Thermodynamic Cost Due to Changing the Initial Distribution Over States

Artemy Kolchinsky
David H. Wolpert

SFI WORKING PAPER: 2016-07-014

SANTA FE INSTITUTE

Thermodynamic cost due to changing the initial distribution over states

Artemy Kolchinsky and David H. Wolpert
Santa Fe Institute, Santa Fe, NM

We consider nonequilibrium systems that obey local detailed balance and are driven by an external system such that no work is dissipated for some initial distribution over states $x \in X$, $q(X)$. We show that in general work is dissipated under that driving if instead the initial distribution is some $r(X) \neq q(X)$, calculating that amount of dissipated work. We then use this result to investigate the thermodynamics of computation. Specifically, we suppose a Markov partition of X into a set of coarse-grained "computational states" labelled by $v \in V$, and identify the dynamics over those computational states as a (possibly noisy) "computation" that runs on the system. We identify the initial distribution over computational states as the distribution over inputs to the computer. We calculate the work that such a computer dissipates, if it is designed to dissipate no work for some input distribution $q(V)$ but is instead used with a different input distribution $r(V)$. This dissipated work is an extra thermodynamic cost of computation, in addition to the well-known Landauer's cost. We also calculate the extra expected dissipated work if there is a probability distribution over the possible input distributions rather than a single one.

Introduction.—Recent research has made great progress in understanding the statistical physics of systems far from equilibrium [1–6]. This has resulted in many novel predictions, some of which are now being experimentally confirmed [7–9].

Much of this recent research is based on a modern formulation of the second law of thermodynamics [10, 11]. To present this modern formulation, first define the *nonequilibrium free energy* of any Hamiltonian H and distribution p as

$$\mathcal{F}(H,p) := \langle H \rangle_p - kTS(p)$$

where $S(.)$ indicates Shannon entropy (measured in nats). Consider a system with microstates $x \in X$ whose Hamiltonian depends on an evolving external parameter $\lambda(t)$ for $t \in [0,1]$. While its Hamiltonian evolves, the system is connected to a heat bath at temperature T. The system moves between states according to a sequence of transition matrices that all obey local detailed balance for the associated Hamiltonian [12–15]. We refer to such a sequence of Hamiltonians and transition matrices as a *thermodynamic process* operating on the system.

As shorthand write $H_t \equiv H(\lambda(t))$. Suppose that the system starts with microstate distribution p_0, and that the thermodynamic process sends the initial Hamiltonian H_0 to H_1 and transforms p_0 to p_1. (In this paper we take the units of time to be arbitrary, so time intervals can be arbitrarily long.) Then the modern statement of the second law is

$$\langle W \rangle \geq \Delta\mathcal{F} := \mathcal{F}(H_1,p_1) - \mathcal{F}(H_0,p_0) \qquad (1)$$

where $\langle W \rangle$ is the expected work on the system between $t=0$ and $t=1$. Importantly, Eq. 1 holds even if both p_0 and p_1 are far from equilibrium for the associated Hamiltonians, H_0 and H_1, respectively.

The dissipated work (which we call *dissipation*) is

$$W_d := \langle W \rangle - \Delta\mathcal{F}$$

This is the work done on a system that cannot be thermodynamically recovered. (It should not be confused with the dissipated heat, which is the total energy transferred to the heat bath). Dissipation is zero iff the process is thermodynamically reversible. In principle, it is always possible to construct such a thermodynamically reversible process transforming any given p_0, H_0 to any p_1, H_1. For instance, this can be done via an appropriate "quench-and-relax" process [11, 16, 17].

However, as we show in this paper, in general any thermodynamic process $\mathscr{P}$ that is thermodynamically reversible when run with one initial distribution will not be thermodynamically reversible for a different initial distribution. More precisely, let the random variable X_t indicate the state of the system at time t, and assume that $\mathscr{P}$ is thermodynamically reversible for some initial distribution over states $q(X_0)$, which we call the *prior* distribution. Define $q(X_1)$ as the associated distribution over states produced after running $\mathscr{P}$ on the system. Suppose though that the initial states are drawn from some *environment distribution* $r(X_0) \neq q(X_0)$, and define $r(X_1)$ as the associated distribution produced after running $\mathscr{P}$ on the system.

Our first result is that the dissipation in this scenario is given by

$$W_d = kT\left[D\left(r(X_0)\|q(X_0)\right) - D\left(r(X_1)\|q(X_1)\right)\right] \qquad (2)$$

where k is Boltzmann's constant and $D(\cdot||\cdot)$ is the Kullback-Leibler divergence [18]. We refer to W_d as the *incorrect prior dissipation*. Intuitively, it reflects the contraction between the environment and prior distributions under the action of $\mathscr{P}$ [19, 20]. Note that this contraction depends only on $r(X_0), q(X_0), r(X_1)$ and $q(X_1)$. All physical details of how $\mathscr{P}$ manages to transform $q(X_0) \to q(X_1)$ and $r(X_0) \to r(X_1)$ are irrelevant.

In many situations there is not a single, fixed environment distribution. Rather there is a distribution $p(\theta)$ where each θ labels a different environment distribution,

and the actual environment distribution is formed by randomly sampling p. As we show below, this increases the expected dissipation, by the (non-negative) amount that the mutual information between the environment θ and the microstate x drops as the system evolves from $t = 0$ to $t = 1$.

These results have important consequences for our understanding of the thermodynamics of computation. Ever since the pioneering analyses of Landauer and Bennett [21, 22], a *computer* has been defined as any physical system with microstates $x \in X$ undergoing a thermodynamic process $\mathscr{P}$ as described above, together with a coarse-graining of X into a set of Computational States (CSs) with labels $v \in V$ (sometimes called "information bearing degrees of freedom" [23]). $\mathscr{P}$ induces a stochastic dynamics over X, and the *computation* is identified with the associated stochastic dynamics over CSs. We write this dynamics as $\pi(v_1|v_0)$, and say that it maps *inputs* v_0 to *outputs* v_1. Here we add the requirement that the computer be time-translation-invariant, in that for any CS $v \in V$, the conditional distribution $P(x_t|v_t)$ is the same at the end of an iteration of the computer ($t = 1$) and the beginning ($t = 0$). (This can be viewed as requiring that each iteration of a computer completes a full thermodynamic cycle over the set of states within each CS, only modifying the distribution over those CSs.) The canonical example of a computer is a laptop, with the computation it runs being the firmware modifying bit patterns in its memory [23].

The early work by Landauer et al. focused on *bit-erasure*. This is the simplest type of computation; a 2-to-1 map that sends the two initial inputs of a binary space V to a single output. Using semi-formal reasoning, early researchers argued that any thermodynamic process $\mathscr{P}$ that implements bit-erasure has a thermodynamic cost of at least $kT \ln 2$, a result known as *Landauer's bound*.

Recent research has significantly clarified and expanded this early work [10–14, 24–31]. For instance, early work equated *logical* irreversibility (such as bit-erasure) with thermodynamic irreversibility (dissipation), suggesting that the $kT \ln 2$ cost of bit-erasure reflects dissipated work. It is now known that logically irreversible processes can be carried out thermodynamically reversibly [11, 32]. Rather than concern dissipation, it is now known that Landauer's bound is the minimal work needed by any thermodynamically reversible process that implements bit-erasure, when the distribution over inputs is uniform.

Bit-erasure is an example of a computation in which the outputs do not depend on input values. Most of the recent research on the thermodynamics of computation has been concerned with such input-independent computations. However this excludes most computations of interest. Other recent research allows the output to depend on the input [33, 34], but imposes other strong restrictions on the kind of computations being modeled [16].

We recently analyzed the thermodynamics of arbitrary computations taking inputs to outputs, in particular including those in which the outputs depend on the inputs [16, 17]. We showed how, given any desired computation $\pi(v_1|v_0)$ and input distribution $q(V_0)$, to construct a computer $\mathscr{C}$ that computes π and that results in no dissipation when inputs are drawn from $q(V_0)$. However in general this computer $\mathscr{C}$ would dissipate work if applied with some different input distribution. It was not known if such dissipation arising due to changing the input distribution is peculiar to the kinds of computers we constructed, or would apply to *any* computer.

Below we answer this question, proving that no matter how the computer is constructed, it is subject to incorrect prior dissipation over its computational states. Precisely, consider a computer running a computation $\pi(v_1|v_0)$ that results in no dissipation when input values are drawn from prior distribution $q(V_0)$. Suppose that we instead run the computer with inputs drawn from some $r(V_0)$. (We will sometimes refer to this as a *user* distribution.) Then the same formula as Eq. 2 again applies, if we substitute distributions over computational states for distributions over microstates. Moreover, as for the case of general thermodynamic processes discussed above, when there is a probability distribution P over user distributions, the computer's expected dissipation will equal the drop in the mutual information between the user and the computational state as the computation proceeds.

This formula for the dissipation of a computer depends only on q_0, r_0 and π. It is independent of how the computer works, holding for laptops, ratchet-based information processing systems [33, 34], any physical instantiation of a "Maxwell's demon" [35] like a Szilard engine, systems that evolve semi-statically, etc.

As an example, suppose user Alice wishes to repeatedly run some computation π, with inputs drawn from a distribution $q(V_0)$. It is possible to construct a computer $\mathscr{C}_{Alice}$ that implements π without any dissipation, as long as its inputs are truly sampled from $q(V_0)$. However in general, that same computer *would* dissipate work if it were instead run with its inputs selected by user Bob according to some distribution $r(V_0)$. The total work required by Bob when using computer $\mathscr{C}_{Alice}$ to run the computation π is the sum of this dissipated work and the work required to run π with no dissipation on input distribution $r(V_0)$ (given by the nonequilibrium free energy difference between starting and ending distributons).

Dissipation due to incorrect priors is a novel thermodynamic cost of computation, over and above the cost given by Landauer's bound. Indeed, in its modern, full formulation, the Landauer bound is the drop in nonequilibrium free energy in an iteration of the computer. This can be negative for some user distributions and (noisy) computations [21, 22]. In contrast, incorrect prior dissipation is a true cost, in the sense that it is always non-negative. Another difference is that while the Lan-

dauer bound reflects loss of information about input values during the course of a computation, incorrect prior dissipation reflects uncertainty at the time when one constructs the computer, about what the computer's input distribution will be.

Note that if a computation π is an invertible deterministic map, then the drop in KL divergence given by Eq. 2 is always zero, regardless of r and q. In other words, if the computation is logically reversible, and $\mathscr{C}$ is thermodynamically reversible for some user distribution q, then dissipation (reflecting irreversible work due to designing the computer for the wrong user distribution) is zero for all user distributions. Having a distribution over users makes no difference. (Similarly, the Landauer cost, reflecting reversible work, is zero for such situations.)

In the following sections we first introduce notation, and then derive Eq. 2 for thermodynamic processes. Next we analyze the dissipation when there is a distribution over environment distributions. We end by deriving our results for computers. Supplementary Material (SM) [36] contains miscellaneous proofs, along with an analysis of incorrect prior dissipation for feedback processes [11].

Background and notation.—We use lower case letters (e.g., z) to indicate an outcome of a random variable, and upper case letters (e.g., Z) to indicate either a random variable, or the set of possible outcomes, depending on the context. We also use bold italics $\boldsymbol{z}$ and $\boldsymbol{Z}$ to refer to outcomes / variables that correspond to trajectories of states rather than to (single-instant) states. Typically p, q, and r refer to probability distributions. We sometimes refer to a conditional distribution $p(b|a)$ as a "map" if it defines a single-valued function from a to b.

As discussed above, we consider both thermodynamic processes over a space of microstates X, and also computers, consisting of the pair of such a thermodynamic process and a coarse-graining of X into computational states V. The random variables X_t and V_t will be used to indicate the microstate and computational state of the computer at time t. We will sometimes use notation like p_0 to indicate the distribution $p(X_0)$ or $p(V_0)$, as will be clear from context (and similarly for p_1).

Following [37], we refer to the thermodynamic process's sequence of Hamiltonians $\{H_0, \ldots, H_1\}$ as the *forward protocol*. When the system is started from initial microstate x_0, the process stochastically produces a trajectory of states $\boldsymbol{x}_{0..1} := \{x_0, \ldots, x_1\}$ with conditional probability $p\left(\boldsymbol{x}_{0..1}|x_0\right)$. We refer to this conditional probability as the (forward) *driven dynamics*.

