Biological molecular machines can process information to reduce energy losses

Michal Kurzynski$^{1,*}$ and Przemyslaw Chelminiak$^1$

$^1$Faculty of Physics, A. Mickiewicz University, Umultowska 85, 61-614 Poznan, Poland
*kurzphys@amu.edu.pl

ABSTRACT

Biological molecular machines are enzymes that simultaneously catalyze two processes, one donating free energy and second accepting it. Recent studies show that stochastic dynamics of many protein enzymes is characterized by a long-time memory and the possibility of realizing catalyzed processes in various randomly selected ways. Employing a critical complex network, a sample model of such dynamics is specified and investigated using computer simulations. For models of this type, we prove the generalized fluctuation theorem, in which the entropy reduction is possible at the expense of some information creation. Creation and storage of information take place in the transient nonergodic stages of the dynamic process before the completion of free energy transduction cycle. The free energy supply at the beginning of the next cycle erases this information. The transient reduction of entropy reduces energy losses and may prove to be crucial for the movement of proteins on their tracks and the reason for most protein machines to operate as dimers or higher organized assemblies. From a broader physical perspective, the division of free energy into the useful energy and a second function of thermodynamic state, we refer to as organization, is worth emphasizing. Information can be assigned a physical meaning of a change in the value of both these functions.

The basic task of statistical physics is to combine the dynamics of the microstates of a studied system with the dynamics of the macrostates. The transition from the microstates to macrostates consists on averaging over time$^1$. In the statistical physics of simple systems, there are no intermediate levels of organization between microscopic mechanics and macroscopic thermodynamics. Living matter is, however, a complex system with the entire hierarchy of organization levels$^{2,3}$. The new achievements at the turn of the century allowed us to understand more accurately the nature of the micro and macrostates of the biological molecular machines, which belong to the level currently being referred to as nanoscopic.

First, the tertiary structure of the proteins, formerly identified with a particular conformational state, has been extended to the whole network of conformational substates$^{2,4–6,8–13}$. The Markov stochastic process on such a network is to be considered as the starting microscopic dynamics for a statistical treatment. As a symbol of the progress made recently, the study of conformational transitions in the native phosphoglycerate kinase could be quoted, in which a network of 530 nodes was found in the long, 17 µs molecular dynamics simulation$^{14}$.

Secondly, new methods of stochastic thermodynamics have been applied to the description of the non-equilibrium behavior of single nanoobjects in a finite time perspective. Work, dissipation and heat in nanoscopic machines are random variables and their fluctuations, proceeding forward and backward in time, proved to be related to each other by the fluctuation theorem$^{15–18}$. For nanoscopic machines, the notion of information production can be defined and the relationships between entropy and information lead to the generalized fluctuation theorem$^{19–27}$. It strengthened an almost universal consensus regarding the view on the operation of Maxwell’s demon consistent with thermodynamics$^{28–30}$. So, the demon must have a memory and in order to use fluctuations to do the work, it reduces entropy at the expense of gathering the necessary information on fluctuations in this memory. The information must be erased sooner or later, and for this the same or greater work must be used.

The purpose of this paper is to answer the intriguing question of whether biological molecular machines can act as Maxwell’s demons in the sense of a possibility of an entropy change into information. The formal approach presented is a continuation of our former trials$^{5,31}$. Based on the results of recent studies$^{6,8–14}$, we assume that the stochastic dynamics of the biological molecular machine is characterized by a long-time memory and the possibility to realize the work in various randomly selected ways. For such dynamics we calculate both the production of entropy and information, and show that they are of the opposite sign, hence one can partially compensate the other. The reduction of entropy reduces energy losses. