Performance limits and trade-offs in entropy-driven chemical computers

Dominique Chu

School of Computing, University of Kent, CT2 7NF, Canterbury, UK
d.f.chu@kent.ac.uk

Abstract

The properties and fundamental limits of chemical computers have recently attracted significant interest as a model of computation, an unifying principle of cellular organisation and in the context of bio-engineering. As of yet, research in this topic is based on case-studies. There exists no generally accepted criterion to distinguish between chemical processes that compute and those that do not. Here, the concept of entropy driven computer (EDC) is proposed as a general model of chemical computation. It is found that entropy driven computation is subject to a trade-off between accuracy and entropy production, but unlike many biological systems, there are no trade-offs involving time. The latter only arise when it is taken into account that the observation of the state of the EDC is not energy neutral, but comes at a cost. The significance of this conclusion in relation to biological systems is discussed. Three examples of biological computers, including an implementation of a neural network as an EDC are given.

Keywords: biological computing, entropy, computational performance

1 Introduction

Computing architectures based on biochemistry, rather than semi-conductor technologies, are attracting increasing interest as alternative models of computation. Researchers have engineered DNA based computers [1, 2] or robots controlled by slime molds [3]. It has also been demonstrated experimentally that logical gates can be implemented in living cells [4, 5, 6] and it is thus possible, in principle, to implement any algorithm in wetware. Beyond these engineering aspects, the study of biochemical computation also gives theoretical insights into biology [7, 8]. There are a number of biosystems that have been studied as *in vivo* computers, including kinetic proofreading [9, 10] during translation, gene regulatory networks, chemotaxis [11], and bacterial sensing [12, 13, 14, 15] and most recently even bacterial growth dynamics [16, 17].

For cell-based computers it is found almost ubiquitously that they are subject to performance trade-offs. Typically, speed and accuracy of the computation can only be increased if its metabolic cost is also increased (see for example [18, 19]). Considerations of the energetic cost of computing has deep historical roots in science. A classical result is Landauer's limit [20] which states that erasing a bit of information requires a minimal amount of work. One would think that this imposes a lower cost limit on computation, but this is not so. Bennett demonstrated that logically reversible computers can avoid this cost and that there is no minimal cost limit to computation. In the context of stochastic or "Brownian" computers [21] Bennett also showed that in the limit of zero energy dissipation the computation takes infinitely long and will be inaccurate (see also [22]). These classical findings seem to correspond well to the computational performance trade-offs found in biological systems.

A general understanding of the fundamental limits of chemical computers is hindered by the lack of an agreed model of chemical computation. The current modus operandi in the field is to identify a biological system (such as sensing or proof-reading) as a computation when it performs a function that appears computational. This approach enables deep insights into specific examples, but is likely to miss most instantiations of chemical computation. It would be much more useful to have a concept of chemical computation that is independent of its function, just as in computer science computation is defined with respect to a number of specific mathematical models, not by reference to what is computed.

The best known model of computation is the *Turing machine*. This is a mathematical construct consisting of a "reading head" which is reading/writing to/from a tape until it reaches a "halting state," at which point the computation stops. It is believed that for every computable function there is a corresponding Turing machine that computes it. Based on this, one could be tempted to define a chemical process as a computation if there is a Turing machine that simulates this process. This does not work, however: The natural equivalent of a halting state in chemical systems is the equilibrium state, i.e. the state of the chemical system where reactions are in detailed balance. Unlike the halting state of a Turing machine, the equilibrium state is of a statistical nature.

This means that on average there are no net-fluxes across the network of reactions [23, 24], but reaction events are still continuing. In equilibrium the sequence of reaction events is symmetric in time [25]. Computation, on the other hand, is necessarily time directed, mapping a particular input to a particular output. Equilibrium systems are therefore not able to compute. Sample paths of equilibrium chemical systems can still be simulated and are thus computable by Turing machines. This demonstrates that not all processes that can be simulated by Turing machines are also themselves processing information.

In order to establish a general concept of chemical computing, we adopt here the working hypothesis that the equilibrium state is the only halting state of chemical computers. This entails that every chemical system that is not in equilibrium is in the process of computing. Stated differently: Every chemical system that is out of equilibrium is computing.

The purpose of this article is twofold. Firstly, we will introduce the concept of *entropy driven computers* (EDC) as a general model of chemical computation. Secondly, we will probe the computational performance of EDCs (i.e. speed, accuracy) in relation to their energy dissipation. We will find that in chemical computers there is no minimal energy consumption, but a trade-off arises between the cost of the computation and its accuracy. The time of the computation is independent of the cost, at least in first order approximation and determined by the biophysical parameters of the reacting species. This result is consistent with the conclusions reached by Bennett, but it does not explain the time-energy trade-offs that are often found in biological computers. In order to understand the origin of these, we expand the analysis to take into account the cost of measuring the result of the computation. While usually ignored, measurement is essential so as to be able to know the result of the computation. Measurement is not energy neutral, but requires a minimal amount of work to be applied that is related to Landauer's principle. The measurement step leads to performance trade-offs involving time. In the Discussion section, we argue that the cost of measuring the outcome of the computation is a major cost contribution to some biological computers.

2 Results

2.1 Entropy driven computation

In this section we define an EDC as a closed chemical system that is coupled to an infinite heat reservoir which is fixed at some temperature T. No exchange of particles with the environment is allowed. An EDC is initialised by a particular chemical composition at time $t = 0$, i.e. a specified abundance for each of its constituent chemical species. After a transient period the chemical system approaches an equilibrium state characterised by detailed balance. Strictly speaking, the transition to equilibrium takes an infinite amount of time. In practice, EDCs will be very close to equilibrium after a finite, possibly very short, time. We will model EDCs here as continuous time Markov chains. Then the time scale for the approach to equilibrium depends on the kinetic parameters appearing in the master-equation that defines the chemical model. A solution to the master equation will define a characteristic time scale over which the influence of initial conditions is forgotten. We will take this as the "speed" of the computation.

2.2 Basic definitions

Definition 1 (Microstate)**.** *Let the system S be a finite (although potentially very large), continuous time, reversible Markov chain (ctmc) defined on the states $s_1, s_2, \ldots, s_n$ and fixed transition rates $\omega_{ij} := \omega(s_i \to s_j)$. We define the set of states $\{s_1, s_2, \ldots, s_n\}$ as the microstates of S.*

For the sake of simplicity, but in abuse of notation, we will henceforth not distinguish between the chemical system of which S is a model and the model itself, and refer to both as the "system." The requirement that the ctmc model S is reversible entails that the chemical system S is a closed system.

The microstates of ctmcs are state labels that are not usually endowed with a direct physical meaning. They need to be translated to observables that correspond to measurable entities in the real world.

Definition 2 (Observable)**.** *Let o_i be the number of particles of species i of chemical system S. The function $f(s_j)$ maps microstates s_j of S to the corresponding o_i. We denote by $\Omega(t) = \{o_1(t), \ldots, o_l(t)\}$ a complete set of observables corresponding to the abundances of all species in the system, such that there is a one-to-one mapping from Ω to the microstates of S.*

Transitions between microstates are typically very fast, making them of little use to gain insight into the state of the system. More interesting is the behaviour of the system when averaged over a finite time Δt.

Definition 3 (Macrostate)**.** *Let S be a system and $P(s_i, t)$ the probability that S is in microstate s_i at time t and let $\Omega(t) = o_1(t), \ldots, \ldots, o_l(t)$ be a complete set of observables of S. The time-averaged observable at time $t > 0$ is given by*

$$m^{o_i}_{t,\Delta t} = \frac{1}{\min(\Delta t, t)} \int_{\max(0,t-\Delta t)}^{t} o_i(t')dt'.$$

The macrostate is defined as $M^{\Omega}_{t,\Delta t} = (m^{o_1}_{t,\Delta t}, \ldots, m^{o_l}_{t,\Delta t})$.

