Generalized Continuous Maxwell Demons

Juan P. Garrahan[1, 2] and Felix Ritort[3, 4]

[1]*School of Physics and Astronomy, University of Nottingham, Nottingham, NG7 2RD, UK*
[2]*Centre for the Mathematics and Theoretical Physics of Quantum Non-Equilibrium Systems,*
University of Nottingham, Nottingham, NG7 2RD, UK
[3]*Small Biosystems Lab, Condensed Matter Physics Department,*
Universitat de Barcelona, C/ Martí i Franquès 1, E-08028, Barcelona (Spain)
[4]*Institut de Nanociència i Nanotecnologia (IN2UB),*
Universitat de Barcelona, E-08028 Barcelona, Spain
(Dated: February 12, 2023)

We introduce a family of Generalized Continuous Maxwell Demons (GCMDs) operating on idealized single-bit equilibrium devices that combine the single-measurement Szilard and the repeated measurements of the Continuous Maxwell Demon protocols. We derive the cycle distributions for extracted work, information-content, and time and compute the power and information-to-work efficiency fluctuations for the different models. We show that the efficiency at maximum power is maximal for an opportunistic protocol of continuous type in the dynamical regime dominated by rare events. We also extend the analysis to finite-time work extracting protocols by mapping them to a three-state GCMD. We show that dynamical finite-time correlations in this model increase the information-to-work conversion efficiency, underlining the role of temporal correlations in optimizing information-to-energy conversion. The effect of finite-time work extraction and demon memory resetting is also analyzed. We conclude that GCMD models are thermodynamically more efficient than the single-measurement Szilard and preferred for describing biological processes in an information-redundant world.

I. INTRODUCTION

Information-to-work conversion is a fundamental process in physics and biology. Paradigmatic examples are the Maxwell demon, and the Szilard engine, small operating devices that fully convert heat into work using measurement information [1–3]. According to Landauer and Bennett such devices do not violate the second law as the erasure procedure required to restore the system's initial state increases the overall entropy offsetting Demon's gain [4]. The field of thermodynamics of information is witnessing major progress [5–7] as these ideas are expanding into new directions [8–12], being also experimentally tested [13–22].

Recently, a continuous version of the Maxwell demon (CMD) has been introduced and implemented in a Szilard information-to-work engine operating on single DNA hairpins [23, 24]. The CMD offers a view complementary to the standard Szilard engine that is analytically solvable and can be implemented experimentally. In the Szilard engine (hereafter referred to as SZ), a gas particle occupies a vessel with two compartments in contact with a thermal bath at temperature T. At a given time, an observation is made of the compartment occupied by the particle. A work extraction process follows by inserting a pulley mechanism with a movable wall such that the compartment reversibly expands under the particle collisions. The combination of measurement, work extraction, and demon resetting defines a cycle in the SZ engine. For a given compartment, the average extracted work per cycle equals $-k_B T \log p$ where p is the probability of observing the particle in that compartment. For two compartments of probability p_0, p_1 the average extracted work per cycle in the SZ engine equals $W_{\mathrm{SZ}} = -k_B T (p_0 \log p_0 + p_1 \log p_1)$ which is bounded from above by $W_{\mathrm{SZ}} \leq k_B T \log 2$ for the case $p_0 = p_1 = 1/2$. For an irreversible work extraction protocol, the average work extracted is further bounded by the information-content of the single-bit measurement expressed in nats, $I_{\mathrm{SZ}} = -(p_0 \log p_0 + p_1 \log p_1)$ also called the Landauer limit, $W \leq k_B T I$. The CMD setting is the same as for SZ; the only difference is the work extraction condition and the cycle length. In the CMD, an observation of the compartment occupied by the particle is made, but work is not extracted right away. Instead, measurements are repeatedly made every time τ from the initial observation until a change in the compartment occupied by the particle is observed. Note that in the CMD a decision to extract work is taken based on a series of observations made at every τ-consecutive measurement, irrespective of any unobserved transition in between. Therefore, any cycle in the CMD contains at least two bits (for the first compartment observation and the next observed compartment change) rather than the single-bit cycle of SZ. The information-content of the multiple-bit stored sequences of the CMD cycle is always larger than for SZ, permitting the extraction of more work per cycle in the former. The average work per cycle extracted in the CMD is τ-independent and given by, $W_{\mathrm{CMD}} = -k_B T (p_0 \log p_1 + p_1 \log p_0)$. In contrast to SZ, W_{CMD} is now bounded from below by $k_B T \log 2$, $W_{\mathrm{CMD}} \geq k_B T \log 2$, with $W_{\mathrm{CMD}} = k_B T \log 2$ for $p_0 = p_1 = 1/2$. The average work per cycle is bounded from above by the information-content of the multiple-bit sequences generated by the CMD, $W_{\mathrm{CMD}} \leq k_B T I_{\mathrm{CMD}}(\tau)$ where τ is the time between consecutive measurements. For uncorrelated

observations ($\tau \gg 1/R$ with R systems' relaxation rate), the information-content of the stored sequences is minimal and equals $I_{CMD}(\tau \to \infty) = -(p_0 \log p_0)/p_1 - (p_1 \log p_1)/p_0 + W_{CMD}/k_B T$ defining the Landauer limit in this case, $W_{CMD} < k_B T I_{CMD}$. Given the simplicity of the CMD protocol and its complementarity to SZ, it is surprising that the CMD was not conceived before. The motivation in Ref. [23] was to devise a model for which large amounts of work could be extracted. The average work extracted $-k_B T \log p$ is large for compartments of low p, so a measurement protocol designed to seek for low p events maximizes work extraction. Indeed, in the limits $p_0 \to 0$ and $p_0 \to 1$, W_{CMD} diverges while W_{SZ} vanishes. What is the biological relevance of these models? There have been several applications of Maxwell demons to biological systems, particularly in the problem of signal transduction in regulatory pathways with feedback loops. However, there is no work extraction in such cases but rather what has been termed transfer entropy and a second law relating the transfer entropy to the system's entropy production. It is no surprise there are no direct applications yet of Maxwell demons to work extraction processes in biological systems but rather entropy transfer calculations in regulatory pathways.

In the single-molecule experimental realization of the SZ and CMD, a DNA hairpin hops between two states, folded and unfolded, of probabilities p_0 and $p_1 = 1 - p_0$, and a discrete or continuous feedback protocol was implemented to extract work depending on the measurement outcome [23, 24]. Work extraction is operated at the first measurement time in the SZ engine. In contrast, in the CMD, measurements are repeatedly made at intervals $\tau \sim 10^{-3}s$, and work is extracted only when the system is observed to switch state for the first time (folded to unfolded or unfolded to folded). As previously said, the CMD exhibits novel features as compared to the SZ, such as a larger average work per cycle (with the Landauer limit being a lower bound rather than an upper bound) and maximum efficiency $\eta < 1$ in the regime dominated by rare transition events ($p_0 \to 0, 1$).

Here we combine the SZ and CMD models into a new work extraction protocol, where either one of the options (SZ or CMD) is implemented with a finite probability depending on the measurement outcome (0 or 1). The so-called generalized Continuous Maxwell demon (GCMD) defines a family of models interpolating between SZ and CMD.

What is the biological relevance of such kinds of models? There have been efforts to translate Maxwell demons to biological systems, particularly for signal transduction in regulatory pathways with feedback loops. In these cases, there is no apparent work extraction but a second law relating transfer entropy to the system's entropy production. The GMCD model is an example of a double-protocol model for information-to-energy conversion. It is akin to molecular folding models with two productive folding pathways: one pathway ends in a low-free energy molecular state while the other ends in a higher-free energy molecular state. Similarly, one can envision a regulatory process with two distinct feedback pathways with different transfer entropies. The applicability of the Maxwell demon to biology remains a long-debated and open question.

This paper investigates power fluctuations in the GCMD, establishing the fundamental differences between the two classes (discrete SZ versus continuous CMD) of information machines. The paper is divided into six main sections. In Section II we introduce the generalized continuous Maxwell demon (GCMD) as a two-pathway model for information-to-energy conversion that combines features of SZ and CMD. Section III analyzes cycle-power and cycle-efficiency fluctuations in the GCMD, the main distinctive feature of continuous models compared to the SZ model. Section IV addresses the temporal correlations introduced in Ref.[19] on the power and efficiency at finite times. Finally, Sections V and VI are devoted to discussion and conclusions.

II. THE GENERALIZED CONTINUOUS MAXWELL DEMON (GCMD) MODEL

We introduce a new family of generalized continuous Maxwell demons (GCMD) that expands the previously studied work extraction protocols to multiple repeated measurements and has the SZ and the CMD as particular cases. Figure 1 illustrates the GCMD in the case of a single-bit measurement outcome (0,1).

A GCMD operates as follows. A system hops between two states $\sigma = 0, 1$ generating a two-level dichotomous signal (Fig. 1a) of probability p_σ ($p_0 + p_1 = 1$). Examples are: (i) a free molecule in a volume V that is subdivided into two compartments (V_0, V_1) such that $V = V_0 + V_1$; (ii) a biomolecule in solution with two conformations, folded ($\sigma = 0$) and unfolded ($\sigma = 1$). For a system in equilibrium, we have $p_0 = 1/(1 + \exp(\Delta G/k_B T))$ with ΔG the free energy difference between states 0 and 1 (hereafter, we set $k_B T = 1$). In example (a), $\Delta G = \ln(V_1/V_0)$ and $p_\sigma = V_\sigma/V$ (Fig. 1b,top). In example (b), ΔG is the folding free energy and $p_0/p_1 = \exp(-\Delta G)$ (Fig. 1b,bottom). A cycle in the GCMD starts with an observation of the system's state σ and two possible actions: with probability q_σ the SZ protocol is operated, and the amount of work per cycle ($W_\sigma^{SZ} = -\log p_\sigma$) readily extracted; with probability $1 - q_\sigma$ the CMD protocol is operated, meaning that new observations are made every τ until the system switches state $\sigma \to 1 - \sigma$ and the work per cycle ($W_\sigma^{CMD} = -\log p_{1-\sigma}$) extracted. The q_0, q_1 define two independent processes with $0 \leq q_0, q_1 \leq 1$ (i.e. $q_0 + q_1$ must not be equal to 1). The model interpolates between SZ ($q_0 = q_1 = 1$) and CMD ($q_0 = q_1 = 0$). The GCMD model is symmetric with respect to the transformation $p_0 \leftrightarrow 1 - p_0$ and $0 \leftrightarrow 1$, so we can restrict the analysis to $0 < p_0 \leq 1/2$.
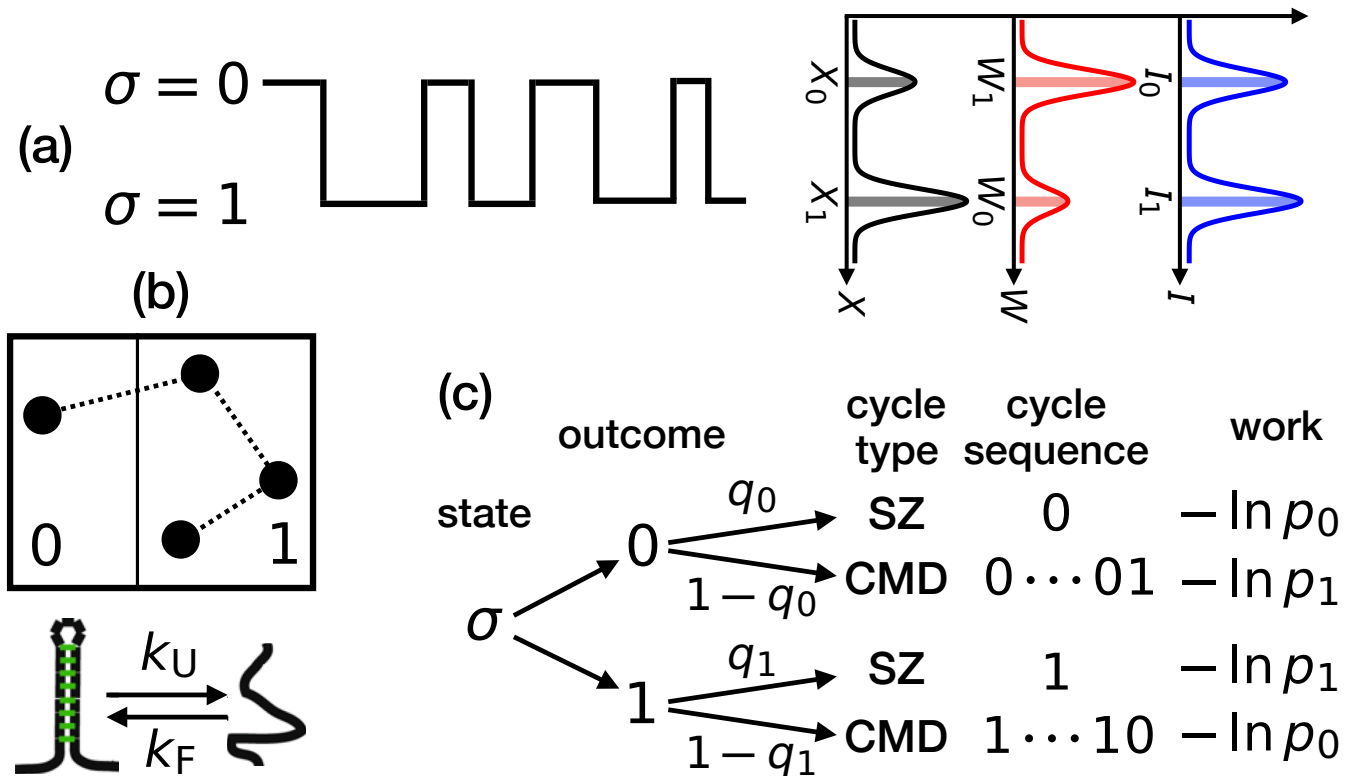
FIG. 1: **Illustration of the GCMD.** (a) Single-bit dichotomous signal (left) and position, work and information distributions (right). In a noisy environment fluctuations widen the single-peaked distributions into Gaussians. (b) Two examples of the model: a single particle moving in a compartmentalized vessel and a molecule hopping between two states. (c) Schematics of the work-extracting protocol in the GCMD. (SZ: Szilard-type protocol; CMD: Continuous-type protocol).

A. Average work, information-content, and cycle-time

The average work per cycle in the GCMD (in $k_B T$ units) is given by,

$$\overline{W} = \sum_{\sigma=0,1} p_\sigma \left(q_\sigma W_\sigma^{\text{SZ}} + (1 - q_\sigma) W_\sigma^{\text{CMD}} \right) \tag{1}$$

$$= - \sum_{\sigma=0,1} p_\sigma \left(q_\sigma \ln p_\sigma + (1 - q_\sigma) \ln(1 - p_\sigma) \right) .$$

Similarly, we analyze the information-content of the stored sequences of cycles $\mathcal{C}$, defined as $I(\mathcal{C}) = -\ln P(\mathcal{C})$ (nat units). SZ cycles are one-bit sequences, $\mathcal{C} = \{\sigma\}$, with $P(\mathcal{C}) = p_\sigma q_\sigma$. In contrast CMD cycles consist of $n + 1$ ($n \geq 1$) bit-sequences that start at the first bit (σ) which is repeated n times, until the bit outcome switches ($\sigma \to 1 - \sigma$) at the $(n+1)^{\text{th}}$ time . Therefore, a CMD cycle contains at least two-bits, $\mathcal{C} = \{\overbrace{\sigma, \sigma, ..., \sigma}^{n}, 1 - \sigma\}$, where $1 - \sigma$ indicates state switching. Note that the stopping time n is stochastic, and varies from cycle to cycle. The probability of the CMD cycle is given by $P(\mathcal{C}) = p_\sigma (1 - q_\sigma) T_{\sigma\sigma}^{n-1} (1 - T_{\sigma\sigma})$, where $T_{\sigma\sigma}$ is the conditional probability of a repeated measurement outcome σ after time τ. It has been shown [23, 24] that the lowest sequence information-content (i.e. minimum redundancy) is obtained for fully uncorrelated bit sequences, that is, when the relaxation kinetic rate of the system R is such that $R\tau \gg 1$, in which case $T_{\sigma\sigma} = p_\sigma$. In this limit, the probability of a GCMD cycle is given by,

$$P_{\text{GCMD}}(\mathcal{C}) = \begin{cases} q_\sigma p_\sigma & (n = 1, \text{SZ}) \\ (1 - q_\sigma) p_{1-\sigma} p_\sigma^n & (n \geq 1, \text{CMD}) \end{cases} , \tag{2}$$

and the corresponding average information-content

$$\overline{I} = \sum_{\mathcal{C}} P(\mathcal{C}) I(\mathcal{C}) = - \sum_{\sigma=0,1} p_\sigma q_\sigma \ln(p_\sigma q_\sigma) - \tag{3}$$

$$\sum_{\sigma=0,1} (1 - q_\sigma)p_{1-\sigma} \sum_{n=1}^{\infty} p_\sigma^n \ln((1 - q_\sigma)p_{1-\sigma}p_\sigma^n)$$

where we used $p_{1-\sigma} = 1 - p_\sigma$.

In the context of the thermodynamics of feedback, the thermodynamic efficiency of information-to-energy conversion is defined as the ratio between the average work extracted ($\overline{W}$) relative to the average energy converted into heat ($\overline{Q}$) that is necessary to erase the stored sequences: $\eta_{\text{th}} = \overline{W}/\overline{Q}$. According to the second law, $\overline{W} \leq \overline{Q}$ and $\eta_{\text{th}} \leq 1$. Following Landauer, the minimum energy $\overline{Q}$ for erasure (assuming a zero-work measurement process) is given by the Shannon information-content $\overline{I}$ of the stored sequences [19, 25] as given in Eq. (3) in the limit case of decorrelated measurements, $R\tau \gg 1$. Therefore, from Eqs.(1,3) we define the *thermodynamic efficiency* as $\eta_{\text{th}} = \overline{W}/\overline{I}$. The second law ensures that $\eta_{\text{th}} \leq 1$ which for SZ is saturated ($\eta_{\text{th}} = 1$) while for CMD, $\eta_{\text{th}} > 1/3$ approaching 1 in the limit $p_0 \to 0$ where rare events dominate dynamics [23].

We also consider the thermodynamic power $P_{\text{th}} = \overline{W}/\overline{t_C}$, as the ratio of $\overline{W}$, Eq. (1), and the average cycle time, $\overline{t_C}$. Combining SZ cycles (duration $n = 1$) and CMD cycles ($n + 1$ duration, with $n \geq 1$), we get for $\overline{t_C}$ (in τ units),

$$
\begin{aligned}
\overline{t_C} &= \sum_{\sigma=0,1} \left[1 \cdot p_\sigma q_\sigma + (1 - q_\sigma) \sum_{n \geq 1} (n+1)p_\sigma^n(1 - p_\sigma) \right] \\
&= 1 + (1 - q_0)\frac{p_0}{p_1} + (1 - q_1)\frac{p_1}{p_0}
\end{aligned}
\tag{4}
$$

which reduces to $\overline{t_C} = 1$ for SZ ($q_\sigma = 1$) and to $\overline{t_C} = (p_0(1 - p_0))^{-1} - 1 \geq 3$ [23] for CMD ($q_\sigma = 0$). While the CMD does extract more work than the SZ [23], this is at the price of a larger $\overline{t_C}$: in both cases, P_{th} vanishes asymptotically like $-p_0 \ln p_0$ when $p_0 \to 0$.

B. Family of GMCD models

Besides the SZ and CMD particular cases, we now consider other physically relevant choices defined by the q_σ values. These are:

(i) The *maximum work model* (MaxW) where $q_0 = 1, q_1 = 0$ if $p_0 \leq p_1$ ($q_0 = 0, q_1 = 1$ if $p_0 \geq p_1$) which ensures maximum extracted work $\overline{W} = -\ln(\min_\sigma(p_\sigma))$. This model may be called *opportunistic* because it extracts maximum work for all measurement outcomes.

(ii) The *constant work model* (ConstW) where $\overline{W} = \ln(2), \forall p_0$. From Eq. (1) we see that for this to occur q_0 and q_1 have to obey,

$$\text{ConstW:} \ \ p_0 q_0 + p_1(1 - q_1) = \ln(2p_1)/\ln(p_1/p_0). \tag{5}$$

(iii) The *thermostable model* (TS) where both $\overline{W}$ and $\overline{I}$ are independent of p_0 in which case $\eta_{\text{th}} = \text{const}$. This is obtained by applying Eq. (5) to solve for q_1 in terms of q_0 and p_0, replacing in $\overline{I}$, and then finding a dependence of q_0 on p_0 which makes $\overline{I}$ independent of p_0 [32].

Figure 2 shows $\overline{W}, \overline{I}, \overline{t_C}, \eta_{\text{th}}, P_{\text{th}}$ for the different models. We observe that SZ beats all models in terms of $P_{\text{th}}, \eta_{\text{th}}$; however, SZ also gives the lowest average work per cycle (Fig. 2a). While the models exhibit different features (CMD: lowest P_{th} and largest $\overline{I}$; MaxW: largest $\overline{W}$; ConstW and TS: large P_{th}) all GCMD variants show the same trend: $\eta_{\text{th}} \to 1$ and $P_{\text{th}} \to 0$ when $p_0 \to 0, 1$.

C. Physical interpretation of the GMCD model

The SZ protocol describes what might be called a *store-take strategy* where the most probable state is the one preferentially chosen for work extraction. Instead, in the CMD the first measurement of the system's state is followed by repeated measurements (store) until a specific condition is met (check) and the work extraction is operated (take). The *store-check-take strategy* of the CMD extracts work from the less probable state, maximizing average work extraction compared to SZ. One might call the two strategies low risk (SZ) and high risk (CMD). Here risk stands for repetitive measurement operations that increase the information content of the stored sequences and extract work only at the end. High-risk strategies rely on rare events of low probability p and large information content that deliver a large amount of work, $-k_B T \log p$. High-risk strategies rely on rare events of low probability p and large information content that deliver a large amount of work, $-k_B T \log p$. Instead, low-risk strategies rely on high probability p and low
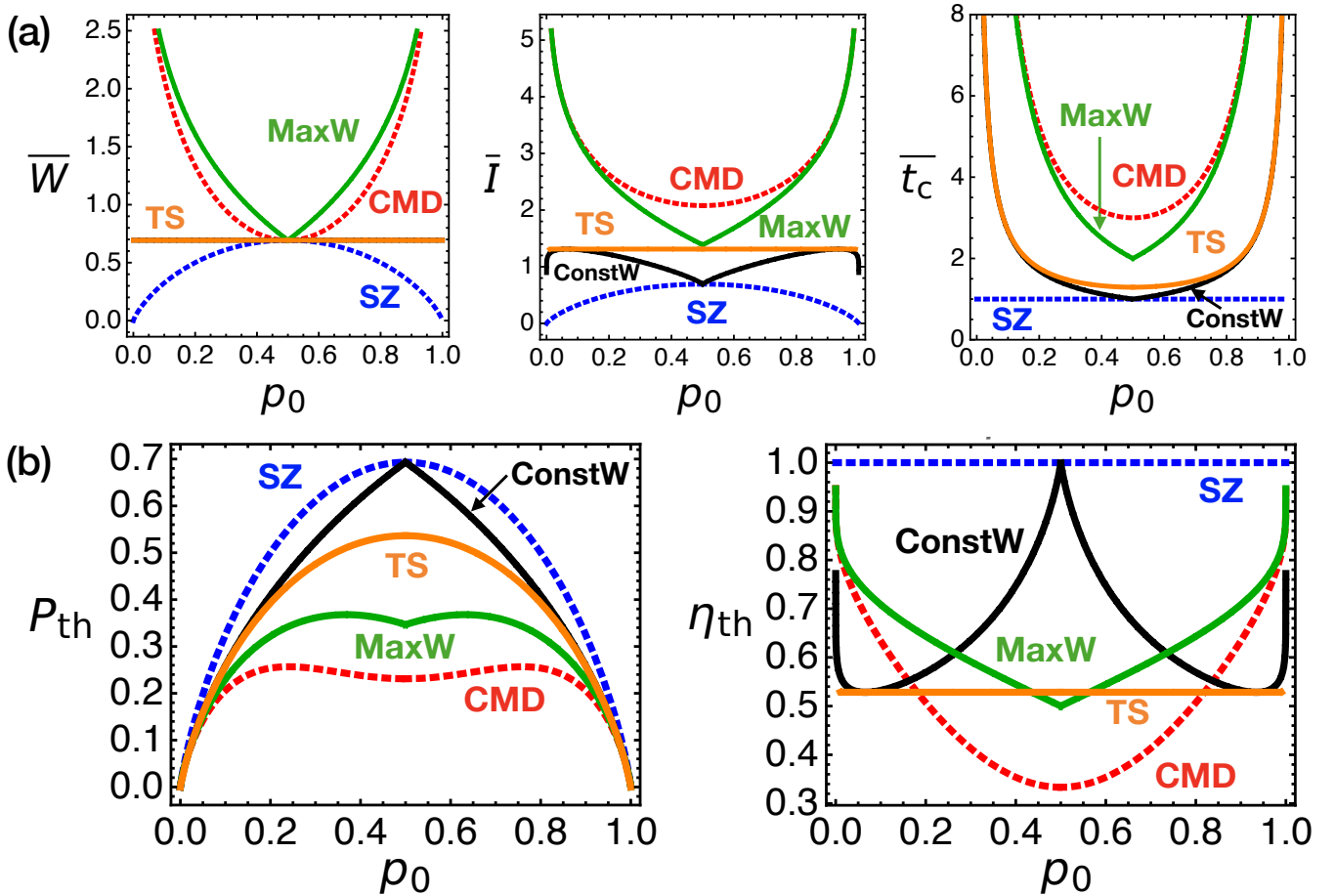
FIG. 2: **Thermodynamic parameters for the different GCMD models.** (a) Average work ($\overline{W}$), information-content ($\overline{I}$) and cycle-time ($\overline{t_\mathcal{C}}$) versus p_0. (b) Average power (P_{th}) and efficiency (η_{th}) versus p_0. In the limit $p_0 \to 0, 1$ all models reach maximum efficiency, $\eta_{\text{th}} \to 1$. Note the symmetry $p_0 \to 1 - p_0$. (SZ: Szilard model, blue; CMD: Continuous Maxwell Demon, red; MaxW: Maximum work, green; ConstW: Constant work, black; TS: Thermostable, orange).

information-content events that deliver a small amount of work, $-k_B T \log p \simeq k_B T (1 - p)$ for p close to 1. High-risk strategies (CMD) trade large amounts of work and information, whereas low-risk strategies (SZ) trade small amounts of work and information.

III. CYCLE-POWER AND CYCLE-EFFICIENCY FLUCTUATIONS

A main feature of the GCMD is the stochastic nature of W, I and $t_\mathcal{C}$, which leads to large fluctuations in the power and efficiency when measured over individual cycles [26],[27]. From Eq. (2) we can readily derive the corresponding $W, I, t_\mathcal{C}$ distributions from which the distributions for cycle-power $P = W/t_\mathcal{C}$ (in $k_B T/\tau$ units) and cycle-efficiency $\eta = W/I$ (denoted as cycle-P and cycle-η) follow. For cycle-P we get a discrete distribution

$$
\begin{aligned}
\mathcal{P}(P) = \sum_{\sigma=0,1} \Bigg[& p_\sigma q_\sigma \delta(P + \ln p_\sigma) \\
& + \frac{p_{1-\sigma}(1 - q_\sigma)}{p_\sigma^{1+(\ln p_{1-\sigma})/P}} \theta(P_{1-\sigma}^* - P) \Bigg] ,
\end{aligned} \tag{6}
$$

where θ is the Heaviside function, $P_\sigma^* = -(\ln p_\sigma)/2$ are thresholds, and the values that P can take in the argument of the Heaviside functions for each σ are discrete: $P = P_{1-\sigma}^*/(n+1)$ $(n \geq 1)$. Similarly, for cycle-η we get,

$$\mathcal{P}(\eta) = \sum_{\sigma=0,1} \left[p_\sigma q_\sigma \delta \left(\eta - \frac{\ln p_\sigma}{\ln(p_\sigma q_\sigma)} \right) + p_{1-\sigma}^{\frac{1}{\eta}} \theta(\eta_\sigma^* - \eta) \right] , \tag{7}$$

where $\eta_\sigma^* = (1 + \ln(p_\sigma(1-q_\sigma))/\ln p_{1-\sigma})^{-1}$ are thresholds, and in the Heaviside functions η takes discrete values $\eta = (1 + \ln(p_\sigma^n(1-q_\sigma))/\ln p_{1-\sigma})^{-1}$ $(n \geq 1)$. As expected, SZ $(q_0 = q_1 = 1)$ does not fluctuate: $\mathcal{P}(P) = p_0\delta(P + \ln p_0) + p_1\delta(P + \ln p_1)$ and $\mathcal{P}(\eta) = \delta(\eta - 1)$.

In the large fluctuation regime of continuous-type models $(p_0 \to 0)$, distributions can be approximated by:

$$\mathcal{P}_{p_0 \to 0}(P) \approx q_1\delta(P) + \frac{p_0(1-q_1)}{(1-p_0)^{1+(\ln p_0)/P}}, \tag{8}$$

$$\mathcal{P}_{p_0 \to 0}(\eta) \approx q_1\delta(\eta) + p_0^{1/\eta}, \tag{9}$$

(where the values of P and η in the second terms of the rhs, are discretized as above).

We find an interesting asymptotic behavior for power and efficiency fluctuations in the small p_0 limit. In Fig. 3a we show $\mathcal{P}(P)$ and $\mathcal{P}(\eta)$ for various values of $p_0 \ll 1$ for the MaxW model $(q_0 = 1, q_1 = 0)$: the probability of the power becomes increasingly uniform with decreasing p_0, while that of the efficiency peaks at $\eta = 1$. The insets in Fig. 3a show the second, third, and fourth cumulants of the distributions as a function of p_0 (all cumulants vanishing for SZ). In Fig.3b we show the first moments of (6) and (7), the average cycle-power $(\overline{P})$ and cycle-efficiency $(\overline{\eta})$ for several models. These are larger than the corresponding thermodynamic values, $\overline{P} \geq P_{\text{th}}$ and $\overline{\eta} \geq \eta_{\text{th}}$, for a wide range of p_0.

