Agential approach to quantum thermodynamics

Ruo Cheng Huang

June 25, 2025

Abstract

Interacting with dynamic, uncertain environments requires AI agents to perform predictive inference under informational and physical constraints. We present a prototype thermodynamic agent—a pattern engine—that extracts useful work from a non-Markovian quantum process, leveraging principles from computational mechanics. The environment is modeled as a classical hidden Markov model (HMM) with quantum outputs, and the agent maintains an internal belief state that synchronizes to the latent dynamics of this process. Critically, the agent's performance is governed by the meta-dynamics of its belief updates, which capture the interaction between the environment's hidden evolution and the agent's internal representation. We demonstrate that belief-informed policies consistently outperform memoryless and classical strategies, and identify phase transitions in performance linked to bifurcations in the belief dynamics. These findings suggest that alignment failures can emerge not solely from policy design flaws, but from structural limitations in the agent's ability to accurately track latent information embedded in the data.

1 Introduction

The relationship between classical thermodynamics and information theory has long been established from thought experiments such as Maxwell's demon and Szilard's engine, which was later resolved by Landauer and his famous principle, stating that to erase 1 bit of information, at least $k_B T \ln 2$ units of work is required [1, 2, 3, 4, 5, 6] where T is the temperature of the surrounding environment. This was then extended into quantum thermodynamics, where instead of ensemble of particles, the focus are on quantum states, characterized by positive semi-definite symmetric matrices, ρ [7, 8, 9, 10, 11]. Maximal amount of energy that equals to the non-equilibrium free energy can be extracted from these quantum objects, assuming that the agent knows the identity of such objects. Given a state ρ, if the identity is known, it is possible to extract work,

$$W = k_B T \mathrm{D}(\rho \| \gamma_\beta) \ , \tag{1}$$

where $\mathrm{D}(\rho\|\sigma) = \mathrm{Tr}(\rho \log \rho) - \mathrm{Tr}(\rho \log \sigma)$ is the quantum analog of the KL-divergence known as quantum relative entropy. γ is known as the thermal state, the quantum analog of the Maxwell Boltzmann distribution in classical thermodynamics. However, one cannot expect a quantum state to remain identical over long time period. Things such as decoherence or interaction with environment will cause the state to change, furthermore, unlike classical counterparts, quantum states cannot be observed as it will result in measurement collapse. Fortunately, nothing is truly random in nature, there usually exist some temporal correlation between the quantum state at t and that at $t+1$. Here we model such correlation using a classical hidden markov model (HMM) but with a quantum output, example of such process can be seen in Fig 1. These memoryful sources of quantum states can be represented by a hidden Markov model (HMM) $\mathcal{M} = \left(\mathcal{S}, \left(\sigma^{(x)}\right)_{x\in\mathcal{X}}, \left(T^{(x)}\right)_{x\in\mathcal{X}}\right)$. Here, $\mathcal{S}$ is the HMM's set of classical latent states.

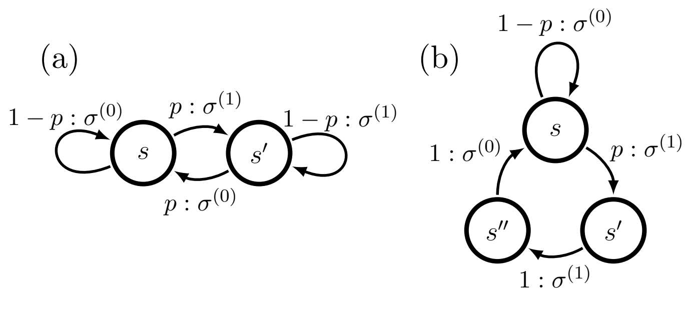


Figure 1: Latent-state sources of correlated quantum processes. Each arrow represents a transition between latent states; the label $p : \sigma^{(x)}$ indicates that the transition happens with probability p and produces a quantum state $\sigma^{(x)}$. (a) Perturbed-coin process. (b) 2-1 golden-mean process.

The random variable representing the latent state at time t shall be denoted by S_t. The transition-matrix element $T^{(x)}_{s,s'} = \Pr(S_t = s', X_t = x | S_{t-1} = s)$ represents the probability of transitioning from latent state s to s' and emitting the d-dimensional quantum state $\sigma^{(x)}$. For simple HMMs, the memoryful transition structure can be visualized as an annotated directed graph, as in Fig. 1. In this graphical representation, nodes correspond to latent states while the directed edges correspond to latent-state-to-state transitions that produce a certain quantum output with a prescribed probability. In the example of Fig. 1(a), the 'biased-coin process' has only two latent states s and s'. At each timestep, the latent state switches with probability p. The quantum output is $\sigma^{(0)}$ whenever the resultant state is s; The quantum output is $\sigma^{(1)}$ whenever the resultant state is s'. As the switching-probability p approaches 1, the process approaches a period-two output, alternating between the two quantum states $\ldots \sigma^{(0)} \otimes \sigma^{(1)} \otimes \sigma^{(0)} \otimes \sigma^{(1)} \ldots$. At the other extreme, as p approaches 0, the latent state remains the same for increasingly long epochs that produce long strings of the same quantum output $\ldots \sigma^{(0)} \otimes \sigma^{(0)} \otimes \sigma^{(0)} \ldots$ or $\ldots \sigma^{(1)} \otimes \sigma^{(1)} \otimes \sigma^{(1)} \ldots$ switching between the two behaviors only rarely. More generally, for any $p \in (0, 1)$, the process interpolates between these two extreme behaviors. More complicated HMMs can generate more complex memoryful structure in the quantum-output process. See Fig. 1(b) for a hint of this possible richness.

The HMM specifies the statistics of the non-Markovian classical variables X_t across time, which, in turn, induce the quantum outputs indexed by $x \in \mathcal{X}$. The quantum output process is described by the (formal) density operator

$$\rho_{\overleftrightarrow{A}} = \sum_{\overleftrightarrow{x}} \Pr(\overleftrightarrow{x}) \bigotimes_{t \in \mathbb{Z}} \sigma^{(x_t)}_{A_t} \ , \tag{2}$$

where each time step t is associated with a unique elementary physical system A_t, and $\overleftrightarrow{x} = \ldots x_{-1} x_0 x_1 \ldots$ denotes a bi- infinite string over $\mathcal{X}$. The joint quantum state has no entanglement across A_ts, but can have non-classical correlations in the form of quantum discord [12]. We assume that the joint state in Eq. 2 is known exactly, but not of which specific string $\overleftrightarrow{x}$ is instantiated.

2 Synchronizing to a quantum source

To fully leverage the structure of the pattern, the agent must dynamically incorporate information from past interactions with the elementary quantum systems, so that the engine's memory becomes correlated with the latent state of the source. This can be done by tracking—within the internal memory M of the engine—an interaction, induced *belief state* $\boldsymbol{\eta}_t$ about the latent state of the source. The type of interaction between engine and fuel at each time can depend on the memory of the engine. The general interaction and observation at time t can be described by a positive operator-valued measure (POVM) on the Hilbert space of the elementary fuel system A_t; Let O_t denote the random variable for the observed outcome thereof. The optimal belief state $\boldsymbol{\eta}_t$ at time t is an observation- induced probability distribution over the latent states $\mathcal{S}$ of the source with probability elements

$$\boldsymbol{\eta}_t(s) = \Pr(S_t = s | O_1 \ldots O_t = o_1 \ldots o_t) \ . \tag{3}$$

This is the best knowledge that a local classical memory can have, as it represents the actual distribution over latent states as one would calculate via Bayes rule. It is convenient to treat $\boldsymbol{\eta}_t$ as a length-$|\mathcal{S}|$ row vector for linear algebraic manipulation. Given a sequence of observations, what is the probability $P_t(x)$ that the source will next produce quantum state $\sigma^{(x)}$? It is $P_t(x) = \boldsymbol{\eta}_t T^{(x)} \mathbf{1}$, where $\mathbf{1}$ is the column vector of all ones. The belief state $\boldsymbol{\eta}_t$ thus determines an expectation of the next quantum state

$$\xi_t = \sum_{x \in \mathcal{X}} P_t(x) \, \sigma^{(x)} \ , \tag{4}$$

with more free energy than the local reduced state ξ_0 of $\rho_{\overleftrightarrow{A}}$. This memory enhancement to free energy is proven in App. B.

The transition rules between these belief states are determined via Bayesian inference, based on the anticipated distribution of the observable O_t. If the source is known, but no observations have yet been made, then the optimal belief state is simply the stationary distribution over source states: $\boldsymbol{\eta}_0 = \boldsymbol{\pi}$, which satisfies $\boldsymbol{\pi} = \boldsymbol{\pi} \sum_{x \in \mathcal{X}} T^{(x)}$.

Theorem 1. *For any POVM on the quantum state of the system at time t, the optimal belief state—about the latent state of the quantum source—updates iteratively according to*

$$\boldsymbol{\eta}_t = z_t^{-1} \sum_{x \in \mathcal{X}} \Pr\big(O_t = o_t | X_t = x, K_{t-1} = \boldsymbol{\eta}_{t-1}\big) \, \boldsymbol{\eta}_{t-1} T^{(x)} \tag{5}$$

where K_t is the random variable for the belief state, and $z_t = \sum_{x' \in \mathcal{X}} \Pr\big(O_t = o_t | X_t = x', K_{t-1} = \boldsymbol{\eta}_{t-1}\big) \, \boldsymbol{\eta}_{t-1} T^{(x')} \mathbf{1}$ is a normalizing factor.

We derive Thm. 1 in Appendix C, generalizing the so-called mixed-state presentation [13, 14, 15, 16, 17].

The probability $\Pr(O_t = o_t | X_t = x, K_{t-1} = \boldsymbol{\eta}_{t-1})$ appearing in Eq. (5) is typically a straightforward physics calculation of the probability that the observed POVM outcome o_t should be obtained, given that the system was prepared as $\sigma^{(x)}$. Conditioning on the previous state of knowledge K_{t-1} is important to the extent that it can influence the choice of POVM applied at time t. Notice that, for each o, the update rule for the belief state can be interpreted as a nonlinear return map. If the return map enables prediction, then we have some access to the nonequilibrium free energy in correlations.

Observation 1. *Memory can be thermodynamically advantageous when the belief-state update rule has an attractor other than the stationary fixed point $\boldsymbol{\pi}$ for at least one observable outcome.*

3 Engine construction

The engine is equipped with an internal classical memory M and access to a heat reservoir R at some fixed temperature T. Its objective is to extract work by raising the internal energy of a work reservoir or 'battery' B. To accomplish this, it can bring each system A_t closer to its equilibrium state $\gamma = e^{-H/k_B T}/Z$, k_B is Boltzmann's constant, and $Z = \mathrm{tr}(e^{-H/k_B T})$ is the associated equilibrium partition function that yields the equilibrium free energy $F = -k_B T \ln Z$. For simplicity of presentation, we assume that each subsystem A_t is subject to the same Hamiltonian H, but our results generalize in an obvious way if we allow different Hamiltonians for each subsystem. The construction of the pattern engine then requires only the description of the HMM $\mathcal{M}$ and the Hamiltonian H.