Note that $m^{o_i}_{t,0}$ is simply the value of the observable o_i at time t. $M_{\infty,\infty}$ is the vector of mean particle abundances at equilibrium. For sufficiently long choices of Δt the macrostate in equilibrium will often be independent of Δt and t. In systems that are characterised by meta-stable states where the probability distribution of some of the macro-observables is multi-modal, the choice of Δt may have to be very long before the macro-state becomes independent of Δt and t. To simplify notation, we will henceforth not explicitly indicate the observables Ω and assume Δt to be long, but finite, so that the macrostate is approximately independent of time.

With the notation being clarified, we can now define the concept of entropy driven computation.

Definition 4 (Entropy driven computation)**.** *$\mathcal{M}$ is the set of all macro-states of S and $M_t \in \mathcal{M}$ is the macrostate at time t and S is initialised in the macro-state $M_0 \in \mathcal{M}$ at time $t = 0$. After a time period T the system is in the macro-state $M_T \in \mathcal{M}$. An* entropy driven computation (EDC) *is the transition from an initial macro-state $M_0 \to M_\infty$. We say that S has computed the state M_∞ from the input M_0.*

Once an EDC computation has been performed, the results need to be read and recorded. This can be done by measuring the macrostate of S.

Definition 5 (Measurement)**.** *Let S and $\bar{S}'$ be two separate systems in macrostates M_∞ and $\bar{M}_T$ respectively. When brought into contact, they form the joint system $S\bar{S}$ which is in the new state $M_0^{S\bar{S}_T}$. A measurement consists of two parts. Firstly, a measurement step $M_T^{S\bar{S}} \mapsto M_{T+\Delta T}^{S\bar{S}}$. Secondly, a separation step of the systems, resulting in the recovery of the original system S in state M_∞ and the state $\bar{S}$ in state $\bar{M}'_{T+\Delta T}$.*

We require here that the system S is returned to its original state, while system $\bar{S}$ should record the outcome of the measurement. Therefore, $\bar{M}$ need not be the same as $\bar{M}'$.

Note that we implicitly defined computation as a system-level phenomenon. An individual reaction cannot be interpreted as a computation. Instead, information processing is an aggregate phenomenon that depends on the statistics of the behaviour of the system as a whole. This systems view of computation does not preclude a minimal chemical computer consisting of 2 molecules that can engage in a single reversible reaction.

2.3 Cost of the EDC proper

2.3.1 The model

We assume that the stochastic system S is defined by the master equation, which is a differential equation for the probability to observe a particular combination of values $\mathbf{n}$ assigned to observable $o_i, o_2, \ldots, o_N$ at time t [26].

$$\dot{P}(\mathbf{n}, t) = \sum_i (\omega_i(\mathbf{n} - \sigma_i)P(\mathbf{n} - \sigma_i, t) - \omega_i()P(\mathbf{n}, t))$$

where $\omega_i(\mathbf{n}) := k_i h_i(\mathbf{n})$ is the total rate of reaction i and h is the multiplier indicating how many combinations of molecules can realise the reaction, $\mathbf{n}$ and σ_i are the particle vector and the stoichiometric vector of reaction i. Given such a stochastic model, we can scale the total number of particles in the initial conditions by a factor c and the reaction rate constants of bimolecular reactions by $1/c$. This amounts to scaling the volume of the system while keeping the concentration fixed. In this way we construct an equivalence class $\mathcal{S}$ of systems S. In general, the members in this class will behave differently. Most importantly, they show different amounts of fluctuations and approach equilibrium at different speeds. For large (but finite) volumes, however, the behaviour of the master equation is increasingly well approximated by the first order linear noise approximation [25] whereby the mean behaviour of the system is invariant to scaling. Scaling the system size in the regime of the linear noise approximation only affects the noise around the mean behaviour, which scales with $V^{-1/2}$. It does not affect the speed with which the systems approach equilibrium. In the limiting case of an infinite volume the noise goes to zero and the system is described by a differential equation whose trajectory corresponds (for monostable systems) to the mean of the linear noise approximation. In this contribution we will only consider stochastic systems.

2.3.2 Cost

An EDC starts with the stochastic system S initialised in some macrostate M_0. The system then relaxes into an equilibrium state M_∞. We will identify the cost of the computation with the amount of entropy generated during the relaxation. There are two components to the entropy. Firstly, the systems entropy or heat dissipated to the environment which is determined by the ratio of the forward rate and the backward rate [27].

$$\dot{E} = k_B \ln \frac{\omega_-}{\omega_+}$$

Using this definition the heat dissipated to the environment is positive, whereas heat extracted from the reservoir is negative.

The total entropy production also contains a contribution intrinsic to the system. The entropy of a microstate s_i can be written as $E_{\text{int}} = -k_B \ln p_i$, where p_i is the probability to observe the system S in microstate s_i. On average this gives the Shannon or intrinsic entropy $\langle E_{\text{int}} \rangle = -k_B \sum_i p_i \ln(p_i)$. The change of system entropy

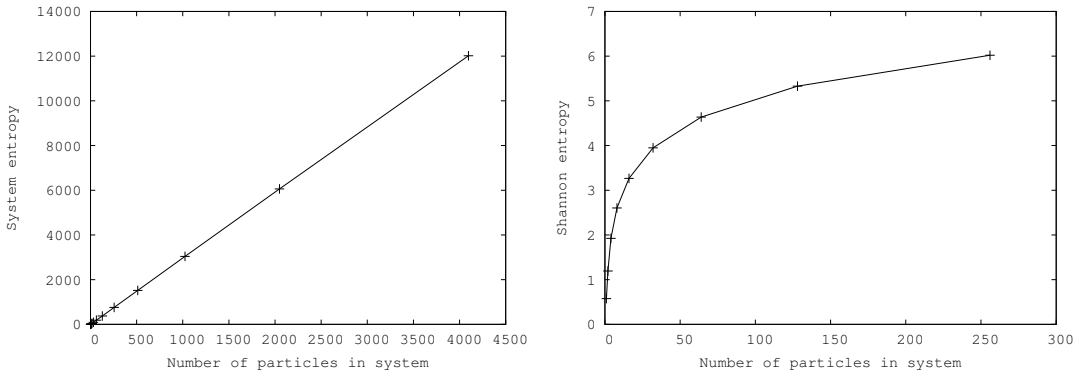

Fig. 1: Numerical calculations of the systems entropy (**left**) and the Shannon entropy (**right**) for the system $A + B \rightleftharpoons C + D \rightleftharpoons E$. The system size is the number of A and B at time $t = 0$. The concentration was kept constant as the particle size was increased. The entropy here was calculated assuming $k_B = 1$. The forward rate constant was set to 0.3 and the backwards rate constants from E and $C + D$ were 0.05 and 0.1 respectively. The entropy was calculated using Prism [28].

during the computation is then given by $\langle E_{\text{int}}(\infty)\rangle - \langle E_{\text{int}}(0)\rangle$. In the case of the chemical systems we are interested here, the microstates are given by a vector that allocates particle numbers to each chemical species. Evaluating the system entropy exactly can be problematic, because it is not normally possible to solve the master equation even in steady state.

Figure 1 shows numerical calculations for the system entropy dissipated during the approach to equilibrium. For large systems one expects the system entropy (or heat dissipated) during the approach to equilibrium to scale linearly with the system size. This is because the system entropy only depends on the distribution in equilibrium. For large systems the distribution will narrowly peak around the mean behaviour concentrating the contributions to the entropy to a few configurations only. The size of the configurations scales linearly with the system size. Hence, the entropy will depend linearly on system size. For small systems this is not true and there can be a non-linear relationship between the system size and heat dissipation. Note that the system entropy can be both negative and positive, depending on the initial conditions of the system. See appendix section B for an explicit calculation of the heat dissipation for a minimal example.

The ability to determine the result of the EDC is limited by the noise of the system in equilibrium. In the linear noise regime the noise of the system scales like $V^{-\frac{1}{2}}$. In this regime, we find the speed with which equilibrium is approached to be independent of V. The cost of the computation scales linearly with the size of the system. Hence there is a trade-off between entropy production (and thus energy cost) and the accuracy of the EDC. Notably, there are no trade-offs involving the computation time, i.e. the characteristic time-scale with which the system approaches equilibrium.