We define the *reverse protocol* as the aforementioned sequence of Hamiltonians taken in reverse order. For notational convenience, we indicate that reverse sequence of Hamiltonians by $\tilde{H}_t$ for $0 \leq t \leq 1$ with $\tilde{H}_t := H_{1-t}$. We also write the reverse of a trajectory $\boldsymbol{x}_{0..1} = \{x_0, \ldots, x_1\}$ as $\tilde{\boldsymbol{x}}_{1..0} := \{x_1, \ldots, x_0\}$. We write the conditional probability of observing the trajectory $\tilde{\boldsymbol{x}}_{1..0}$ under the reverse protocol as $\tilde{p}\left(\tilde{\boldsymbol{x}}_{1..0}|x_1\right)$.

Crooks [37] derived a relation between the work done to implement the trajectory $\boldsymbol{x}_{0..1}$ under the forward protocol, indicated by $W\left(\boldsymbol{x}_{0..1}\right)$, the conditional probability of that trajectory under the forward protocol, and the conditional probability of its reverse under the reverse protocol:

$$\exp\left(\beta\left(W\left(\boldsymbol{x}_{0..1}\right) - \Delta H\right)\right) = \frac{p\left(\boldsymbol{x}_{0..1}|x_0\right)}{\tilde{p}\left(\tilde{\boldsymbol{x}}_{1..0}|x_1\right)} \tag{3}$$

where $\beta = \frac{1}{kT}$ and $\Delta H := H_1\left(x_1\right) - H_0\left(x_0\right)$.

Dissipation due to incorrect priors.—Eq. 3 only involves conditional probabilities, leaving us free to choose any initial distribution of the reverse protocol, $\tilde{p}(x_1)$. We set it to be the same as the final distribution of the forward protocol, writing $\tilde{p}\left(x_1\right) = p\left(x_1\right)$. So for any initial distribution $p(x_0)$, the (non-conditional) probabilities of the trajectories under the forward and reverse protocols are $p\left(\boldsymbol{x}_{0..1}\right) = p\left(x_0\right) p\left(\boldsymbol{x}_{0..1}|x_0\right)$ and $\tilde{p}\left(\tilde{\boldsymbol{x}}_{1..0}\right) = \tilde{p}\left(x_1\right) \tilde{p}\left(\tilde{\boldsymbol{x}}_{1..0}|x_1\right) = p\left(x_1\right) \tilde{p}\left(\tilde{\boldsymbol{x}}_{1..0}|x_1\right)$, and so

$$\begin{aligned}
D\left(p\left(\boldsymbol{X}_{0..1}\right) \| \tilde{p}\left(\tilde{\boldsymbol{X}}_{1..0}\right)\right) &= \left\langle \ln \frac{p\left(X_0\right) p\left(\boldsymbol{X}_{0..1}|X_0\right)}{\tilde{p}\left(X_1\right) \tilde{p}\left(\tilde{\boldsymbol{X}}_{1..0}|X_1\right)} \right\rangle \\
&= \left\langle \ln \frac{p\left(X_0\right)}{\tilde{p}\left(X_1\right)} + \beta\left(W(\boldsymbol{X}_{0..1}) - \Delta H\right) \right\rangle \\
&= S\left(p\left(X_1\right)\right) - S\left(p\left(X_0\right)\right) + \beta\left(\langle W(\boldsymbol{X}_{0..1})\rangle - \langle \Delta H \rangle\right) \\
&= \beta\left(\langle W(\boldsymbol{X}_{0..1})\rangle - \Delta \mathcal{F}\right) \\
&= \beta W_d\left(p\right)
\end{aligned} \tag{4}$$

where angle-brackets indicate expectations under $p\left(\boldsymbol{X}_{0..1}\right)$, and $W_d\left(p\right)$ is the average work dissipated under the forward protocol when initial states are drawn from $p\left(X_0\right)$ (see also [15, 38]).

KL divergence – and hence dissipation – is always non-negative [39]. Furthermore, KL divergence is zero iff its two arguments are equal. So dissipation is 0 iff the probability of every forward trajectory $\boldsymbol{x}_{0..1}$ under the forward protocol is the same as the probability of $\tilde{\boldsymbol{x}}_{1..0}$ under the reverse protocol. Loosely speaking, no work is dissipated iff it is impossible to tell if one is "watching a movie of the system evolving forward in time" or instead "watching the movie in reverse".

As discussed in the introduction, we hypothesize that there is some *prior* initial distribution $q\left(X_0\right)$ that the protocol is "designed for", in the sense that there is zero dissipation when that distribution evolves according to the driven dynamics $p\left(\boldsymbol{x}_{0..1}|x_0\right)$ (the reason for calling q a 'prior' is given below). However we allow the protocol to be used instead with an initial distribution $r\left(X_0\right)$.

Define the $t = 1$ distributions for the reverse protocol as $\tilde{q}\left(x_1\right) = q\left(x_1\right)$ and $\tilde{r}\left(x_1\right) = r\left(x_1\right)$ as above. Since the process is dissipationless for q by hypothesis, $p\left(\boldsymbol{x}_{0..1}|x_0\right) q(x_0) = \tilde{p}\left(\tilde{\boldsymbol{x}}_{1..0}|x_1\right) \tilde{q}(x_1)$ for all trajectories. However, by Eq. 4, when the dynamics are started with

initial distribution r_0,

$$\begin{aligned}\beta W_d(r) &= D\left(r(\boldsymbol{X}_{0..1}) \,\|\tilde{r}\left(\tilde{\boldsymbol{X}}_{1..0}\right)\right) \\ &= \left\langle \ln \frac{r(X_0)\, p(\boldsymbol{X}_{0..1}|X_0)}{\tilde{r}(X_1)\,\tilde{p}\left(\tilde{\boldsymbol{X}}_{1..0}|X_1\right)} \right\rangle_{r(\boldsymbol{X}_{0..1})} \\ &= \left\langle \ln \frac{r(X_0)\,\tilde{q}(X_1)}{\tilde{r}(X_1)\,q(X_0)} \right\rangle_{r(\boldsymbol{X}_{0..1})} \\ &= D(r(X_0)\|q(X_0)) - D(r(X_1)\|q(X_1)) \end{aligned} \quad (5)$$

So the dissipation equals the drop in our ability to distinguish whether the initial distribution was $q(X_0)$ or $r(X_0)$, as the system evolves from that initial distribution.