To harvest the free energy locked up in correlations, the engine's internal memory should somehow become correlated with the latent state of the source during its energy-harvesting operation. However, directly measuring each quantum system would disturb the state and potentially cost energy. Rather, our engine updates its memory conditioned on the extracted-work value W_t at each time. The change in the energy of the battery thus serves as the observable $O_t = W_t$ for updating the belief state. At each time step, conditioned on its memory state, the engine performs a *work-extraction protocol*—a unitary transformation of the composite A_t, B, and R supersystem, designed to transfer energy from A_t to B. At best, the work-extraction protocol at time step t would extract all the nonequilibrium addition to free energy $k_B T \mathrm{D}[\rho_t^* \| \gamma]$ when it acts on a chosen state ρ_t^*, where $\mathrm{D}[\rho \| \gamma] = \mathrm{tr}(\rho \ln \rho) - \mathrm{tr}(\rho \ln \gamma)$ is the quantum relative entropy. These are the *ρ^*-ideal work-extraction protocols*, which we define as any work-extraction protocol that satisfies the following:

1. When the initial state of the system is ρ^*, it achieves zero entropy production, transferring on-average all nonequilibrium addition to free energy $k_B T \mathrm{D}[\rho^* \| \gamma]$ to a work reservoir, B;
2. It conserves energy globally among the reduced states of A_t, B, and R when acting on any eigenstate of ρ^*.

Our next theorem fully characterizes the set of extracted values and their probabilities for any such protocol (see Appendix D for the derivation).

Theorem 2. *Each ρ^*-ideal work-extraction protocol exhibits at most d distinct extracted-work values. These extracted-work values can be expressed, in terms of the ideal input's spectral decomposition $\rho^* = \sum_n \lambda_n |\lambda_n\rangle \langle \lambda_n|$, as*

$$w^{(n)} := \langle \lambda_n | H | \lambda_n \rangle + k_B T \ln \lambda_n - F \tag{6}$$

with associated probabilities

$$\Pr(W = w^{(n)} | \sigma) = \sum_m \langle \lambda_m | \sigma | \lambda_m \rangle \, \delta_{w^{(n)}, w^{(m)}} \tag{7}$$

This set of values is independent of the actual d-dimensional quantum state σ input to the protocol, although the input state determines the probabilities of each outcome. Notably, this yields the probability distribution for work extracted when the protocol optimized for ρ_t^* actually operates on $\sigma^{(x)}$. Regardless of how the belief state influences the choice of ρ_t^*, we can now leverage Thms. 1 and 2 to rewrite the belief update as

$$\boldsymbol{\eta}_{t+1} = \frac{\sum_{x \in \mathcal{X}} \sum_n \delta_{w_{t+1}, w^{(n)}} \langle \lambda_n | \sigma^{(x)} | \lambda_n \rangle \, \boldsymbol{\eta}_t T^{(x)}}{\sum_n \delta_{w_{t+1}, w^{(n)}} \langle \lambda_n | \xi_t | \lambda_n \rangle} \, . \tag{8}$$

With Eq. (8), the belief-state return maps now reflect the physics of the work-extraction protocol.

From Eqs. (4) and (7), we find that the work-induced transitions between belief states have probabilities

$$\Pr(W_{t+1} = w | K_t = \boldsymbol{\eta}_t) = \sum_n \langle \lambda_n | \xi_t | \lambda_n | \lambda_n | \xi_t | \lambda_n \rangle \, \delta_{w, w^{(n)}} \, . \tag{9}$$

Finally, to take thermodynamic advantage of this knowledge, the work-extraction protocol at each step[1] is optimized for the expected state, so that $\rho_t^* = \xi_t$. Indeed, extracting all work from the expected state ξ_t requires a

[1] Except at the first step, where we use some $\rho^* \neq \xi_0$, to avoid an unstable fixed point in the knowledge update.

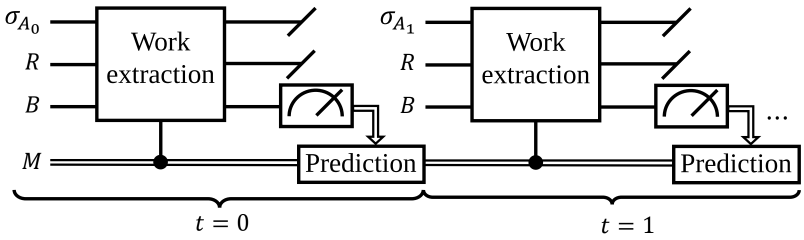


Figure 2: Schematic diagram of a quantum-pattern engine. At each time step, the process will take a quantum system, σ_{A_t}, from the "fuel" tape, reservoir qudit, R, battery, B, and memory, M, as input. The 'Work extraction' box should be interpreted as a memory-dependent unitary. States of battery and memory are recycled.

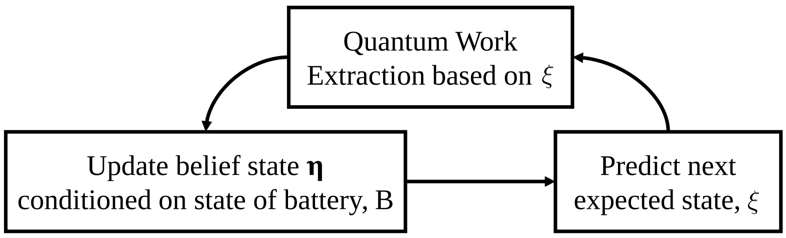


Figure 3: The protocol proceeds cyclically to fine-tune the belief state.

protocol designed around this expectation [18]. In this case, the denominator in Eq. (8) simplifies to $\sum_n \lambda_n \, \delta_{w_{t+1},w^{(n)}}$. Similarly, Eq. (9) simplifies to $\sum_n \lambda_n \, \delta_{w,w^{(n)}}$. Combining Eqs. (6) and (9), we compute

$$\langle W_t \rangle = \sum_{\boldsymbol{\eta},w} w \Pr(K_{t-1} = \boldsymbol{\eta}) \Pr(W_t = w | K_{t-1} = \boldsymbol{\eta}) \tag{10}$$

and find

$$\langle W_t \rangle = k_B T \langle \mathrm{D}[\xi_t \| \gamma] \rangle_{\Pr(K_t)} \ . \tag{11}$$

Note that the expectation value on the right-hand side is taken over the instantaneous distribution over belief states. The meta-dynamic over belief states thus determines both the transient and asymptotic work-extraction rate. In particular, the stationary distribution over recurrent belief states allows closed-form expressions for the asymptotic work-extraction rate.

Both the belief states and transitions between them derive from the HMM of the known source. Belief states can thus be explicitly represented in the memory M of an autonomous work-harvesting device. The memory states $\{(\boldsymbol{\eta}, \varepsilon)\}_{\boldsymbol{\eta},\varepsilon}$ should also store the last measured energy state ε of the battery. A memory-controlled unitary can implement memory-assisted quantum work extraction, as depicted in the circuit diagram of Fig. 2. Subsequent measurement of the battery state then gives access to the work extracted and allows an autonomous update of the memory, according to the above-outlined rules of Bayesian prediction. This prediction-extraction cycle continues repeatedly, as suggested in Fig. 3.

4 Example processes and alternative approaches

To demonstrate our *memory-assisted quantum* approach—using a quantum work-extraction protocol designed for the work-observation-induced expected state $\rho_t^* = \xi_t$—we apply it to the quantum perturbed-coin and golden-mean processes depicted in Fig. 1. We compute our engine's long-term work output from this approach (1) and compare its performance with those of three alternative approaches (2 - 4):

1. *memory-assisted quantum*, where the quantum work-extraction protocol is optimized for the work-observation-induced expected state $\rho_t^* = \xi_t$;

2. *memory-assisted classical*, where the protocol, unable to extract work from quantum coherences, is optimized for the energy-dephased version of the expected state $\rho_t^* = \xi_t^{\mathrm{dec}} := \sum_E |E\rangle \langle E| \langle E| \xi_t |E\rangle$;

3. *memoryless quantum processing*, where memory is never updated by observations, and the quantum work-extraction protocol is simply optimized for the time-averaged quantum state $\rho^* = \xi_0 = \sum_x \boldsymbol{\pi} T^{(x)} \mathbf{1} \, \sigma^{(x)}$; and

