receiver stores the $8L^2$-bit generator matrix, $8L$ quantized observations of bit-width $O(\zeta_0)$, and running accumulator state, while iterating over 2^L candidates, which yields (4.4). □

Combined with Corollary 3.6, Theorem 4.2 resolves the fault-free energy complexity for arbitrary $\varepsilon \in (0, 1/8]$ and shows that neither interactive feedback nor shared randomness can improve the asymptotic energy scaling.

4.3 Committee agreement

Now suppose $f/n \le 1/3 - \kappa$ for a fixed margin $\kappa \in (0, 1/3)$, and assume the fault set is fixed before a public committee of size m is sampled uniformly without replacement from

$\{1, \ldots, n\}$, where

$$m = \min\left\{n, \max\left\{4, \left\lceil \frac{\ln(3/\varepsilon)}{2\kappa^2} \right\rceil\right\}\right\}, \qquad t = \left\lfloor \frac{m-1}{3} \right\rfloor. \tag{4.5}$$

When $m < n$, Hoeffding's inequality for sampling without replacement [10] bounds the probability that the sampled committee contains more than t faulty processes by

$$\mathbb{P}(\text{committee fault count} > t) \leq e^{-2m\kappa^2} \leq \varepsilon/3,$$

whereas for $m = n$ the condition $f \leq t$ holds deterministically.

The committee executes the classical Phase King protocol [4] across $t+1$ sequential phases, operating on all L coordinates in parallel. Each phase $p \in \{1, \ldots, t+1\}$ designates a distinct committee member as king and consists of three broadcast rounds within the committee:

1. In Round 1, each member broadcasts its current binary state vector. In each coordinate, a member sets its proposal to $b \in \{0, 1\}$ if it receives at least $m - t$ copies of b (counting its own state locally), and to $\perp$ otherwise.

2. In Round 2, each member broadcasts its proposal vector in $\{0, 1, \perp\}^L$. For each coordinate, a member assigns grade 2 to a bit b supported by at least $m - t$ proposals, grade 1 to a bit supported by at least $t + 1$ proposals, and grade 0 otherwise.

3. In Round 3, the phase king broadcasts its candidate vector, using its graded bit in coordinates with grade 1 or 2 and a fixed default bit in coordinates with grade 0. In each coordinate, a member with grade 2 retains its own graded bit, whereas a member with grade 1 or 0 adopts the king's bit (using the same fixed default if the king sends an invalid symbol).

For a complete protocol on every channel history, ties between grade-1 bits are resolved by a fixed deterministic rule, and the unused two-bit encoding of a ternary proposal is interpreted as $\perp$. Each committee member outputs its own final state after the last phase.

To verify correctness on the event that the committee contains at most $h \leq t$ faulty members and all honest-to-honest committee packets are decoded correctly, note that proposing b in Round 1 requires at least $m - t - h$ honest Round-1 messages equal to b. Since $2(m - t - h) > m - h$ for $m > 3t$, two honest members cannot propose opposite bits 0 and 1 in the same coordinate. Consequently, if any honest member assigns grade 2 to b in Round 2 (requiring at least $m - t - h$ honest proposals for b), then every honest member receives at least $m - t - h \geq t + 1$ proposals for b and assigns at least grade 1 to the unique bit b. Thus, in any phase led by an honest king, all honest members converge to the same bit in Round 3, and any coordinate on which honest members already agree at the start of a phase remains unanimous thereafter. Because $t + 1 > h$, at least one phase has an honest king, guaranteeing both agreement and coordinatewise validity across the committee.

Across the $t + 1$ phases, the committee exchanges a total of

$$P_C = (t + 1)(m - 1)(2m + 1) \tag{4.6}$$

directed packets. Encoding each packet (including the ternary Round-2 proposals) into $b = 2L$

bits via Lemma 4.1 with per-packet error budget $\delta = \varepsilon/(3P_C)$, the total internal committee signal energy satisfies

$$E_C \le \frac{N_0 P_C}{c_*}[2L\ln 2 + \ln(6P_C/\varepsilon)] = O\Big(N_0 m^3[L + \ln(m/\varepsilon)]\Big). \tag{4.7}$$

With probability at least $1 - 2\varepsilon/3$, the committee contains at most t faults and every honest-to-honest internal packet is decoded correctly, leaving all honest committee members with a common decision vector $D \in \{0,1\}^L$.

4.4 Robust joint reception

When $v = n - m > 0$, the committee must disseminate D to the v external processes without allowing faulty committee members to corrupt the outcome or inflating the payload energy by a factor of m. To achieve this, each committee member transmits a single length-$8L$ codeword (from Lemma 4.1 with $b = L$) to each of the v external processes. Each honest committee member transmits the codeword of its own final state; on the good internal event all these codewords equal $c(D)$. Faulty members may transmit arbitrary waveforms.

Each external receiver clips and quantizes its m incoming length-$8L$ vectors and performs a tournament over the 2^L candidate messages: for every ordered pair of candidates (u, u'), each of the m received branches casts a binary vote based on which candidate achieves the higher quantized correlation score (breaking ties deterministically via a fixed lexicographic rule so that votes are strictly antisymmetric). The receiver outputs the unique candidate u that wins a strict majority of at least $\lfloor m/2 \rfloor + 1$ votes against every competitor $u' \neq u$, or a fixed default if no such Condorcet winner exists.