2.4 Trade-offs arising from the measurement

The equilibration of the EDC is only one part of the computation. In order, for the outcome of the computation to be known a measurement must be performed on the macrostate M_∞. The result also needs to be recorded. This comes at a cost. We assume that the measurement is done by the measurement device $\bar{S}$, which is itself an EDC. Initially, we will assume that $\bar{S}$ performs binary measurements on observables of S, i.e. it determines whether a particular species L of S is above or below a threshold abundance. Extensions to the continuous case are straightforward and will be discussed below.

Measurement is only possible if S and $\bar{S}$ are temporarily brought into contact. We will assume that the contact remains limited to specific interfaces. In particular, we assume that S and $\bar{S}$ do not mix, as it would be difficult to separate the systems after the measurement. Instead, we consider here a protocol whereby S and $\bar{S}$ remain separated by a wall. A single "trans-membrane" receptor acts as a sensor for the device $\bar{S}$. The external part of the sensor contains a number of binding sites for molecules of type L which are assumed to be part of S. The inside of the receptor can interact with $\bar{S}$, as will be described in the next section. The receptor itself can be in either of two states, namely "activated" or "deactivated." Binding of substrate at the S-facing side will result in changes to the state at the inside of $\bar{S}$. We will assume a Monod-Wyman-Changeux (MWC) [29] receptor (but other models are possible too). When its external sites are bound then the receptor is heavily biased towards the "active" state. Otherwise it is heavily biased towards being deactivated. When there are several binding sites, then the receptor can support ultra-sensitivity. This means that for low abundances of L the binding sites are almost never bound; above a threshold abundance the sites are almost always bound. In this way, MWC receptors can be used for binary measurements of external concentrations.

Upon contact, the system and the measurement device form a new system $S\bar{S}$. The joint system will therefore evolve from an initial macro-state $M_t^{S\bar{S}}$ at time t of the measurement, to a macrostate $M_{t+T_{\text{meas}}}^{S\bar{S}}$, where T_{meas} is the time required for the measurement. Upon separation of S and $\bar{S}$ the new macrostate of the measurement device is a (transient) record of the state of S. Once the measurement is completed it is necessary to restore S as an independent system, i.e. to separate S and $\bar{S}$.

Conceptually, there are thus two aspects to the measurement process that are relevant for the question we consider here. Firstly, the process of measurement itself and secondly, the separation of the measurement device from the system S. We find that the binary measurement itself, while it cannot be done for free, does not imply any trade-offs. In contrast, the separation step leads to a trade-off between energy, cost and time.

2.4.1 The measurement proper

In order to perform a measurement we use a measurement device $\bar{S}$ that is bistable. This means that for a suitably long integration time Δt there are two macrostates M^m and M^g characterised by a high amount of m or g respectively. Bistability can occur in equilibrium chemical systems. Stochastic bistable systems will switch spontaneously between the two macrostates. The expected time between switching events will depend on the system size and the "well depth." The latter is essentially determined by the kinetic parameters of the system and indicates the difficult of escaping a meta-stable state. Larger systems are more stable but even relatively small systems can be sufficient to virtually guarantee stability over any practically relevant time-scale.

Switching between the macrostates can occur spontaneously in equilibrium. In order to switch the system at a *particular time*, it is necessary to introduce a net probability flux to the desired state of the system. This implies breaking detailed balance and hence dissipates heat. A controlled switch of the bistable system can therefore only be achieved at a certain expense of work. Formally, it would be sufficient to connect a source of free energy (a "chemical battery") to $\bar{S}$ during the switch only and disconnect it afterwards. In practice, due to the nature of the chemical systems it is difficult to remove such a battery without extra work input. It is therefore much better to operate the device permanently from a battery away from equilibrium.

2.4.2 A possible design of a measurement device

Here we consider a concrete measurement device $\bar{S}$. We chose a design based on bi-stability. C briefly describes a different measurement device that can record continuous signals. The bistable system $\bar{S}$ consists of unstable molecules c that are converted into d, catalysed by g. We assume that d themselves are reversibly converted into g. Altogether, g is thus autocatalytic. To achieve bistability, we add symmetrically a second autocatalytic circuit consisting of e which is produced from c catalysed by m. Molecules of e are also reversibly converted into m. The reactions can be summarised as follows:

$$\begin{aligned} c &\overset{g}{\rightleftharpoons} d \\ c &\overset{m}{\rightleftharpoons} e \\ d + d &\rightleftharpoons g \\ e + e &\rightleftharpoons m \end{aligned}$$

To create bi-stability, this system needs to be extended by an antagonistic interaction between g and m.

$$\begin{aligned} d + m &\rightleftharpoons m + c \\ e + g &\rightleftharpoons g + c \end{aligned}$$

Finally, we also assume that g and m interconvert reversibly, albeit at a very low rate. Forward and backwards rates are equal.

$$g \rightleftharpoons m$$

If initialised with a sufficient number of c and one g and m each, then the system will, after a transient period, evolve to a macrostate M_∞^g where $g > m$ or M_∞^m characterised by $m > g$. The probability of either of these states will be equal if the parameters are symmetric between the autocatalytic pathways of m and g.

2.4.3 Switching

In order to switch the state of $\bar{S}$, we couple the chemistry of the device to the internal part of the MWC receptor. Here we assume that the active receptor catalyses the interconversion of m and g greatly speeding up the inter-conversion of m and g.

$$\begin{aligned} r &\rightleftharpoons r^* \\ m &\overset{r^*}{\rightleftharpoons} g \end{aligned}$$

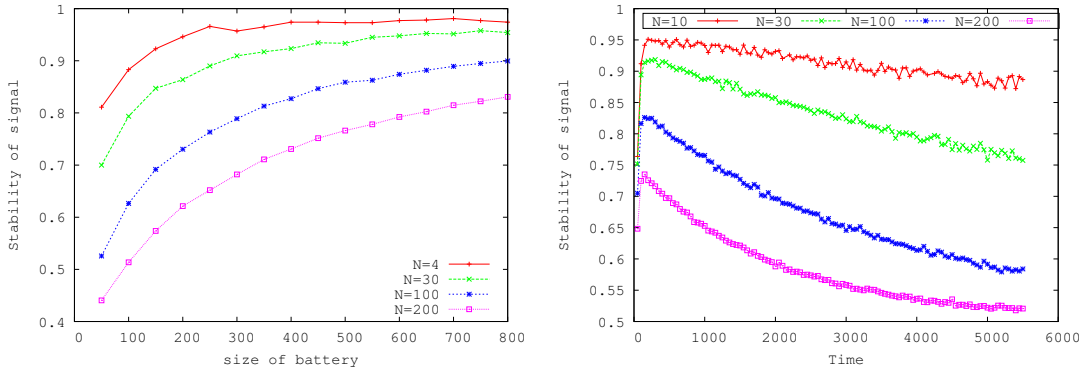

Fig. 2: **Left:** The stability of the memory defined as the ratio $m/(g+m)$ was taken after 600 time units of simulation. Each point is the average over 500 simulations. Simulations were started with g dominant. For large systems and small battery there may not be sufficient power to switch the system, resulting in a stability below 0.5. **Right:** The stability over time. The abundance of atp at the start of the simulation 300 and a high abundance of g. At small times the systems have not yet switched to m. Immediately following the switch the system reaches maximal stability, but then decays over time. Small systems perform better at a fixed cost. The spontaneous interconversion rate is set to 0.001.

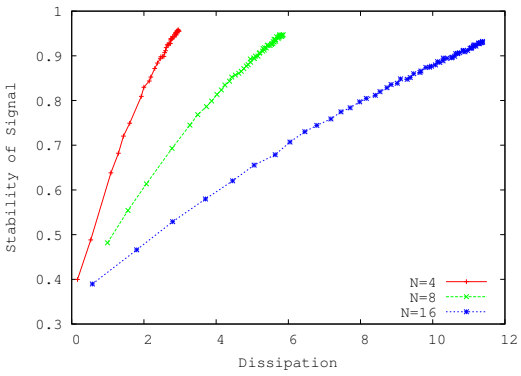

Fig. 3: The trade-off between dissipation and stability. For networks of different size the energy dissipation after 600 time units was recorded against the stability of the memory (defined as $m/(m+g)$.). Small systems are performing better than large systems.