Notice that while P_{th} is always maximum for SZ, $\overline{P}$ is maximum for SZ in the intermediate regime $0.3 \lesssim p_0 \lesssim 0.7$ [33], while in the strongly fluctuating regime $p_0 \to 0$ (or $p_1 \to 0$), $\overline{P}$ is maximum for the MaxW model, see left panel of Fig. 3b (orange envelope). Although $\eta_{\text{th}} = \overline{\eta} = 1$ for SZ, this is at the cost of the lowest $\overline{P}$ in the limit $p_0 \to 0$ where GCMDs are most relevant. In the right panel of Fig. 3b we show the efficiency at maximum power, η^*, among all possible models (q_σ). In the intermediate regime $0.3 \lesssim p_0 \lesssim 0.7$ $\eta^* = 1$ is maximal for SZ, whereas if $p_0 \lesssim 0.3$ (or $p_1 \lesssim 0.3$) we get $\eta^* = \overline{\eta}_{\text{MaxW}}$. These results demonstrate: 1) the efficiency at maximum power is strongly sensitive to fluctuations and; 2) in the rare events regime, $p_0 \lesssim 0.3$, an opportunistic (MaxW) model maximizes the efficiency.

A. *Store-take* **versus** *store-check-take* **strategies**

One may ask which average (thermodynamical or cycle-average) is physically more relevant. In small biological systems (such as molecular machines operating in a cellular environment), we envision dynamics as a sequence of repeated regulatory-feedback cycles of different time duration. During each regulatory cycle, a physical variable is continuously monitored until a specific condition is met, at which point information-to-energy conversion occurs. For example, neuronal transmission consists of three phases [28]: first, a stimulus drives cell depolarization and a rise in membrane potential (*storage* phase); the membrane potential reaches a threshold value (*check* phase); the action potential (spike) is triggered followed by stimulus propagation and energy release at the terminal axon (*take* phase). Upon cycle termination, the system is reset, and the stored information is erased. The duration of these elementary cycles cannot be too long as the robustness of the stored information is at stake in the noisy cellular environment. In other words, the transduction and consumption of the energy accumulated in regulatory-feedback cycles inside the cell must occur over time intervals sufficiently short for the stored information to persist before the thermal forces erase it. In a strategy of *store-check-take*, it seems more appropriate to consider cycle quantities rather than thermodynamic quantities averaged over long times. The efficiency at maximum power η^* in the region $0 < p_0 \lesssim 0.3$ where continuous GCMD models show a higher power (Fig. 3b, right panel) suggests that the *store-check-take* strategy of repeated measurement protocols is better than the single-measurement action *store-take* of the SZ model.

IV. ENHANCED POWER AND EFFICIENCY BY TEMPORAL CORRELATIONS

We can further assess the fluctuating nature of cycle power and efficiency by calculating the average power and efficiency measured over a finite time t encompassing multiple work-extracting cycles. This defines a work-extracting engine that operates along consecutive cycles until time t, where the engine stops and work is extracted for the last time. To address this type of multi-cycle engine, we have extended the GCMD model of Fig. 1 to sequences of cycles,
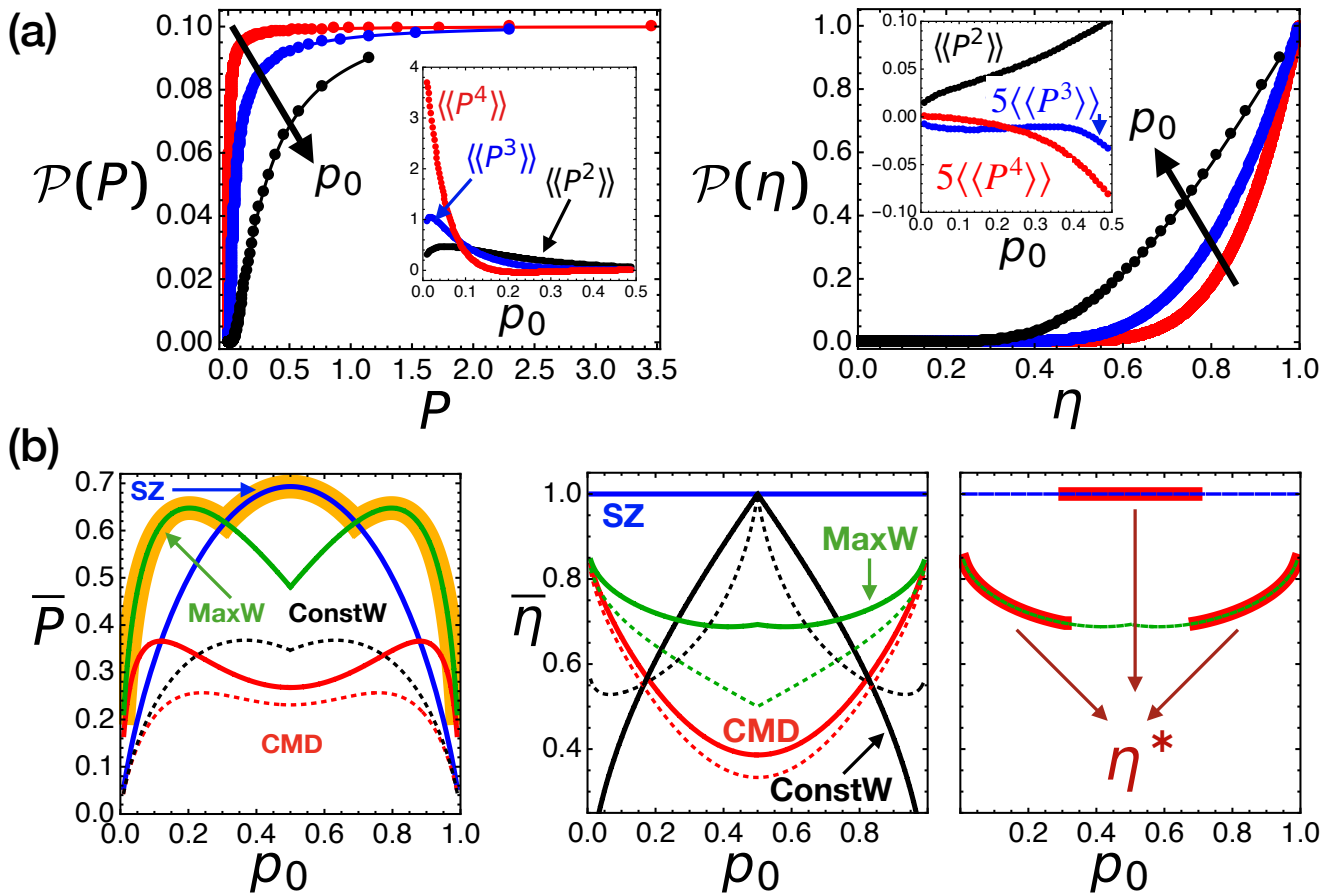
FIG. 3: **Statistics of cycle-power and cycle-efficiency for the different GCMD models.** (a) Cycle-power (left) and cycle-efficiency (right) distributions for decreasing values of $p_0 = 0.1$ (black), 0.01 (blue), 0.001 (red) for the MaxW model. Insets show the second, third, and fourth cumulants of the distributions versus p_0. (b) Average cycle-power ($\overline{P}$, left) and cycle-efficiency ($\overline{\eta}$, middle) versus p_0 (continuous lines). Dashed lines are the thermodynamic efficiencies of Fig.2. Orange thick line in the left panel is the maximum power as given by MaxW ($p_0 \lesssim 0.3$) and SZ ($0.3 \lesssim p_0 \leq 1/2$). Information-to-work efficiency at maximum power (right panel, red). Models abbreviations and colors as in Fig.2.

each cycle defined as before, i.e. a series of measurements terminating in a work extraction process, see Fig. 4a. We calculate the average work, information, power, efficiency, and number of cycles over a finite time t ($\overline{W}_t, \overline{I}_t, \overline{P}_t, \overline{\eta}_t, \overline{K}_t$). A sequence of measurements of time duration t consists of a series of SZ-cycles and CMD-cycles that are selected with probabilities q_σ and $1 - q_\sigma$, respectively, depending on the measurement outcome σ at the beginning of each cycle. The full sequence of measurements terminates at state σ_t from which the last work $-\ln p_{\sigma_t}$ is extracted. The end state σ_t determines three possible situations: it is the end of a CMD cycle, with $\sigma_t = 1 - \sigma_{t-1}$; it is an SZ cycle; or it is the end time reached before a cycle of the CMD-type has completed, in which case $\sigma_t = \sigma_{t-1}$.