Set

$$r = \lceil m/2 \rceil - t \ge m/6. \tag{4.8}$$

When the committee contains at most t faulty members, there are at least $m - t$ honest branches. For a competitor $u' \neq D$ to deny D a strict majority, D can receive at most $\lfloor m/2 \rfloor$ votes in the pairwise comparison against u', which implies that at least $(m - t) - \lfloor m/2 \rfloor = r$ honest branches must vote in favor of u' over D. We set the per-branch codeword energy to

$$w = \frac{N_0}{c_* r}\left[L\ln 2 + \ln(6v/\varepsilon) + \ln\binom{m}{r}\right]. \tag{4.9}$$

Condition on the committee selection and a good completed internal execution, so the honest decision D is fixed and the external channel noises are fresh and mutually independent. For any fixed competitor $u' \neq D$ and any fixed set of r honest branches, the joint event that these branches all favor u' and none of their samples saturates implies the intersection of the r adverse Gaussian projection events used to prove (4.2). These projection events depend on disjoint fresh noise blocks, so their joint probability is at most $e^{-rc_* w/N_0}$. A union bound over $2^L - 1$ competitors, at most $\binom{m}{r}$ sets of honest branches, and all v external receivers therefore bounds the probability of a majority-tournament failure without honest-sample saturation by $\varepsilon/6$. This argument bounds an intersection event without conditioning on the absence of saturation, and it requires no independence or power restriction on the Byzantine waveforms.

With $k = 8L$ and per-symbol energy $a = w/k$, choosing the quantization saturation radius

$$R_D = \sqrt{a} + \sqrt{N_0 \ln(12kmv/\varepsilon)}, \qquad J_D = \lceil R_D/(2\eta\sqrt{a}) \rceil ,$$

ensures that saturation of any honest external sample across all v receivers occurs with probability at most $\varepsilon/6$. Consequently, the external dissemination phase fails with probability at most $\varepsilon/3$ and expends a total honest signal energy of at most

$$E_D \le mvw \le \frac{6N_0 v}{c_*}[L \ln 2 + \ln(6v/\varepsilon) + m \ln 2], \tag{4.10}$$

where we used $m/r \le 6$ and $\binom{m}{r} \le 2^m$. Because the external receiver combines evidence across all m branches in a joint pairwise tournament rather than requiring each individual branch to be decoded reliably in isolation, the factor of $1/r$ in (4.9) cancels the committee size m in the total energy mvw.

4.5 Resources and the optimum

Combining the $\varepsilon/3$ failure budgets for committee sampling, internal Phase King decoding, and external joint reception guarantees ε-correctness (Definition 2.1) with a deterministic pathwise honest energy bound of

$$E_{\mathrm{up}} = O\Big(N_0\{m^3[L + \ln(m/\varepsilon)] + n[L + \ln(n/\varepsilon) + m]\}\Big) . \tag{4.11}$$

The total number of scalar channel uses M_{up} and per-process storage B_{up} satisfy

$$M_{\mathrm{up}} = 16LP_C + 8Lmv, \tag{4.12}$$

$$B_{\mathrm{up}} = O\Big(L^2 + mL \ln(2 + nL/\varepsilon) + m \ln n\Big) , \tag{4.13}$$

which accounts for the generator matrices, input vector, committee roster, quantized observation buffers, and decoding accumulators. With word length $\zeta = O(\ln(2 + nLm/\varepsilon))$ for the integer correlation scores, the total number of bit operations across the network is bounded by

$$G_{\mathrm{up}} = O\Big((P_C + nm)4^L(L+1)^2\zeta\Big) , \tag{4.14}$$

so any deadline $\tau \ge M_{\mathrm{up}}/\nu + G_{\mathrm{up}}t_{\mathrm{bit}}$ suffices for a serial implementation.

Theorem 4.3 (Optimal signal-energy order)**.** *Fix $\kappa \in (0, 1/3)$ and $a > 0$. Suppose $0 \le f \le (1/3 - \kappa)n$ and $n^{-a} \le \varepsilon \le 1/8$, with the fault set fixed prior to independent public committee selection. If the prescribed storage, channel-use, and deadline budgets accommodate* (4.12)–(4.14)*, then for all sufficiently large n,*

$$\mathcal{E}^{\mathrm{av}}_{\star,L} = \Theta_{\kappa,a}(N_0 n[L + \ln(n/\varepsilon)]) , \qquad \mathcal{E}^{\mathrm{hard}}_{\star,L} = \Theta_{\kappa,a}(N_0 n[L + \ln(n/\varepsilon)]) , \tag{4.15}$$

with implied constants uniform over all $L \ge 1$.

Proof. Corollary 3.6 provides the lower bound. For $n^{-a} \le \varepsilon \le 1/8$, (4.5) gives $m = O_{\kappa,a}(\ln n)$. Consequently, the internal committee energy in (4.11) is $O_{\kappa,a}(N_0[L \ln^3 n + \ln^4 n]) = o(N_0 n[L +$

$\ln(n/\varepsilon)]$) uniformly in $L \geq 1$, while the external dissemination energy is $O_{\kappa,a}(N_0 n[L + \ln(n/\varepsilon)])$. Since the upper bound holds pathwise, it establishes both $\mathcal{E}^{\mathrm{av}}_{\star,L}$ and $\mathcal{E}^{\mathrm{hard}}_{\star,L}$. □

As an immediate consequence, the amortized signal energy per coordinate in a batch of size L is

$$\frac{\mathcal{E}_{\star,L}}{L} = \Theta_{\kappa,a}\left(N_0 n\left[1 + \frac{\ln(n/\varepsilon)}{L}\right]\right). \tag{4.16}$$

Choosing a moderate batch size $L = \Theta(\ln(n/\varepsilon))$ eliminates the logarithmic per-coordinate reliability overhead while keeping the exhaustive decoding time $2^{O(L)} = \mathrm{poly}(n/\varepsilon)$ polynomial and the local storage polylogarithmic in n.

5 Thermal interfaces and energy recovery

5.1 Thermodynamic boundary conditions

We now embed the Gaussian channel model into a self-contained stochastic-thermodynamic framework. Let W denote the mean work supplied to the composite system, Q_{env} the mean heat transferred to a thermal bath at temperature T, and Σ the total entropy production in units of Boltzmann's constant k_{B}. By the first and second laws of thermodynamics for a fully accounted system state,

$$W = \Delta U + Q_{\mathrm{env}} = \Delta \mathcal{F} + k_{\mathrm{B}} T \Sigma, \qquad \mathcal{F} = U - k_{\mathrm{B}} T H, \tag{5.1}$$

where $\mathcal{F}$, U, and H denote the non-equilibrium free energy, mean internal energy, and differential or Shannon entropy relative to a fixed phase-space reference measure. All bath modes are initialized in independent canonical Gibbs states, and any mode returned to the bath is treated as inaccessible for subsequent work extraction or information storage.

To model a single scalar channel use physically, consider a classical harmonic mode $Q = (x, y) \in \mathbb{R}^2$ with Hamiltonian $H_Q(x, y) = x^2 + y^2$ initialized in thermal equilibrium at temperature T, with density

$$\pi(x, y) = \frac{1}{\pi k_{\mathrm{B}} T} \exp\left(-\frac{x^2 + y^2}{k_{\mathrm{B}} T}\right) \tag{5.2}$$

independent of the device state. Conditioned on a retained digital register K that specifies the transmitted amplitude $A = A(K)$, an isolated, entropy-preserving displacement $x \mapsto x + A$ requires mean mechanical work $\mathbb{E}A^2$ and prepares the quadrature $x = A + Z$ with thermal noise $Z \sim N(0, k_{\mathrm{B}} T/2)$, matching the channel law (2.2) with noise spectral density $N_0 = k_{\mathrm{B}} T$.

Next, the receiver performs a finite-record measurement that updates its local digital state while leaving the conditional mode marginal $p_{Q|K}$ undisturbed. Let R denote the complete retained digital state of the system (including K and the newly acquired measurement record). To prepare the port for the next transmission, the displaced mode Q is swapped out into the bath and replaced by a fresh equilibrium mode drawn from π, while R is held fixed. Under a non-recovering interface, this replacement occurs directly without prior work extraction from Q. We analyze the fundamental entropy production of this port reset primitive, which lower-bounds any physical implementation realizing these operations.

Theorem 5.1 (Non-recovering reset identity)**.** *Under the non-recovering port replacement primitive, the entropy production of a single mode reset is*

$$\Sigma_{\mathrm{reset}} = D(p_{RQ}\|p_R\pi) = \frac{\mathbb{E}A^2}{k_{\mathrm{B}}T} + I(Q;R \mid K). \tag{5.3}$$

In particular, $\Sigma_{\mathrm{reset}} \geq \mathbb{E}A^2/(k_{\mathrm{B}}T)$.

Proof. Before replacement, the retained memory and the mode have joint distribution p_{RQ}; replacement by an independent equilibrium mode changes it to $p_R\pi$. Since the memory marginal p_R is unchanged and the mean heat absorbed by the bath is $Q_{\mathrm{env}} = U_Q - U_\pi$, the entropy production is

$$\Sigma_{\mathrm{reset}} = H(R) + H(\pi) - H(R,Q) + \frac{U_Q - U_\pi}{k_{\mathrm{B}}T} = D(p_{RQ}\|p_R\pi).$$

Since K is a deterministic function of R, the chain rule for relative entropy yields

$$D(p_{RQ}\|p_R\pi) = D(p_{Q|K}\|\pi \mid p_K) + D(p_{R|Q,K}\|p_{R|K} \mid p_{QK}) = \mathbb{E}_K D(p_{Q|K}\|\pi) + I(Q;R \mid K).$$

Evaluating the Gaussian relative entropy $D(p_{Q|K}\|\pi) = A(K)^2/(k_{\mathrm{B}}T)$ completes the proof. □

Physically, while an idealized energy-conserving swap between the port mode and a bath mode preserves the fine-grained joint Gibbs entropy of the universe, discarding the outgoing mode into an inaccessible bath dissipates both its nonequilibrium displacement free energy $\mathbb{E}A^2$ and the measurement correlation $k_{\mathrm{B}}T\, I(Q;R \mid K)$ [9].

5.2 Recovery with specified information access

Suppose now that prior to discarding the mode Q, a recovery controller is permitted to apply a feedback displacement to extract work, subject to both a speed constraint and an informational constraint [12]. Specifically, let $\mathscr{G} \subseteq \sigma(R)$ be the sub-σ-field of the retained state R accessible to the controller, and write $\mu_{\mathscr{G}} = \mathbb{E}[x \mid \mathscr{G}]$. Over a recovery interval of duration t_r with maximum drift speed $v_{\max}$, the controller may apply an entropy-preserving translation $x \mapsto x + d$, where d is $\mathscr{G}$-measurable and satisfies $|d| \leq d_{\max} = v_{\max}t_r$, while R remains frozen.

Theorem 5.2 (Information-constrained displacement recovery)**.** *Suppose the non-recovering reset entropy production* Σ_0 *and* $\mathbb{E}x^2$ *are finite. The minimum entropy production of the port replacement following an admissible* $\mathscr{G}$*-measurable recovery displacement is*

$$\Sigma_{\min}(\mathscr{G}, d_{\max}) = \Sigma_0 - \frac{1}{k_{\mathrm{B}}T}\mathbb{E}\Big[\mu_{\mathscr{G}}^2 - (|\mu_{\mathscr{G}}| - d_{\max})_+^2\Big], \tag{5.4}$$

which is uniquely attained, up to almost-sure equality, by the clipped estimator $d^* = -\operatorname{clip}(\mu_{\mathscr{G}}, -d_{\max}, d_{\max})$. *Furthermore, if* $\mathscr{G}_1 \subseteq \mathscr{G}_2 \subseteq \sigma(R)$, *then* $\Sigma_{\min}(\mathscr{G}_2, d_{\max}) \leq \Sigma_{\min}(\mathscr{G}_1, d_{\max})$.

Proof. Because d is $\mathscr{G}$-measurable and $\mathscr{G} \subseteq \sigma(R)$, the conditional translation $(x,y) \mapsto (x+d, y)$ is invertible given R and preserves both $H(Q \mid R)$ and $H(R)$. It alters the mean mode energy

$\mathbb{E}(x^2+y^2)$ by $\mathbb{E}[(x+d)^2-x^2] = \mathbb{E}(2dx+d^2) = \mathbb{E}(2d\mu_{\mathscr{G}}+d^2)$. The post-recovery replacement entropy production is therefore

$$\Sigma(d) = \Sigma_0 + \frac{1}{k_{\mathrm{B}}T}\mathbb{E}(2d\mu_{\mathscr{G}}+d^2) = \Sigma_0 + \frac{1}{k_{\mathrm{B}}T}\mathbb{E}\Big[(d+\mu_{\mathscr{G}})^2 - \mu_{\mathscr{G}}^2\Big].$$

For almost every realization, the quadratic $(d+\mu_{\mathscr{G}})^2$ subject to $|d| \le d_{\max}$ is minimized by projecting $-\mu_{\mathscr{G}}$ onto $[-d_{\max}, d_{\max}]$, which gives the $\mathscr{G}$-measurable minimizer $d^* = -\operatorname{clip}(\mu_{\mathscr{G}}, -d_{\max}, d_{\max})$ and the minimum value (5.4). Monotonicity in $\mathscr{G}$ follows immediately from the inclusion of admissible control sets (or equivalently from Jensen's inequality applied to the convex function $u \mapsto u^2 - (|u|-d_{\max})_+^2$). □

To illustrate how control access shapes dissipation without altering the communication channel law, consider binary antipodal signaling $A \in \{\pm a\}$ where the receiver records a 1-bit sign threshold $B = \mathbf{1}_{\{x\ge 0\}}$ and the retained state is $R = (K, B)$. Setting

$$u = \frac{a^2}{k_{\mathrm{B}}T}, \qquad p(u) = \tfrac{1}{2}\operatorname{erfc}\sqrt{u}, \tag{5.5}$$

we have $I(Q; R \mid K) = H(B \mid K) = h(p(u))$ and $\Sigma_0 = u + h(p(u))$. If the recovery controller can access only the transmitter's amplitude label ($\mathscr{G} = \sigma(K)$), then $\mu_{\mathscr{G}} = A$ and (5.4) reduces to

$$\Sigma_{\min}(\sigma(K), d_{\max}) = \frac{(|a|-d_{\max})_+^2}{k_{\mathrm{B}}T} + h(p(u)). \tag{5.6}$$

Even as $d_{\max} \to \infty$, an amplitude-only controller can recover at most the injected signal energy a^2, leaving the full measurement correlation $h(p(u))$ to be dissipated.

By contrast, suppose the controller also has access to the receiver's sign bit, $\mathscr{G} = \sigma(K, B)$. Even at zero signal amplitude ($a = 0$, where K is trivial and B is a fair bit), the conditional mean of the thermal fluctuation with variance $\sigma^2 = k_{\mathrm{B}}T/2$ is $\mathbb{E}[x \mid B] = (2B-1)\sqrt{k_{\mathrm{B}}T/\pi}$. For $d_{\max} \ge \sqrt{k_{\mathrm{B}}T/\pi}$, Theorem 5.2 yields a strict separation between the two control regimes:

$$\Sigma_{\min}(\sigma(K), d_{\max}) = \ln 2, \qquad \Sigma_{\min}(\sigma(K, B), d_{\max}) = \ln 2 - \frac{1}{\pi}. \tag{5.7}$$

Here, conditioning the recovery shift on the measured sign B extracts $k_{\mathrm{B}}T/\pi$ of thermal work from the correlated mode prior to reset (while the cost of subsequently erasing the digital bit B is accounted for separately in the memory cycle). More generally, condition on an amplitude label K with $A(K) = a \ge 0$, and let $z = a/\sigma$, $q = \Phi(z)$, and $p = 1-q$, where φ and Φ are the standard normal density and cumulative distribution function. The conditional means of the truncated Gaussian are

$$\mu_+ = \mathbb{E}[x \mid K, B=1] = a + \sigma\frac{\varphi(z)}{q}, \qquad \mu_- = \mathbb{E}[x \mid K, B=0] = a - \sigma\frac{\varphi(z)}{p}. \tag{5.8}$$

The negative-amplitude branch follows by reflecting $x \mapsto -x$ and exchanging the two sign events. Since $q\mu_+^2 + p\mu_-^2 = a^2 + \sigma^2\varphi(z)^2/(pq)$, whenever $d_{\max} \ge \max\{|\mu_+|, |\mu_-|\}$, sample-controlled recovery achieves the residual entropy production

$$\Sigma_{\min}(\sigma(K, B), d_{\max}) = h(p) - \frac{\varphi(z)^2}{2pq}. \tag{5.9}$$

5.3 A sampling cost distinct from signal energy

When a receiver retains an estimate of each observation's sign through its non-recovering mode reset, with bit-error probability at most a fixed $\delta_{\mathrm{adc}} < 1/2$, the measurement correlation term $I(Q; R \mid K)$ in Theorem 5.1 imposes a per-use dissipation floor that does not vanish at low signal-to-noise ratio.

Proposition 5.3 (Obligatory sign sampling)**.** *Suppose that at a scheduled channel use, the receiver stores an estimate of* $\operatorname{sgn}(x)$ *in* R *with conditional bit-error probability at most* $\delta_{\mathrm{adc}} < 1/2$*, retains the amplitude label* K*, and performs a non-recovering mode reset. Then*

$$\Sigma_{\mathrm{reset}} \geq \ln 2 - h(\delta_{\mathrm{adc}}), \tag{5.10}$$

even when the transmitted amplitude is $A = 0$*.*

Proof. Fix $|A| = a$ and assign an auxiliary fair random sign to $A = \pm a$. By the data-processing inequality and the Gaussian relative-entropy bound on mutual information,

$$\ln 2 - h(p(u)) = I(A; \operatorname{sgn} x) \leq I(A; x) \leq \mathbb{E}_A D(p_{x|A} \| p_{x|0}) = u,$$

which implies $u + h(p(u)) \geq \ln 2$ for all $u \geq 0$. Conditioning on any realization of K with $|A(K)| = a$, Fano's inequality for the stored sign estimate gives $I(Q; R \mid K) \geq H(\operatorname{sgn} x \mid K) - h(\delta_{\mathrm{adc}}) = h(p(u)) - h(\delta_{\mathrm{adc}})$. Substituting this into (5.3) yields $\Sigma_{\mathrm{reset}} \geq \mathbb{E}_K[u(K) + h(p(u(K)))] - h(\delta_{\mathrm{adc}}) \geq \ln 2 - h(\delta_{\mathrm{adc}})$. □