In the absence of a chemical driving force, the active receptor accelerates the approach to $m = g$. Given the right range of parameters it will thus temporarily disrupt bistability and bring $\bar{S}$ into a mono-stable regime. Upon removal of the catalysing reaction, i.e. when the internal receptor reverts to the inactive state, $\bar{S}$ will relax quickly to M^g or M^m, both with equal probability. No directed switch is possible.

A directed switch is only possible if the catalytic reaction is driven preferentially into one direction. This can be done by coupling it to a chemical battery, i.e. a reaction that is not in equilibrium. A modified reaction scheme to achieve this could then be the following:

$$m + atp \overset{r^*}{\rightleftharpoons} g + adp$$

If there is a large excess of *atp* over *adp* then this would result in a net drive towards g whenever the receptor is activated. Assuming the excess of g is sufficient, the activation of the receptor leads to a switch of the macrostate to M^g.

2.4.4 Measurement and recording

The measurement device $\bar{S}$ needs to have two receptors. The first one is as described above and is used to determine the abundance of molecules of type L of the system S. A second receptor measures the abundance of some molecular species I, which may not be part of S. This receptor is used to reset the measurement device before usage and works analogously to the receptor for L only that it switches $\bar{S}$ to the state $\bar{M}^g$ by measuring the abundance of I in some reference system that was prepared to be high in I.

A measurement procedure is as follows: (*i*) Reset $\bar{S}$. (*ii*) Initiate the measurement proper on S by bringing S and $\bar{S}$ into contact mediated by the MWC receptor. (*iii*) Wait for a fixed amount of time $T_{\text{meas.}}$ (*iv*) Separate S and $\bar{S}$. We stipulate that after separation the system S is restored to a state that is statistically equivalent to the state before the measurement. If S contains the ligand L then S will be in macrostate M^m with a probability that depends on the abundance of L in S, the kinetic binding constants and the number of receptor binding sites.

In principle it is possible to remove the battery from the bistable system after the switch has been done. In practice this will be difficult because it entails extracting the *atp* and *adp* molecules from $\bar{S}$ and requires chemical work. Therefore, it is better not to remove the battery. The memory remains functional for as long as the battery is sufficiently charged, i.e. there is an excess of *atp* over *adp*, but once the battery has run out $\bar{S}$ will no longer be bistable (fig. 2). The chemical battery is discharged by two processes. Firstly, during switching the receptor is active with high probability and the battery will drive the conversion from g to m while using *atp*. When the external binding sites of the MWC receptor are occupied the receptor will be active most of the time, resulting in a high rate of discharge. Secondly, there is spontaneous activation of the receptor even if no ligands are binding to the outside receptor. The rate of spontaneous activation determines the base-rate of *atp* usage/battery discharge.

The details of the rate of *atp* usage depend on the kinetic parameter settings. There are no trade-offs between energy, system size and accuracy. In the particular model we present here, small systems perform better than those with a large number of molecules, while also dissipating substantially less energy (see fig. 2). Similarly, there are no trade-offs between the speed of the measurement and its accuracy. The speed is related to the scale of the reaction rates inside of $\bar{S}$ which does not affect the heat dissipation of the system.

There is, however, a minimal amount of work required to switch the device. The recording of information of S by $\bar{S}$ entails a correlation between the system and the measurement device. This correlation increases the free energy and could be exploited to extract usable work from the system [30]. To remain consistent with the second law, the measurement process itself must come at an energy expense [31, 32]. Furthermore, according to Landauer's limit, the recording is a dissipative step which requires a minimum amount of work.

Finally, the accuracy of the binary measurement corresponds to the ability of $\bar{S}$ to indicate whether or not the concentration of the system is below or above a threshold. For abundances of L that are far from the threshold this can be done accurately by increasing the number of receptor binding sites. The dynamics of the MWC can then be ultra-sensitive. If the concentration of L is close to the threshold, then measurements will necessarily remain inaccurate, although a higher number of receptors will again increase accuracy. The receptor number does not influence cost of switching, but it does increase the cost and time of separating S and $\bar{S}$, as discussed in the next section.

2.5 The separation step

Once the measurement of the system has been completed the coupled system $S\bar{S}$ needs to be separated, while restoring S. We use here a model inspired by the Szilard engine: We stipulate that contact between S and $\bar{S}$ is mediated by a straight wall. The system S is fixed in space, but $\bar{S}$ can be moved to the right, which opens up a volume V_0 between the two systems (see fig. 4). We also assume a number of removable walls that can be inserted at arbitrary (but fixed) points to further sub-partition V_0. We now describe two different protocols to restore S.

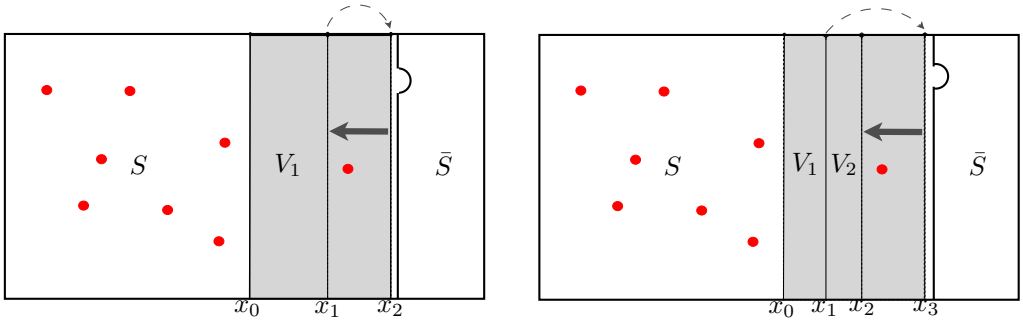

Fig. 4: A schematic representation of the resetting procedure. The volume V_0 is shaded in grey. **Left:** The simplified scheme of protocol 1. **Right:** The scheme with three walls of protocol 2.

2.5.1 Protocol 1

The first protocol is as follows: Initially the system and the measurement device are in contact forming $S\bar{S}$. The ligand may or may not be bound to the receptor at this stage. (i) Initiate the separation process by inserting a wall (wall 1) at point x_0 to separate S from the wall of $\bar{S}$. The receptor of $\bar{S}$ will be to the right of wall 1. Ligands may still be bound to the receptor. Now move system $\bar{S}$ to the right, thus creating a volume V_0. At this stage wall 1 is located at x_0 and separates V_0 from S. (ii) Wait for T units of time to allow the ligand to unbind from the receptor. Then insert wall 2 close to the receptor at point x_2 immediately to the left of the membrane of $\bar{S}$ and the receptor. This traps the ligand (should it indeed have unbound) in V_0. (iii) Move wall 2 to the left up to the point x_1. (iv) Remove wall 1, which was originally inserted at x_0 and move wall 2 to position x_0. This step restores the ligand to S. This cycle should be repeated for as long as necessary to achieve the restoration of S.

Step (i) does not require any work. Step (ii) does not require work either, but entails a waiting time T, so comes at an execution time cost. During step (iii) the work is zero if the ligand is still bound after the waiting time T. Positive work is only required if the ligand has unbound from the receptor. A lower bound for this work is given by the quasi-static compression of the one particle gas.

$$W_{\min} = \begin{cases} \beta^{-1} \int_{V_0}^{V_1} \frac{1}{V'} dV' & \text{with probability } p* \\ 0 & \text{with probability } (1 - p*) \end{cases}$$

Here $p^* = 1 - \frac{k_b}{k_u V_0 + k_b}$ is the equilibrium probability that a single ligand has unbound from the receptor and the ligand binding/unbinding rate constants are given by k_b/V_0 and k_u respectively. We set the constant $\beta = 1$ to simplify notation.

Finally, during step (iv) the minimal amount of work required is:

$$W_2 = \int_{V_S + V_1}^{V_S} \frac{N}{V'} dV'$$

Here V_S is the volume of system S. This latter contribution to the work scales linearly with the number of particles N contained in the system S. It has to be applied at each iteration of the separation procedure.