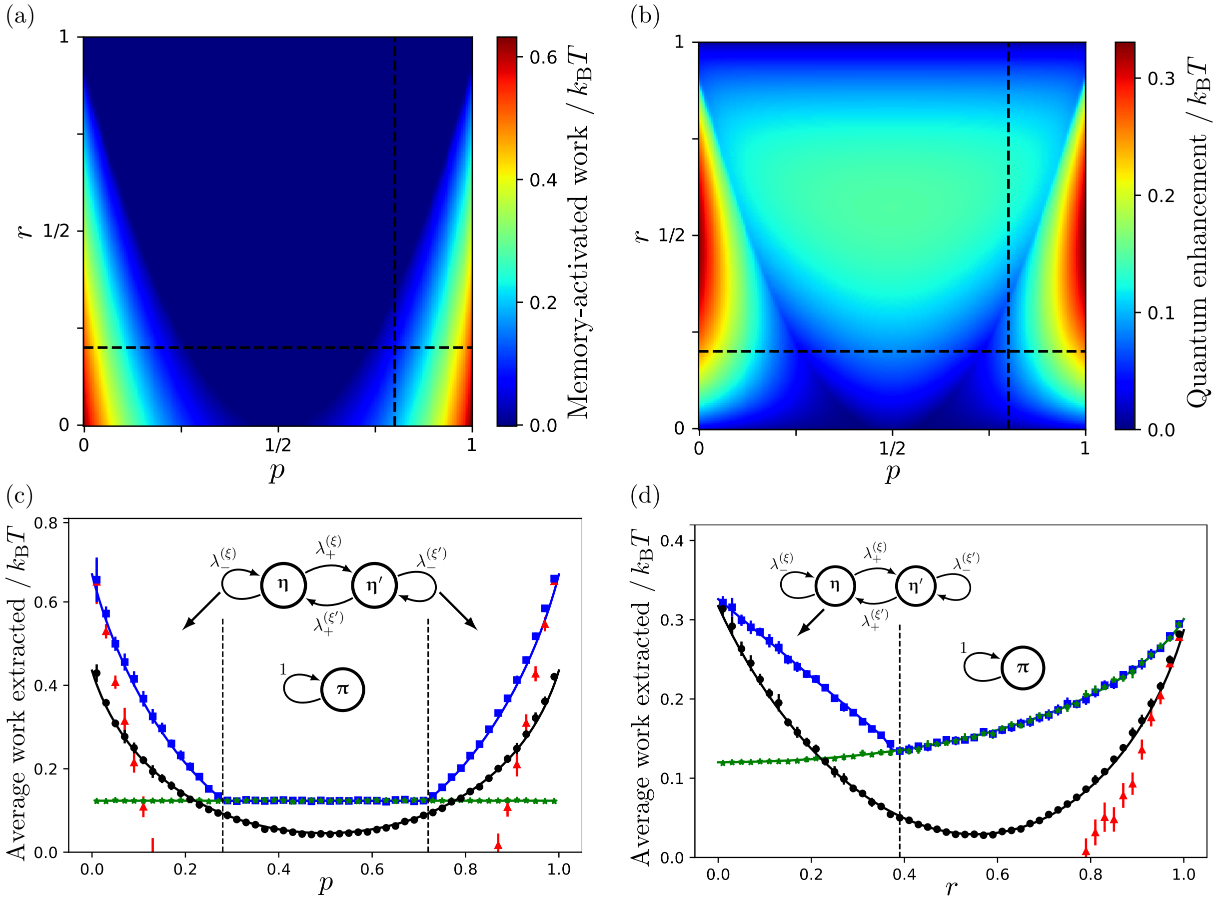


Figure 4: Comparison between average work-extraction rates of various approaches. p characterizes the transition probability between the two latent states of the perturbed-coin process, and r quantifies the overlap between the two quantum outputs. (a) Memory enhancement of work extraction. (b) Quantum enhancement of work extraction. Panels (c) and (d) reveal phase transitions in memory enhancement through cross-sections of parameter space. Analytic results (solid lines) and simulations (markers) are shown. Blue (squares) represents approach 1; black (circles) represents approach 2; green (stars) represents approach 3; red (triangles) represents approach 4.

4. *overcommitment to most probable quantum state*, with protocol optimized for $\rho_t^* = \sigma^{(\mathrm{argmax}_x \boldsymbol{\eta}_t T^{(x)} \mathbf{1})}$.

The asymptotic work-extraction rates $\lim_{t\to\infty}\langle W_t\rangle$ from these approaches are compared in Fig. 4, both analytically and with numerical simulations, for the perturbed-coin process. In Fig. 4(a) memory-activated work is defined as the difference between memory-assisted quantum approach 1 and memoryless quantum approach 3. In Fig. 4(b) quantum enhancement is defined as the difference between memory-assisted quantum approach 1 and memory-assisted classical approach 2.

In this demonstration, the system extracts work from a sequence of qubits, each with Hamiltonian $H = E_0 |0\rangle\langle 0| + E_1 |1\rangle\langle 1|$, with $E_1 - E_0 = k_B T$ in this case. At each time step, the source produces one of two nonorthogonal quantum states: $\sigma^{(0)} = |0\rangle\langle 0|$ or $\sigma^{(1)} = |\psi\rangle\langle\psi|$, where $|\psi\rangle = \sqrt{r}\,|0\rangle + \sqrt{1-r}\,|1\rangle$, according to the labeled transition matrices of the perturbed-coin process. As seen in Fig. 4, our memory-assisted quantum approach 1 always performs at least as well as all other approaches, and strictly outperforms them in many regions of parameter space. In these examples, the ideal input ρ_t^* is a qubit density matrix with eigenvalues $\lambda_\pm^{(\rho_t^*)}$, where $\lambda_+^{(\rho_t^*)} \geq \lambda_-^{(\rho_t^*)} \geq 0$. Due to Eqs. (6) and (7), we thus expect to observe one of two possible work values, $w^{(\pm)}$ from each distinct belief state. Approaches 1-3 share some nice features. From Eq. (7), assuming distinct work values $w^{(+)} \neq w^{(-)}$, we find that the probability of observing each possible work value is simply given by the corresponding eigenvalue of the optimal input:

$$\Pr\Big(W_{t+1} = w^{(\pm)} | K_t = \boldsymbol{\eta}_t\Big) = \lambda_\pm^{(\rho_t^*)} \ . \tag{12}$$

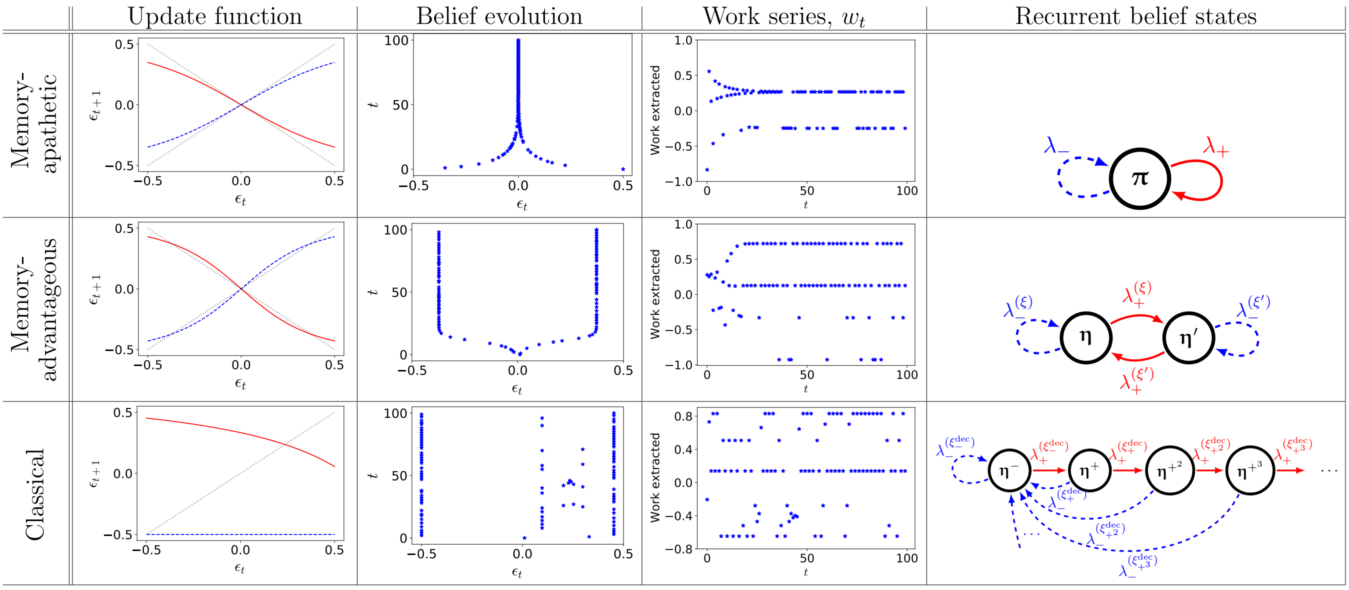
Combining this with Eq. (6), we find that the rate of work extraction can be expressed as

$$\langle W_t\rangle = k_B T \langle \mathrm{D}[\rho_t^* \| \gamma]\rangle_{\Pr(K_t)} \tag{13}$$

for approaches 1-3—although, notably, both ρ_t^* and the set of belief states will be different in each approach—with $\rho_t^* = \xi_t$, ξ_t^{dec}, and ξ_0, respectively. The non-negativity of relative entropy thus guarantees the non-negativity of expected work extraction from these three approaches. Note that Eq. (13) is more general than Eq. (11) and allows us to compare work-extraction performance of the three approaches. Without any memory, the extractable structure is limited to the time-averaged statistical bias of the output [19], which explains why memoryless work extraction varies with r but not p. Further details of the analytic solution for expected work extraction can be found in Appendix H.

Our numerical simulations use the quantum work-extraction protocol of Ref. [7] at each time-step, and agree very well with our more general analytical predictions. Indeed, the quantum work-extraction protocol of Ref. [7] provides an example of a ρ^*-ideal work-extraction protocol, in the limit of many bath interactions. More details can be found in Appendices F and G. It is tempting to commit to the most likely outcome. However, the overcommitment approach 4 performs the worst, since any reset operation (to γ in this case) with minimal entropy production for a pure-state input leads to infinite heat dissipation when operating on any other input [18, 20]. This translates to infinite negative work extraction $\langle W_t | W_t \rangle = -\infty$ in this case of $\rho_t^* \in \{|0\rangle\langle 0|, |\psi\rangle\langle\psi|\}$. This divergence can alternatively be seen from Eq. (6) as $w^{(-)} \sim \ln \lambda_- \to \ln 0 = -\infty$. In our numerical simulations, following Ref. [7], this minimal eigenvalue λ_- is inversely proportional to the number N of bath interactions, so that the overcommitment work penalty diverges as $-k_B T(1-r)\min(p, 1-p)\ln N$.

Table 1: Summary of metadynamics in different regimes. The update function shows the nonlinear relationship between ϵ_t and ϵ_{t+1}. The belief evolution shows the evolution of ϵ_t over iterations, which give rise to the corresponding work series, with two possible work values per belief state. The recurrent belief states show the recurrent metadynamic of the different regimes.



4.1 Phase transitions in efficacy of knowledge

Surprisingly, there exists a blue inner region of panel 4(a) where the memoryless quantum approach 3 achieves the same performance as our memory-assisted quantum approach 1. *There exists a sharp phase boundary within which the use of memory does not boost performance.* As seen clearly in panels 4(c) and 4(d), the phase boundary exhibits a discontinuity in the first derivative of work extraction with respect to process parametrization. This phase boundary is not unique to the perturbed-coin process and, indeed, also occurs in the 2-1 golden-mean process. Such phase transitions originate from bifurcations of the attractors of the belief-state update maps. This is illustrated in Table 1, which shows the nonlinear return maps along with the consequences for belief dynamics and work-extraction dynamics. We focus for now on the first two rows of Table 1, which illustrate the nonlinear dynamics of our memory-assisted quantum approach 1, in the memory-apathetic and memory-advantageous regimes, before and after bifurcation respectively.

Recall from Fig. 1(a) that a two-state machine generates the perturbed-coin process. Hence, a scalar $\epsilon_t \in [-\frac{1}{2}, \frac{1}{2}]$ suffices to describe the time-dependent belief state $\boldsymbol{\eta}_t = (\frac{1}{2} + \epsilon_t, \frac{1}{2} - \epsilon_t)$. The magnitude of this scalar ϵ_t indicates the strength of evidence that the process is in a particular hidden state. The first column of Table 1 shows the return maps for $\epsilon_t \mapsto \epsilon_{t+1}$ induced by either $w^{(+)}$ (red solid graph) or $w^{(-)}$ (blue dashed graph), when the work-extraction protocol is optimized for $\rho_t^* = \xi_t$ (for the first two rows) or $\rho_t^* = \xi_t^{\text{dec}}$ (for the last row). To aid the visual bifurcation and stability analysis, we include a dotted diagonal line with slope one—representing the identity map—and a dotted diagonal line with slope minus one—representing the swap map. Intersections between a return map and the dotted identity line would indicate a fixed point of the map upon its repeated application. If the magnitude of the slope at the intersection is less than unity, then it is a stable fixed point; if the magnitude of the slope at the intersection is greater than unity, then it is an unstable fixed point. Within the memory-apathetic region of parameter space, Work extraction does not supply enough evidence to nudge an observer out of a state of complete ignorance. In this regime, memory does not enhance quantum work extraction.