Proposition 5.3 highlights an important physical distinction: when the sign estimates are retained through their respective non-recovering resets, M channel uses dissipate free energy of at least $M[\ln 2 - h(\delta_{\mathrm{adc}})]\, k_{\mathrm{B}} T$, regardless of how small the total signal energy $\sum_t A_t^2$ is. Because the multi-round committee protocol of Section 4 spreads its energy across M_{up} low-SNR uses, minimizing total signal energy and minimizing total port reset dissipation are distinct physical objectives unless analog accumulation or pre-reset recovery is employed.

6 Cyclic work, storage, and finite rates

6.1 Data transitions and a common input distribution

We now combine the port reset cost with the thermodynamic cost of updating physical data registers across a cyclic sequence of consensus executions. Suppose the n processes store their inputs in nL designated physical data bits, and the protocol overwrites these same registers with the final decisions $O \in \{0,1\}^{nL}$. The data registers evolve via continuous-time Markov jump processes satisfying local detailed balance, where each transition flips at most one of the nL data bits and, in every system microstate, the aggregate intensity of transitions flipping any given data bit is at most γ. Port replacements and data-bit flips occur as disjoint elementary events, and all remaining internal operations consist of entropy-preserving Hamiltonian evolutions or locally detailed-balanced thermal relaxations with nonnegative entropy production.

To track the physical distance traversed by the data registers, define the coordinatewise minority weight of a configuration $x \in \{0,1\}^{nL}$ and its uniform binomial expectation as

$$g(x) = \sum_{\ell=1}^{L} \min\{|x_{\cdot\ell}|, n - |x_{\cdot\ell}|\}, \qquad d_n = \frac{n}{2} - \frac{n}{2^n}\binom{n-1}{\lfloor (n-1)/2 \rfloor}, \tag{6.1}$$

where $|x_{\cdot\ell}| = \sum_{i=1}^{n} x_{i\ell}$ and $d_n = \mathbb{E}[\min\{J, n-J\}]$ for $J \sim \mathrm{Bin}(n, 1/2)$.

To evaluate all thermodynamic contributions simultaneously under a single input law, we introduce the protocol-independent mixture distribution on $\{0,1\}^{nL}$ given by

$$\mu = \tfrac{1}{2}\,\mathrm{Unif}(\{0,1\}^{nL}) + \tfrac{1}{4}\mu_{\mathrm{rel}} + \tfrac{1}{4}\mu_{\mathrm{info}}. \tag{6.2}$$

Here, μ_{rel} places equal mass $1/3$ on the all-zero array, the all-one array, and a fixed balanced array $z \otimes 1_L$ (where each node's bit in z is replicated across all L coordinates), while μ_{info} assigns independent uniform L-bit vectors U and V to the two node populations of sizes s and $n - s$ from Theorem 3.5. Associated with μ, we define the communication, Landauer erasure, and finite-rate transport constants

$$\mathcal{A} = \frac{s}{96} d_{\mathrm{Ber}}(q_\varepsilon \| \varepsilon/s) + \frac{s c_\varepsilon L}{4}, \tag{6.3}$$

$$\mathcal{J} = (1/2 - \varepsilon)(n-1)L \ln 2 + \tfrac{1}{4}\ln 3 - h(\varepsilon), \tag{6.4}$$

$$\ell = L\left[\frac{d_n}{2} + (5/24 - \varepsilon)s\right], \qquad \mathcal{V} = 2\ell \operatorname{artanh}\frac{\ell}{nL\gamma\tau}, \tag{6.5}$$

where $\mathcal{V}$ is defined whenever $0 < \ell < nL\gamma\tau$.

6.2 The cyclic lower bound

Consider a device that processes K sequential, independent consensus batches $X^{(1)}, \ldots, X^{(K)}$ drawn i.i.d. from μ. We evaluate these batches in fault-free executions, even if the protocol is required to tolerate Byzantine faults. A lower bound for these executions also lower-bounds the worst-case cost over admissible adversaries. Each batch executes within duration τ and satisfies the ε-correctness guarantee conditioned on the state retained from prior batches. To enforce a closed thermodynamic cycle without hidden entropy sinks, we specify the boundary conditions as follows:

1. Unread future input batches and archived past output batches reside on inert external data tapes. Loading a new batch into the active nL registers and archiving the completed decision are reversible tape permutations that preserve the sum of g across all tapes; only the active nL registers undergo Markovian bit flips during a batch.

2. Across the K batches, the device may retain at most B_{tot} bits of cumulative input-dependent working memory outside the archived output tapes.

3. All remaining auxiliary degrees of freedom are initially independent of the new inputs and return at the end of the K-batch cycle to their exact initial *joint* marginal distribution. Requiring restoration of the joint auxiliary distribution (rather than merely

single-component marginals) ensures that any consumption of pre-shared correlations is properly charged within the cycle.