By the KL data processing inequality [40, Lemma 3.11], this drop is non-negative. It achieves its minimum value of 0 if the driven dynamics $p(x_1|x_0)$ form an invertible map, regardless of the prior $q(X)$. Moreover, in the special case that $q(X)$ has full support, there is a always some environment distribution $r(X)$ with strictly positive dissipation iff the driven dynamics are not an invertible map (proof in SM [36]). Furthermore, if the dynamics over X obeys the strong data processing inequality for KL divergence [41], then dissipation will always be strictly positive whenever $r \neq q$. Concretely, if we design a physical device to implement such a dynamics $p(\boldsymbol{x}_{0..1}|x_0)$ with no dissipation for initial distribution $q(x_0)$, then there will be dissipation if the device is instead run in an environment with initial distribution $r(x_0) \neq q(x_0)$. In this sense, it is typically the case that there is at most one environment that will result in no dissipation for any given device.

For another perspective on Eq. 5, note that by the chain rule for KL divergence [39, Eq. 2.67],

$$\begin{aligned} &D(r(X_0, X_1)\|q(X_0, X_1)) \\ &\quad = D(r(X_0)\|q(X_0)) + D(r(X_1|X_0)\|q(X_1|X_0)) \\ &\quad = D(r(X_1)\|q(X_1)) + D(r(X_0|X_1)\|q(X_0|X_1)) \end{aligned} \quad (6)$$

However, since $r(x_1|x_0) = q(x_1|x_0) = p(x_1|x_0)$, $D(r(X_1|X_0)\|q(X_1|X_0)) = 0$. So by Eq. 5,

$$\beta W_d = D(r(X_0|X_1)\|q(X_0|X_1))$$

(See also [17].) In this expression $r(x_0|x_1)$ and $q(x_0|x_1)$ are Bayesian posterior probabilities of the initial state conditioned on the final state, for the assumed priors $r(X_0)$ and $q(X_0)$ respectively, and the shared likelihood function $p(x_1|x_0)$.

Dissipation due to an uncertain environment.—Suppose we have a set of identical copies of a physical device which all implement the dynamics $p(\boldsymbol{x}_{0..1}|x_0)$, and are all dissipationless for the prior $q(X_0)$. Suppose we also have a set of different environments, indicated by the set Θ. Each environment $\theta \in \Theta$ gets one of the identical copies of our device, and then runs the process $p(\boldsymbol{x}_{0..1}|x_0)$ with an environment-specific initial distribution, indicated by $r(X_0|\theta)$.

Running the copy of the device in each environment will result in an associated amount of dissipation, determined by that environment's initial state distribution. So a probability distribution $w(\theta)$ over environments specifies an expected amount of dissipation, given by

$$\begin{aligned} &\beta \langle W_d \rangle = \\ &\sum_\theta w(\theta) \left[D(r(X_0|\theta)\|q(X_0)) - D(r(X_1|\theta)\|q(X_1))\right] \\ &\quad = D(r(X_0|\Theta)\|q(X_0)) - D(r(X_1|\Theta)\|q(X_1)) \end{aligned} \quad (7)$$

In SM [36], we provide a lower bound on this expected dissipation,

$$\beta \langle W_d \rangle \geq I(X_0;\Theta) - I(X_1;\Theta) \quad (8)$$

where I is the mutual information. The RHS is the drop in the mutual information between the environment and the system microstate from $t = 0$ to $t = 1$. It reflects loss of information in the system state about what the environment is under the action of the stochastic dynamics (see also [17]). These two mutual informations equal the initial and final Jensen-Shannon divergences of the distribution over environments [42]. So the RHS can also be interpreted as the degree to which all environment distributions become indistinguishable as the system unfolds.

One way that the bound in Eq. 8 is reached is if the prior distribution $q(x_0)$ equals the expected environment distribution, $\sum_\theta w(\theta) r(x_0|\theta)$. It is also reached for certain thermodynamic processes no matter what the initial distribution is. For example, if $p(x_1|x_0)$ is an invertible single-valued function, the lower bound is always 0 and is reached even if $q(x_0) \neq \sum_\theta w(\theta) r(x_0|\theta)$.

There is an important subtlety in this analysis, involving the ways expectations are formed. The calculation of Eq. 8 concerns a scenario where each environment gets its own thermodynamic process (i.e., its own physical system with state space X, its own external system acting on X, and its own heat bath). It just so happens that all those thermodynamical processes are mathematically (though not physically) identical. The average dissipation in this scenario is the expectation, over environments θ, of the gap between the average work required to run those copies of the same process with an x_0 sampled from $r(x_0|\theta)$, and the associated change in non-equilibrium free energy.

As an example of this scenario, suppose that the thermodynamic process, $\mathscr{P}$, performs bit-erasure over a binary X, not dissipating any work if the initial distribution over X is uniform. So the prior for this $\mathscr{P}$ is the distribution $q(x_0 = 0) = q(x_0 = 1) = 1/2$. Suppose as well that there are two environments, θ_1 and θ_2, and take $r(x_0 = 0|\theta_1) = 1$ while $r(x_0 = 0|\theta_2) = 0$. Then for a uniform $w(\theta)$, the expected dissipated work is $kT \ln[2]$.

In contrast, consider the closely related scenario where we use a single thermodynamic process $\mathscr{P}$, whose initial states are generated by first sampling $w(\theta)$ and then sampling $r(X_0|\theta)$. If the prior distribution for $\mathscr{P}$ is $\sum_\theta w(\theta) r(x_0|\theta)$, there is no dissipation in this scenario, even though there is dissipation in the first scenario. In particular, for the same $\mathscr{P}$, X, $w(\theta)$ and pair of environment distributions as in our example of the first scenario, which results in dissipation of $kT \ln[2]$ in that scenario, no work is dissipated in this second scenario.

Ultimately, the reason for this distinction in the amount of dissipation in the two scenarios is that in the first scenario there are multiple physical systems but in the second scenario there is only one. In the first scenario, the dissipated work for each instance of θ is the difference between the amount of work used to drive that environment's system and the minimal amount of work that *could* have been used if we had used a different system, optimized for that environment's input distribution. In the second scenario, there is no such difference, since there is only a single device, and it is optimized for the distribution of inputs that it receives.