At the phase boundary in the process' parameter space, the fixed point at $\boldsymbol{\pi}$ becomes unstable, and new attractors emerge for each map. However, the coexistence of the maps introduces competition between the attractors, as the maps are selected stochastically with probabilities $\lambda_{\pm}$. The two maps (red solid and blue dashed) interact to induce a steady-state metadynamic over recurrent belief states, shown as a Markov process in each row of Table 1's last column. In this memory-advantageous region, work extraction supplies sufficient evidence to inform an observer about the hidden state of the process, which in turn avails more extractable work.

The elegance of memory-assisted quantum work extraction is reflected in the simple one and two state recurrent memory structures. In comparison, the classical extractor not only harvests less work, but requires more memory to achieve its relatively meager returns. Note the infinite number of recurrent memory states in the classical-processing case, in the last row of Table 1.

5 Discussion

We have introduced a theoretical model of a belief-state-driven thermodynamic agent that operates in a quantum environment with structured temporal correlations. By grounding the agent's operation in computational mechanics, we showed how internal belief states can be updated through local interactions, enabling the agent to predict and exploit latent structure for useful work extraction. This framework unifies concepts from statistical physics, quantum information, and computational mechanics. The agent's capacity to align with its environment depends not only on its physical capabilities, but also on the dynamics of its internal representations.

Our results highlight that the efficacy of the agent is governed by the agent's belief meta-dynamics—i.e., how its internal model evolves in response to observations. We demonstrated that such agents can outperform both memoryless and classical counterparts, and that sudden phase transitions in performance emerge when belief synchronization breaks down. These transitions underscore a core challenge in alignment: optimal behavior may be sharply sensitive to model fidelity, or limit of the model's representation.

This work suggests that alignment failures in physically agents—whether in thermodynamic, quantum, or AI systems—could stem not just from flawed objectives or policies, but from fundamental constraints on what can be known and predicted from observations alone. Understanding and quantifying these limits will be crucial for designing agents that remain robust and aligned across diverse, uncertain environments.

Future work may explore agents equipped with quantum memory or non-greedy strategies, and investigate whether more general quantum belief-state dynamics can further improve long-term alignment and energetic efficiency.

References

[1] Rolf Landauer. Irreversibility and heat generation in the computing process. *IBM journal of research and development*, 5(3):183–191, 1961.

[2] Leo Szilard. On the decrease of entropy in a thermodynamic system by the intervention of intelligent beings. *Behavioral Science*, 9(4):301–310, 1964.

[3] Charles H Bennett. The thermodynamics of computation—a review. *International Journal of Theoretical Physics*, 21(12):905–940, 1982.

[4] Dibyendu Mandal and Christopher Jarzynski. Work and information processing in a solvable model of Maxwell's demon. *Proceedings of the National Academy of Sciences*, 109(29):11641–11645, 2012.

[5] Dibyendu Mandal, HT Quan, and Christopher Jarzynski. Maxwell's refrigerator: an exactly solvable model. *Physical review letters*, 111(3):030602, 2013.

[6] Juan MR Parrondo, Jordan M Horowitz, and Takahiro Sagawa. Thermodynamics of information. *Nature physics*, 11(2):131–139, 2015.

[7] Paul Skrzypczyk, Anthony J Short, and Sandu Popescu. Work extraction and thermodynamics for individual quantum systems. *Nature communications*, 5, 2014.

[8] Johan Åberg. Catalytic coherence. *Physical review letters*, 113(15):150402, 2014.

[9] Kamil Korzekwa, Matteo Lostaglio, Jonathan Oppenheim, and David Jennings. The extraction of work from quantum coherence. *New Journal of Physics*, 18(2):023045, 2016.

[10] Andrew JP Garner, Jayne Thompson, Vlatko Vedral, and Mile Gu. Thermodynamics of complexity and pattern manipulation. *Physical Review E*, 95(4):042140, 2017.

[11] Matteo Lostaglio. Thermodynamic laws for populations and quantum coherence: A self-contained introduction to the resource theory approach to thermodynamics. *arXiv preprint arXiv:1807.11549*, 2018.

[12] Kavan Modi, Tomasz Paterek, Wonmin Son, Vlatko Vedral, and Mark Williamson. Unified view of quantum and classical correlations. *Phys. Rev. Lett.*, 104:080501, Feb 2010.

[13] James P Crutchfield. The calculi of emergence: computation, dynamics and induction. *Physica D: Nonlinear Phenomena*, 75(1-3):11–54, 1994.

[14] Christopher J Ellison, John R Mahoney, and James P Crutchfield. Prediction, retrodiction, and the amount of information stored in the present. *Journal of Statistical Physics*, 136(6):1005–1034, 2009.

[15] Sarah E Marzen and James P Crutchfield. Nearly maximally predictive features and their dimensions. *Physical Review E*, 95(5):051301, 2017.

[16] Paul M Riechers and James P Crutchfield. Spectral simplicity of apparent complexity. I. The nondiagonalizable metadynamics of prediction. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 28(3):033115, 2018.

[17] Alexandra M Jurgens and James P Crutchfield. Shannon entropy rate of hidden Markov processes. *Journal of Statistical Physics*, 183(2):1–18, 2021.

[18] Paul M Riechers and Mile Gu. Initial-state dependence of thermodynamic dissipation for any quantum process. *Physical Review E*, 103(4):042145, 2021.

[19] Alexander B. Boyd, Dibyendu Mandal, and James P. Crutchfield. Correlation-powered information engines and the thermodynamics of self-correction. *Phys. Rev. E*, 95:012152, Jan 2017.

[20] Paul M Riechers and Mile Gu. Impossibility of achieving Landauer's bound for almost every quantum state. *Phys. Rev. A*, 104:012214, Jul 2021.

[21] Patryk Lipka-Bartosik, Paweł Mazurek, and Michał Horodecki. Second law of thermodynamics for batteries with vacuum state. *Quantum*, 5:408, 2021.

[22] F. L. Curzon and B. Ahlborn. Efficiency of a Carnot engine at maximum power output. *American Journal of Physics*, 43(1):22–24, 01 1975.

[23] John P. S. Peterson, Tiago B. Batalhão, Marcela Herrera, Alexandre M. Souza, Roberto S. Sarthour, Ivan S. Oliveira, and Roberto M. Serra. Experimental characterization of a spin quantum heat engine. *Phys. Rev. Lett.*, 123:240601, Dec 2019.

[24] Tan Van Vu and Keiji Saito. Finite-time quantum Landauer principle and quantum coherence. *Phys. Rev. Lett.*, 128:010602, Jan 2022.

[25] Philip Taranto, Faraj Bakhshinezhad, Andreas Bluhm, Ralph Silva, Nicolai Friis, Maximilian P.E. Lock, Giuseppe Vitagliano, Felix C. Binder, Tiago Debarba, Emanuel Schwarzhans, Fabien Clivaz, and Marcus Huber. Landauer versus Nernst: What is the true cost of cooling a quantum system? *PRX Quantum*, 4:010332, Mar 2023.

A Decomposition of free energy in quantum patterns

Consider a finite portion of the quantum pattern:

$$\rho^{(1:L)} = \mathrm{tr}_{\mathbb{Z}\setminus\{\ell\}_1^L}(\rho_{\dots A_{-1}A_0A_1\dots}) \ . \tag{14}$$

If each subsystem is non-interacting and the ℓ^{th} subsystem has a reference equilibrium Gibbs state $\gamma^{(\ell)}$, then the nonequilibrium free energy for this portion of the pattern is given by

$$\mathcal{F}^{(1:L)} = F_{\text{eq}}^{(1:L)} + k_B T \mathrm{D}\Big[\rho^{(1:L)} \| \bigotimes_{\ell=1}^{L} \gamma^{(\ell)}\Big] \tag{15}$$

$$= F_{\text{eq}}^{(1:L)} + k_B T \Big[\mathrm{tr}\Big(\rho^{(1:L)} \ln \rho^{(1:L)}\Big) - \sum_{\ell=1}^{L} \mathrm{tr}\Big(\rho^{(\ell)} \ln \gamma^{(\ell)}\Big)\Big] \tag{16}$$

$$= F_{\text{eq}}^{(1:L)} + k_B T \underbrace{\mathrm{D}[\rho^{(1:L)} \| \bigotimes_{\ell=1}^{L} \rho^{(\ell)}]}_{\text{total correlation}} + k_B T \sum_{\ell=1}^{L} \mathrm{D}[\rho^{(\ell)} \| \gamma^{(\ell)}] \ , \tag{17}$$

where $F_{\text{eq}}^{(1:L)}$ is the equilibrium free energy.

We recognize that $\mathrm{D}[\rho^{(1:L)} \| \bigotimes_{\ell=1}^{L} \rho^{(\ell)}]$ is the *total correlation* within the quantum pattern, while $\mathrm{D}[\rho^{(\ell)} \| \gamma^{(\ell)}]$ is the *local nonequilibrium addition to free energy.* Each of these factors contributes uniquely to the free energy. When operating sequentially on each subsystem, quantum pattern engines must leverage past information to harvest the free energy in the correlations.

In the main text, we suppose each non-interacting subsystem has the same local Hamiltonian, which implies that the reference Gibbs states are also all the same: $\gamma^{(\ell)} = \gamma$. The general decomposition here shows that both quantum and classical correlations contribute to the extractable nonequilium addition to free energy. However, in the main text, the quantum pattern is assumed to be classically-generated, despite having non-orthogonal states. The inter-time quantum correlations are thus restricted to quantum discord, with no inter-time entanglement [12].