Dynamics can be represented by three states, $\sigma = 0, 1$ (corresponding to mid-CMD cycle states), and an extra state $\sigma \equiv W$ for the work extraction steps. We denote this as the 3-GCMD model. The dynamics between the states are encoded in the three-state (discrete time) Markov network of Fig. 4b. For instance, for t=9 a possible sequence in the 3-GCMD model could be (Fig. 4b): $\{ \underbrace{0}_{\text{S}}, \underbrace{0,0,0,1}_{\text{C}}, \underbrace{1}_{\text{S}}, \underbrace{1}_{\text{S}}, \underbrace{1,0}_{\text{C}} \}$. This sequence has $K_t = 5$ cycles (three SZ -S- and two CMD -C- ones) and the work extracted equals $W_t = -2 \ln p_0 - 3 \ln p_1$. In the 3-GCMD model, the sequence reads $\{W, 0, 0, 0, W, W, W, 1, W\}$, and for calculating the average quantities ($\overline{W}_t, \overline{I}_t, \overline{P}_t, \overline{\eta}_t, \overline{K}_t$) it is not necessary to distinguish between the W states (all that matters are transitions ending in W, Fig. 4b).

For $t \to \infty$ it is easy to prove that $\overline{P}_t \to P_{\text{th}}$ and $\overline{\eta}_t \to \eta_{\text{th}}$, while $\overline{W}_t / \overline{K}_t \to W_{\text{th}}$ and $\overline{I}_t / \overline{K}_t \to I_{\text{th}}$, see Fig. 4c. Significantly, in some cases, the 3-GCMD performs better at finite times than it does at long times: in the bottom panel Fig. 4c we show that in the MaxW model $\overline{P}_t \geq P_{\text{th}}$ for all times, while both in the CMD and MaxW the finite-time efficiency $\overline{\eta}_t \geq \eta_{\text{th}}$ for all times.

Throughout the paper, we neglected temporal correlations in the Demon state (i.e., $R\tau \gg 1$), making the statistics
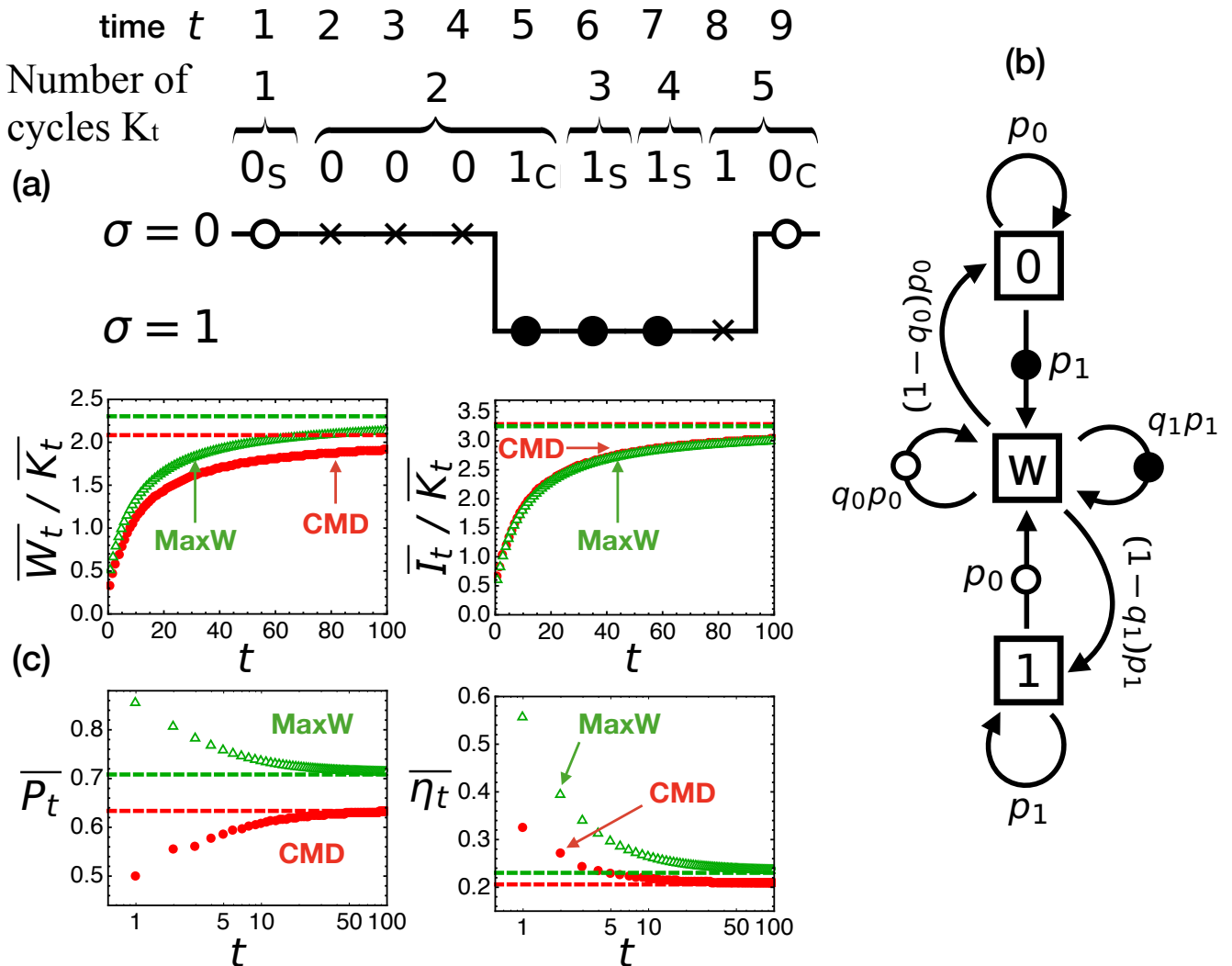
FIG. 4: **Power and efficiency at finite times: the 3-GCMD model.** (a) A series of repeated measurements terminating at a given finite time t ($t = 9$ in the illustration). Work is extracted at specific times (empty and filled circles) depending on the chosen protocol (SZ or CMD) and the measurement outcome ($\sigma = 0, 1$). Work is always extracted at the last measurement time t even if a CMD-type cycle has not terminated. (b) Three-state (3-GCMD) model ($\sigma = 0,1,W$) to calculate the work and information-content up to t. W denotes the third state corresponding to work extraction. (c) Average cycle-power and cycle-information (top) and average power and efficiency (bottom) versus t. Dashed lines are the $t \to \infty$ thermodynamic limit.