Theorem 6.1 (Work and entropy production in the reset model)**.** *Under the above fault-free K-batch cyclic boundary conditions, the non-recovering port interface with $N_0 = k_{\mathrm{B}}T$, and $0 < \ell < nL\gamma\tau$, the cumulative ensemble-averaged entropy production Σ_K and supplied work W_K satisfy*

$$\frac{\Sigma_K}{K} \geq \mathcal{A} + \mathcal{V}, \tag{6.6}$$

$$\frac{W_K - \Delta U_K}{K} \geq k_{\mathrm{B}}T \left[\mathcal{A} + \mathcal{J} + \mathcal{V} - \frac{B_{\mathrm{tot}} \ln 2}{K}\right], \tag{6.7}$$

where ΔU_K is the mean change in total internal energy between the start and end of the K-batch cycle (in particular, $\Delta U_K = 0$ when the data registers have degenerate energy levels and the cyclic auxiliaries have equal endpoint mean energy).

Proof. We bound the port reset entropy production, the data register transition entropy production, and the endpoint Shannon entropy change in turn.

First, for each batch, applying the proof of Theorem 3.2 to the replicated inputs in μ_{rel} gives $E_{z\otimes 1_L} + E_{(1-b)^{nL}} \geq \frac{N_0 s}{8} d_{\mathrm{Ber}}(q_\varepsilon \| \varepsilon/s)$, so the expected signal energy under μ_{rel} is at least $\frac{N_0 s}{24} d_{\mathrm{Ber}}(q_\varepsilon \| \varepsilon/s)$. Under μ_{info}, Theorem 3.5 gives an expected signal energy of at least $N_0 s c_\varepsilon L$. Weighting by the mixture coefficients $1/4$ and $1/4$ in (6.2), each batch requires expected signal energy $\mathbb{E}_\mu E_{\mathrm{tx}} \geq k_{\mathrm{B}}T\mathcal{A}$, which by Theorem 5.1 generates at least $\mathcal{A}$ of port reset entropy production per batch, or $K\mathcal{A}$ over K batches.

Second, we evaluate the net Shannon entropy reduction of the data and memory degrees of freedom. By the concavity of entropy applied to the mixture (6.2), each input batch has entropy

$$H(\mu) \geq \tfrac{1}{2} H(\mathrm{Unif}) + \tfrac{1}{4} H(\mu_{\mathrm{rel}}) + \tfrac{1}{4} H(\mu_{\mathrm{info}}) = \tfrac{1}{2} nL \ln 2 + \tfrac{1}{4} \ln 3 + \tfrac{1}{2} L \ln 2.$$

On the event of agreement (probability at least $1 - \varepsilon$), the output array O takes at most 2^L distinct values; on the complementary failure event (probability at most ε), it takes at most 2^{nL} values. Fano's splitting inequality therefore bounds the single-batch output entropy by

$$H(O) \leq L \ln 2 + h(\varepsilon) + \varepsilon(n-1)L \ln 2,$$

so $H(X) - H(O) \geq \mathcal{J}$ per batch. Because the K input batches are i.i.d., subadditivity of entropy across the archived outputs, the restoration of the joint distribution of the cyclic auxiliaries, and the B_{tot}-bit cap on residual working memory yield the global endpoint entropy drop

$$H_{\mathrm{initial}} - H_{\mathrm{final}} \geq K\mathcal{J} - B_{\mathrm{tot}} \ln 2. \tag{6.8}$$

Third, we lower-bound the entropy production Σ_{data} of the continuous-time Markovian data-bit flips via a thermodynamic speed limit [8, 19]. Under μ, the initial expected minority weight per batch is $\mathbb{E}_\mu g(X) = \frac{1}{2} L d_n + \frac{1}{4}(\frac{1}{3} Ls) + \frac{1}{4}(\frac{1}{2} Ls) = L(d_n/2 + 5s/24)$. Since $g(O) = 0$ whenever all nodes agree and $g(O) \leq Ls$ otherwise, $\mathbb{E}_\mu g(O) \leq \varepsilon Ls$, so each batch decreases the expected value of g by at least $\ell = L[d_n/2 + (5/24 - \varepsilon)s]$. Because each allowed data

transition flips a single bit, it changes g by at most 1. For each undirected transition edge e that flips a data bit, let $a_e(t)$ and $b_e(t)$ denote its forward and reverse probability fluxes, with net current $j_e = a_e - b_e$ and dynamical activity $v_e = a_e + b_e$. Local detailed balance gives the instantaneous edge entropy production rate

$$(a_e - b_e) \ln \frac{a_e}{b_e} = 2|j_e| \operatorname{artanh} \frac{|j_e|}{v_e}.$$

Let $\mathscr{L} = \int_0^{K\tau} \sum_e |j_e(t)|\, dt$ be the total integrated variation and $\mathscr{T} = \int_0^{K\tau} \sum_e v_e(t)\, dt$ the total dynamical activity across all K batches. By the joint convexity of $(x, y) \mapsto 2x \operatorname{artanh}(x/y)$ for $0 \le x < y$, Jensen's inequality gives

$$\Sigma_{\text{data}} \ge 2\mathscr{L} \operatorname{artanh} \frac{\mathscr{L}}{\mathscr{T}}.$$

By the triangle inequality, the total variation $\mathscr{L}$ is at least the net expected change in g across the K batches, so $\mathscr{L} \ge K\ell$. Meanwhile, since each of the nL active data bits has total flip intensity at most γ, the total activity satisfies $\mathscr{T} \le KnL\gamma\tau$. Monotonicity of $2x \operatorname{artanh}(x/y)$ (increasing in x and decreasing in y) yields $\Sigma_{\text{data}} \ge K\mathcal{V}$.