B Memory-enhanced free energy

By self-consistency, the local reduced state of $\rho_{\overleftrightarrow{A}}$ will simply be a probabilistic mixture of $\sigma^{(x)}$ in the form of $\langle \xi_t | \xi_t \rangle_{K_t} = \sum_{x \in \mathcal{X}} \boldsymbol{\pi} T^{(x)} \mathbf{1} \, \sigma^{(x)} = \xi_0$. The predicted quantum state ξ_t thus has more free energy than the reduced state on average since

$$\langle \mathrm{D}[\xi_t \| \gamma] | \mathrm{D}[\xi_t \| \gamma] \rangle_{K_t} - \mathrm{D}[\xi_0 \| \gamma] = \langle \mathrm{D}[\xi_t \| \gamma] | \mathrm{D}[\xi_t \| \gamma] \rangle_{K_t} - \mathrm{D}[\langle \xi_t | \xi_t \rangle_{K_t} \| \gamma] \tag{18}$$

$$= S(\langle \xi_t | \xi_t \rangle_{K_t}) - \langle S(\xi_t) | S(\xi_t) \rangle_{K_t} \tag{19}$$

$$\geq 0 \ . \tag{20}$$

Thus,

$$\langle \mathrm{D}[\xi_t \| \gamma] | \mathrm{D}[\xi_t \| \gamma] \rangle_{K_t} \geq \mathrm{D}[\xi_0 \| \gamma] \ , \tag{21}$$

and so more work can be extracted on average when memory is leveraged to predict sequential quantum states. The non-negativity of Eq. (20) can be seen either from the convexity of relative entropy in Eq. (18) or from the concavity of entropy in Eq. (19).

A nearly identical argument also shows that the classically predicted state has more free energy than the average classical state: $\big\langle \mathrm{D}[\xi_t^{\text{dec}} \| \gamma] \big| \mathrm{D}[\xi_t^{\text{dec}} \| \gamma] \big\rangle_{K_t} \geq \mathrm{D}[\xi_0^{\text{dec}} \| \gamma]$, where the states of knowledge K_t are now the classically induced ones.

C Synchronizing to a memoryful quantum source

Inferring the latent state of a known memoryful quantum source allows maximal work extraction when operating serially on the quantum states of the process. The optimal state of knowledge, given a sequence of observations $o_1 o_2 \ldots o_t$ obtained via interventions on the sequence of quantum systems $\sigma^{(x_1)}, \sigma^{(x_2)}, \ldots \sigma^{(x_t)}$ is the conditional probability distribution induced by these interventions,

$$\boldsymbol{\eta}_t := \Pr(S_t | O_1 \ldots O_t = o_1 \ldots o_t, S_0 \sim \boldsymbol{\pi}) \ . \tag{22}$$

The last condition $S_0 \sim \boldsymbol{\pi}$ means that the initial latent state of the generator is distributed as $\boldsymbol{\pi}$. This can be rewritten as $\boldsymbol{\eta}_t = \sum_s \boldsymbol{\pi}(s) \Pr(\S_t | O_1 \ldots O_t = o_1 \ldots o_t, S_0 = s)$ for $t > 0$. Recall that $\boldsymbol{\pi} = \boldsymbol{\pi} \sum_{x \in \mathcal{X}} T^{(x)}$ is the stationary distribution over the states of the generator. Thus, $\boldsymbol{\eta}_0 = \boldsymbol{\pi}$.

If we introduce a new random variable K_t to denote the optimally updated state of knowledge about the latent state of the pattern generator, then we can replace the condition $S_{t-1} \sim \boldsymbol{\eta}_{t-1}$ with $K_{t-1} = \boldsymbol{\eta}_{t-1}$. The condition on the state of knowledge is relevant to the extent that the choice of POVM is influenced by the state of knowledge. We remind the reader that in our framework the POVM on the current quantum output is chosen as a function of the state of knowledge K_t.

Note that the current quantum output only depends on the current latent state of the process. Accordingly, the next observation—which is the outcome of the POVM on the current quantum output—is conditionally independent of all previous outputs, given the current latent state and given the state of knowledge induced by all previous outputs.

We will now show that the optimal state of knowledge is recursive. I.e., we will show that:

$$\boldsymbol{\eta}_t = \Pr\big(S_t | O_t = o_t, S_{t-1} \sim \boldsymbol{\eta}_{t-1}\big) \ . \tag{23}$$

This follows from marginalizing over intervening latent states, employing Bayes' rule, and recognizing that the belief state $\boldsymbol{\eta}_t$ is a function of the observations up to that time $o_1 \dots o_t$. Starting from Eq. (22), we find:

$$\begin{aligned}
\boldsymbol{\eta}_t &:= \Pr(S_t | O_1 \dots O_t = o_1 \dots o_t, S_0 \sim \boldsymbol{\pi}) \\
&= \sum_s \Pr(S_t, S_{t-1} = s | O_1 \dots O_t = o_1 \dots o_t, S_0 \sim \boldsymbol{\pi}) && (24) \\
&= \frac{\sum_s \Pr(S_t, O_t = o_t, S_{t-1} = s | O_1 \dots O_{t-1} = o_1 \dots o_{t-1}, S_0 \sim \boldsymbol{\pi})}{\Pr(O_t = o_t | O_1 \dots O_{t-1} = o_1 \dots o_{t-1}, S_0 \sim \boldsymbol{\pi})} && (25) \\
&= \frac{\sum_s \Pr(S_{t-1} = s | O_1 \dots O_{t-1} = o_1 \dots o_{t-1}, S_0 \sim \boldsymbol{\pi}) \Pr(S_t, O_t = o_t | O_1 \dots O_{t-1} = o_1 \dots o_{t-1}, S_0 \sim \boldsymbol{\pi}, S_{t-1} = s)}{\sum_{s'} \Pr(O_t = o_t, S_{t-1} = s' | O_1 \dots O_{t-1} = o_1 \dots o_{t-1}, S_0 \sim \boldsymbol{\pi})} && (26) \\
&= \frac{\sum_s \boldsymbol{\eta}_{t-1}(s) \Pr\big(S_t, O_t = o_t | K_{t-1} = \boldsymbol{\eta}_{t-1}, S_{t-1} = s\big)}{\sum_{s'} \boldsymbol{\eta}_{t-1}(s') \Pr\big(O_t = o_t | K_{t-1} = \boldsymbol{\eta}_{t-1}, S_{t-1} = s'\big)} && (27) \\
&= \Pr\big(S_t | O_t = o_t, K_{t-1} = \boldsymbol{\eta}_{t-1}\big) \ . && (28)
\end{aligned}$$

Hence, we have obtained Eq. (23) from Eq. (22) as promised.

$$\begin{tikzcd}
S_0 \arrow[rr] \arrow[dr] & & S_1 \arrow[rr] \arrow[dr] & & S_2 \arrow[rr] \arrow[dr] & & S_3 \\
& X_1 \arrow[ur] \arrow[d] & & X_2 \arrow[ur] \arrow[d] & & X_3 \arrow[ur] \arrow[d] & & \cdots \\
& O_1 \arrow[dr] & & O_2 \arrow[dr] & & O_3 \arrow[dr] & \\
K_0 \arrow[rr] \arrow[ur] & & K_1 \arrow[rr] \arrow[ur] & & K_2 \arrow[rr] \arrow[ur] & & K_3
\end{tikzcd}$$

Figure 5: Bayesian network showing the structure of conditional independencies among latent states S_t of the quantum source, the type X_t of quantum state produced, the observable O_t attained from interaction, and the state of knowledge K_t that influences the work extraction protocol.

Further manipulations, using the rules of probability and the conditional independencies indicated in the Bayesian network depicted in Fig. 5, allow us to express the optimal state of knowledge in terms of both conditional work distributions and simple linear algebraic manipulations of the generative HMM representing the memoryful source. We find

$$\begin{aligned}
\boldsymbol{\eta}_t &= \Pr\big(S_t | O_t = o_t, S_{t-1} \sim \boldsymbol{\eta}_{t-1}\big) && (29) \\
&= \sum_{x \in \mathcal{X}} \Pr\big(S_t, X_t = x | O_t = o_t, S_{t-1} \sim \boldsymbol{\eta}_{t-1}\big) && (30) \\
&= \sum_{x \in \mathcal{X}} \Pr\big(X_t = x | O_t = o_t, S_{t-1} \sim \boldsymbol{\eta}_{t-1}\big) \Pr\big(S_t | X_t = x, S_{t-1} \sim \boldsymbol{\eta}_{t-1}\big) && (31) \\
&= \sum_{x \in \mathcal{X}} \Pr\big(X_t = x | O_t = o_t, S_{t-1} \sim \boldsymbol{\eta}_{t-1}\big) \frac{\boldsymbol{\eta}_{t-1} T^{(x)}}{\boldsymbol{\eta}_{t-1} T^{(x)} \mathbf{1}} && (32) \\
&= \frac{\sum_{x \in \mathcal{X}} \Pr\big(X_t = x, O_t = o_t | S_{t-1} \sim \boldsymbol{\eta}_{t-1}\big) \boldsymbol{\eta}_{t-1} T^{(x)} / \boldsymbol{\eta}_{t-1} T^{(x)} \mathbf{1}}{\sum_{x' \in \mathcal{X}} \Pr\big(X_t = x', O_t = o_t | S_{t-1} \sim \boldsymbol{\eta}_{t-1}\big)} && (33) \\
&= \frac{\sum_{x \in \mathcal{X}} \Pr\big(O_t = o_t | X_t = x, S_{t-1} \sim \boldsymbol{\eta}_{t-1}\big) \, \boldsymbol{\eta}_{t-1} T^{(x)}}{\sum_{x' \in \mathcal{X}} \Pr\big(O_t = o_t | X_t = x', S_{t-1} \sim \boldsymbol{\eta}_{t-1}\big) \, \boldsymbol{\eta}_{t-1} T^{(x')} \mathbf{1}} \ . && (34)
\end{aligned}$$

C.1 Using this to build a predictive work-extraction engine

Rather than repeatedly calculating these ideal belief states on the fly for a specific realization of the process, we can alternatively systematically build up the set of all such belief states, together with the observation-induced transitions among them, to inform the design of an autonomous engine. There will be both a set of transient belief states and a set of recurrent belief states. Both of these sets may be either finite or infinite. In the case that only finitely many belief states are induced by observations, we can explicitly build out the transition structure among them. If there are infinitely many such states, then we would need to truncate unlikely states in the design of our finite physical engine [15].

The physical memory system of our proposed engine should have at least one distinguishable state corresponding to every observation-induced belief state. In fact, the memory must encode both the belief state and the most recent energy of the battery, so that conditioning on the new state of the battery is sufficient to supply the change in battery energy. These will likely be encoded with some finite precision, to avoid storing real numbers. Conditioned on the state of the memory encoding $\boldsymbol{\eta}$, the work extraction protocol will operate jointly on the quantum system, thermal reservoirs, and battery, to optimally extract work from the expected state $\xi = \sum_{x \in \mathcal{X}} \boldsymbol{\eta} T^{(x)} \mathbf{1} \sigma^{(x)}$.

The subsequently observed work value w uniquely updates the memory from the state encoding $\boldsymbol{\eta}$ to the state encoding $\boldsymbol{\eta}' = \frac{\sum_{x \in \mathcal{X}} \Pr(W_t = w | X_t = x, S_{t-1} \sim \boldsymbol{\eta}) \, \boldsymbol{\eta} T^{(x)}}{\sum_{x' \in \mathcal{X}} \Pr(W_t = w | X_t = x', S_{t-1} \sim \boldsymbol{\eta}) \, \boldsymbol{\eta} T^{(x')} \mathbf{1}}$. Once the next quantum system arrives, the predictive quantum work extraction cycle begins again.

