Speed Limit for a Highly Irreversible Process and Tight Finite-Time Landauer's Bound

Jae Sung Lee[1,*,†] Sangyun Lee[1,*] Hyukjoon Kwon[2,‡] and Hyunggyu Park[1,3,§]

[1]*School of Physics, Korea Institute for Advanced Study, Seoul 02455, Korea*
[2]*School of Computational Sciences, Korea Institute for Advanced Study, Seoul 02455, Korea*
[3]*Quantum Universe Center, Korea Institute for Advanced Study, Seoul 02455, Korea*

Landauer's bound is the minimum thermodynamic cost for erasing one bit of information. As this bound is achievable only for quasistatic processes, finite-time operation incurs additional energetic costs. We find a tight finite-time Landauer's bound by establishing a general form of the classical speed limit. This tight bound well captures the divergent behavior associated with the additional cost of a highly irreversible process, which scales differently from a nearly irreversible process. We also find an optimal dynamics which saturates the equality of the bound. We demonstrate the validity of this bound via discrete one-bit and coarse-grained bit systems. Our Letter implies that more heat dissipation than expected occurs during high-speed irreversible computation.

Introduction.—Memory erasure is an elementary operation in irreversible computation. As the erasing operation incurs a thermodynamic cost and takes a finite physical time, low energy consumption and a short process time are critical requirements for efficient computation. The fundamental limitation of the energetic cost is given by Landauer's principle [1,2], which states that at least $k_B T \ln 2$ of work is necessary to erase a single bit memory, where k_B is the Boltzmann constant and T is the environment temperature. Landauer's bound is universal in the sense that it is independent of memory device type or physical platform. This bound has been confirmed experimentally using various physical setups, including a double-well potential realized by optical tweezers [3,4] and virtual potential [5,6], an electric-circuit system [7], and a nanomagnetic memory bit [8,9].

In real-world situations, however, this fundamental bound is less practical, as it requires a quasistatic process, which far exceeds the system's relaxation timescale. The reported experimental timescale of the quasistatic erasure process ranges from a few hundred milliseconds [10] to several tens [4] or hundreds [5] of seconds, which is far from the time required for practical computation. Therefore, it is important to understand the finite-time effect on thermodynamic cost, which generally increases as a process becomes faster and more irreversible [11–15]. Several experimental studies have suggested that the minimum energetic cost should increase by an additional cost inversely proportional to the erasing time τ, i.e., $k_B T \ln 2 + C/\tau$, with a system-dependent constant C [3–6]. This behavior has also been investigated theoretically for the classical stochastic system described by the overdamped Langevin equation [16,17] and an open quantum system described by the Lindblad equation [18].

These studies suggest that a trade-off relation plays a central role in understanding the overhead cost of the Landauer's bound. Over the past decade, various types of trade-off relations have been reported in stochastic systems and also for open quantum systems such as thermodynamic uncertainty relations [19–27], kinetic uncertainty relations [28–32], and speed limits for state change [33–42]. Recently, Zhen *et al.* [43] showed that the $1/\tau$ behavior of the minimum work bound of erasing processes is governed by the speed-limit inequality associated with the thermodynamic cost.

In this Letter, we first present a simpler way to derive the general form of the speed limit introduced in Ref. [42], which can have various functional forms. Two different speed-limit regimes are considered in terms of the degree of irreversibility. For a nearly reversible process, we retrieve the previous speed limits [33,34] by taking a simple functional form that provides a tight bound on the operation time in terms of entropy production (EP) and dynamical activity. However, this bound gradually loosens as the process becomes more irreversible. We find a tight bound for a highly reversible process from the general speed limit with a different functional form. In the limit of high irreversibility, this new bound becomes finite, depending solely on dynamical activity.

We apply this general speed limit to calculate the tight bound of the additional cost associated with a finite-time erasing operation. We demonstrate that this bound scales as $1/\tau$ for a nearly reversible process, while much stronger divergence appears for a fast or highly irreversible process. As a result, high-speed irreversible computation requires much more heat dissipation and, thus, associated supporting cooling architecture. We also find an explicit protocol which meets the equality condition of the bound.

Speed limit for a highly irreversible process.—Suppose that the time evolution of a probability distribution $\boldsymbol{p}(t) = \{p_n(t)\}$ at time t is described by

$$\dot{p}_n(t) = \sum_m R_{nm}(t) p_m(t),$$

where the transition rate matrix $\boldsymbol{R}(t)$ satisfies the condition $R_{nm}(t) \geq 0$ for $m \neq n$ and $R_{nn}(t) = -\sum_{m(\neq n)} R_{mn}(t)$. The statistical distance between the initial and the final distributions after time τ can be measured by the total variation distance

$$\ell = d_{\mathrm{TV}}[\boldsymbol{p}(\tau), \boldsymbol{p}(0)] \equiv \frac{1}{2} \sum_n |p_n(\tau) - p_n(0)|.$$

We first establish a general form of the speed limit, in terms of the distance ℓ, the EP Σ^*, and the total activity (number of jumps) A_{tot} during τ, given by

$$\frac{\ell}{A_{\mathrm{tot}}} \leq f\left(\frac{\Sigma^*}{A_{\mathrm{tot}}}\right), \tag{1}$$

with an appropriate choice of a monotonically increasing concave function f as listed in Table I. Hereafter, we set $k_B = 1$ for convenience. The EP $\Sigma^* = \int_0^\tau (\dot{\Sigma}^*) dt$ is characterized by the transition matrix of the adjoint process $\boldsymbol{R}^*(t)$ [44–47], where the corresponding EP rate is defined as

$$\dot{\Sigma}^* \equiv \sum_{n \neq m} R_{nm}(t) p_m(t) \ln\left[\frac{R_{nm}(t) p_m(t)}{R^*_{mn}(t) p_n(t)}\right]. \tag{2}$$

We assume that the adjoint process is also stochastic with the same *escape* rate: $R^*_{nn}(t) = -\sum_{m(\neq n)} R^*_{mn}(t) = R_{nn}(t)$. A trivial choice of the adjoint process is taking it the same as the time-reversal process $R^*_{mn}(t) = R_{mn}(t)$, which leads to the total EP Σ, i.e., $\Sigma^* = \Sigma$. Alternatively, by considering the instantaneous steady-state $\boldsymbol{p}^{\mathrm{ss}}(t)$ such that $\dot{\boldsymbol{p}}^{\mathrm{ss}}(t) = 0$,

TABLE I. Various choices of the concave function $f(x)$ satisfying Eq. (7) and its inverse (convex) function $h(v) = f^{-1}(v)$.

	$f(x)$	$h(v)$
Pinsker [49]	$\sqrt{(x/2)}$	$2v^2$
Bretagnolle-Huber [50]	$\sqrt{1 - e^{-x}}$	$-\ln(1 - v^2)$
Vajda [51]	$\cdots$[a]	$\ln\left[(1+v)/(1-v)\right] - \left[2v/(1+v)\right]$
Gilardoni [52]	$\cdots$[a]	$\ln\left[(1+v)^{-1+v}/(1-v)\right]$
Symmetric KLD[b][52,53]	$\cdots$[a]	$v \ln\left[(1+v)/(1-v)\right]$

[a]Analytic compact expression is not available.
[b]This bound is valid only when the KLD is symmetric.

the adjoint process defined as $R^*_{mn}(t) = R_{nm}(t)[p^{\mathrm{ss}}_m(t)/p^{\mathrm{ss}}_n(t)]$ leads to the Hatano-Sasa (excess) EP $\Sigma^* = \Sigma_{\mathrm{HS}}$. Since $0 \leq \Sigma_{\mathrm{HS}} \leq \Sigma$ [44,48], Σ_{HS} always gives a tighter bound than Σ in Eq. (1).