Because port resets and data-bit transitions are disjoint operations, their entropy productions add to give $\Sigma_K \ge K(\mathcal{A} + \mathcal{V})$, proving (6.6). Combining Σ_K with the endpoint entropy reduction (6.8) in the first-law identity $W_K - \Delta U_K = k_{\mathrm{B}}T(H_{\text{initial}} - H_{\text{final}} + \Sigma_K)$ establishes (6.7). □

Notice that Theorem 6.1 accounts for each physical dissipation source without double counting: $\mathcal{A}$ captures the irreversible port resets, $\mathcal{V}$ captures the finite-time Markovian relaxation of the data registers, and $\mathcal{J}$ captures the entropy contribution to the nonequilibrium free-energy change of compressing n distributed inputs into a unanimous decision. For fixed B_{tot}, let $K \to \infty$ through devices satisfying these conditions. If in addition $\Delta U_K = 0$ for each cycle, the per-batch entropy production and supplied work satisfy

$$\begin{aligned} \liminf_{K\to\infty} \frac{\Sigma_K}{K} &= \Omega(n[L + \ln(n/\varepsilon)]), \\ \liminf_{K\to\infty} \frac{W_K}{K} &= \Omega(k_{\mathrm{B}}Tn[L + \ln(n/\varepsilon)]). \end{aligned} \tag{6.9}$$

These bounds hold under the fixed mixture μ, hence also for worst-case costs. Without the equal-energy endpoint condition, the second bound applies to $(W_K - \Delta U_K)/K$. If $\ell > nL\gamma\tau$, the required mean transport is infeasible; at equality it requires infinite entropy production. Thus no finite-production implementation is omitted by the theorem's strict domain condition.

6.3 Deadline and peak-energy obstructions

Finite transition rates and per-use expected-energy constraints also impose hard feasibility thresholds on the deadline τ and schedule length M.

First, suppose $f \ge 1$, and consider an input configuration in which f minority processes hold 1^L and the remaining $n - f$ majority processes hold 0^L. In a fault-free execution on this input, validity in an indistinguishable scenario where the f minority nodes are Byzantine

(but behave honestly) forces the honest majority to output 0^L with probability at least $1-\varepsilon$; combining this with fault-free agreement shows that all n nodes output 0^L with probability at least $1-2\varepsilon$. Consequently, all fL physical data bits of the minority nodes must flip from 1 to 0 by time τ with probability at least $1-2\varepsilon$. Because a Markov jump process with per-bit flip intensity bounded by γ can be coupled to fL independent rate-γ Poisson proposal clocks, all fL clocks must ring at least once by time τ, which occurs with probability at most $(1-e^{-\gamma\tau})^{fL}$. Thus, any ε-correct implementation in this register model requires

$$\gamma\tau \geq -\ln\left[1-(1-2\varepsilon)^{1/(fL)}\right]. \tag{6.10}$$

Second, if each scheduled channel use is subject to a per-use expected-energy constraint $\mathbb{E}A_t^2 \leq e_{\max}$ uniformly over every admissible fixed input and adversarial strategy, and the aggregate network channel-use rate is ν, then $E_\Pi \leq Me_{\max}$ and $M \leq \nu\tau$. Substituting these bounds into the finite-use capacity inequality of Theorem 3.5 and the reliability bound of Theorem 3.2 yields the necessary timing condition

$$\nu\tau \geq M \geq \max\left\{\frac{2sc_\varepsilon L}{\ln(1+2e_{\max}/N_0)}, \frac{N_0 s}{16e_{\max}} d_{\mathrm{Ber}}(q_\varepsilon\|\varepsilon/s)\right\}, \tag{6.11}$$

alongside the trivial register storage requirement $B \geq L$ when full L-bit blocks are buffered locally.

7 Consequences

Our analysis demonstrates that enforcing a joint error budget ε across n processes and L binary coordinates cleanly decouples the signal-energy complexity of noisy consensus into an $\Omega(N_0 nL)$ payload dissemination cost and an $\Omega(N_0 n\ln(n/\varepsilon))$ reliability cost. In the fault-free setting, Theorems 3.2, 3.5, and 4.2 establish this exact order across all $0<\varepsilon\leq 1/8$. Against a static Byzantine adversary corrupting $f \leq (1/3-\kappa)n$ processes, Theorem 4.3 achieves the same optimal scaling in the polynomial-error regime via public committee sampling and majority-tournament joint reception. In both regimes, deciding a batch of $L=\Theta(\ln(n/\varepsilon))$ coordinates simultaneously reduces the amortized signal energy per coordinate from $\Theta(N_0 n\ln(n/\varepsilon))$ to $\Theta(N_0 n)$.

To place these results in the broader context of distributed systems, we emphasize that our formulation addresses coordinatewise binary agreement among a fixed set of n authenticated participants. In full state-machine replication or blockchain protocols, agreeing on a batch of bits does not by itself resolve transaction ordering conflicts or external ledger validity. Moreover, while our static Byzantine upper bound leverages a single public random sample to select a committee, permissionless environments require additional cryptographic or Sybil-resistance mechanisms to establish participant identities and verify resource ownership and use [13, 16], where economic security budgets are distinct from physical thermodynamic work [6].

On the physical side, Theorems 5.1 and 6.1 show how our communication converse translates into lower bounds on entropy production and, under equal endpoint mean internal energies, cyclic work for thermal ports that discard displaced modes without recovery, while Theorem 5.2 characterizes the exact work recoverable by the permitted translations prior to

reset under bounded drift speed and restricted controller information. Notably, conditioning the recovery displacement on the receiver's quantized measurement allows the controller to extract more free energy than conditioning on the transmitted amplitude alone, even though both schemes realize the exact same forward communication channel. Conversely, Proposition 5.3 reveals that when sign estimates are retained through their respective non-recovering resets, protocols using many channel observations incur a port-reset entropy cost proportional to M. Bridging the gap between signal-energy optimality and end-to-end physical work optimality therefore motivates the joint design of low-complexity codes, analog pre-detection accumulation, and reversible port interfaces.

Data and code availability

The accompanying source package contains this English manuscript, its bibliography, and the numerical verification scripts in `anc/`. Running `python anc/run_checks.py` with Python, NumPy, and SciPy reproduces fixed-seed checks of the causal-path interchange identity, the two-population information inequality, small linear codes, small-instance committee executions, and the displacement-recovery formulas. The tested environment and outputs are recorded in `anc/verification.json`; a new run writes to `results/`. These finite checks provide consistency evidence and do not replace the proofs or establish correctness against every admissible adversary.