A new thermodynamic cost of computation.—Recall from the introduction that we define a computer that runs π as any combination of a thermodynamic process over X and coarse-graining of X into CSs V with several extra conditions. We write those conditions as follows:

1) We overload notation to write the initial distribution over macrostates (sometimes called the *user* or *input* distribution) as $q(v_0)$;

2) We write the initial distribution over microstates x_0 given any initial CS v_0 as the distribution s, i.e., as $s(x = x_0|v = v_0)$;

3) The microstate dynamics $p(x_1|x_0)$ implements π in the sense that for any v_0, v_1,

$$\sum_{x_0, x_1} p(x_1|x_0)\, s(x_0|v_0) = \pi(v_1|v_0)$$

As a result the *output* distribution at $t = 1$ is given by $q(v_1) = \sum_{v_0} q(v_0)\, \pi(v_1|v_0)$;

4) The process is *cyclic*, i.e., for all v_1,

$$q(x_1|v_1) = s(x = x_1|v = v_1) \tag{9}$$

(This is slightly stronger than requiring that the coarse-graining be a Markov partition.)

In previous work [17], we showed that for any given π and $q(v_0)$, a computer can be designed that implements π with zero dissipation for user distribution $q(v_0)$. In analogy with the case of dynamics over X, we say that such a q is the *prior* for the computer. To calculate how much work is dissipated if the initial macrostate of a computer with prior $q(x_0)$ is formed by sampling a different user distribution $r(x_0)$, use Eq. 5 to write

$$\begin{aligned}
\beta W_d &= D(r(X_0) \| q(X_0)) - D(r(X_1) \| q(X_1)) \\
&= D(r(V_0, X_0) \| q(V_0, X_0)) - D(r(V_1, X_1) \| q(V_1, X_1)) \\
&= D(r(V_0) \| q(V_0)) - D(r(V_1) \| q(V_1)) \\
&\quad + D(r(X_0|V_0) \| q(X_0|V_0)) - D(r(X_1|V_1) \| q(X_1|V_1))
\end{aligned}$$

where the second line follows because v_0 and v_1 are deterministic functions of x_0 and x_1, and the third line from the chain rule for KL divergence.

Due to condition 2 in the definition of computers, $r(x_0|v_0) = s(x_0|v_0) = q(x_0|v_0)$. So $D(r(X_0|V_0) \| q(X_0|V_0)) = 0$. Similarly, requirement 4 gives $D(r(X_1|V_1) \| q(X_1|V_1)) = 0$. Therefore

$$\beta W_d = D(r(V_0) \| q(V_0)) - D(r(V_1) \| q(V_1)) \tag{10}$$

This minimal dissipation arises no matter how the computer operates, and in particular even if the driving protocol is different from that analyzed in [17]. Recall that the Landauer bound is the work required to run the computer without any dissipation. The dissipation given by Eq. 10 is an extra thermodynamic computation cost, over and above that minimal work given by the Landauer bound, that arises by running a computer with the "wrong prior" for the user. Note though that just as with dynamics over X, incorrect prior dissipation does not occur with all computations. For instance, computations that implement invertible maps have no incorrect prior dissipation.

Similarly to the analysis that led to Eq. 8, suppose we have multiple copies of the same computer, each of which is paired with a different user. If each user $u \in U$ is selected with probability $w(u)$ and has an associated input distribution $r(v_0|u)$, then expected dissipation is bounded by (see SM [36]):

$$\beta \langle W_d \rangle \geq I(V_0; U) - I(V_1; U) \tag{11}$$

This bound is reached when the prior of the computer is $q(v_0) = \sum_u r(v_0|u)\, w(u)$.

Discussion.—In this paper we consider driven nonequilibrium systems with states $x \in X$ that are coupled to a heat bath. We show that if such a process dissipates no work when run with an initial distribution $q(x_0)$, then in general if *will* dissipate work if run with a different distribution $r(x_0) \neq q(x_0)$, providing equations for the precise amount of such dissipated work.

We then extend our analysis to apply to computers by introducing a coarse-graining of the microstate space X into a set of computational states (CSs). In its modern formulation, Landauer's bound gives the minimal work needed to implement a given dynamics over the CSs (i.e., a given computation). This bound is reached with a thermodynamically reversible process *which is designed for an assumed distribution over the initial values of the CSs*. We add the requirement that the computer

execute a complete thermodynamic cycle over the microstates within each CS, and show that as a result, if the computer is used with any initial distribution over the CSs other than the one it is designed for, then in general it will dissipate work. (Similar results hold if we relax the requirement that the computer execute a thermodynamic cycle within each CS.) Physically, if you design a computer to achieve the Landauer bound when it is used by user Alice, in general it will no longer achieve that bound, instead dissipating work, if used by user Bob. This is a new kind of thermodynamic cost of information processing, over and above the cost of Landauer's bound.

There are several relatively straight-forward extensions of these analyses. For example, rather than assume that the computational states comprise a partition of X, we could assume that there is some noise relating microstates and macrostates, i.e., allow a distribution $p(v|x)$ that is nonzero for multiple v's for any given x. As another example, there has been interest in the literature in what are sometimes called "feedback processes" [11] (which differ somewhat from what is meant by the term in the control theory literature). Our analyses can also be extended to such processes (see SM).

Acknowledgements — We would like to thank the Santa Fe Institute for helping to support this research. This paper was made possible through the support of Grant No. TWCF0079/AB47 from the Templeton World Charity Foundation and Grant No. FQXi-RHl3-1349 from the FQXi foundation. The opinions expressed in this paper are those of the authors and do not necessarily reflect the view of Templeton World Charity Foundation.