To derive Eq. (1), we first note that

$$\ell = \frac{1}{2} \sum_n |p_n(\tau) - p_n(0)| \leq \frac{1}{2} \int_0^\tau dt \sum_n |\dot{p}_n(t)|, \tag{3}$$

from the triangle inequality. Using $R^*_{nn}(t) = R_{nn}(t)$, the instantaneous change of the probability distribution is bounded by

$$\sum_n |\dot{p}_n(t)| = \sum_n \left| \sum_{m(\neq n)} R_{nm}(t) p_m(t) + R_{nn}(t) p_n(t) \right|$$

$$= \sum_n \left| \sum_{m(\neq n)} \{R_{nm}(t) p_m(t) - R^*_{mn}(t) p_n(t)\} \right|$$

$$\leq \sum_{n \neq m} |R_{nm}(t) p_m(t) - R^*_{mn}(t) p_n(t)|$$

$$= 2A(t) d_{\mathrm{TV}}[\boldsymbol{Q}(t), \boldsymbol{Q}^*(t)], \tag{4}$$

where $A(t) = \sum_{n \neq m} R_{nm}(t) p_m(t) = \sum_{m \neq n} R^*_{mn}(t) p_n(t)$ is the jump rate at time t and $\boldsymbol{Q}(t)$ [$\boldsymbol{Q}^*(t)$] denotes the normalized conditional joint probability distribution of the forward (reverse) process which is defined as follows:

$$Q_{mn}(t) = P[m, n|\mathrm{jump}] = \frac{(1 - \delta_{nm}) R_{nm}(t) p_m(t)}{A(t)},$$

$$Q^*_{mn}(t) = P^*[n, m|\mathrm{jump}] = \frac{(1 - \delta_{mn}) R^*_{mn}(t) p_n(t)}{A(t)}.$$

Thus, $d_{\mathrm{TV}}[\boldsymbol{Q}(t), \boldsymbol{Q}^*(t)]$ captures how irreversible the process is at time t. We remark that $A_{\mathrm{tot}} = \int_0^\tau dt A(t)$ has a meaning of the total number of jumps during the entire process. By combining Eqs. (3) and (4), we have the following inequality:

$$\ell \leq \int_0^\tau dt A(t) d_{\mathrm{TV}}[\boldsymbol{Q}(t), \boldsymbol{Q}^*(t)]. \tag{5}$$

It is worth noting that the EP rate can be expressed in terms of the conditional joint distributions as

$$\dot{\Sigma}^* = A(t) D[\boldsymbol{Q}(t) \| \boldsymbol{Q}^*(t)], \tag{6}$$

where $D(\boldsymbol{p} \| \boldsymbol{q}) \equiv \sum_x p_x \ln(p_x/q_x)$ is the Kullback-Leibler divergence (KLD) between two probability distributions $\boldsymbol{p}$ and $\boldsymbol{q}$. Note that the KLD corresponding to the total EP rate ($\dot{\Sigma}$) is symmetric [$D(\boldsymbol{p} \| \boldsymbol{q}) = D(\boldsymbol{q} \| \boldsymbol{p})$], while the KLD is generally asymmetric for other choices such as $\dot{\Sigma}_{\mathrm{HS}}$. There exist various choices of a monotonic concave function f

(see Table I) that connects the total variation distance and the KLD to obey the following inequality [49–52]:

$$d_{\text{TV}}(\boldsymbol{p}, \boldsymbol{q}) \le f[D(\boldsymbol{p}\|\boldsymbol{q})]. \tag{7}$$

The speed limit is obtained by plugging in Eqs. (6) and (7) to Eq. (5) and then dividing both sides with A_{tot}, which leads to

$$\frac{\ell}{A_{\text{tot}}} \le \frac{\int_0^\tau dt A(t) f(\frac{\dot{\Sigma}^*}{A(t)})}{\int_0^\tau dt A(t)} \le f\left(\frac{\Sigma^*}{A_{\text{tot}}}\right), \tag{8}$$

from the concavity of f.

As $g(x) \equiv h(x)/(2x)$ is a monotonically increasing function for all h's in Table I, where $h = f^{-1}$, we can rewrite Eq. (8) as

$$\tau \ge \frac{\ell}{\langle A \rangle_\tau g^{-1}\left(\frac{\Sigma^*}{2\ell}\right)}, \tag{9}$$

by defining $\langle A \rangle_\tau \equiv A_{\text{tot}}/\tau$ and $g^{-1}(x)$ the inverse function of $g(x)$. Equation (9) is the general form of the speed limit, where various types of bounds can be obtained based on the choice of $h(x)$. The previous speed limit $\tau \ge 2\ell^2/(\langle A \rangle_\tau \Sigma^*)$ in Refs. [33,34] is readily obtained by taking $h(x) = x^2$ (Pinsker [49]), which is tight only for a nearly reversible (slow) process but yields a very loose bound for a highly irreversible process. We note that, for any $\Sigma^* \ge 2\ell$, this bound is even worse than the *fundamental* bound $\tau \ge \ell/\langle A \rangle_\tau$ obtained from the minimum activity to change the probability distribution regardless of the EP, $A_{\text{tot}} = \langle A \rangle_\tau \tau \ge \ell$ [42].

We find that a speed limit can be tightened for a highly irreversible process with alternative choices of $h(x)$ such as Bretagnolle-Huber, Vajda, Gilardoni, and symmetric KLD as listed in Table I. All these four functions provide speed limits, always tighter than the fundamental bound, which can be accessible only when $\Sigma^* \to \infty$. Therefore, in the highly irreversible limit, time is bounded solely by the dynamical activity but not the EP. The symmetric KLD bound is always the tightest among all $h(x)$, though it is valid only for the symmetric KLD. Otherwise, the Gilardoni bound is the tightest for $\Sigma^*/\ell \ge 1.14$, while the Pinsker bound is the tightest elsewhere. Simple derivation of the symmetric KLD bound is presented in Supplemental Material [54].

Tight finite-time Landauer's bound.—The speed limit in Eq. (1) can be rearranged to bound the EP as

$$\Sigma^* \ge \ell \frac{h(v)}{v} \equiv B_X, \tag{10}$$

where $v = \ell/A_{\text{tot}}$ is the average distance change per jump, which ranges from 0 to 1, measuring the irreversibility of the process. When v is close to 0 (1), the distribution changes gradually (abruptly), so the process is nearly reversible (highly irreversible). The bound B_X monotonically increases with v for all h's, where X denotes a

specific functional form, e.g., $X = P$ (Pinsker) and $X = S$ (symmetric KLD) with $B_P = 2\ell v$ and $B_S = 2\ell \tanh^{-1} v$, respectively.

Now, we use the EP bound to estimate the minimum cost for a finite-time erasing process. Suppose an erasing operation resets a one-bit system composed of 0 and 1 states with the initial probabilities $[p_0(0); p_1(0)] = (1/2; 1/2)$ and yields the final probabilities $[p_0(\tau); p_1(\tau)] = (1 - \epsilon; \epsilon)$ with erasing error ϵ after time τ. The statistical distance between the initial and final states becomes $\ell = 1/2 - \epsilon$, and the Shannon entropy change of the system can be computed as $\Delta S_{\text{sys}} = -\ln 2 - (1 - \epsilon)\ln(1 - \epsilon) - \epsilon \ln \epsilon$. Furthermore, by setting Σ^* as the total EP [55], we have $\Sigma^* = \Delta S_{\text{sys}} + Q/T$, where Q is the heat dissipated into the surrounding environment with temperature T during the erasing process. As the total EP Σ corresponds to the symmetric KLD, we use the symmetric KLD bound to obtain the tightest Landauer's bound.

In the perfect erasing limit $\epsilon \to 0$, we get $\ell \to 1/2$ and $\Delta S_{\text{sys}} \to -\ln 2$, and the finite-time Landauer's bound from Eq. (10) is expressed as

$$\frac{Q}{T} \ge \ln 2 + B_S = \ln 2 + \tanh^{-1} v, \tag{11}$$

where B_S represents the additional cost due to finite-time operation. For small v (nearly reversible), $B_S \simeq v = 1/(2\tau\langle A \rangle_\tau)$, which corresponds to the previously known $1/\tau$ behavior [3–6,43]. As we approach $v = 1^-$ (highly irreversible regime), B_S diverges asymptotically as

$$B_S \simeq -\frac{1}{2}\ln(1 - v) = -\frac{1}{2}\ln\left(1 - \frac{1}{2\tau\langle A \rangle_\tau}\right). \tag{12}$$

This implies that much higher dissipation should occur in a highly irreversible erasing operation.

Practical computation requires a small erasing error as well as a short operation time. To this end, the transition rate from 1 to 0 state has to be large for fast operation, necessitating large driving (large $\langle A \rangle_\tau$). In comparison, the reverse transition (0 to 1) should be suppressed to prevent erasing-error operations. Therefore, the best strategy for a desired erasing operation is that all "particles" initially located at state 1 jump to state 0 once, and no jump occurs afterward; this condition can be read as $A_{\text{tot}} \approx 1/2$ with $p_1(0) = 1/2$. Consequently, the operation for a practical erasing process should be highly irreversible with $v = \ell/A_{\text{tot}} \approx 1$. Thus, Eq. (12) for a highly irreversible process is well deserved for practical computation.