2.5.2 Protocol 2

This fixed cost at each iteration can be avoided by allowing for three walls. Wall 1 is as before, inserted at x_0. It separates S from $\bar{S}$. The measurement device is them moved to the right again opening up the volume V_0 as in protocol 1. Wall 2 is then inserted at some distance to the right of wall 1 at position x_1 so as to form the volume V_1 between it and wall 1. The separation procedure is as follows. (i) After a waiting T time units insert wall 3 at x_3, which is immediately to the left of the membrane of $\bar{S}$. Wall 3 separates the receptor (and possibly ligands bound to it) from the volume V_0. Move wall 3 to the point x_2. Wall 3 now forms a volume V_2 between itself and wall 2 and a volume $V_0 - V_1 - V_2$ between itself and the membrane of $\bar{S}$. (ii) Remove wall 2; this opens the volume $V_1 + V_2$ between S and wall 3. (iv) Move wall 3 from position x_2 to position x_1. At this point the cycle is started again at (i) by inserting wall 2 at x_3.

This scheme is more complicated than the first scheme, but allows work to be independent of the size of S. Step (i) only requires work if there is a particle trapped between the walls, but it always requires a waiting time of T. The second step only requires work if there are particles trapped. Then the quasi-static work during step (i) and (ii) respectively are given by:

$$W_1 = \int_{V_0 - V_1}^{V_2} \frac{1}{V'} dV' \qquad W_2 = \int_{V_1 + V_2}^{V_1} \frac{1}{V'} dV'$$

The total yields:

$$W_1 + W_2 = \ln\left(\frac{V_1 V_2}{(V_0 - V_1)(V_1 + V_2)}\right)$$

Note that this work is required at most once during the resetting procedure.

This above resetting cycle needs to be repeated a number of times to ensure that the ligand has been captured and S can be restored with a given (high) probability. The restoration itself is necessary only once and is done by merging the molecules trapped in V_1 with S. This is done by removing the first wall and expanding the volume of S by V_1. Subsequently, S is then reduced again to its original volume V_S by pushing the remaining wall from x_1 to x_0. This requires the following work:

$$W_3 = \int_{V_S+V_1}^{V_S} \frac{1}{V'} dV' = (N+1)\ln\left(\frac{V_S}{V_S + V_1}\right)$$

This overhead work can be further minimised by splitting the final restoration of S into a two-step process. In the first step, reduce the volume holding the single particle from V_1 to some volume V_c. In the second step remove the separating wall to expand S to $V_S + V_c$. As a third step restore S to its original volume V_S. The total work required is given by

$$W_3' = \left[(N+1)\ln\left(\frac{V_S}{V_S + V_c}\right) + \ln\left(\frac{V_c}{V_1}\right)\right]$$

This work is optimised for $V_c = V_S/N$. Substituting this optimal choice of V_c into the work formula yields

$$\begin{aligned} W_3' &= -N\ln(N+1) + N\ln(N) - \ln(N+1) + \ln(V_1) - \ln(N) \\ &\approx -\ln(N+1) + \ln\left(\frac{V_S}{V_1}\right) - 1 \\ &\approx \ln\left(\frac{V}{V_S + N\kappa}\right) - 1 \end{aligned}$$

In the last line we substituted $V_1 = (V_S/N) + \kappa$, where the $\kappa > 0$ is a constant. We also assumed $N \gg 1$ by setting $\ln(N+1) \approx \ln(N)$ and $N\ln(N/(N+1)) = -1$. The combined work necessary to separate the systems is then given by

$$\begin{aligned} W_{\text{tot}} &= \left(W_3' + (W_1 + W_2)_{|_{V_1 = \frac{V_S}{N} + \kappa}}\right) \\ &= \ln\left(\frac{V_S}{\kappa N + V_S}\right) + \ln\left(\frac{V_2\left(\frac{V_S}{N} + \kappa\right)}{\left(V_0 - \frac{V_S}{N} - \kappa\right)\left(\frac{V_S}{N} + \kappa + V_2\right)}\right) - 1 \end{aligned}$$

This has an extremum at $\kappa = (V_0\,N - V_2\,N - 2\,V_S)/2N$ corresponding to the highest work expended. To the left of this extremum, for $\kappa \to 0$ the contribution from $W_1 + W_2$ becomes dominant, whereas $W_3' = -1$ in this limit. Similarly, on the right side, $\kappa \to V_0 - V_2$, the contribution from W_3' becomes important. To reduce the total work required, κ should be chosen near one of these two points. The scaling behaviour of this expression can be made more apparent by substituting the ratio of the number of particles in S and its volume V_S by the concentration x.

$$W_{\text{tot}} = \ln\left(\frac{x^{-1}}{\kappa + x^{-1}}\right) + \ln\left(\frac{V_2\left(x^{-1} + \kappa\right)}{\left(V_0 - x^{-1} - \kappa\right)\left(x^{-1} + \kappa + V_2\right)}\right) - 1$$

This shows that for high concentrations $x \gg 1$ the work scales as the logarithm of the concentration. More importantly though, within any particular equivalence class of systems $\mathcal{S}$ where the concentration does not change, the separation work is independent of the system size.

2.5.3 Time required for protocol 2

The work estimates calculated above are valid for quasi-static processes. Finite time processes will always require more work [33]. This is a source of a speed-time trade-off. Another contribution to the separation time comes from waiting time for the ligand-receptor bond to equilibrate during step (i) in protocol 2. The time scale for equilibration is $T \sim k_u + k_b/V_0$. Recalling that the equilibrium probability to find the particle unbound is $p_u = V_0/(V_0 + \eta)$ where $\eta = k_b/k_u$ we obtain the number of iterations of the resetting procedure necessary to ensure that the average number of bound ligands bound ϵ is:

$$\text{min number of steps to succeed with confidence } \epsilon = \left\lceil \frac{\ln\left(\frac{\epsilon}{l}\right)}{\ln(\eta) + \ln V_0 + \eta} \right\rceil,$$

where $\lceil x \rceil$ denotes the smallest integer $> x$ and l is the number of ligand binding sites at the external receptor. Multiplying this by the time scale T gives a time-scale indicator for the restoration of S (see fig. 5). A trade-off between the measurement time and the work required arises here, because both quantities depend on the volume V_0. A larger volume makes it more likely for the unbinding receptor to be captured, but equally increases the amount of work that needs to be done (see fig. 5).

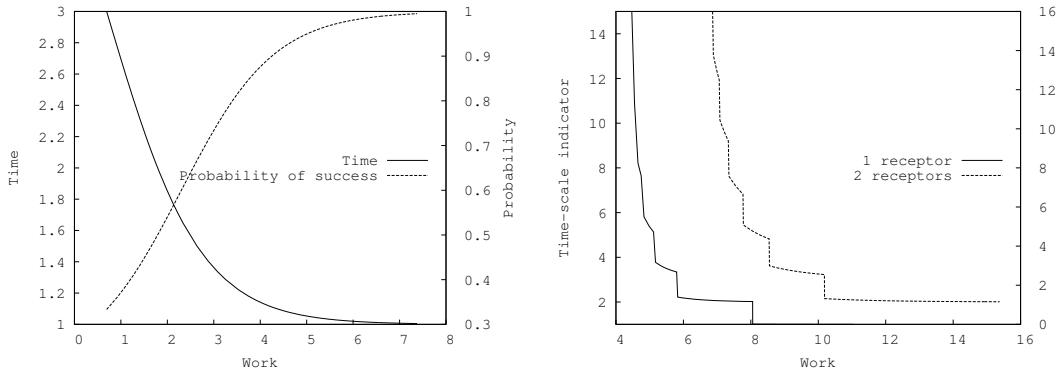

Fig. 5: **Left:** Trade-off between the time and the work spent on the restoration. Here we chose $k = k_u = 1$ and $V_1 = V_2 = 0.25$. The reaction volume V_0 was varied from 0.5 to 200. **Right:** Trade-off between the time spent on computation and the work, for 1 and 2 receptors. Here we chose $\epsilon = 0.01$, $k/k_u = 1, N = 200, V_S = 10$ and $V_2 = 1$.