of all relevant quantities solvable for finite times in the 3-GCMD model. Interestingly, the intrinsic correlations of the 3-GCMD model (Fig. 4b) make $\overline{I}_t/\overline{K}_t$ decrease faster than $\overline{W}_t/\overline{K}_t$ for lower t (Fig. 4c,top) enhancing efficiency (Figure 4c,bottom). Such intrinsic correlations model are apparent from the fact that direct transitions $1 \to 0$ and $0 \to 1$ are forbidden in the 3-GCMD model (Fig. 4b). They are on top of those that would appear for correlated measurements $R\tau \sim 1$, in which case $T_{\sigma'\sigma} \neq p_{\sigma'}$ ($T_{\sigma'\sigma}$ being the probability of measuring σ' conditioned to measuring σ at the previous time τ). In general, $T_{\sigma'\sigma}$ also depends on σ, and such extra correlations should further increase efficiency. The role of temporal correlations in maximizing power and information-to-work efficiency has been considered in Ref.[19] where a wall is repeatedly moved to rectify the motion of a diffusive colloidal particle against a flow. It would be interesting to extend the present analysis to finite τ, and different stopping conditions [29, 30], finding optimal regimes that maximize efficiency. Combining SZ and CMD-type protocols appears to be a promising route to develop improved protocols for information-to-energy conversion. These can be readily implemented in currently available experimental setups.
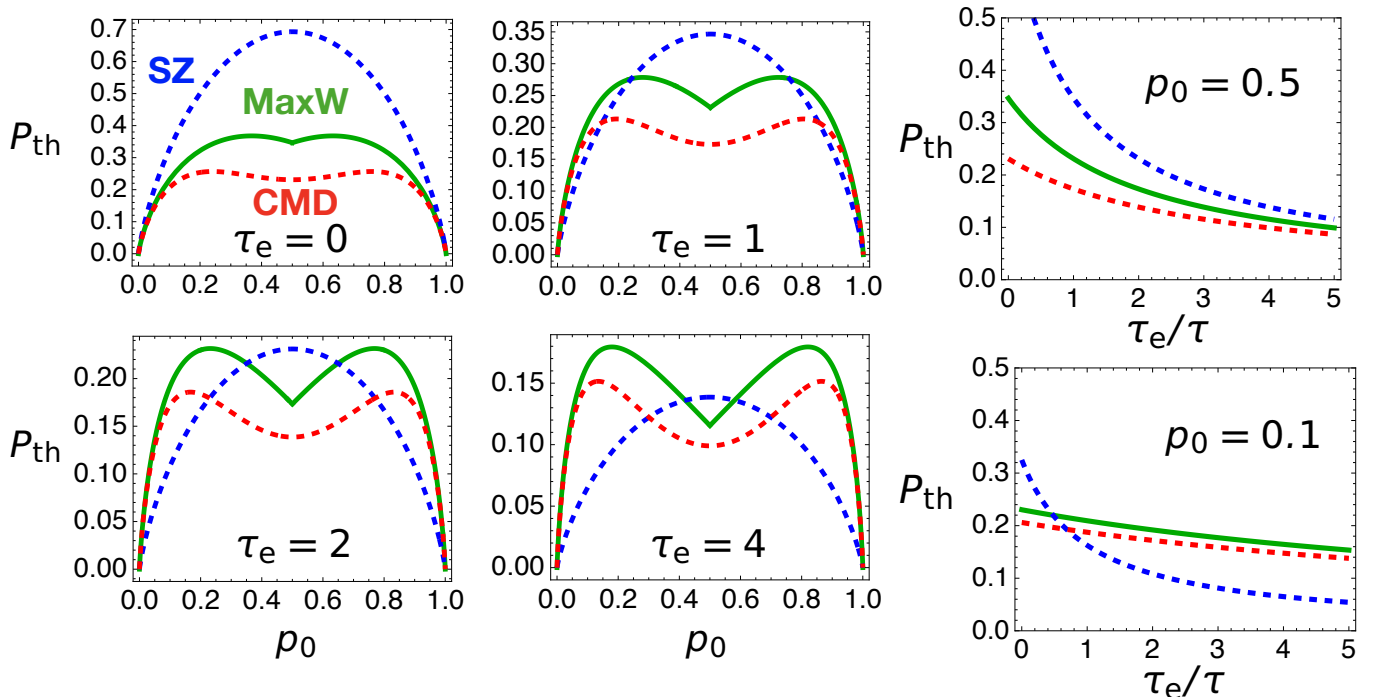
FIG. 5: **Effect of the finite-time τ_e for work extraction and demon resetting.** (a) Thermodynamic power P_{th} versus p_0 for the three models (SZ, blue dashed line; CMD, red dashed line; MaxW, green continuous line) for different values of τ_e with $\tau = 1$. (b) P_{th} versus τ_e/τ for $p_0 = 0.5$ (top) and $p_0 = 0.1$ (bottom). GCMD models become more efficient than SZ for $\tau_e/\tau > 1$. Same color and line codes as in (a)

A. Effect of finite-time work extraction and demon resetting

Throughout the paper, we assumed that the time of the work extraction and the resetting processes of the demon (erasure of the stored sequences) is negligible. This is a justified approximation in the CMD in the limit $R\tau \gg 1$ of largely-separated and uncorrelated measurements. Also, in this limit, redundancy of the stored sequences' information content is minimized, while the average extracted work is τ-independent, making the information-to-work efficiency maximum for all GMCD cases. However, a finite-time duration τ_e for the combination of the work extraction and reset-erasure processes of the demon might change the thermodynamic power and information-to-work efficiency of generic GCMD models relative to SZ. Indeed, a finite time τ_e would be detrimental to the performance of the SZ where the total experimental time t is increased by $K_t\tau_e$, where K_t is the number of measurement cycles. In contrast, for the same t, the number of cycles of the CMD-type K_t would be lower, and the average extracted work per cycle larger, increasing the thermodynamic power of the CMD relative to SZ. Therefore, for $\tau_e/\tau > 1$, the thermodynamic power of SZ will decrease relative to the other models (CMD,MaxW,ConstW,TS). In figure 5 we plot P_{th} as a function of p_0 (panel a) and τ_e/τ (panel b) for the CMD (red), MaxW (green) and SZ (blue) models. GCMD models become thermodynamically more efficient than SZ for $\tau_e/\tau > 1$.

V. DISCUSSION

The CMD-type and SZ-type cycles might be put in analogy with the anabolic pathways of cell metabolism. In anabolic or biosynthetic pathways, high-energy complex molecules are built in a sequence of steps. Conversely, high-energy molecules are broken into smaller components in catabolic or degradation pathways. In this analogy, the high information-content sequences (multiple-bit sequences) stored in the CMD might be seen as the equivalent of the high-energy molecules produced in the anabolic pathway. Instead, the low information-content sequences (single-bit sequences) in the SZ, are analogous to the much lower free energy molecules that can be produced in biosynthetic reactions with a fixed number of steps. A typical example of an anabolic reaction of the SZ-type is glucose production from carbon dioxide in photosynthetic cells. Here the SZ-type cycle is envisioned as a well-defined reaction that produces a molecule (glucose) with a well-defined amount of energy. In contrast, an anabolic reaction of

the CMD-type might be realized in the biosynthesis of a polypeptide chain. In protein synthesis, the amino acids are sequentially linked in a three-step process (initiation, elongation, and termination) where termination occurs when a specific condition (stop codon) is met. The analogy between the CMD and protein synthesis lies in the proteins of variable length and free energy that can be assembled during ribosomal translation. This is analogous to the varying information-content of the (multiple bits) sequences in the CMD. The interest of the GCMD model is that it allows for two information-content pathways, reminiscent of the different biosynthetic pathways of anabolism. An efficient metabolism requires synthesizing many molecules of variable length and free energy, each specific molecular type for a given metabolic step. Analogously, the GMCD model considers two information-to-energy conversion pathways (SZ and CMD) that generate bit sequences of variable information-content (single-bit sequences for SZ versus multiple-bit sequences for CMD).