We find explicitly the *optimal* dynamics which minimizes the EP, satisfying the equality of Eq. (10) with $h = h_S$. Its sufficient condition is

$$\frac{R_{01}(t)p_1(t)}{R_{10}(t)p_0(t)} = c(\text{const}) \quad \forall\ 0 \le t \le \tau, \tag{13}$$

along with a monotonic change of $p_n(t)$ in time. The detailed derivation is presented in Supplemental Material [54], where we also show that a process with $v \approx 1^-$ ($v \approx 0$) is realized with large c and short τ ($c \approx 1$ and long τ).

The finite-time Landauer's bound, Eqs. (10) and (11), is also applicable to a bit system made by coarsening, such as a Langevin system with a double-well potential [3–6]. This can be verified by the fact that the EP of a coarse-grained bit system is equal to or smaller than that of its original system without coarse-graining. See Eq. (10) in Ref. [43] and Supplemental Material [54] for details. Therefore, the original EP is also bounded by the same additional cost term in Eq. (10).

Numerical confirmation.—Here, we investigate two examples. The first one is a discrete one-bit system consisting of states 0 and 1 with energy levels $E_0 = 0$ (fixed) and $E_1(t)$ (time varying), respectively. Its dynamics is described by the following master equation:

$$\dot{p}_0(t) = R_{01}(t)p_1(t) - R_{10}(t)p_0(t),$$
$$\dot{p}_1(t) = R_{10}(t)p_0(t) - R_{01}(t)p_1(t). \qquad (14)$$

The transition rate $R_{nm}(t)$ satisfies the detailed balance condition, that is, with $\gamma(t) \equiv e^{-E_1(t)/T}/(1 + e^{-E_1(t)/T})$

$$R_{10}(t) = \mu(t)\gamma(t), \qquad R_{01}(t) = \mu(t)[1 - \gamma(t)], \qquad (15)$$

where $\mu(t)$ is an overall transition rate. The erasing process is illustrated in Fig. 1(a). The system is prepared with the initial probabilities $(p_0; p_1) = (1/2; 1/2)$ with $E_1 = 0$ and $\mu = 0$ for $t < 0$. Here, $\mu = 0$ indicates that the transition is blocked. E_1 and μ are abruptly raised to E_{eras} and μ_{eras} at time $t = 0$, respectively, and maintained up to $t = \tau$. Then, both E_1 and μ are immediately lowered to 0 at $t = \tau$. The final probability at $t = \tau$ is $(p_0; p_1) = (1 - \epsilon; \epsilon)$. This protocol is the simplest one in the n-step energy-raising procedure [43,56]. The exact solution of this model is

$$p_1(t) = e^{-\mu t}p_1(0) + (1 - e^{-\mu t})p_{1,E_{\text{eras}}}^{\text{eq}}, \qquad (16)$$

where $p_{1,E_{\text{eras}}}^{\text{eq}} = 1/[1 + \exp(E_{\text{eras}}/T)]$. Using Eq. (16), we explicitly calculate the entropy change of the system $\Delta S_{\text{sys}} = -\sum_i[p_i(\tau)\ln p_i(\tau) - p_i(0)\ln p_i(0)]$ and heat $Q = E_{\text{eras}}[p_1(\tau) - p_1(0)]$, which leads to the total EP $\Sigma = \Delta S_{\text{sys}} + Q/T$, as well as the total activity A_{tot}. We also construct an optimal time-dependent control of $R_{nm}(t)$ satisfying the saturation condition Eq. (13), of which the explicit form can be found in Supplemental Material [54].

The second example is a coarse-grained bit system consisting of a one-dimensional Brownian particle trapped in a double-well potential. Dynamics of the particle is governed by the following overdamped Langevin equation:

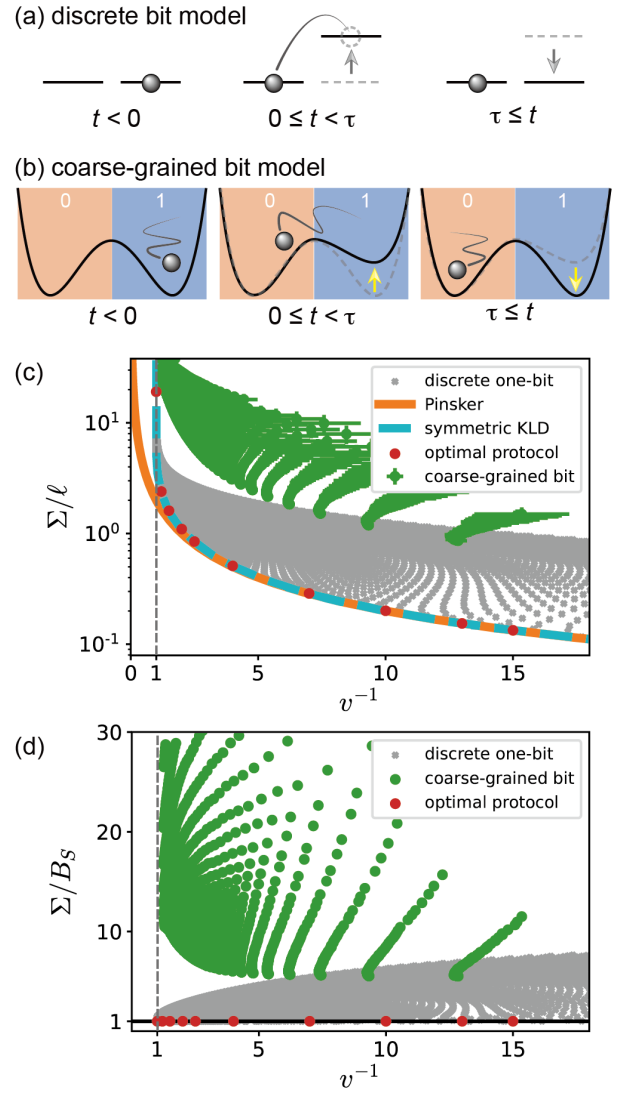$$\gamma\dot{x} = -\frac{\partial V_{\text{DW}}(x,t)}{\partial x} + \sqrt{2\gamma T}\xi(t), \qquad (17)$$

(a) discrete bit model

(b) coarse-grained bit model

(c)

(d)

FIG. 1. (a) and (b) Schematics of the two erasure models: discrete one-bit system (a) and coarse-grained bit system (b). (c) Plot of Σ/ℓ versus v^{-1} for discrete system (gray cross) and continuous system (green circle). The orange solid (sky-blue dashed) curve denotes the Pinsker (symmetric KLD) bound. The result of the optimal protocol is denoted by a red dot circle. (d) Plot of Σ divided by the symmetric KLD bound B_S versus v^{-1}. For a fixed ℓ and $A(t)$, v^{-1} can be simply regarded as a scaled time. The same data are used for (c) and (d).

where x is the position of the particle, $\xi(t)$ is a Gaussian white noise satisfying $\langle\xi(t)\xi(t')\rangle = \delta(t - t')$, and the double-well potential $V_{\text{DW}}(x,t)$ is given as

$$V_{\text{DW}}(x,t) = E_b\left[\left(\frac{x}{x_m}\right)^4 - 2\left(\frac{x}{x_m}\right)^2\right] + \Theta(t)\frac{x}{x_m}, \qquad (18)$$

where $\Theta(t)$ provides a time-dependent protocol. This model corresponds to the experimental setup in Ref. [57]. The system can be treated as a coarse-grained bit memory by regarding the system being in state "0" ("1") when $x \leq 0$

($x > 0$). Then, the probabilities for the coarse-grained state i ($i \in \{0, 1\}$) are

$$p_0^{\mathrm{cg}}(t) = \int_{x \leq 0} dx P(x, t) \quad \text{and} \quad p_1^{\mathrm{cg}}(t) = 1 - p_0^{\mathrm{cg}}(t), \quad (19)$$

where $P(x, t)$ is the probability distribution of the original continuous system. The erasing process of this model is presented in Fig. 1(b). An initial state is prepared as the equilibrium state determined by the double-well potential with $\Theta(t) = 0$ for $t < 0$. As the potential is symmetric with respect to $x = 0$, $p_0^{\mathrm{cg}}(0) = p_1^{\mathrm{cg}}(0) = 1/2$. At $t = 0$, we immediately raises $\Theta(t)$ to Θ_{eras} and maintain it up to $t = \tau$. $\Theta(t)$ then returns to 0 at $t = \tau$. Because of the nonlinearity of the potential force, an analytic solution is not available. Instead, the total EP $\Sigma = \Delta S_{\mathrm{sys}} + Q/T$ is estimated by numerically evaluating $\Delta S_{\mathrm{sys}} = -\int dx [P(x, \tau) \ln P(x, \tau) - P(x, 0) \ln P(x, 0)]$ and $Q = \int_0^\tau dt (-\partial_x V_{\mathrm{DM}}) \circ \dot{x}(t)$, where $\circ$ is the Stratonovich product. ℓ and A_{tot} are estimated by using $p_i^{\mathrm{cg}}(t)$ and by counting the number of transitions between the different coarse-grained states [58].

Figure 1(c) shows the plot of Σ/ℓ against v^{-1} for the discrete and the coarse-grained bit models. The data for the discrete model are obtained by varying parameters E_{eras} and τ within the ranges $10^{-5} \leq E_{\mathrm{eras}} \leq 10$ and $10^{-10} \leq \tau \leq 20$ with fixed $\mu_{\mathrm{eras}} = 1$ and $T = 1$. The data of the coarse-grained bit model are the simulation results for the parameter ranges (used in real experiment [57]) of $0.1 k_B T \leq \Theta_{\mathrm{eras}} \leq 10 k_B T$ and 0.1 ms $\leq \tau \leq 110$ ms with fixed $x_m = 50$ nm, $k_B T = 4.1$ pN · nm ($T = 300$ K), $E_b = 3 k_B T$, and $\gamma = 24\sqrt{2} k_B T \cdot$ ms$/\pi x_m^2$. Each point of the coarse-grained bit model in the plot is obtained by averaging 10^6 realizations. The Pinsker and the symmetric KLD bounds are presented along with the result of the optimal erasing process in the figure, and the comparison with other bounds is shown in Supplemental Material [54].

Indeed, the symmetric KLD tightly bounds the EP of the discrete bit model for all v. This tightness can be also checked in Fig. 1(d), which presents the total EP divided by B_S [see Eq. (10)]. Note that the Pinsker bound is quite tight for nearly reversible processes (small v); however, it becomes extremely loose near $v = 1$. The data of the coarse-grained model are also well bounded by B_S. However, the bound is not tight due to the "intra-EP" induced by transitions between *microstates* inside the same coarse-grained state. The detailed explanation is presented in Supplemental Material [54]. Thus, it is also important to reduce the intra-EP for lowering the thermodynamic cost for a coarse-grained system.

Conclusion.—We find the finite-time Landauer's bound, which is tight for an erasing process with any irreversibility and any error rate, from the general form of the speed limit. We also find an optimal dynamics which saturates the equality of the bound. This bound is applicable to a coarse-grained bit system as well as an intrinsically two-state system. We demonstrate that, for a highly irreversible process, the diverging behavior of the additional cost is much steeper than that of a nearly reversible process. This indicates that, in a practical computation, which belongs to a highly irreversible regime, reducing the operation time and error rate gives rise to much more heat dissipation than expected. Thus, enhancing the cooling power or heat tolerance of a memory device to maintain a proper device temperature is more critical when computation becomes more irreversible. Our formula is also directly applicable for estimating the proper bound of cooling power for a given computation speed, and the density of memory. Furthermore, to save thermodynamic costs, it is important to reduce the dissipation produced inside the same coarse-grained state. Subsequent experimental studies in various physical systems are anticipated in the future.

*These authors contributed equally to this work.
†jslee@kias.re.kr
‡hjkwon@kias.re.kr
§hgpark@kias.re.kr

[1] R. Landauer, Irreversibility and heat generation in the computing process, IBM J. Res. Dev. **5**, 183 (1961).

[2] Juan M. R. Parrondo, Jordan M. Horowitz, and Takahiro Sagawa, Thermodynamics of information, Nat. Phys. **11**, 131 (2015).

[3] Antoine Bérut, Artak Arakelyan, Artyom Petrosyan, Sergio Ciliberto, Raoul Dillenschneider, and Eric Lutz, Experimental verification of Landauer's principle linking information and thermodynamics, Nature (London) **483**, 187 (2012).

[4] Antoine Bérut, Artyom Petrosyan, and Sergio Ciliberto, Information and thermodynamics: Experimental verification of Landauer's erasure principle, J. Stat. Mech. (2015) P06015.

[5] Yonggun Jun, Mom čilo Gavrilov, and John Bechhoefer, High-Precision Test of Landauer's Principle in a Feedback Trap, Phys. Rev. Lett. **113**, 190601 (2014).

[6] Momčilo Gavrilov and John Bechhoefer, Arbitrarily slow, non-quasistatic, isothermal transformations, Europhys. Lett. **114**, 50002 (2016).

[7] Alexei O. Orlov, Craig S. Lent, Cameron C. Thorpe, Graham P. Boechler, and Gregory L. Snider, Experimental test of Landauer's principle at the sub-kBT level, Jpn. J. Appl. Phys. **51**, 06FE10 (2012).

[8] Jeongmin Hong, Brian Lambson, Scott Dhuey, and Jeffrey Bokor, Experimental test of Landauer's principle in single-bit operations on nanomagnetic memory bits, Sci. Adv. **2**, e1501492 (2016).

 [9] L. Martini, M. Pancaldi, M. Madami, P. Vavassori, G. Gubbiotti, S. Tacchi, F. Hartmann, M. Emmerling, S. Höfling, L. Worschech, and G. Carlotti, Experimental and theoretical analysis of landauer erasure in nano-magnetic switches of different sizes, Nano Energy **19**, 108 (2016).

[10] Salambô Dago, Jorge Pereda, Nicolas Barros, Sergio Ciliberto, and Ludovic Bellon, Information and Thermodynamics: Fast and Precise Approach to Landauer's Bound in an Underdamped Micromechanical Oscillator, Phys. Rev. Lett. **126**, 170601 (2021).

[11] Giovanni Diana, G. Baris Bagci, and Massimiliano Esposito, Finite-time erasing of information stored in fermionic bits, Phys. Rev. E **87**, 012111 (2013).

[12] Erik Aurell, Krzysztof Gawędzki, Carlos Mejía-Monasterio, Roya Mohayaee, and Paolo Muratore-Ginanneschi, Refined second law of thermodynamics for fast random processes, J. Stat. Phys. **147**, 487 (2012).

[13] A. B. Boyd, A. Patra, C. Jarzynski, and J. P. Crutchfield, Shortcuts to thermodynamic computing: The cost of fast and faithful erasure, arXiv:1812.11241.

[14] Paolo Muratore-Ginanneschi and Kay Schwieger, An application of Pontryagin's principle to brownian particle engineered equilibration, Entropy **19**, 379 (2017).

[15] Tim Schmiedl and Udo Seifert, Optimal Finite-Time Processes in Stochastic Thermodynamics, Phys. Rev. Lett. **98**, 108301 (2007).

[16] Karel Proesmans, Jannik Ehrich, and John Bechhoefer, Finite-Time Landauer Principle, Phys. Rev. Lett. **125**, 100602 (2020).

[17] Karel Proesmans, Jannik Ehrich, and John Bechhoefer, Optimal finite-time bit erasure under full control, Phys. Rev. E **102**, 032105 (2020).

[18] Tan Van Vu and Keiji Saito, Finite-Time Quantum Landauer Principle and Quantum Coherence, Phys. Rev. Lett. **128**, 010602 (2022).

[19] Naoto Shiraishi, Keiji Saito, and Hal Tasaki, Universal Trade-Off Relation Between Power and Efficiency for Heat Engines, Phys. Rev. Lett. **117**, 190601 (2016).

[20] Jae Sung Lee, Jong-Min Park, Hyun-Myung Chun, Jaegon Um, and Hyunggyu Park, Exactly solvable two-terminal heat engine with asymmetric onsager coefficients: Origin of the power-efficiency bound, Phys. Rev. E **101**, 052132 (2020).

[21] Andre C. Barato and Udo Seifert, Thermodynamic Uncertainty Relation for Biomolecular Processes, Phys. Rev. Lett. **114**, 158101 (2015).