[1] K. Wiesner, M. Gu, E. Rieper, and V. Vedral, Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science **468**, 4058 (2012).
[2] S. Still, D. A. Sivak, A. J. Bell, and G. E. Crooks, Physical review letters **109**, 120604 (2012).
[3] M. Prokopenko, J. T. Lizier, and D. C. Price, Entropy **15**, 524 (2013).
[4] M. Prokopenko and J. T. Lizier, Nature Scientific reports **4** (2014).
[5] W. H. Zurek, Nature **341**, 119 (1989).
[6] W. H. Zurek, Phys. Rev. A **40**, 4731 (1989).
[7] J. Dunkel, Nature Physics **10**, 409 (2014).
[8] É. Roldán, I. A. Martinez, J. M. Parrondo, and D. Petrov, Nature Physics (2014).
[9] A. Bérut, A. Arakelyan, A. Petrosyan, S. Ciliberto, R. Dillenschneider, and E. Lutz, Nature **483**, 187 (2012).
[10] M. Esposito and C. Van den Broeck, EPL (Europhysics Letters) **95**, 40004 (2011).
[11] J. M. Parrondo, J. M. Horowitz, and T. Sagawa, Nature Physics **11**, 131 (2015).
[12] G. E. Crooks, Physical Review E **60**, 2721 (1999).
[13] G. E. Crooks, Journal of Statistical Physics **90**, 1481 (1998).
[14] C. Jarzynski, Physical Review Letters **78**, 2690 (1997).
[15] U. Seifert, Reports on Progress in Physics **75**, 126001 (2012).
[16] D. H. Wolpert, Entropy **18**, 138 (2016).
[17] D. H. Wolpert, "Extending landauer's bound from bit erasure to arbitrary computation," (2016), submitted. arXiv:1508.05319 [cond-mat.stat-mech].
[18] T. Cover and J. Thomas, *Elements of Information Theory* (Wiley-Interscience, New York, 1991).
[19] R. Ahlswede and P. Gács, The Annals of Probability , 925 (1976).
[20] J. E. Cohen, Y. Iwasa, G. Rautu, M. Beth Ruskai, E. Seneta, and G. Zbaganu, Linear Algebra and its Applications **179**, 211 (1993).
[21] R. Landauer, IBM journal of research and development **5**, 183 (1961).
[22] C. H. Bennett, International Journal of Theoretical Physics **21**, 905 (1982).
[23] C. H. Bennett, Studies In History and Philosophy of Science Part B: Studies In History and Philosophy of Modern Physics **34**, 501 (2003).
[24] P. Faist, F. Dupuis, J. Oppenheim, and R. Renner, arXiv preprint arXiv:1211.1037 (2012).
[25] H. Touchette and S. Lloyd, Physica A: Statistical Mechanics and its Applications **331**, 140 (2004).
[26] T. Sagawa and M. Ueda, Physical review letters **102**, 250602 (2009).
[27] R. Dillenschneider and E. Lutz, Physical review letters **104**, 198903 (2010).
[28] T. Sagawa and M. Ueda, Physical review letters **109**, 180602 (2012).
[29] F. Chejne Janna, F. Moukalled, and C. A. Gómez, International Journal of Thermodynamics **16**, 97 (2013).
[30] M. Esposito and C. Van den Broeck, Physical Review E **82**, 011143 (2010).
[31] B. S. Pollard, Open Systems & Information Dynamics **23**, 1650006 (2016).
[32] T. Sagawa, Journal of Statistical Mechanics: Theory and Experiment **2014**, P03025 (2014).
[33] D. Mandal and C. Jarzynski, Proceedings of the National Academy of Sciences **109**, 11641 (2012).
[34] A. C. Barato and U. Seifert, Physical Review E **90**, 042150 (2014).
[35] W. Thomson, in *R. Soc. Proc*, Vol. 9 (1879) pp. 113–114.
[36] See Supplemental Material at [URL will be inserted by publisher].
[37] G. E. Crooks, Journal of Statistical Physics **90**, 1481 (1998).
[38] R. Kawai, J. M. R. Parrondo, and C. Van den Broeck, Physical review letters **98**, 080602 (2007).
[39] T. M. Cover and J. A. Thomas, *Elements of information theory* (John Wiley & Sons, 2012).
[40] I. Csiszar and J. Körner, *Information theory: coding theorems for discrete memoryless systems* (Cambridge University Press, 2011).
[41] Y. Polyanskiy and Y. Wu, arXiv preprint arXiv:1508.06025 (2015).
[42] D. M. Endres and J. E. Schindelin, IEEE Transactions on Information theory (2003).
[43] M. Esposito, Physical Review E **85**, 041125 (2012).
[44] F. J. Cao and M. Feito, Physical Review E **79**, 041118 (2009).
[45] J. M. Horowitz and S. Vaikuntanathan, Physical Review E **82**, 061120 (2010).
[46] M. Ponmurugan, Physical Review E **82**, 031129 (2010).

[47] J. M. Horowitz and J. M. Parrondo, New Journal of Physics **13**, 123019 (2011).
[48] D. Abreu and U. Seifert, Physical review letters **108**, 030601 (2012).
[49] T. Sagawa and M. Ueda, Physical Review E **85**, 021104 (2012).

SUPPLEMENTARY MATERIAL

A. Proof: Strictly positive dissipation for non-invertible maps

Suppose the driven dynamics $p(X_1|X_0)$ is a stochastic map from $X \to X$ that results in no dissipation for some prior $q(X_0)$. By Eq. 5 in the main text, if the environment distribution is instead $r(X_0)$, then

$$\begin{aligned}\beta W_d(r) &= D(r(X_0)\|q(X_0)) - D(r(X_1)\|q(X_1)) \\ &= D(r(X_0|X_1)\|q(X_0|X_1))\end{aligned} \tag{12}$$

We provide necessary and sufficient conditions for no dissipation:

Theorem 1. *Suppose that the prior distribution $q(X_0)$ has full support. Then, there exists $r(X_0)$ for which $W_d(r) > 0$ iff $p(X_1|X_0)$ is not an invertible map.*

Proof. KL divergence is invariant under invertible transformations. Therefore, if $p(X_1|X_0)$ is an invertible map, then $D(r(X_0)\|q(X_0)) = D(r(X_1)\|q(X_1))$ for all possible $r(X_0)$. Hence, $W_d(r) = 0\ \forall r$.

We now prove that if $p(X_1|X_0)$ is not an invertible map, then there exists $r(X_0)$ such that $W_d(r) > 0$. For simplicity, write the dynamics $p(X_1|X_0)$ as the right stochastic matrix M. Because M is a right stochastic matrix, it has a right (column) eigenvector $\mathbf{1}^T = (1, \ldots, 1)^T$ with eigenvalue 1. Since $p(X_1|X_0)$ is not an invertible map, M also has a non-zero left eigenvector $\mathbf{s}$ with eigenvalue $|\lambda| < 1$ such that $\mathbf{s}\mathbf{1}^T = 0$. We use $s(x)$ to refer to elements of $\mathbf{s}$ indexed by $x \in X$. Without loss of generality, assume $\mathbf{s}$ is scaled such that $\max_x |s(x)| = \min_{x_0} q(x_0)$ (which is greater than 0, by assumption that $q(X_0)$ has full support).

We now define $r(X_0)$ as:

$$r(x_0) := q(x_0) + s(x_0)$$

Due to the scaling of $\mathbf{s}$ and because $\mathbf{s}\mathbf{1}^T = 0$, $r(X_0)$ is a valid probability distribution.