3 Discussion

The conclusion following the classical work by Bennett is that there is no lower cost limit to computation, and that costs due to Landauer's limit can always be avoided by constructing logically reversible computers (see also [22]). Bennett also hints at performance trade-offs as the energy dissipation of computers is decreased, but his examples, instructive as they may be, remain difficult to interpret in a wider context of chemical computers. Here, the model of EDCs becomes useful, in that it makes it very clear what the origins of performance trade-offs in chemical computers are. The model, albeit abstract, can also be related to non-equilibrium steady state cases, as they are common in biology. An easy conclusion to draw from EDCs is that the computation time is not subject to trade-offs in EDCs. The only way to vary system parameters in EDCs is to scale the volume. As is well known, a larger system entails a higher entropy production, but also reduces the fluctuations in the system. In the context of EDCs this can be interpreted as a trade-off between accuracy and cost of the computation, but there is no trade-off involving time.

This seems to contradict many of the previous contributions that found time-accuracy trade-offs in various biological computers. This is resolved once one acknowledges the significance of the measurement step. The state of the EDC cannot be known unless it is probed. The act of measurement necessarily entails cost. There are two parts to it. (*i*) The cost of recording the result of the computation and (*ii*) the separation cost. In the literature the observation cost is often ignored because it does not seem to form part of the computation *per se*, and is perhaps also thought to be negligible in comparison to the cost of computation.

In principle, once could argue, the observation costs are avoidable. Computer scientists would never worry about how one can know the result of a particular algorithm, because the method of output is irrelevant to the algorithmic complexity. Then again, theoretical computer science is not normally concerned with the thermodynamics of computation. In chemical computers, however, the situation is different. Thermodynamics is important and observation costs may be significant. Unless we accept not to know the outcome of the computation, or more generally, that the computation is not allowed to impact the world around it at all, the costs of observation become due.

The restoration step introduces a four-way trade-off. If the probability for a successful restoration and the required measurement accuracy are fixed then there is a trade-off between time and work (see fig. 5). In order to have more accurate binary measurements, the receptor number needs to be increased. This leads to a deteriorating trade-off, because more restoration cycles are required in order to restore the system. Equally, in order to have a higher confidence that the restoration is achieved successfully, i.e. a lower parameter ϵ, more cycles must be completed and hence the restoration procedure takes longer.

Conveniently, the restoration step yields the trade-off involving time that one often observes. In reality though, restoration is likely of subordinate importance. Restoration of the system is not strictly necessary. For large systems S the impact of the measurement may be minimal and restoration can simply be ignored without introducing significant errors in the computation. More importantly, biological systems operate at a non-equilibrium steady state. Computations in cells therefore do not proceed in distinct cycles, but measurement devices are often tethered permanently to the system to be measured. The decoupling then happens via stochastic receptor loss. The trade-offs involving time, that are frequently observed in biological systems, are therefore unlikely to be due to the separation step.

The recording step of measurement is the more interesting origin of trade-offs. As shown above, recording of a binary measurement comes at a fixed cost bounded from below by Landauer's limit and does not lead to trade-offs at least not for binary measurements. Still, trade-off arise when a higher accuracy is desired. Assume that $d_{\max}$ is the maximal abundance of L in S. Given a set of perfect binary measurement devices $\bar{S}_i$ one could

use half-interval search in order to determine the abundance of the particles with an accuracy of $d_{\max}/N$ by using $\lceil \log(N) \rceil$ different binary measurements. Each binary measurement comes at a fixed cost bounded from below by the Landauer limit and takes a finite amount of time. This entails a trade-off between the measurement cost and time against the accuracy. A second origin of trade-offs are stochastic fluctuations in the system S, which can be dominant for small volumes. In this case even binary measurements need to be repeated several times in order to be able to estimate the probability that the system is below or above a certain threshold concentration (see for example [18]). This leaves the user who wishes to increase the accuracy of the computation with the following choice: Either she scales up the EDC, and thus increases accuracy and cost, or, alternatively, she keeps the EDC constant but increases the number of measurements, which increases the time and cost required. Many time trade-offs that appear in biological cells are precisely due to measurement [18].

Recording of external information is an identifiable source of cost in biological sensing systems, even at steady state. For example, Govern *et al.* [13] show that push-pull phosphorylation mechanisms are used to record sensor data and show that this essential so that the cell is able to process, i.e. average, the sensory data downstream. Here, the recording step is the origin of the biological cost and it cannot be ignored in the metabolic accounting. It is instructive to compare this to the de Palo - Endres system [34], a (hypothetical) change detector of external molecular concentrations. The authors show that their sensor operates at no internal cost to the cell, but feeds entirely from the free energy changes in the environment. Crucially, their change detector does not record information and consequently cannot average out noise from the external system. Indeed, the hypothetical cell that they propose is not a measurement device either, but rather part of an unrecorded EDC.

A final remark concerns the relationship between EDCs and non-equilibrium steady state chemical computers. While there are a few EDC schemes that have attracted the attention of the research community, most notably DNA computers [1], biological computers typically operate far from equilibrium near a steady state. The question is whether non-equilibrium steady state information processing is fundamentally different from EDCs. We believe that there is no fundamental difference. At the same time, biological systems could not work as EDCs because the individual operation steps of the EDC require an external handler or agent to connect and disconnect devices. In contrast, non-equilibrium steady state systems do not operate in separate cycles and are permanently connected to an external power source. They can compute changes to their their own configurations, including the preparation of novel inputs, without the need for a handler. An equilibrium state would mean death for the biological cell. Yet, this is mostly an implementation detail, but not of fundamental scientific significance for the properties of chemical computers. We therefore conjecture that every steady state computer can be implemented as a series of EDCs. Future research needs to clarify this.

References

[1] Seelig, G., Soloveichik, D., Zhang, D. Y. & Winfree, E., 2006 Enzyme-free nucleic acid logic circuits. *Science* **314**, 1585–1588. (doi:10.1126/science.1132493).

[2] Lakin, M. & Phillips, A., 2011 Modelling, simulating and verifying turing-powerful strand displacement systems. In *Proceedings of the 17th International Conference on DNA Computing and Molecular Programming*, DNA'11, pp. 130–144. Berlin, Heidelberg: Springer-Verlag. ISBN 978-3-642-23637-2.

[3] Tsuda S., A. S., Zauner K.P., 2009 *The Phi-Bot: A Robot Controlled by a Slime Mould*, chapter 10, pp. 213–232. Springer-Verlag London.

[4] Friedland, A., Lu, T., Wang, X., Shi, D., Church, G. & Collins, J., 2009 Synthetic gene networks that count. *Science* **324**, 1199–1202. (doi:10.1126/science.1172005).

[5] Sole, R. & Macia, J., 2013 Expanding the landscape of biological computation with synthetic multicellular consortia. *Natural Computing* pp. 1–13. ISSN 1567-7818. (doi:10.1007/s11047-013-9380-y).

[6] Silva-Rocha, R. & de Lorenzo, V., 2011 Implementing an or-not (orn) logic gate with components of the sos regulatory network of escherichia coli. *Molecular Biosystems* **7**, 2389–2396. (doi:10.1039/c1mb05094j).

[7] Walker, S., Kim, H. & Davies, P., 2016 The informational architecture of the cell. *Phil. Trans. R. Soc. A* **374**, 20150057. (doi:10.1098/rsta.2015.0057).

[8] Davies, P. & Walker, S., 2016 The hidden simplicity of biology. *Reports on Progress in Physics* **79**, 102601. (doi:10.1088/0034-4885/79/10/102601).

[9] Ninio, J., 1975 Kinetic amplification of enzyme discrimination. *Biochimie* **57**, 587–595.

[10] Fluitt, A., Pienaar, E. & Viljoen, H., 2007 Ribosome kinetics and aa-trna competition determine rate and fidelity of peptide synthesis. *Computional Biology and Chemistry* **31**, 335–346. (doi:10.1016/j.compbiolchem.2007.07.003).

[11] Alon, U., Surette, M., Barkai, N. & Leibler, S., 1999 Robustness in Bacterial Chemotaxis. *Nature* **397**, 168–171.

[12] Gregor, T., Tank, D., Wieschaus, E. & Bialek, W., 2007 Probing the limits to positional information. *Cell* **130**, 153–164. (doi:10.1016/j.cell.2007.05.025).