[22] Andreas Dechant and Shin-ichi Sasa, Fluctuation–response inequality out of equilibrium, Proc. Natl. Acad. Sci. U.S.A. **117**, 6430 (2020).

[23] Jordan M. Horowitz and Todd R. Gingrich, Thermodynamic uncertainty relations constrain non-equilibrium fluctuations, Nat. Phys. **16**, 15 (2020).

[24] Yoshihiko Hasegawa and Tan Van Vu, Fluctuation Theorem Uncertainty Relation, Phys. Rev. Lett. **123**, 110602 (2019).

[25] Jae Sung Lee, Jong-Min Park, and Hyunggyu Park, Universal form of thermodynamic uncertainty relation for langevin dynamics, Phys. Rev. E **104**, L052102 (2021).

[26] Yoshihiko Hasegawa, Quantum Thermodynamic Uncertainty Relation for Continuous Measurement, Phys. Rev. Lett. **125**, 050601 (2020).

[27] Yoshihiko Hasegawa, Thermodynamic Uncertainty Relation for General Open Quantum Systems, Phys. Rev. Lett. **126**, 010602 (2021).

[28] Juan P. Garrahan, Simple bounds on fluctuations and uncertainty relations for first-passage times of counting observables, Phys. Rev. E **95**, 032134 (2017).

[29] Ivan Di Terlizzi and Marco Baiesi, Kinetic uncertainty relation, J. Phys. A **52**, 02LT03 (2018).

[30] Ken Hiura and Shin-ichi Sasa, Kinetic uncertainty relation on first-passage time for accumulated current, Phys. Rev. E **103**, L050103 (2021).

[31] Arnab Pal, Shlomi Reuveni, and Saar Rahav, Thermodynamic uncertainty relation for systems with unidirectional transitions, Phys. Rev. Research **3**, 013273 (2021).

[32] Arnab Pal, Shlomi Reuveni, and Saar Rahav, Thermodynamic uncertainty relation for first-passage times on markov chains, Phys. Rev. Research **3**, L032034 (2021).

[33] Naoto Shiraishi, Ken Funo, and Keiji Saito, Speed Limit for Classical Stochastic Processes, Phys. Rev. Lett. **121**, 070601 (2018).

[34] Van Tuan Vo, Tan Van Vu, and Yoshihiko Hasegawa, Unified approach to classical speed limit and thermodynamic uncertainty relation, Phys. Rev. E **102**, 062132 (2020).

[35] Tan Van Vu and Yoshihiko Hasegawa, Geometrical Bounds of the Irreversibility in Markovian Systems, Phys. Rev. Lett. **126**, 010601 (2021).

[36] Kohei Yoshimura and Sosuke Ito, Thermodynamic Uncertainty Relation and Thermodynamic Speed Limit in Deterministic Chemical Reaction Networks, Phys. Rev. Lett. **127**, 160601 (2021).

[37] Sosuke Ito and Andreas Dechant, Stochastic Time Evolution, Information Geometry, and the CramÉr-Rao Bound, Phys. Rev. X **10**, 021056 (2020).

[38] Gianmaria Falasco and Massimiliano Esposito, Dissipation-Time Uncertainty Relation, Phys. Rev. Lett. **125**, 120604 (2020).

[39] Schuyler B. Nicholson, Luis Pedro García-Pintos, Adolfo del Campo, and Jason R. Green, Time–information uncertainty relations in thermodynamics, Nat. Phys. **16**, 1211 (2020).

[40] Farita Tasnim and David H. Wolpert, Thermodynamic speed limits for co-evolving systems, arXiv:2107.12471.

[41] Andreas Dechant, Minimum entropy production, detailed balance and wasserstein distance for continuous-time markov processes, J. Phys. A **55**, 094001 (2022).

[42] Jean-Charles Delvenne and Gianmaria Falasco, Tight classical speed limits, arXiv:2110.13050.

[43] Yi-Zheng Zhen, Dario Egloff, Kavan Modi, and Oscar Dahlsten, Universal Bound on Energy Cost of Bit Reset in Finite Time, Phys. Rev. Lett. **127**, 190602 (2021).

[44] Massimiliano Esposito and Christian Van den Broeck, Three Detailed Fluctuation Theorems, Phys. Rev. Lett. **104**, 090601 (2010).

[45] Carlos Pérez-Espigares, Alejandro B. Kolton, and Jorge Kurchan, Infinite family of second-law-like inequalities, Phys. Rev. E **85**, 031135 (2012).

[46] Richard E. Spinney and Ian J. Ford, Nonequilibrium Thermodynamics of Stochastic Systems with Odd and Even Variables, Phys. Rev. Lett. **108**, 170603 (2012).

[47] Hyun Keun Lee, Chulan Kwon, and Hyunggyu Park, Fluctuation Theorems and Entropy Production with Odd-Parity Variables, Phys. Rev. Lett. **110**, 050602 (2013).

[48] Takahiro Hatano and Shin-ichi Sasa, Steady-State Thermodynamics of Langevin Systems, Phys. Rev. Lett. **86**, 3463 (2001).

[49] M. S. Pinsker, Information and information stability of random variables and processes (in russian), Moscow: Izv. Akad. Nauk (1960).

[50] J. Bretagnolle and C. Huber, Estimation des densites: Risque minimax, Probab. Theory Relat. Fields **47**, 119 (1979).

[51] I. Vajda, Note on discrimination information and variation, IEEE Trans. Inf. Theory **16**, 771 (1970).

[52] Gustavo L. Gilardoni, An improvement on Vajda's inequality, in *In and Out of Equilibrium 2*, edited by Vladas Sidoravicius and Maria Eulália Vares (Birkhäuser Basel, Basel, 2008), pp. 299–304.

[53] Gustavo L Gilardoni, On the minimum f-divergence for given total variation, C. R. Math. **343**, 763 (2006).

[54] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.129.120603 for details of the simple derivation of the symmetric KLD bound, the derivation of the optimal protocol, and application to the coarse-grained bit system.

[55] Note that for a two-state model, where $R_{nm}(t)p_m^{ss}(t) = R_{mn}(t)p_n^{ss}(t)$ is satisfied, the Hatano-Sasa EP coincides with the total EP, i.e., $\Sigma_{HS} = \Sigma$; it can be easily checked by plugging $R_{mn}^*(t) = R_{nm}(t)[p_m^{ss}(t)/p_n^{ss}(t)] = R_{mn}(t)$ into Eq. (2).

[56] Cormac Browne, Andrew J. P. Garner, Oscar C. O. Dahlsten, and Vlatko Vedral, Guaranteed Energy-Efficient Bit Reset in Finite Time, Phys. Rev. Lett. **113**, 100603 (2014).

[57] Govind Paneru, Tsvi Tlusty, and Hyuk Kyu Pak, New type of stochastic resonance in an active bath, arXiv:2106.12443.

[58] Note that the calculation of A_{tot} from the original (not-coarse-grained) dynamics leads to the meaningless result as $A(t)$ diverges in the Langevin dynamics with a continuous space.

Supplemental material for "Speed limit for a highly irreversible process and tight finite-time Landauers bound"

Jae Sung Lee and Sangyun Lee

School of Physics, Korea Institute for Advanced Study, Seoul 02455, Korea

Hyukjoon Kwon

School of Computational Sciences, Korea Institute for Advanced Study, Seoul 02455, Korea

Hyunggyu Park

School of Physics, Korea Institute for Advanced Study, Seoul 02455, Korea and

Quantum Universe Center, Korea Institute for Advanced Study, Seoul 02455, Korea

I. SIMPLE DERIVATION OF THE SYMMETRIC KLD (KULLBACK-LEIBLER DIVERGENCE) BOUND

The total variational distance is bounded by the Le Cam distance $d_{\mathrm{LC}}(\boldsymbol{p}, \boldsymbol{q})$ as follows:

$$d_{\mathrm{TV}}(\boldsymbol{p}, \boldsymbol{q})^2 = \left(\frac{1}{2} \sum_n \frac{|p_n - q_n|}{\sqrt{p_n + q_n}} \cdot \sqrt{p_n + q_n} \right)^2 \leq \frac{1}{4} \sum_n \frac{|p_n - q_n|^2}{p_n + q_n} \cdot 2 \equiv d_{\mathrm{LC}}(\boldsymbol{p}, \boldsymbol{q})^2. \qquad \text{(S1)}$$