We use the notation $s(x_1) := \sum_{x_0} s(x_0)\, p(x_1|x_0)$ and $r(x_1) := \sum_{x_0} r(x_0)\, p(x_1|x_0) = q(x_1) + s(x_1)$. We also use the notation $\mathcal{C} := \operatorname{supp} r(X_1)$. The fact that $q(X_0)$ has full support also means that $\mathcal{C} \subseteq \operatorname{supp} q(X_1)$.

The proof proceeds by contradiction. Assume that $W_d(r) = 0$. Using Eq. 12 and due to properties of KL divergence, this means that for each $x_0 \in X$ and $x_1 \in \mathcal{C}$:

$$\begin{aligned}
q(x_0|x_1) &= r(x_0|x_1) \\
\frac{q(x_0)\, p(x_1|x_0)}{q(x_1)} &= \frac{r(x_0)\, p(x_1|x_0)}{r(x_1)} \\
\frac{r(x_1)}{q(x_1)} p(x_1|x_0) &= \frac{r(x_0)}{q(x_0)} p(x_1|x_0) \\
\frac{q(x_1) + s(x_1)}{q(x_1)} p(x_1|x_0) &= \frac{q(x_0) + s(x_0)}{q(x_0)} p(x_1|x_0) \\
\frac{s(x_1)}{q(x_1)} p(x_1|x_0) &= \frac{s(x_0)}{q(x_0)} p(x_1|x_0) \\
s(x_1)\, q(x_0|x_1) &= s(x_0)\, p(x_1|x_0)
\end{aligned}$$

Taking absolute value of both sides gives:

$$|s(x_1)|\, q(x_0|x_1) = |s(x_0)|\, p(x_1|x_0)$$

Summing over $x_0 \in X$ and $x_1 \in \mathcal{C}$:

$$\begin{aligned}
\sum_{x_1 \in \mathcal{C}} \sum_{x_0 \in X} |s(x_1)|\, q(x_0|x_1) &= \sum_{x_1 \in \mathcal{C}} \sum_{x_0 \in X} |s(x_0)|\, p(x_1|x_0) \\
\sum_{x_1 \in X} |s(x_1)| - \sum_{x_1 \notin \mathcal{C}} |s(x_1)| &= \sum_{x_1 \in X} \sum_{x_0 \in X} |s(x_0)|\, p(x_1|x_0) - \sum_{x_1 \notin \mathcal{C}} \sum_{x_0 \in X} |s(x_0)|\, p(x_1|x_0)
\end{aligned} \tag{13}$$

Note that for all $x_1 \notin \mathcal{C}$, $r(x_1) = 0$, meaning that $s(x_1) = -q(x_1)$. Thus,

$$\sum_{x_1 \notin \mathcal{C}} |s(x_1)| = \sum_{x_1 \notin \mathcal{C}} q(x_1)$$

Furthermore, for all $x_1 \notin \mathcal{C}$, $r(x_1) = \sum_{x_0} r(x_0)\, p(x_1|x_0) = 0$. Thus, for all $x_0 \in X$ where $p(x_1|x_0) > 0$ for some $x_1 \notin \mathcal{C}$, $r(x_0) = 0$, meaning $s(x_0) = -q(x_0)$. This allows us to rewrite the last term in Eq. 13 as:

$$\sum_{x_1 \notin \mathcal{C}} \sum_{x_0 \in X} |s(x_0)|\, p(x_1|x_0) = \sum_{x_1 \notin \mathcal{C}} \sum_{x_0 : p(x_1|x_0) > 0} |s(x_0)|\, p(x_1|x_0) = \sum_{x_1 \notin \mathcal{C}} \sum_{x_0 : p(x_1|x_0) > 0} q(x_0)\, p(x_1|x_0) = \sum_{x_1 \notin \mathcal{C}} q(x_1)$$

Cancelling terms that equal $\sum_{x_1 \notin \mathcal{C}} q(x_1)$ from both sides of Eq. 13, we rewrite:

$$\sum_{x_1} |s(x_1)| = \sum_{x_1} \sum_{x_0} |s(x_0)|\, p(x_1|x_0) = \sum_{x_0} |s(x_0)| \tag{14}$$

In matrix notation, Eq. 14 states that:

$$\|\mathbf{s}M\|_1 = \|\mathbf{s}\|_1 \tag{15}$$

where $\|\cdot\|_1$ indicates the vector ℓ_1 norm. However, by definition $\mathbf{s}M = \lambda \mathbf{s}$. Hence:

$$\|\mathbf{s}M\|_1 = \|\lambda \mathbf{s}\|_1 = |\lambda|\, \|\mathbf{s}\|_1 < \|\mathbf{s}\|_1$$

meaning that Eq. 15 cannot be true and the original assumption $W_d(r) = 0$ is incorrect. We have shown that for non-invertible maps, there exists $r(X_0)$ for which $W_d(r) > 0$. □

B. Proof: Dissipation in different environments

Assume some non-equilibrium thermodynamic process which maps initial states x_0 to final states x_1 according to probability distribution $p(x_1|x_0)$, and that this process is dissipationless for some initial distribution $q(x_0)$. This process is evaluated across a set of "environments" Θ, with each environment $\theta \in \Theta$ initializing the process with initial state distribution $r(x_0|\theta)$. Each environment θ occurs with probability $w(\theta)$.

The expected dissipation across the set of environments is

$$\beta \langle W_d \rangle = D(r(X_0|\Theta) \| q(X_0)) - D(r(X_1|\Theta) \| q(X_1)) \tag{16}$$

Theorem 2. *$\beta \langle W_d \rangle \geq I(X_0;\theta) - I(X_1;\theta)$. This bound is achieved if $q(x_0) = \sum_\theta w(\theta)\, r(x_0|\theta)$.*

Proof. Let $r(x_0) := \sum_\theta w(\theta)\, r(x_0|\theta)$, $r(x_0, x_1|\theta) := p(x_1|x_0)\, r(x_0|\theta)$, $r(x_1) := \sum_{\theta, x_0} w(\theta)\, r(x_0, x_1|\theta)$, and $q(x_0, x_1) := p(x_1|x_0)\, q(x_0)$. We also use $r(x_0|x_1, \theta) := \frac{r(x_0, x_1|\theta)}{\sum_{x_0'} r(x_0', x_1|\theta)}$, $q(x_0|x_1) := \frac{q(x_0, x_1)}{\sum_{x_0'} q(x_0', x_1)}$.