[13] Govern, C. & ten Wolde, P., 2014 Optimal resource allocation in cellular sensing systems. *Proceedings of the National Academy of Science USA* **111**, 17486–17491. (doi:10.1073/pnas.1411524111).

[14] Berg, H. C. & Purcell, E. M., 1977 Physics of chemoreception. *Biophys J* **20**, 193–219. (doi:10.1016/S0006-3495(77)85544-6).

[15] Mehta, P. & Schwab, D., 2012 Energetic costs of cellular computation. *Proceedings of the National Academy of Science USA* **109**, 17978–17982. (doi:10.1073/pnas.1207814109).

[16] Chu, D. & Barnes, D., 2016 The lag-phase during diauxic growth is a trade-off between fast adaptation and high growth rate. *Scientific Reports* **6**, 25191. (doi:10.1038/srep25191).

[17] Chu, D., 2015 In silico evolution of diauxic growth. *BMC Evolutionary Biology* **15**, 211. (doi:10.1186/s12862-015-0492-0).

[18] Zabet, N. & Chu, D., 2010 Computational limits to binary genes. *Journal of the Royal Society Interface* **7**, 945–954. (doi:10.1098/rsif.2009.0474).

[19] Chu, D., Zabet, N. & Hone, A., 2011 Optimal parameter settings for information processing in gene regulatory networks. *BioSystems* **104**, 99–108. (doi:10.1016/j.biosystems.2011.01.006).

[20] Landauer, R., Bennett, C., Laing, R. & Zurek, W. *Maxwells Demon, Information Erasure, and Computing*, pp. 187–288. Princeton University Press.

[21] Bennett, C., 1982 The thermodynamics of computation. A review. *International Journal of Theoretical Physics* **21**, 905–940. ISSN 0020-7748. (doi:10.1007/bf02084158).

[22] Strasberg, P., Cerrillo, J., Schaller, G. & Brandes, T., 2015 Thermodynamics of stochastic turing machines. *Physical Review E* **92**. (doi:10.1103/physreve.92.042104).

[23] Qian, H. & Beard, D., 2005 Thermodynamics of stoichiometric biochemical networks in living systems far from equilibrium. *Biophysical Chemistry* **114**, 213220. ISSN 0301-4622. (doi:10.1016/j.bpc.2004.12.001).

[24] Beard, D., Liang, S. & Qian, H., 2002 Energy balance for analysis of complex metabolic networks. *Biophys J* **83**, 79–86. (doi:10.1016/S0006-3495(02)75150-3).

[25] van Kampen, N., 2007 *Stochastic Processes in Physics and Chemistry*. Amsterdam: Elsevier. Third edition.

[26] Gillespie, D., 1992 A rigorous derivation of the chemical master equation. *Physica A: Statistical Mechanics and its Applications* **188**, 404–425. (doi:10.1016/0378-4371(92)90283-v).

[27] Seifert, U., 2005 Entropy Production along a Stochastic Trajectory and an Integral Fluctuation Theorem. *Physical Review Letters* **95**, 040602. (doi:10.1103/PhysRevLett.95.040602).

[28] Kwiatkowska, M., Norman, G. & Parker, D., 2001 PRISM: Probabilistic symbolic model checker. In *Proc. Tools Session of Aachen 2001 International Multiconference on Measurement, Modelling and Evaluation of Computer-Communication Systems* (ed. P. Kemper), pp. 7–12. Available as Technical Report 760/2001, University of Dortmund.

[29] Marzen, S., Garcia, H. & Phillips, R., 2013 Statistical mechanics of monod–wyman–changeux (MWC) models. *Journal of Molecular Biology* **425**, 1433–1460. (doi:10.1016/j.jmb.2013.03.013).

[30] Parrondo, J., Horowitz, J. & Sagawa, T., 2015 Thermodynamics of information. *Nature Physics* **11**, 131–139. (doi:10.1038/nphys3230).

[31] Granger, L. & Kantz, H., 2011 Thermodynamic cost of measurements. *Phyical Review E* **84**, 061110. (doi:10.1103/PhysRevE.84.061110).

[32] Kawai, R., Parrondo, J. M. R. & Van den Broeck, C., 2007 Dissipation: the phase-space perspective. *Physical Review Letters* **98**, 080602. (doi:10.1103/PhysRevLett.98.080602).

[33] Bena, I., den Broeck, C. V. & Kawai, R., 2005 Jarzynski equality for the jepsen gas. *Europhysics Letters (EPL)* **71**, 879–885. (doi:10.1209/epl/i2005-10177-0).

[34] De Palo, G. & Endres, R., 2013 Unraveling adaptation in eukaryotic pathways: Lessons from protocells. *PLoS computational biology* **9**, e1003300.

A Random walk

The simplest example of an EDC is a discrete random walk in 1D. The forward rate of the chain at site n is k_n^+; the backwards rate is k_n^-. We assume that there are altogether N sites. The long-term probability to be at a site n is then given by

$$P_n = \frac{\prod_{i=0}^{n-1} k_i^+ \prod_{j=n+1}^{N} k_j^-}{\sum_{l=0}^{N} \prod_{i=0}^{l-1} k_i^+ \prod_{j=l+1}^{N} k_j^-}$$

the simplest choice one can make here is to choose all forward rates to be the same and the backward rates to be equal as well. We can then set $k^- = \epsilon k^+$. In equilibrium it must be the case that $\epsilon = 1$. The expression for the long-term probability then reduces to a simple expression:

$$P_n = \frac{(k^+)^N \epsilon^{N-n}}{\sum_{l=0}^{N} (k^+)^N \epsilon^{N-l}} = \left(\sum_{l=0}^{N} \epsilon^{n-l} \right)^{-1} = \epsilon^{-n} \frac{\epsilon - 1}{\epsilon - \epsilon^{-N}}$$

For this system, the heat transferred to the bath per step to the right is given by $-\ln(\epsilon)$ and similarly for each step to the left the heat generated is given by $\ln(\epsilon)$. Hence, when the system is at site n the total amount of heat generated is then given by $-(n+1)\ln(\epsilon)$. The average heat dissipation then is given by:

$$\begin{aligned} Q_{\text{diss}} &= -\sum_{n=0}^{N} P_n (n+1) \ln(\epsilon) = -\sum_{n=0}^{N} \epsilon^{-n} \frac{\epsilon - 1}{\epsilon - \epsilon^{-N}} (n+1) \ln(\epsilon) \\ &\approx \frac{1 + N - (2+N)\epsilon}{(1-\epsilon)} \ln\left(\frac{1}{\epsilon}\right) \\ &\approx N \ln\left(\frac{1}{\epsilon}\right) \end{aligned}$$

Here, the second line is an approximation obtained from the closed form solution by ignoring higher order terms in ϵ. This approximation is only good for $\epsilon < 1$ when N is large.[1] The equilibrium system is characterised by $\epsilon = 1$. In this case no systems entropy is produced. For other choices of ϵ the entropy production is proportional to N.

B Example system $A + B \rightleftharpoons C + D$.