The CauchySchwarz inequality is used to derive the inequality of Eq. (S1). By using the equality $|p - q|/(p + q) = \tanh\left|\frac{1}{2} \ln \frac{p}{q}\right|$ for $p, q \geq 0$ and the normalized probability $\tilde{p}_n \equiv \frac{|p_n - q_n|}{2d_{\mathrm{TV}}(\boldsymbol{p}, \boldsymbol{q})}$, we can show that

$$\begin{aligned} d_{\mathrm{LC}}(\boldsymbol{p}, \boldsymbol{q})^2 &= d_{\mathrm{TV}}(\boldsymbol{p}, \boldsymbol{q}) \sum_n \tilde{p}_n \cdot \frac{|p_n - q_n|}{p_n + q_n} \\ &\leq d_{\mathrm{TV}}(\boldsymbol{p}, \boldsymbol{q}) \tanh\left(\sum_n \tilde{p}_n \left| \frac{1}{2} \ln \frac{p_n}{q_n} \right| \right) \\ &= d_{\mathrm{TV}}(\boldsymbol{p}, \boldsymbol{q}) \tanh\left[\frac{1}{4d_{\mathrm{TV}}(\boldsymbol{p}, \boldsymbol{q})} \sum_n (p_n - q_n) \ln \frac{p_n}{q_n} \right] \end{aligned} \qquad \text{(S2)}$$

The concavity of the tanh function is used for the inequality of Eq. (S2). Combining Eqs. (S1) and (S2) and using the symmetrized KLD, $\frac{1}{2} \sum_n (p_n - q_n) \ln \frac{p_n}{q_n} = \frac{1}{2} \left[D(\boldsymbol{p}||\boldsymbol{q}) + D(\boldsymbol{q}||\boldsymbol{p}) \right] \equiv D_{\mathrm{S}}(\boldsymbol{p}||\boldsymbol{q})$, we have

$$2d_{\mathrm{TV}}(\boldsymbol{p}, \boldsymbol{q}) \tanh^{-1} d_{\mathrm{TV}}(\boldsymbol{p}, \boldsymbol{q}) \leq D_{\mathrm{S}}(\boldsymbol{p}, \boldsymbol{q}). \qquad \text{(S3)}$$

When the KLD is symmetric, that is, $D(\boldsymbol{p}, \boldsymbol{q}) = D(\boldsymbol{q}, \boldsymbol{p})$, Eq. (S3) can be written as

$$h_{\mathrm{S}}(d_{\mathrm{TV}}(\boldsymbol{p}, \boldsymbol{q})) \leq D(\boldsymbol{p}, \boldsymbol{q}). \qquad \text{(S4)}$$

where $h_{\mathrm{S}}(x) \equiv x \ln \frac{1+x}{1-x}$. For Eq. (S4), $\tanh^{-1}(x) = \frac{1}{2} \ln \frac{1+x}{1-x}$ is used.

II. OPTIMAL PROTOCOL FOR A SINGLE-BIT SYSTEM

We consider the time evolution of a probability distribution $\boldsymbol{p}(t) = \{p_0(t), p_1(t)\}$ at time t for a single-bit (two-state) system

$$\dot{p}_0(t) = R_{01}(t)p_1(t) - R_{10}(t)p_0(t),$$
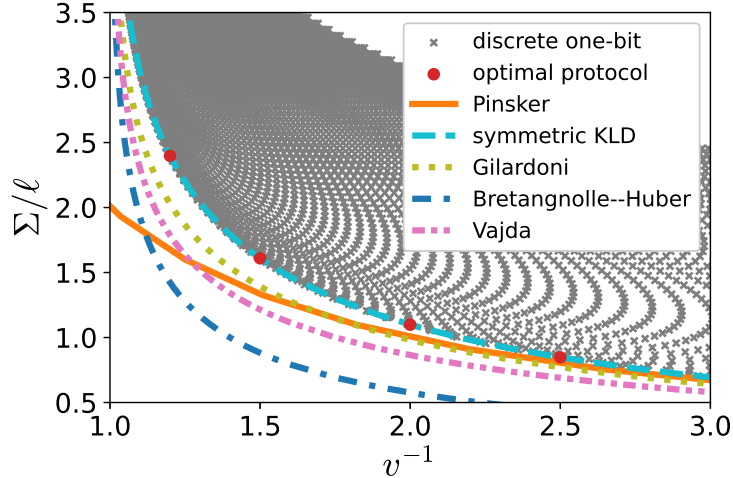$$\dot{p}_1(t) = R_{10}(t)p_0(t) - R_{01}(t)p_1(t).$$

FIG. S1. Comparison between various bounds. The symmetric KLD bound is the tightest: the result of the optimal protocol is on the symmetric KLD bound.

Note that $R_{nm} = R^*_{nm}$ for a two-state system and the total EP rate and the dynamical activity are written as

$$\dot{\Sigma}(t) = \dot{\Sigma}_{\text{HS}}(t) = (R_{01}(t)p_1(t) - R_{10}(t)p_0(t)) \ln \left[\frac{R_{01}(t)p_1(t)}{R_{10}(t)p_0(t)} \right],$$

$$A(t) = R_{01}(t)p_1(t) + R_{10}(t)p_0(t).$$
(S5)

We provide an explicit form of dynamics for a bit system that saturates the equality of the following speed limit bound

$$\frac{\ell}{A_{\text{tot}}} \leq f\left(\frac{\Sigma}{A_{\text{tot}}}\right) \Longleftrightarrow h\left(\frac{\ell}{A_{\text{tot}}}\right) \leq \frac{\Sigma}{A_{\text{tot}}},$$
(S6)

for the tightest bound, i.e., the symmetric KLD bound $h(x) = h_{\text{S}}(x) = x \ln \frac{1+x}{1-x}$ [1, 2]. We first show that the following condition,

$$\frac{d}{dt}\left[\frac{R_{01}(t)p_1(t)}{R_{10}(t)p_0(t)}\right] = 0 \Longleftrightarrow \frac{R_{01}(t)p_1(t)}{R_{10}(t)p_0(t)} = c \text{ (const.)}, \ \forall \ 0 \leq t \leq \tau \ ,$$
(S7)

with monotonic change of $p_n(t)$ in time is sufficient to saturate the symmetric KLD bound.

Proof. Equation (S6) is derived by using the four inequalities in the main text: Eqs. (3), (4), (7), and (8). Thus, we should find the equality conditions for the four inequalities to find the saturation condition of Eq. (S6). Without loss of generality, we consider the case of $p_0(\tau) \geq p_0(0)$ and $p_1(\tau) \leq p_1(0)$. We first observe that the inequality, Eq. (3) of the main

3

text,

$$\ell = \frac{1}{2} \sum_n |p_n(\tau) - p_n(0)| \leq \frac{1}{2} \int_0^\tau \sum_n |\dot{p}_n(t)| dt, \tag{S8}$$

is saturated when p_0 (p_1) is monotonically increasing (decreasing) in time, i.e.,

$$\dot{p}_0(t) = -R_{10}(t)p_0(t) + R_{01}(t)p_1(t) \geq 0,$$
$$\dot{p}_1(t) = R_{10}(t)p_0(t) - R_{01}(t)p_1(t) \leq 0. \tag{S9}$$

For a bit system, we then have the following formula,

$$\frac{1}{2} \sum_n |\dot{p}_n(t)| = -R_{10}(t)p_0(t) + R_{01}(t)p_1(t). \tag{S10}$$