D Proof of Thm. 2: Work extraction, in the limit of zero entropy production

Work extraction in the limit of zero entropy production is important since it extracts all extractable work from a quantum state. It thus indicates the best possible scenario, against which other efforts can be compared.

In the limit of zero-entropy-production work extraction from ρ^*, the net unitary time evolution of the system–battery–baths supersystem must take a special form. In particular, the state of the battery will change deterministically when the initial state of the system is an eigenstate $|\lambda_n\rangle$ of $\rho^* = \sum_n \lambda_n |\lambda_n\rangle \langle \lambda_n|$, almost-surely independent of the initial realization of the reservoirs. This implies that the net unitary time evolution will be of the form

$$U = \sum_{\varepsilon, n, r} \left(\left| \varepsilon + w^{(n)} \right\rangle \otimes |f_\varepsilon(n, r)\rangle \right) \left(\langle \varepsilon | \otimes \langle \lambda_n | \otimes \langle r | \right) \tag{35}$$

for some $w^{(n)} \in \mathbb{R}$. Above, $|\varepsilon\rangle$ and $\left|\varepsilon + w^{(n)}\right\rangle$ are energy eigenstates of the work reservoir, while $|r\rangle$ is an energy eigenstate of the thermal baths. It will be useful in the following to note that $\langle f_\varepsilon(n,r) | f_\varepsilon(n', r') | f_\varepsilon(n,r) | f_\varepsilon(n', r') \rangle = \delta_{n,n'} \delta_{r,r'}$ since unitary operations map orthogonal states to orthogonal states

The form of the unitary Eq. (35) effectively assumes that the energy of the battery is well above its ground state. Some interesting nuances have recently been explored for batteries close to their ground state (see, e.g., Ref. [21]), which would affect the statistics of the work-extraction values, but we avoid that regime here to instead focus on the best-possible scenario.

One way to determine $w^{(n)}$ is via the initial-state dependence of entropy production. Let $\langle \Sigma | \Sigma \rangle_\rho$ denote the expectation value for entropy production, given initial system-state ρ, under the fixed work-extraction protocol optimized for ρ^*. In our case with a single heat bath at temperature T, the expected entropy production can be defined as usual as $\langle \Sigma \rangle_\rho = \left(\langle \tilde{W} \rangle_\rho - \Delta \mathcal{F}_t \right) / T$. This is the entropy production for a fixed protocol operating on the initial state ρ, where $\mathcal{F}_t$ is the nonequilibrium free energy at time t, while $\Delta \mathcal{F}_t$ is the change in nonequilibrium free energy over the course of the protocol, and $\tilde{W}$ is the work *exerted*, which is just the negative of the extractable work [6]. Since all initial states map to γ by the end of the work-extraction protocol, we know from Ref. [18] that

$$\langle \Sigma \rangle_\sigma - \langle \Sigma \rangle_{\rho^*} = k_B \mathrm{D}[\sigma \| \rho^*] \ . \tag{36}$$

In this case, $\langle \Sigma | \Sigma \rangle_{\rho^*} = 0$ and $T \langle \Sigma \rangle_\sigma = \langle \tilde{W} \rangle_\sigma - \Delta \mathcal{F}_t = \langle \tilde{W} \rangle_\sigma + k_B T \mathrm{D}[\sigma \| \gamma]$. Hence, with $\beta = (k_B T)^{-1}$,

$$\begin{aligned} \beta \langle \tilde{W} \rangle_\sigma &= \mathrm{D}[\sigma \| \rho^*] - \mathrm{D}[\sigma \| \gamma] & (37) \\ &= \mathrm{tr}(\sigma \ln \gamma) - \mathrm{tr}(\sigma \ln \rho^*) \ . & (38) \end{aligned}$$

In particular, let $\sigma = |\lambda_n\rangle \langle \lambda_n|$, and note that $\ln \gamma = \ln\!\left(e^{-\beta H} / Z\right) = \beta (F - H)$. This yields $\langle \tilde{W} \rangle_{|\lambda_n\rangle\langle\lambda_n|} = F - \langle \lambda_n | H | \lambda_n | \lambda_n | H | \lambda_n \rangle - k_B T \ln \lambda_n$. The deterministic work-extraction value, given initial pure state $|\lambda_n\rangle$, must be the same as its expected value $w^{(n)} = -\langle \tilde{W} \rangle_{|\lambda_n\rangle\langle\lambda_n|}$, and is thus given by

$$w^{(n)} = \langle \lambda_n | H | \lambda_n \rangle + k_B T \ln \lambda_n - F \ . \tag{39}$$

The probability of obtaining the work-extraction value w, given any input state $\sigma = \sum_{n,m} |\lambda_n\rangle\langle\lambda_m| \langle\lambda_n|\sigma|\lambda_m\rangle$, can be calculated as

$$\Pr(W = w|\sigma) = \mathrm{tr}\Big[\big(|\varepsilon_0 + w\rangle\langle\varepsilon_0 + w| \otimes I\big) U \big(|\varepsilon_0\rangle\langle\varepsilon_0| \otimes \sigma \otimes |r\rangle\langle r|\big) U^\dagger\Big] \tag{40}$$

$$= \sum_{n,m} \langle\lambda_n|\sigma|\lambda_m\rangle \, \mathrm{tr}\Big[\big(|\varepsilon_0 + w\rangle\langle\varepsilon_0 + w| \otimes I\big) U \big(|\varepsilon_0\rangle\langle\varepsilon_0| \otimes |\lambda_n\rangle\langle\lambda_m| \otimes |r\rangle\langle r|\big) U^\dagger\Big] \tag{41}$$

$$= \sum_{n,m} \langle\lambda_n|\sigma|\lambda_m\rangle \, \mathrm{tr}\Big[\big(|\varepsilon_0 + w\rangle\langle\varepsilon_0 + w| \otimes I\big) \Big(\Big|\varepsilon_0 + w^{(n)}\Big\rangle \otimes |f_{\varepsilon_0}(n,r)\rangle\Big) \Big(\Big\langle\varepsilon_0 + w^{(m)}\Big| \otimes \langle f_{\varepsilon_0}(m,r)|\Big)\Big] \tag{42}$$

$$= \sum_{n,m} \langle\lambda_n|\sigma|\lambda_m\rangle \, \delta_{w,w^{(n)}} \delta_{w,w^{(m)}} \, \langle f_{\varepsilon_0}(m,r)|f_{\varepsilon_0}(n,r)|f_{\varepsilon_0}(m,r)|f_{\varepsilon_0}(n,r)\rangle \tag{43}$$

$$= \sum_{n,m} \langle\lambda_n|\sigma|\lambda_m\rangle \, \delta_{w,w^{(n)}} \delta_{w,w^{(m)}} \delta_{n,m} \tag{44}$$

$$= \sum_{n} \langle\lambda_n|\sigma|\lambda_n\rangle \, \delta_{w,w^{(n)}} \tag{45}$$

independent of the initial energy state of the battery $|\varepsilon_0\rangle$, and almost-surely independent of the initial realization $|r\rangle$ of the thermal reservoirs in the probability theoretic sense. We see that $\Pr(W = w|\sigma) = 0$ unless $w \in \{w^{(n)}\}_n$. The probability distribution over these allowed work-extraction values is

$$\Pr\Big(W = w^{(n)}|\sigma\Big) = \sum_m \langle\lambda_m|\sigma|\lambda_m\rangle \, \delta_{w^{(n)},w^{(m)}} \ . \tag{46}$$

When there is some entropy production, the probability density of work extraction will have more diffuse peaks. However, for sufficiently low entropy production, the peaks will still be well separated and, so, effectively discrete for the purpose of Bayesian updating.

The above derivation is valid whether or not ρ^* has degenerate eigenvalues. Notably, the above sums are taken over the eigenstates and their associated eigenvalues, rather than summing over the eigenvalues directly.

E Power and inefficiency at rapid operation

It is a familiar concept in the design of any engine: that maximal thermodynamic efficiency requires sufficiently slow operation. Clearly, this has implications for the power output of the engine [22, 23]. However, the relaxation timescales of an engine depend on particular material properties of the system and baths, as well as the particular interaction Hamiltonian, so there is no implementation-independent timescale that determines the practical operation speed of an engine.

Nevertheless, we can apply rather general principles to assess how power typically scales with increasingly fast operation. For example, under assumptions of a Lindblad master equation, there will be a contribution to entropy production (and a corresponding decrease in extracted work per operation) that scales as $1/\tau_0$, where τ_0 is the duration of the work-extraction protocol [24]. Or, for unitary interactions with the bath, a similar statement can be made but with τ_0 proportional to the number of interactions with bath degrees of freedom (i.e., the 'circuit complexity') [25]. In either case, we expect entropy production to scale as $\Sigma \approx c/(\kappa + \tau_0) \approx c/\tau_0$ for $\tau_0 \gg \kappa > 0$ where c and κ are implementation-dependent positive quantities.

In a fixed time τ, the number of work-extraction operations t determines the maximal allowed time $\tau/t \geq \tau_0$ for each work-extraction protocol. Let $\langle W_{\mathrm{ext}}|W_{\mathrm{ext}}\rangle$ be the steady-state work-extraction rate per operation in the limit of very slow operation (i.e., in the limit of infinitely many relaxation steps per work-extraction operation). The power P achieved by finite-time operation is then sandwiched by

$$\tfrac{t}{\tau} \langle W_{\mathrm{ext}}|W_{\mathrm{ext}}\rangle \geq P = \tfrac{t}{\tau}\big(\langle W_{\mathrm{ext}}|W_{\mathrm{ext}}\rangle - cT/\tau_0\big) \geq \tfrac{t}{\tau} \langle W_{\mathrm{ext}}|W_{\mathrm{ext}}\rangle - cT/\tau_0^2 \ . \tag{47}$$

We focus on the regime where each work-extraction protocol is of sufficiently long duration $\tau_0 \gg c/k_B$, such that cT/τ_0 is negligible. In this regime, the power trivially scales with the number of operations per second, $P = \frac{t}{\tau} \langle W_{\mathrm{ext}}|W_{\mathrm{ext}}\rangle \propto t/\tau$. In the main text, we focus on results about $\langle W_{\mathrm{ext}}|W_{\mathrm{ext}}\rangle$ obtained in this simple regime, since it highlights the thermodynamic role of correlations and quantum discord.

F The work extraction protocol of Skrzypczyk *et al.*

For self containment, here we give a brief summary of the Skrzypczyk *et al.* work-extraction protocol [7], utilized in our numerical simulations. Ref. [7] should be consulted for further details.