This is a minimal chemical system consisting of a single reversible, bimolecular reaction. If a is the number of molecules of type A, $p(a,t)$ the probability to observe a molecules of A at time t, and noting that the total number of particles is conserved; $N := a + c$ is half the number of particles in the system. We can write the master equation corresponding to this system.

$$\dot{p}(a,t) = k^-(a+1)^2 p(a+1,t) + k^+(N-a+1)^2 p(a-1,t) - \left(a^2 k^- + (N-a)^2 k^+\right) p(a,t)$$

We can solve this master equation at the equilibrium state by using the detailed balanced condition which gives for $p(a) := p(a,\infty)$:

$$p(a) = p(a-1) \underbrace{\frac{k^+}{k^-}}_{:=K} \frac{(N-a+1)^2}{(a)^2}$$

From this one obtains

$$p(1) = p(0) K \left(\frac{N}{1}\right)^2$$

and subsequently one can see easily see that

$$p(n) = p(0) K^n \left(\frac{N!}{N!(N-n)!}\right)^2$$

[1] The exact expression is

$$\frac{\left(\left((-N-2)\,\epsilon + N + 1\right)\epsilon^{-N} + \left(-\epsilon^{N+2} + (N+3)\,\epsilon - N - 1\right)\epsilon\right)\epsilon}{(-1+\epsilon)\left(\epsilon - \epsilon^{N+2}\right)\left(-\epsilon^{-N} + \epsilon\right)} \ln\left(\frac{1}{\epsilon}\right)$$

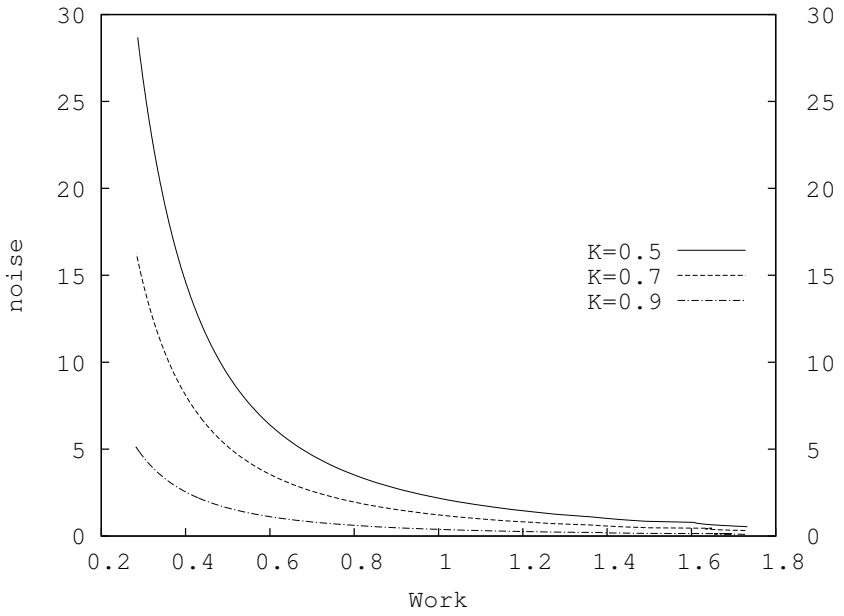

Fig. 6: Trade-off between the noise and the work required for the chemical computer $A + B \rightleftharpoons C + D$. The noise is given by the $\sigma^2/\langle n \rangle^2$.

Where $p(0)$ is obtained by the normalisation condition

$$p(0) = \left(1 + \sum_{n=1}^{N} K^n \left(\frac{N!}{n!(N-n)!}\right)^2\right)^{-1}$$

Assuming that we started in a state $a = 0$ then the system entropy at equilibrium is given by $E(a) = a \ln(K)$ and the average entropy at equilibrium is given by:

$$\begin{aligned}
\langle E \rangle &= \sum_{n=0}^{N} E(n)p(n) = \ln(K) \sum_{n=0}^{N} np(n) \\
&= \ln(K) \frac{\sum_{n=0}^{N} K^n n \left(\frac{N!}{n!(N-n)!}\right)^2}{1 + \sum_{i=1}^{C} K^i \left(\frac{N!}{i!(N-i)!}\right)^2} \\
&\approx \ln(K) \frac{\sum_{n=0}^{N} K^n n \left(\frac{N!}{n!(N-n)!}\right)^2}{\sum_{i=1}^{N} K^i \left(\frac{N!}{i!(N-i)!}\right)^2}
\end{aligned} \tag{1}$$

This formula cannot be evaluated any further. Some understanding can be gained by observing that this is up to a factor $\ln K$ the average number of particles contained in A and B. It is clear that for moderately large N the contribution will concentrate on a few terms only. For $K = 1$ the dominant contribution will come from $n \approx N/2$, corresponding to the expected result that half of molecules are A and B on average. For different values of K the peak shifts. This represents an average and therefore it is clear that a system twice as large has in steady state twice as many particles in A and B. Note that linearity is not true for small systems, where the sum is not dominated around a narrow range of n

Using the steady state probabilities it possible to evaluate the variance, which scales approximately linear in N, hence the variation of the particle number scales as $\sqrt{N}$ as assumed by the van Kampen linear noise approximation. Hence, there is a trade-off between the noise and the system size.

C Neural network

Artificial neural networks (ANN) are networks of typically binary automata— or "neurons — update their state according to inputs they receive from connected neurons. A common update rule is to take a weighted sum of

the input from connected neurons . The neuron will be set into state 1, if this sum is greater than a threshold θ. Otherwise it will be set to 0. Given the right network topology, neural networks are universal function approximators. Finding the weights that implement a particular function is generally non-trivial, but not the concern here. In this section we indicate how a cost estimate for neural network computation can be obtained by using EDC. Specifically, we will describe the update procedure for a single neuron that is connected to $N-1$ other neurons.

We model a neuron as a bi-stable chemical system (equivalent in nature to the measuring device discussed in the main text). Each neuron can be characterised by its macrostate $M^i \in 0, 1$, corresponding to m and g respectively being dominant. We also stipulate that each neuron has two interfaces (receptors), one of which is used to reset it.

Updating the state according to the rules of the neural network requires a thresholding operation that depends on the states of $N-1$ other systems. The MWC receptor can only threshold over concentration [29], but cannot take weighted sums over a number of samples. In order to translate the weights of the system into a chemical concentration, we introduce an intermediate measurement system Σ, which can add up the contributions from all the connected neurons. Note that Σ is fundamentally different in operation from the measurement device $\bar{S}$ from the main text. It contains two species A and B and a reversible reaction that interconverts them with backwards and forward rates k^+ and k^-. The precise values of the rate constants are unimportant, but $k^-, k^+ \ll 1$; this entails that the system relaxes to equilibrium very slowly. Ideally, the forward rate from A to B is higher than the backwards rate, but this is not essential. We assume that Σ starts in a standard state far from equilibrium, say with only molecules of type A. We further assume that at the outside Σ has a MWC type receptor with n binding sites; the inside of the receptor preferentially switches from the inactive state r to the active state r^* when all n binding sites are bound, otherwise it switches preferentially to the deactivated state. The active state catalyses the interconversion of A and B substantially speeding up the conversion. This can be summarised as follows:

$$r \rightleftharpoons r^*$$
$$a \overset{r^*}{\rightleftharpoons} b$$
$$a \rightleftharpoons b$$

We now describe the update procedure for neuron 1, assuming that it has weights $w_2, w_3, \ldots, w_N$ connecting it to neurons $2, \ldots N$ respectively. The external receptor of Σ binds to g and hence detects state 1 of a neuron.

1. Initialise Σ to be in a state $b = 0$, $a = N_\Sigma$
2. Initialise neuron 1 to be in macrostate 0, indicated by the predominant presence of m.
3. For all neurons $l = 2, \ldots, N$, let Σ measure the state of neuron l for a time period $T_0 + \eta \cdot w_l$, where η is a user-chosen proportionality constant and T_0 a fixed time period to allow the receptor to equilibrate. After the measurement is finished continue with neuron $l+1$.
4. Let neuron 1 measure the state of Σ.

Initialisation of Σ during step 1 comes at a fixed cost per update round. The cost depends on the size N_Σ of Σ.

During step 3 the device Σ acts like a stochastic memory device that remembers how many systems have been in state 1. The weighting is achieved in this model by allowing interactions with each system only for a time period that is proportional to the weight of the connection between the two neurons. This outcome of the measurement will be stochastic. The accuracy of the measurement depends on the choice of parameters. One source of error is the persistence of ligand binding to the receptor of Σ even after the separation from the neuron. A consequence of the finite unbinding time is that the catalytic action inside the cell continues even when the neuron is removed from Σ. This error can be reduced by decreasing the efficiency of the catalyst r^* while at the same time increasing the contact time between neurons and Σ. Thus the gain in accuracy would have to be paid for by an increased operation time. This leads to a trade-off between time and accuracy. A side effect would also be that the neurons themselves require larger batteries so as to be able to keep a memory of their state for longer. Another way to reduce the impact of the residual binding is to increase the unbinding rate from the receptor. This increases the cost of the computation because it requires a higher concentration of ligand (i.e. g molecules) in the neurons. Altogether, the error due to residual binding can be reduced by increasing the time of the measurement, or by increasing the cost of the measurement.