The next inequality we consider is Eq. (4) of the main text, that is,

$$\frac{1}{2} \sum_n |\dot{p}_n(t)| \leq A(t) d_{\mathrm{TV}}(\boldsymbol{Q}, \boldsymbol{Q}^*), \tag{S11}$$

where

$$Q_{mn}(t) = \frac{(1 - \delta_{nm}) R_{nm}(t) p_m(t)}{A(t)} \quad \text{and} \quad Q^*_{mn}(t) = \frac{(1 - \delta_{nm}) R_{mn}(t) p_n(t)}{A(t)},$$

for a two-state model. We note that this inequality is always saturated for a bit system by directly calculating the total variational distance as $d_{\mathrm{TV}}(\boldsymbol{Q}, \boldsymbol{Q}^*) = \frac{1}{A(t)} (-R_{10}(t)p_0(t) + R_{01}(t)p_1(t))$ which is equivalent to $\frac{\sum_n |\dot{p}_n(t)|}{2A(t)}$ from Eq. (S10). We observe that for any bit system, the third inequality (Eq. (7) of the main text),

$$h(d_{\mathrm{TV}}(\boldsymbol{Q}, \boldsymbol{Q}^*)) \leq \frac{\dot{\Sigma}(t)}{A(t)},$$

is also saturated for the tightest bound $h_{\mathrm{S}}(x) = x \ln \frac{1+x}{1-x}$. Using Eqs. (S10) and (S5), this can be directly shown by

$$h_{\mathrm{S}}(d_{\mathrm{TV}}(\boldsymbol{Q}, \boldsymbol{Q}^*)) = h_{\mathrm{S}}\left(\frac{\sum_n |\dot{p}_n(t)|}{2A(t)}\right) = h_{\mathrm{S}}\left(\frac{-R_{10}(t)p_0(t) + R_{01}(t)p_1(t)}{A(t)}\right)$$
$$= \frac{-R_{10}(t)p_0(t) + R_{01}(t)p_1(t)}{A(t)} \ln\left[\frac{R_{01}(t)p_1(t)}{R_{10}(t)p_0(t)}\right] = \frac{\dot{\Sigma}(t)}{A(t)},$$

Finally, in order to saturate the inequality (Eq. (8) of the main text),

$$\ell = \int_0^\tau A(t) f\left(\frac{\dot{\Sigma}(t)}{A(t)}\right) dt \leq A_{\mathrm{tot}} f\left(\frac{\Sigma_{\mathrm{tot}}}{A_{\mathrm{tot}}}\right),$$

4

for a concave function f, we require the condition

$$\frac{\dot{\Sigma}(t)}{A(t)} = \frac{(R_{01}(t)p_1(t) - R_{10}(t)p_0(t))\ln\left[\frac{R_{01}(t)p_1(t)}{R_{10}(t)p_0(t)}\right]}{R_{10}(t)p_0(t) + R_{01}(t)p_1(t)} = \frac{\left(\frac{R_{01}(t)p_1(t)}{R_{10}(t)p_0(t)} - 1\right)\ln\left[\frac{R_{01}(t)p_1(t)}{R_{10}(t)p_0(t)}\right]}{1 + \frac{R_{01}(t)p_1(t)}{R_{10}(t)p_0(t)}} = \text{const.},$$

which leads to Eq. (S7). As $\frac{R_{01}(t)p_1(t)}{R_{10}(t)p_0(t)} = c$ with $c > 1$ directly meets the condition in Eq. (S9), we complete the proof. For the opposite case $p_0(\tau) \leq p_0(0)$, we reach the same conclusion with $c < 1$. $\qquad\square$

Now we find the explicit form of the dynamics of a bit system satisfying the equality of Eq. (S6) with the symmetric KLD bound. Equation (S7) leads to

$$\dot{p}_0(t) = -R_{10}(t)p_0(t) + R_{01}(t)p_1(t) = (c - 1)R_{10}(t)p_0(t). \tag{S12}$$

The solution of Eq. (S12) is

$$p_0(t) = p_0(0)e^{(c-1)\int_0^t R_{10}(t')dt'}. \tag{S13}$$

Consequently, we have

$$R_{01}(t) = cR_{10}(t)\frac{p_0(t)}{p_1(t)} = \frac{cR_{10}(t)p_0(0)}{e^{-(c-1)\int_0^t R_{10}(t')dt'} - p_0(0)}. \tag{S14}$$

The relevant physical quantities can also be simplified as

$$\dot{\Sigma}(t) = [R_{01}(t)p_1(t) - R_{10}(t)p_0(t)]\ln\left[\frac{R_{01}(t)p_1(t)}{R_{10}(t)p_0(t)}\right] = (c - 1)(\ln c)R_{10}(t)p_0(t)$$

$$A(t) = R_{10}(t)p_0(t) + R_{01}(t)p_1(t) = (c + 1)R_{10}(t)p_0(t).$$

For a special case of a bit erasing process $\boldsymbol{p}(0) = (1/2, 1/2)$ and $\boldsymbol{p}(\tau) = (1 - \epsilon, \epsilon)$ with a constant transition rate $R_{10}(t) = 1$, we have

$$p_0(t) = \frac{1}{2}e^{(c-1)t},$$

$$R_{01}(t) = \frac{2c}{2e^{-(c-1)t} - 1},$$

and the entropy production rate and the dynamical activity become

$$\dot{\Sigma}(t) = (c - 1)(\ln c)p_0(t) = \frac{(c-1)(\ln c)}{2}e^{(c-1)t}, \tag{S15}$$

$$A(t) = (c + 1)p_0(t) = \frac{c+1}{2}e^{(c-1)t}. \tag{S16}$$

5

In order to reset the memory with error ϵ, the required process time τ is given as

$$p_0(\tau) = 1 - \epsilon = \frac{1}{2}e^{(c-1)\tau} \implies \tau = \frac{\ln\left[2(1-\epsilon)\right]}{c-1}. \tag{S17}$$

Then, the total entropy production throughout the process is given as

$$\begin{aligned}
\Sigma &= \int_0^\tau \dot{\Sigma}(t)dt \\
&= \frac{(c-1)(\ln c)}{2} \int_0^\tau e^{(c-1)t}dt \\
&= \left(\frac{\ln c}{2}\right)(1 - 2\epsilon),
\end{aligned} \tag{S18}$$

or equivalently, $\Sigma = \frac{(1-2\epsilon)\ln\left[\frac{\ln(2(1-\epsilon))}{\tau}+1\right]}{2}$. From Eqs. (S16) and (S17), the total activity becomes

$$A_{\text{tot}} = \int_0^\tau A(t)dt = \frac{c+1}{2(c-1)}(1 - 2\epsilon), \tag{S19}$$

which leads to $v^{-1} = A_{\text{tot}}/\ell = (c+1)/(c-1)$ with $c > 1$. For fixed ϵ, $v^{-1} \to 1^+$ is achieved in $c \to \infty$ limit, resulting in $\tau \to 0$ from Eq. (S17). This indicates that the process should be highly irreversible and fast for approaching $v^{-1} \to 1^+$ limit. On the other hand, $v^{-1} \to \infty$ is attained via a nearly reversible process $(c \to 1^+)$ with very large τ, which is the quasistatic process. The result of the optimal protocol is plotted in Fig. S1 and Figs. 1(c) and 1(d) of the main text.

III. LANDAUER'S BOUND FOR A COARSE-GRAINED BIT SYSTEM

Here we show the entropy production (EP) of an original system is always larger than or equal to the EP of the associated coarse-grained bit system and also satisfies the finite-time Landauer's bound derived in the main text. We follow the derivation in Ref. [3]. Consider a physical system consisting of N discrete states, where the probability of each state n is denoted by $p_n(t)$ for $n \in \{1, 2, \cdots, N\}$. Dynamics of the system is described by the following master equation:

$$\dot{p}_n(t) = \sum_{m(\neq n)} \left[R_{nm}(t)p_m(t) - R_{mn}p_n(t)\right], \tag{S20}$$

6

where $R_{nm}(t)$ is the transition rate from state m to n. The EP rate $\dot{\Sigma}(t)$ of this system is written as

$$\dot{\Sigma}(t) = \frac{k_{\mathrm{B}}}{2} \sum_{n \neq m} [R_{nm}(t)p_m(t) - R_{mn}(t)p_n(t)] \ln \left[\frac{R_{nm}(t)p_m(t)}{R_{mn}(t)p_n(t)} \right], \tag{S21}$$

where k_{B} is the Boltzmann constant.