Use of generative AI

Generative AI tools were used to assist with mathematical exploration, manuscript drafting and translation, and the generation of numerical verification scripts. The author has checked and understood all parts of the manuscript.

References

[1] Ittai Abraham, T.-H. Hubert Chan, Danny Dolev, Kartik Nayak, Rafael Pass, Ling Ren, and Elaine Shi. Communication complexity of Byzantine agreement, revisited. *Distributed Computing*, 36(1):3–28, 2023. doi: 10.1007/s00446-022-00428-8. URL `https://arxiv.org/abs/1805.03391`.

[2] Ziv Bar-Yossef, T. S. Jayram, Ravi Kumar, and D. Sivakumar. An information statistics approach to data stream and communication complexity. *Journal of Computer and System Sciences*, 68(4): 702–732, 2004. doi: 10.1016/j.jcss.2003.11.006.

[3] Charles H. Bennett. Logical reversibility of computation. *IBM Journal of Research and Development*, 17(6):525–532, 1973. doi: 10.1147/rd.176.0525.

[4] Piotr Berman, Juan A. Garay, and Kenneth J. Perry. Towards optimal distributed consensus. In *Proceedings of the 30th Annual Symposium on Foundations of Computer Science*, pages 410–415. IEEE, 1989. doi: 10.1109/SFCS.1989.63511.

[5] Joseph Bonneau, Benedikt Bünz, Miranda Christ, and Yuval Efron. How much public randomness do modern consensus protocols need? In *7th Conference on Advances in Financial Technologies (AFT 2025)*, volume 354 of *Leibniz International Proceedings in Informatics*, pages 12:1–12:19. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2025. doi: 10.4230/LIPIcs.AFT.2025.12.

[6] Eric Budish, Andrew Lewis-Pye, and Tim Roughgarden. The economic limits of permissionless consensus. *SIAM Journal on Computing*, 55(4):622–670, 2026. doi: 10.1137/24M1677514. URL `https://arxiv.org/abs/2405.09173`.

[7] Andrea Clementi, Luciano Gualà, Emanuele Natale, Francesco Pasquale, Giacomo Scornavacca, and Luca Trevisan. Consensus vs broadcast, with and without noise. In *11th Innovations in Theoretical Computer Science Conference (ITCS 2020)*, volume 151 of *Leibniz International Proceedings in Informatics*, pages 42:1–42:13. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2020. doi: 10.4230/LIPIcs.ITCS.2020.42.

[8] Andreas Dechant. Minimum entropy production, detailed balance and Wasserstein distance for continuous-time Markov processes. *Journal of Physics A: Mathematical and Theoretical*, 55: 094001, 2022. doi: 10.1088/1751-8121/ac4ac0.

[9] Massimiliano Esposito, Katja Lindenberg, and Christian Van den Broeck. Entropy production as correlation between system and reservoir. *New Journal of Physics*, 12:013013, 2010. doi: 10.1088/1367-2630/12/1/013013.

[10] Wassily Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58(301):13–30, 1963. doi: 10.1080/01621459.1963.10500830.

[11] Young-Han Kim, Amos Lapidoth, and Tsachy Weissman. Error exponents for the Gaussian channel with active noisy feedback. *IEEE Transactions on Information Theory*, 57(3):1223–1236, 2011. doi: 10.1109/TIT.2011.2104991.

[12] Artemy Kolchinsky and David H. Wolpert. Work, entropy production, and thermodynamics of information under protocol constraints. *Physical Review X*, 11:041024, 2021. doi: 10.1103/PhysRevX.11.041024.

[13] Jovan Komatovic, Andrew Lewis-Pye, Joachim Neu, Tim Roughgarden, and Ertem Nusret Tas. From permissioned to proof-of-stake consensus. In *7th Conference on Advances in Financial Technologies (AFT 2025)*, volume 354 of *Leibniz International Proceedings in Informatics*, pages 18:1–18:26. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2025. doi: 10.4230/LIPIcs.AFT.2025.18.

[14] Rolf Landauer. Irreversibility and heat generation in the computing process. *IBM Journal of Research and Development*, 5(3):183–191, 1961. doi: 10.1147/rd.53.0183.

[15] Rolf Landauer. Minimal energy requirements in communication. *Science*, 272(5270):1914–1918, 1996. doi: 10.1126/science.272.5270.1914.

[16] Andrew Lewis-Pye and Tim Roughgarden. Permissionless consensus. arXiv:2304.14701, 2023. URL `https://arxiv.org/abs/2304.14701`.

[17] Shachar Meir, Hugo Mirault, David Peleg, and Peter Robinson. Time-optimal and energy-efficient deterministic consensus. In *29th International Conference on Principles of Distributed Systems (OPODIS 2025)*, volume 361 of *Leibniz International Proceedings in Informatics*, pages 15:1–15:16. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2026. doi: 10.4230/LIPIcs.OPODIS.2025.15.

[18] Yury Polyanskiy, H. Vincent Poor, and Sergio Verdú. Minimum energy to send k bits through the Gaussian channel with and without feedback. *IEEE Transactions on Information Theory*, 57 (8):4880–4902, 2011. doi: 10.1109/TIT.2011.2159041.

[19] Naoto Shiraishi, Ken Funo, and Keiji Saito. Speed limit for classical stochastic processes. *Physical Review Letters*, 121:070601, 2018. doi: 10.1103/PhysRevLett.121.070601.

[20] Abhishek Yadav and David Wolpert. Minimal thermodynamic cost of communication. *Physical Review Research*, 7:043324, 2025. doi: 10.1103/qvc2-32xr.

[21] Shou-Cheng Zhang. In Math We Trust: Tsinghua x-lab lecture. Published lecture transcript, Big Data Digest; reproduced by Tencent Cloud Developer Community, May 2018. URL `https://cloud.tencent.com/developer/article/1133688`. In Chinese; republication dated 23 May 2018. Accessed 28 October 2026.