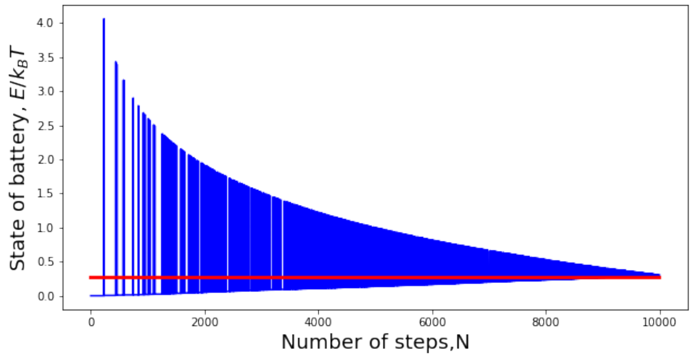


Figure 6: Quasistatic evolution of battery state over a finite number N of bath interactions, under the work-extraction protocol from Skrzypczyk et al.[7] . Red line represents the total nonequilibrium addition to free energy present in the initial input state.

The work extraction protocol of Ref. [7], proceeds in two stages. The first stage (isentropically) puts the system into a mixture of energy eigenstates, using the battery to offset internal energy changes. In the second stage, the system is coupled with the ambient heat bath to slowly (in a sequence of N steps) relax it into a thermal state, again using the battery to absorb all energy offsets.

Consider a system and a weight initially in an uncoupled state $\rho^* \otimes \rho_w$. Let $\rho^* = \sum_n \lambda_n |\lambda_n\rangle\langle\lambda_n|$ be a spectral decomposition of the system state ρ^*, with $\lambda_n \geq \lambda_{n+1}$. Step 1 of the work extraction protocol maps the eigenvectors $|\lambda_n\rangle$ into the energy eigenstates of the system, via the unitary operator

$$V = \sum_n |E_n\rangle\langle\lambda_n| \otimes \Gamma_{h_n}, \tag{48}$$

where $h_n = \langle\lambda_n| H_s |\lambda_n\rangle - E_n$ accounts for the difference in energy in the two states and Γ is an operator to raise the potential energy of the weight. V therefore conserves energy on average. The resultant state will be

$$\rho_{\mathrm{SW}} = \sum_n \lambda_n |E_n\rangle\langle E_n| \otimes \Gamma_{h_n}\rho_w\Gamma^\dagger_{h_n} \tag{49}$$

In Step 2, the resultant state ρ_{SW} will go through a sequence of transformation to reach the Gibbs state, $\gamma = \sum_n e^{-\beta E_n}/Z\,|E_n\rangle\langle E_n|$. At each sub-step of the transformation, the relative probabilities of two levels will be adjusted by δp towards the Gibbs distribution over energy eigenstates. Suppose we now focus on the occupation probability of the ground state and first excited state, $|E_0\rangle$ and $|E_1\rangle$. In the first sub-step of thermalization, the system will interact with a bath qubit, $\rho_B = \frac{q_0}{q_0+q_1}|0\rangle\langle 0| + \frac{q_1}{q_0+q_1}|1\rangle\langle 1|$, where $q_0 = \lambda_0 - \delta p$ and $q_1 = \lambda_1 + \delta p$. A unitary transformation is then used to swap the occupation statistics of the bath qubit and the system qubit; in doing so, the battery's energy level rises or drops to conserve the total energy. The transformation can be described as

$$|E_0\rangle_S |1\rangle_B |x\rangle_W \longleftrightarrow |E_1\rangle_S |0\rangle_B |x+h\rangle_W \tag{50}$$

where $h = k_BT\log\frac{q_0}{q_1} - (E_1 - E_0)$. The process then repeats itself by varying the value of q_0 and q_1 until ρ_{SW} reaches a Gibbs state. It is provable that this protocol is able to extract all the free energy of the system up to $\mathcal{O}(\delta p^2)$.

G Simulation

Here we elaborate on the the method of simulations. We considered a string of output with length $n = 5000$ produced by the the perturbed-coin process. The possible emissions are $\sigma^{(0)} = |0\rangle\langle 0|$ and $\sigma^{(1)} = |\psi\rangle\langle\psi|$ where $|\psi\rangle = \sqrt{r}\,|0\rangle + \sqrt{1-r}\,|1\rangle$. Here we considered two pure states—this however is not necessary: mixed states can be used too.

For the thermalization, the number of SWAP operations (between system and tailored baths) was chosen to be $N = 200$ due to limitation in computational power. One can refer to Fig. 6 to see that as $N \to \infty$ the protocol indeed extracts all the free energy from the system. After the extraction, the change in the battery system is measured and recorded. As mentioned in a previous section, if $N \to \infty$, the work distribution will converge to a set of δ-functions. If N is finite, the probability distribution of the work measured becomes more diffuse, as shown in Fig. 7. The state of knowledge is updated via Bayesian inference, conditioned on the observed work value. This inference step is

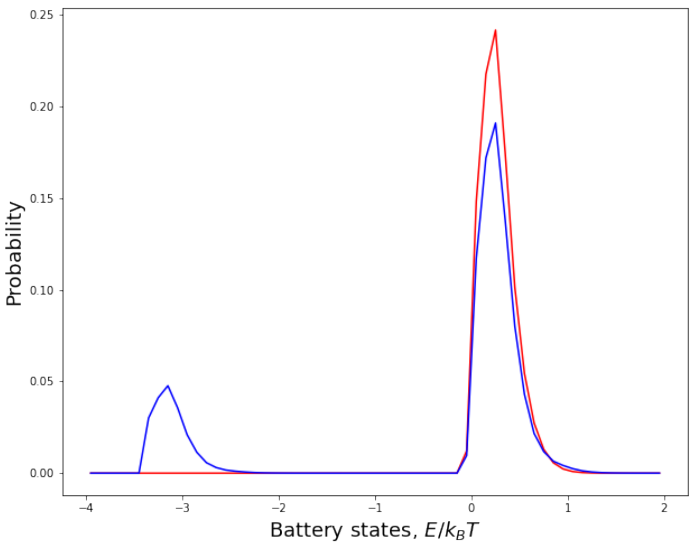


Figure 7: Probability distribution of work extracted, when using the the Skrzypczyk work-extraction protocol with a total of $N = 22$ bath interactions. Red line represents the distribution when the protocol— thermodynamically ideal for some mixed state very close to $|0\rangle$—acts on the pure state $|0\rangle$. Blue represent the same protocol's work distribution when acting on a relatively non-orthogonal pure state $|\psi\rangle$, where fidelity between the two states is $F\big(|0\rangle, |\psi\rangle\big) = |\langle 0|\psi|0|\psi\rangle|^2 = 4/5$.

notably absent from the memoryless approach, where the state of knowledge effectively remains at the initial state of ignorance $\boldsymbol{\eta}_t = \boldsymbol{\pi}$.

For the classical protocol, we assume that energetic coherences are inaccessible. In this case, our simulations utilize the Skrzypczyk work-extraction protocol tailored to the decohered state $\rho_t^* = \xi_t^{\mathrm{dec}}$. However, initial decoherence in any other basis, besides ξ_t's eigenbasis, would likewise underperform compared to the memory-assisted quantum protocol.

The overcommitment protocol differs from the rest as it tailors the extraction protocol for the state with the highest probability of emission. The probability of emission can be calculated from the state of knowledge $\boldsymbol{\eta}_t T^{(x)} \mathbf{1}$. The graphs in Panels (c) and (d) of Fig 4 display only positive work-extraction values on the vertical axis; hence most of the data points for overcommitment are not shown owing to its bad performance.

H Expected work extraction from the four approaches

Here, we derive analytical expressions for the expectation value of work extraction from the various approaches compared in the main text. We derive these expressions for the (p, r)-parametrized family of perturbed-coin processes of classically correlated quantum states discussed in the main text.

Recall that the quantum states $(\sigma^{(x)})_{x\in\mathcal{X}}$ are the outputs of a Mealy HMM with labeled transition matrices $(T^{(x)})_{x\in\mathcal{X}}$. An element of the labeled transition matrix $T^{(x)}_{s\to s'} = \Pr(X_t = x, S_t = s' | S_{t-1} = s)$ gives the joint probability of producing quantum state $\sigma^{(x)}$ and arriving at latent state s', given that the HMM begins in state s.

For the perturbed-coin example, the HMM's labeled transition matrices are

$$T^{(0)} = \begin{bmatrix} 1-p & 0 \\ p & 0 \end{bmatrix} \qquad \text{and} \qquad T^{(1)} = \begin{bmatrix} 0 & p \\ 0 & 1-p \end{bmatrix} . \tag{51}$$

The stationary distribution over the latent states is $\boldsymbol{\pi} = \left[\frac{1}{2}, \frac{1}{2}\right]$ and the two different quantum states created are

$$\sigma^{(0)} = |0\rangle\langle 0| = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad \sigma^{(1)} = |\psi\rangle\langle\psi| = \begin{bmatrix} r & \sqrt{r(1-r)} \\ \sqrt{r(1-r)} & 1-r \end{bmatrix} . \tag{52}$$

For any state of knowledge, $\boldsymbol{\eta}_t = \left[\frac{1}{2} + \epsilon_t, \frac{1}{2} - \epsilon_t\right]$ parameterized by $\epsilon_t \in [-\frac{1}{2}, \frac{1}{2}]$, the induced expected state is

$$\xi_t = \rho^{(\epsilon_t)} := \sum_x \left[\tfrac{1}{2} + \epsilon_t \quad \tfrac{1}{2} - \epsilon_t\right] T^{(x)} \mathbf{1} \sigma^{(x)} = \tfrac{1}{2} \begin{bmatrix} 1 + r + \epsilon_t' \sqrt{1-r} & \sqrt{r(1-r)} - \epsilon_t' \sqrt{r} \\ \sqrt{r(1-r)} - \epsilon_t' \sqrt{r} & 1 - r - \epsilon_t' \sqrt{1-r} \end{bmatrix} , \tag{53}$$

where $\epsilon' := 2\epsilon(1-2p)\sqrt{1-r} \; \propto \epsilon$.