VI.   CONCLUSIONS

An important quest in the field of thermodynamics of information is to search for protocols that maximize the information-to-energy conversion efficiency, a quantity that has been hypothesized to be optimized in small biological systems. Indeed, the large efficiency of many molecular motors and enzymes is remarkable. Light harvesting is an example of how the energy of a single light photon absorbed by chlorophyll powers single electron transfer reactions with almost 100% quantum efficiency. In our paper, we digress about neuronal transmission as an example of a work extraction cycle of three phases under information feedback: information storage (cell depolarization), state check (potential threshold), and energy release (voltage spike). It is argued that these activity cycles resemble more the continuous version of the MD (CMD) than the Szilard protocol (SZ). Indeed, no state-check condition is present in SZ, which is equivalent to a hypothetical neuronal transmission activity without threshold-driven spiking. There is no irrefutable evidence that efficiency is an optimized quantity in small biological systems, although it is clear the high efficiency achieved by biological processes when compared to equivalent man-made machines.

A general result of our study is the more significant thermodynamic efficiency of the SZ relative to the different variants of the GMCD with a CMD component ($q_\sigma < 10$). However, this holds if we assume a negligible time for the work extraction and erasure process (Sec.IV A). Moreover, the high thermodynamic efficiency of SZ comes at the price of a bounded extracted work per cycle ($< k_B T \log 2 \sim 0.69 k_B T$, Figure 2a). In contrast, the average work per cycle is much larger in the CMD, being unbounded in the limit $P_0 \to 0$ and $P_0 \to 1$. The lower thermodynamic efficiency of the CMD stems from the redundancy of the information content encoded in the multiple-bits stored sequences. Such a redundancy led us in Sec.II C to classify SZ as a low-risk strategy (low both average work and information content per cycle) compared to CMD (both large average work and information content per cycle). The large thermodynamic power for SZ is concurrent with the low work per cycle ($< k_B T \log 2 \sim 0.69 k_B T$), which does not even reach $1 k_B T$, poses severe limitations to SZ to perform arbitrary information-to-energy conversion tasks. From the standpoint of biological systems, it suggests that information-to-work efficiency is more relevant than thermodynamic power. The limited amount of extracted work per cycle does not cloud the CMD, a better model for biological processes in an information-redundant world.

Most results presented in this paper are purely mathematical. The message we tried to convey is that CMD realizations store words of multiple-bit sequences that can be converted into large amounts of work. The Szilard engine is the simplest single-bit information-to-energy conversion machine with limited average work and information content per cycle. Yet, there is room for designing innovative information-to-work protocols in systems with a finite number of states Ω that can extract more work than just $k_B T \log \Omega$. In a recent work [31], we have addressed the CMD with N states finding a rich phenomenology depending on the network of kinetic rates. Introducing N states into the GCMD will enrich the families of analytically solvable models for information-to-energy conversion. This will spur the thinking of novel experimental realizations of information-to-work protocols capable of transducing into energy multiple-bit sequences of significant information content. Plausibly, information content is also the main trait of the living entities generated during eons of natural evolution in an information-redundant world.

---

[1] H. S. Leff and A. F. Rex, Computing **349** (1990).
[2] C. H. Bennett, Int. J. Theor. Phys. **21**, 905 (1982).
[3] E. Lutz and S. Ciliberto, Physics Today **68**, 30 (2015).

[4] K. Maruyama, F. Nori, and V. Vedral, Reviews of Modern Physics **81**, 1 (2009).

[5] T. Sagawa and M. Ueda, Phys. Rev. Lett. **104**, 090602 (2010).

[6] T. Sagawa, J. Stat. Mech. **2014**, P03025 (2014).

[7] J. M. Parrondo, J. M. Horowitz, and T. Sagawa, Nature Phys. **11**, 131 (2015).

[8] P. Strasberg, G. Schaller, T. Brandes, and M. Esposito, Physical review letters **110**, 040601 (2013).

[9] É. Roldán, I. A. Martinez, J. M. Parrondo, and D. Petrov, Nature Physics **10**, 457 (2014).

[10] J. V. Koski, A. Kutvonen, I. M. Khaymovich, T. Ala-Nissila, and J. P. Pekola, Physical review letters **115**, 260602 (2015).

[11] I. A. Martínez, É. Roldán, L. Dinis, D. Petrov, J. M. Parrondo, and R. A. Rica, Nature physics **12**, 67 (2016).

[12] J. P. Pekola and I. M. Khaymovich, Annual Review of Condensed Matter Physics **10**, 193 (2019).

[13] S. Toyabe, T. Sagawa, M. Ueda, E. Muneyuki, and M. Sano, Nature Phys. **6**, 988 (2010).

[14] A. Bérut, A. Arakelyan, A. Petrosyan, S. Ciliberto, R. Dillenschneider, and E. Lutz, Nature **483**, 187 (2012).

[15] J. V. Koski, V. F. Maisi, J. P. Pekola, and D. V. Averin, Proc. Natl. Acad. Sci. USA **111**, 13786 (2014).

[16] M. D. Vidrighin, O. Dahlsten, M. Barbieri, M. Kim, V. Vedral, and I. A. Walmsley, Phys. Rev. Lett. **116**, 050401 (2016).

[17] M. Gavrilov and J. Bechhoefer, Phys. Rev. Lett. **117**, 200601 (2016).

[18] K. Chida, S. Desai, K. Nishiguchi, and A. Fujiwara, Nature Comms. **8**, 1 (2017).

[19] T. Admon, S. Rahav, and Y. Roichman, Phys. Rev. Lett. **121**, 180601 (2018).

[20] A. Kumar, T.-Y. Wu, F. Giraldo, and D. S. Weiss, Nature **561**, 83 (2018).

[21] G. Paneru, D. Y. Lee, T. Tlusty, and H. K. Pak, Phys. Rev. Lett. **120**, 020601 (2018).

[22] G. Manzano, D. Subero, O. Maillet, R. Fazio, J. P. Pekola, and É. Roldán, Phys. Rev. Lett. **126**, 080603 (2021).

[23] M. Ribezzi-Crivellari and F. Ritort, Nature Phys. **15**, 660 (2019).

[24] M. Ribezzi-Crivellari and F. Ritort, J. Stat. Mech. **2019**, 084013 (2019).

[25] R. Schmitt, J. Parrondo, H. Linke, and J. Johansson, New Journal of Physics **17**, 065011 (2015).

[26] G. Verley, M. Esposito, T. Willaert, and C. Van den Broeck, Nature communications **5**, 1 (2014).

[27] G. Paneru, S. Dutta, T. Sagawa, T. Tlusty, and H. K. Pak, Nature communications **11**, 1 (2020).

[28] W. Bialek, Biophysics: searching for principles (Princeton University Press, 2012).

[29] I. Neri, É. Roldán, and F. Jülicher, Physical Review X **7**, 011019 (2017).

[30] I. Neri, É. Roldán, S. Pigolotti, and F. Jülicher, Journal of Statistical Mechanics: Theory and Experiment **2019**, 104006 (2019).

[31] P. Raux and F. Ritort, Entropy **25**, DOI:10.3390/e25020321 (2023).

[32] There are multiple solutions of $q_{0,1}$ as functions of p_0 that give $\overline{W}$, $\overline{I}$ constant. We choose the one that simultaneously spans the largest range of p_0 and minimizes $\overline{I}$. This solution exists in the range $0.01 \lesssim p_0 \lesssim 0.99$.

[33] These limits are x and $1-x$, where x is the the solution of the equation $(1-p_0)\ln(1-p_0)+p_0\ln p_0(1+\ln p_0/(1-p_0)) = 0$.