Using chain rule for KL divergence, we write:

$$\begin{aligned} D(r(X_0, X_1|\Theta) \| q(X_0, X_1)) =& D(r(X_0|\Theta) \| q(X_0)) + D(r(X_1|X_0, \Theta) \| q(X_1|X_0)) \\ =& D(r(X_1|\Theta) \| q(X_1)) + D(r(X_0|X_1, \Theta) \| q(X_0|X_1)) \end{aligned}$$

Note that for all θ, $r(x_1|x_0, \theta) = q(x_1|x_0) = p(x_1|x_0)$ for any distribution q. Thus:

$$D(r(X_1|X_0, \Theta) \| q(X_1|X_0)) = 0 \tag{17}$$

Expected dissipation (Eq. 16) can then be rewritten as:

$$\begin{aligned} \beta \langle W_d \rangle &= D(r(X_0|X_1, \Theta) \| q(X_0|X_1)) \\ &= D(r(X_0|X_1, \Theta) \| r(X_0|X_1)) + D(r(X_0|X_1) \| q(X_0|X_1)) \end{aligned} \tag{18}$$

where

$$r(x_0|x_1) := \frac{\sum_\theta w(\theta)\, r(x_0|\theta)\, p(x_1|x_0)}{\sum_{\theta, x_0'} w(\theta)\, r(x_0'|\theta)\, p(x_1|x_0')} = \frac{r(x_0)\, p(x_1|x_0)}{\sum_{x_0'} r(x_0')\, p(x_1|x_0')} \tag{19}$$

We seek the $q(x_0)$ which minimizes expected dissipation. In Eq. 18, the first KL term is independent of q. The second term reaches its minimum value of 0 when $q(x_0|x_1) = r(x_0|x_1)$, since KL divergence is minimized when both

of its arguments are equal. From Eq. 19, it is clear that $q(x_0|x_1) = r(x_0|x_1)$ if $q(x_0) = r(x_0)$. Note, however, that this may also be true for other q, depending on $p(x_1|x_0)$. For instance, it is true for all q if $p(x_1|x_0)$ is an invertible map.

Again using the chain rule for KL divergence, we write:

$$\begin{aligned}\beta \langle W_d \rangle &\geq D(r(X_0|X_1,\Theta) \| r(X_0|X_1)) \\ &= D(r(X_0,X_1|\Theta) \| r(X_0,X_1)) - D(r(X_1|\Theta) \| r(X_1)) \\ &= D(r(X_0|\Theta) \| r(X_0)) - D(r(X_1|\Theta) \| r(X_1)) + D(r(X_1|X_0,\Theta) \| r(X_1|X_0)) \\ &= D(r(X_0|\Theta) \| r(X_0)) - D(r(X_1|\Theta) \| r(X_1)) \\ &= I(X_0;\Theta) - I(X_1;\Theta)\end{aligned}$$

where we used Eq. 17 and the standard definition of I (mutual information) in terms of KL divergence. □

C. Dissipation for feedback processes

Here we analyze dissipation due to incorrect priors in *feedback processes*. Feedback processes [44–49] assume two subsystems coupled to a heat bath, with state spaces X and M (for 'memory'). A feedback process involves two stages:

First, the state $x \in X$ is held fixed while the memory system performs a measurement and updates its state according to some $p(M|x)$. Assuming that this stage completes by time $t = 0$, and that x is initially distributed according to some $q(X_0)$, the joint distribution at $t = 0$ is $q(x_0, m_0) = q(x_0)\, p(m_0|x_0)$.

In the second 'control' stage, the memory $m \in M$ is held fixed while an m-dependent driving protocol is applied to x. At the end of this driving protocol at time 1, x is statistically independent of m. We assume that X is distributed according to some probability distribution $p^\star(x_1)$ that does not depend on starting conditions or the state of M. The joint distribution at $t = 1$ is $q(x_1, m) = p^\star(x_1)\, q(M)$, where $q(m)$ is the marginalized distribution of M: $q(m) := \sum_{x_0} q(x_0)\, p(m|x_0)$. Analyses of feedback processes usually assume that the joint Hamiltonian is decoupled at $t = 0$ and $t = 1$, having the form $H_t(X, M) = H_t(X) + H_t(M)$ [11].

It is possible to use Eq. 1 in the main text to compute the work necessary for the control stage. This gives the *second law for feedback processes* [11]:

$$\langle W \rangle \geq \Delta \mathcal{F}_X - kT \cdot I(X_0; M) \tag{20}$$

where $\Delta \mathcal{F}_X$ is the non-equilibrium free energy change in subsystem X. This result shows that correlations between X and M created during measurement stage reduce the work required during the control stage. The bound in Eq. 20 can be met by designing an 'optimal' feedback process [11, 47].

The optimality of the feedback process depends on knowledge of the distribution $q(x_0)$. Here we evaluate what occurs when a feedback process optimized for $q(x_0)$ is instead evaluated with initial distribution $r(x_0)$. In this case, the joint distribution at $t = 0$ will be $r(x_0, m) = r(x_0)\, p(m|x_0)$, while at $t = 1$ it will be $r(x_1, m) = p^\star(x_1)\, r(m)$, where $r(m) := \sum_{x_0} r(x_0)\, p(m|x_0)$. Inserting these distributions into Eq. 5 in the main text gives:

$$\begin{aligned}\beta W_d =& D(r(X_0, M) \| q(X_0, M)) - D(r(X_1, M) \| q(X_1, M)) \\ =& D(r(X_0) \| q(X_0)) - D(r(M) \| q(M)) + D(p(M|X_0) \| p(M|X_0)) - D(p^\star(X_1) \| p^\star(X_1)) \\ =& D(r(X_0) \| q(X_0)) - D(r(M) \| q(M))\end{aligned}$$

By the data processing inequality for KL, this dissipation is non-negative. As can be seen, memory can minimize dissipation by preserving information about initial states: in fact, dissipation reaches its minimum of 0 when the memory has a perfect copy of the initial condition x_0. It is important to note, however, that resetting (i.e. erasing) the memory after the control stage – necessary, for instance, for the feedback controller to operate in a cyclical manner – will dissipate $\beta^{-1} D(r(M) \| q(M))$ work. Thus, total dissipation for control and reset stages will equal

$$W_d = \beta^{-1} D(r(X_0) \| q(X_0))$$