H.1 Memory-assisted quantum processing

In the memory-assisted quantum approach, we utilize work-extraction protocols that are thermodynamically optimized for the expected quantum state

$$\rho_t^* = \xi_t = \rho^{(\epsilon_t)} \; . \tag{54}$$

Recall that the eigenvalues and eigenstates of ρ_t^* play a prominent role in the work-extraction statistics. We find that the eigenvalues of $\rho^{(\epsilon)}$ are

$$\lambda_{\pm}^{(\rho^{(\epsilon)})} = \tfrac{1}{2} \pm \tfrac{1}{2}\sqrt{r + \epsilon'^2} \; , \tag{55}$$

with corresponding eigenstates

$$\left|\lambda_{\pm}^{(\rho^{(\epsilon)})}\right\rangle = 2^{-1/2}\Big[r + \epsilon'^2 \mp \big(r + \epsilon'\sqrt{1-r}\big)\sqrt{r+\epsilon'^2}\,\Big]^{-1/2} \begin{bmatrix} \sqrt{r(1-r)} - \epsilon'\sqrt{r} \\ -r - \epsilon'\sqrt{1-r} \pm \sqrt{r+\epsilon'^2} \end{bmatrix} \; . \tag{56}$$

For all times after $t = 0$, the update rule for belief states simplifies to the following

$$\boldsymbol{\eta}_{t+1}\Big|_{W_{t+1}=w^{(\pm)}} = \frac{\sum_{x\in\mathcal{X}} \left\langle \lambda_{\pm}^{(\xi_t)} \middle| \sigma^{(x)} \middle| \lambda_{\pm}^{(\xi_t)} \right\rangle \boldsymbol{\eta}_t T^{(x)}}{\lambda_{\pm}^{(\xi_t)}} \; , \tag{57}$$

which can be expressed explicitly in terms of p, r, and ϵ_t.

When $\epsilon_t = 0$, we find that $\boldsymbol{\eta}_{t+1} = \left[\frac{1}{2}, \frac{1}{2}\right] = \boldsymbol{\pi}$. I.e., the stationary distribution is a fixed point for this dynamic over belief states. Because of this, we break the initial symmetry by setting ϵ to a small non-zero value to obtain useful knowledge. In other words, for the very first work-extraction protocol, we choose some $\rho_0^* \neq \xi_0$ to avoid an unstable fixed point of the update rule. However, for all subsequent time steps, we choose $\rho_t^* = \xi_t$.

For the perturbed coin, the metadynamic of the belief state in the long run will yield two different results, depending on which regime the system is in, "memory-apathetic regime" or "memory-advantageous regime".

The reason for this separation comes from the shape of their update function. For the memory-apathetic region, the update function has gradient less than unity, making $\epsilon = 0$ an attractor. For the memory-advantageous region, the gradient of the update function exceeds unity, therefore making $\epsilon = 0$ a repellor, at the same time two other points become part of a new attractor.

In the long run, transient belief states die out, leaving only the steady-state dynamics among the recurrent states of knowledge; any initial distribution over belief states generically converges to the stationary measure $\boldsymbol{\pi}_{\mathcal{K}}$. Hence the steady-state rate of work extraction is given by

$$\lim_{t\to\infty} \langle W_t \rangle = k_B T \langle \mathrm{D}[\xi_t \| \gamma] \rangle_{\Pr(K_t) = \boldsymbol{\pi}_{\mathcal{K}}} \; , \tag{58}$$

The expected extracted work for the memory-apathetic region coincide with that of memoryless extraction and is given by

$$\left\langle W^{\text{apathetic}} \right\rangle = k_B T \mathrm{D}[\xi_0 \| \gamma] = \left\langle W^{\text{memoryless}} \right\rangle \; . \tag{59}$$

On the other hand, in the regime where memory enhances the performance of the protocol, the stationary distribution over the two recurrent belief states $\boldsymbol{\eta}$ and $\boldsymbol{\eta}'$, with corresponding expected quantum states ξ and ξ', is

$$\boldsymbol{\pi}_{\mathcal{K}} = \frac{1}{\lambda_+^{(\xi)} + \lambda_+^{(\xi')}} [\lambda_+^{(\xi')}, \lambda_+^{(\xi)}] \; . \tag{60}$$

Hence, the work extraction rate is given by

$$\left\langle W^{\text{advantage}} \right\rangle = \frac{k_B T}{\lambda_+^{(\xi)} + \lambda_+^{(\xi')}} \Big(\lambda_+^{(\xi')} \mathrm{D}[\xi \| \gamma] + \lambda_+^{(\xi)} \mathrm{D}[\xi' \| \gamma] \Big) \; . \tag{61}$$

H.2 Classical approach

The derivation for the memory-assisted classical approach is similar to that of the memory-assisted quantum approach illustrated above. However rather than operating on the induced expected state ξ_t, the classical approach uses work-extraction protocols that are thermodynamically optimized for the decohered state

$$\rho_t^* = \xi_t^{\text{dec}} = \frac{1}{2} \begin{bmatrix} 1 + r + \epsilon_t'\sqrt{1-r} & 0 \\ 0 & 1 - r - \epsilon_t'\sqrt{1-r} \end{bmatrix} \; . \tag{62}$$

The eigenstates of ρ_t^* are thus $|0\rangle$ and $|1\rangle$, independent of time in this case. In the classical approach, $\boldsymbol{\pi}$ is no longer a fixed point of the belief-state update maps. The transition probabilities between belief states are now given by

$$\lambda_{\pm}^{(\xi_t^{\text{dec}})} = \tfrac{1}{2}\Big[1 \pm \big(r + \epsilon_t'\sqrt{1-r}\,\big)\Big] \; . \tag{63}$$

The metadynamic of belief in the classical case behaves as a reset processes. Unlike the quantum case with only two recurrent belief states, the classical protocol induces an infinite set of recurrent belief states. To construct a finite-state autonomous engine, we could choose to truncate those states within some small δ distance from another recurrent state, or truncate belief states with negligible probability, with vanishing work-extraction penalty.

We find that the work-extraction rate can again be computed by averaging the relative entropy—now between the decohered expected state and thermal state—over all recurrent states of knowledge:

$$\langle W_t^{\text{classical}} \rangle = k_B T \langle \mathrm{D}[\xi_t^{\text{dec}} \| \gamma] \rangle_{\Pr\left(K_t^{\text{classical}}\right)} \; . \tag{64}$$

H.3 Overcommitment to the most likely outcome

The "overcommitment" approach used for comparison in the main text bets exclusively on the most likely outcome in $\{\sigma^{(x)}\}_x$.

The expected thermodynamic cost of misaligned expectations during work extraction can be quantified exactly via the relative entropy $\mathrm{D}[\rho_0 \| \alpha_0]$ between the actual input ρ_0 and the anticipated input α_0 that the protocol is optimal for, if we assume that the final state is independent of the initial state [18, 20]. Hence, if we design the protocol for a pure state, but operate on a mixed state, we will encounter divergent thermodynamic penalties.

Accordingly, we can observe divergent thermodynamic costs when we design the Skrzypczyk work extraction protocol to be optimal for operation on a pure state.

Using the Skrzypczyk protocol (with N relaxation steps) to extract work from the pure state bet upon, we see that the first bath state swapped with the system for energy extraction is not exactly pure, but rather satisfies $\gamma_{\text{B}} = \left(1 - \frac{e^{-\beta E_0}}{N(e^{-\beta E_0} + e^{-\beta E_1})}\right) |0\rangle \langle 0| + \frac{e^{-\beta E_1}}{N(e^{-\beta E_0} + e^{-\beta E_1})} |1\rangle \langle 1|$. (Recall that H is the Hamiltonian for the system, not of the bath.) Any purity of the actual input beyond this initial bath purity is wasted. The input state leading to minimal entropy production under this protocol is thus a unitary rotation of γ_{B}.

Thus, for this use case of the Skrzypczyk protocol, the minimally dissipative state α_0 becomes pure as $N \to \infty$. As $N \to \infty$, we observe the battery's final expected energy diverging (but only logarithmically in N) to negative infinity, when this protocol acts on any other state. I.e., $\langle W \rangle \sim -k_B T \ln N$.

More specifically, we can leverage Eqs. (6) and (7) to calculate the expected value of work for the overcommitment approach. We find that

$$\Pr\Big(W = w^{(-)} | \sigma^{(\operatorname{argmin}_x \boldsymbol{\eta}_t T^{(x)} \mathbf{1})}\Big) = |\, \langle 1|\psi|1|\psi\rangle\, |^2 = 1 - r \; . \tag{65}$$

With $\lambda_- = \frac{e^{-\beta E_1}}{N(e^{-\beta E_0} + e^{-\beta E_1})}$, $w^{(-)} \sim -k_B T \ln N$, and $\min_x \boldsymbol{\eta}_t T^{(x)} \mathbf{1} \sim \min(p, 1-p)$ when $\boldsymbol{\eta}_t$ is close to either latent state, we anticipate that the overcommited work penalty diverges as $-k_B T (1-r) \min(p, 1-p) \ln N$, as observed.

Interestingly, for a finite number of bath interactions, some work can be extracted on average within certain regimes. But other regions of parameter space would yield very negative work-extraction averages.

Unlike the other approaches, the expectation value of work in the overcommitment approach cannot be written as a relative entropy. Hence, whereas the other approaches were guaranteed to have non-negative work extraction on average, the overcommitment approach enjoys no such guarantee of non-negativity. Indeed in the limit of many bath interactions, the overcommitment approach leads to infinitely negative work extraction.

I Non-Markovian generators

Unlike the classical case, the classical control symbols X_t are hidden from direct observation when the process emits non-orthogonal quantum states. The latent-state generators may thus be referred to as 'doubly-hidden Markov models'. Accordingly, even if the intermediary X_t process is Markovian, this would not directly imply a meaningful sense of quantum Markovianity of the outputs. Nevertheless, there is some sense in which processes with non-Markovian control outputs X_t have more deeply hidden structure.

To benchmark the performance of the memory-assisted protocol on a process with higher Markov order of the control symbols X_t, the 2-1 golden-mean process was chosen for comparison. The time-averaged density matrix of the memoryless approach for both models is kept the same, $\xi_0 = \frac{1}{2}(\sigma^{(0)} + \sigma^{(1)})$. The comparison is shown in Fig. 8, where we see that more work is extracted from the non-Markovian generator. This example suggests that memory can become even more important for enabling work extraction from non-Markovian generators of quantum processes, since the extractable structure can be more deeply hidden.

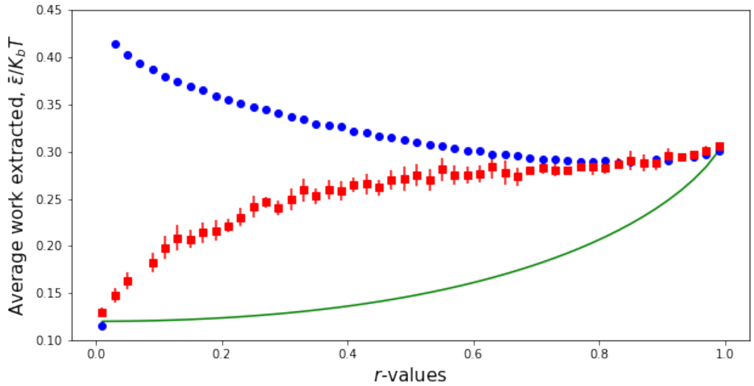


Figure 8: Comparison of average work extracted between 2-1 golden-mean and perturbed-coin processes, varying the nonorthogonality parameter r. Blue and red dots represent the memory-assisted quantum approach on golden mean and perturbed coin respectively; Green line represents the memoryless approach.