These N microstates can be divided into two sets, Ω_0 and Ω_1. Then the system can be treated as a coarse-grained bit system by regarding the system being in state "0" ("1") when its microstate belongs to Ω_0 (Ω_1). The probability of this coarse-grained bit system $p_a^{\mathrm{cg}}(t)$ for $a \in \{0, 1\}$ is given by

$$p_a^{\mathrm{cg}}(t) = \sum_{i \in \Omega_a} p_i(t). \tag{S22}$$

The master equation for the coarse-grained probability $p_a^{\mathrm{cg}}(t)$ can be derived as follows:

$$\begin{aligned}
\dot{p}_a^{\mathrm{cg}} = \sum_{n \in \Omega_a} \dot{p}_i(t) &= \sum_{n \in \Omega_a} \sum_{m(\neq n)} [R_{nm}(t)p_m(t) - R_{mn}(t)p_n(t)] \\
&= \sum_{n \in \Omega_a} \left(\sum_{m \in \Omega_{\bar{a}}} + \sum_{m(\neq n) \in \Omega_a} \right) [R_{nm}(t)p_m(t) - R_{mn}(t)p_n(t)] \\
&= \sum_{n \in \Omega_a} \sum_{m \in \Omega_{\bar{a}}} \left[R_{nm}(t) \frac{p_m(t)}{p_{\bar{a}}^{\mathrm{cg}}(t)} p_{\bar{a}}^{\mathrm{cg}}(t) - R_{mn}(t) \frac{p_n(t)}{p_a^{\mathrm{cg}}(t)} p_a^{\mathrm{cg}}(t) \right] \\
&= R_{a\bar{a}}^{\mathrm{cg}}(t) p_{\bar{a}}^{\mathrm{cg}}(t) - R_{\bar{a}a}^{\mathrm{cg}}(t) p_a^{\mathrm{cg}}(t), \tag{S23}
\end{aligned}$$

where $R_{a\bar{a}}^{\mathrm{cg}}(t) = \sum_{n \in \Omega_a} \sum_{m \in \Omega_{\bar{a}}} R_{nm}(t)p_m(t)/p_{\bar{a}}^{\mathrm{cg}}(t)$. For deriving the third equality of Eq. (S23), we use

$$\sum_{n \in \Omega_a} \sum_{m(\neq n) \in \Omega_a} [R_{nm}(t)p_m(t) - R_{mn}(t)p_n(t)] = \sum_{n \neq m \in \Omega_a} R_{nm}(t)p_m(t) - \sum_{n \neq m \in \Omega_a} R_{mn}(t)p_n(t) = 0. \tag{S24}$$

The first summation in Eq. (S24) is over intra-transitions taking place inside the same Ω_a, which has no contribution to the change of p_a^{cg}, thus it should be zero. From the master equation of p_a^{cg}, Eq. (S23), we can write the expression of the EP rate $\dot{\Sigma}^{\mathrm{cg}}(t)$ for the coarse-grained bit system as

$$\dot{\Sigma}^{\mathrm{cg}}(t) = \frac{k_{\mathrm{B}}}{2} \sum_a [R_{a\bar{a}}^{\mathrm{cg}}(t) p_{\bar{a}}^{\mathrm{cg}}(t) - R_{\bar{a}a}^{\mathrm{cg}}(t) p_a^{\mathrm{cg}}(t)] \ln \left[\frac{R_{a\bar{a}}^{\mathrm{cg}}(t) p_{\bar{a}}^{\mathrm{cg}}(t)}{R_{\bar{a}a}^{\mathrm{cg}}(t) p_a^{\mathrm{cg}}(t)} \right]. \tag{S25}$$

The EP of the original system (S21) can be separated into two parts as follows:

$$\dot{\Sigma}(t) = \frac{k_{\mathrm{B}}}{2} \sum_a \left(\sum_{n \in \Omega_a} \sum_{m \in \Omega_{\bar{a}}} + \sum_{n \neq m \in \Omega_a} \right) [R_{nm}(t)p_m(t) - R_{mn}(t)p_n(t)] \ln \left[\frac{R_{nm}(t)p_m(t)}{R_{mn}(t)p_n(t)} \right]$$

$$= \dot{\Sigma}^{\mathrm{inter}}(t) + \dot{\Sigma}^{\mathrm{intra}}(t), \tag{S26}$$

where $\dot{\Sigma}^{\mathrm{inter}}(t)$ is the EP induced by inter-transitions between microstates n and m belonging to different sets such as $n \in \Omega_a$ and $m \in \Omega_{\bar{a}}$ for $a \neq \bar{a}$ defined as

$$\dot{\Sigma}^{\mathrm{inter}}(t) \equiv \frac{k_{\mathrm{B}}}{2} \sum_a \sum_{n \in \Omega_a} \sum_{m \in \Omega_{\bar{a}}} [R_{nm}(t)p_m(t) - R_{mn}(t)p_n(t)] \ln \left[\frac{R_{nm}(t)p_m(t)}{R_{mn}(t)p_n(t)} \right] \tag{S27}$$

and $\Sigma^{\mathrm{intra}}(t)$ is the EP induced by intra-transitions between states belonging to the same Ω_a defined as

$$\dot{\Sigma}^{\mathrm{intra}}(t) \equiv \frac{k_{\mathrm{B}}}{2} \sum_a \sum_{n \neq m \in \Omega_a} [R_{nm}(t)p_m(t) - R_{mn}(t)p_n(t)] \ln \left[\frac{R_{nm}(t)p_m(t)}{R_{mn}(t)p_n(t)} \right]. \tag{S28}$$

We can easily check that $\dot{\Sigma}^{\mathrm{inter}}(t)$ and $\dot{\Sigma}^{\mathrm{intra}}(t)$ are non-negative from the inequality $(a - b)\ln(a/b) \geq 0$ for $a, b > 0$. Therefore, we get

$$\dot{\Sigma}(t) = \dot{\Sigma}^{\mathrm{inter}}(t) + \dot{\Sigma}^{\mathrm{intra}}(t) \geq \dot{\Sigma}^{\mathrm{inter}}(t)$$

$$\geq \frac{k_{\mathrm{B}}}{2} \sum_a \left[\sum_{n \in \Omega_a} \sum_{m \in \Omega_{\bar{a}}} \{R_{nm}(t)p_m(t) - R_{mn}(t)p_n(t)\} \ln \left\{ \frac{\sum_{n \in \Omega_a} \sum_{m \in \Omega_{\bar{a}}} R_{nm}(t)p_m(t)}{\sum_{n \in \Omega_a} \sum_{m \in \Omega_{\bar{a}}} R_{mn}(t)p_n(t)} \right\} \right]$$

$$= \frac{k_{\mathrm{B}}}{2} \sum_a (R_{a\bar{a}}^{\mathrm{cg}}(t)p_{\bar{a}}^{\mathrm{cg}}(t) - R_{\bar{a}a}^{\mathrm{cg}}(t)p_a^{\mathrm{cg}}(t)) \ln \frac{R_{a\bar{a}}^{\mathrm{cg}}(t)p_{\bar{a}}^{\mathrm{cg}}(t)}{R_{\bar{a}a}^{\mathrm{cg}}(t)p_a^{\mathrm{cg}}(t)}$$

$$= \dot{\Sigma}^{\mathrm{cg}}(t). \tag{S29}$$

The log-sum inequality, $\sum_n u_n \ln(u_n/v_n) \geq (\sum_n u_n) \ln(\sum_n u_n / \sum_n v_n)$ for any $u_n, v_n \geq 0$, is used to derive the second line of Eq. (S29). From Eq. (S29), we have $\Sigma(\tau) - \Sigma(0) \geq \Sigma^{\mathrm{cg}}(\tau) - \Sigma^{\mathrm{cg}}(0)$. As we can set $\Sigma(0) = \Sigma^{\mathrm{cg}}(0) = 0$, we finally have the inequality:

$$\Sigma(\tau) \geq \Sigma^{\mathrm{cg}}(\tau). \tag{S30}$$

As $\Sigma(\tau)$ is always larger than or equal to $\Sigma^{\mathrm{cg}}(\tau)$, $\Sigma(\tau)$ is also bounded by the finite-time Landauer's bound derived in the main text as

$$\Sigma(\tau) \geq \Sigma^{\mathrm{cg}}(\tau) \geq \ell^{\mathrm{cg}} \frac{h(v^{\mathrm{cg}})}{v^{\mathrm{cg}}} \equiv B_h(\ell^{\mathrm{cg}}, v^{\mathrm{cg}}), \tag{S31}$$

8

where ℓ^{cg} and v^{cg} are distribution distance and jump distance, respectively, evaluated in the coarse-grained bit system. It is worthwhile to note that $\Sigma(\tau)$ is not tightly bounded by $B_h(\ell^{\mathrm{cg}}, v^{\mathrm{cg}})$ as long as $\Sigma^{\mathrm{intra}}(\tau)$ is not negligible. The above derivation for a discrete system is straightforwardly applicable to a continuous case [3].

[1] G. L. Gilardoni, Comptes Rendus Mathematique **343**, 763 (2006).

[2] I. Sason, IEEE Transactions on Information Theory **61**, 701 (2014).

[3] Y.-Z. Zhen, D. Egloff, K. Modi, and O. Dahlsten, Phys. Rev. Lett. **127**, 190602 (2021).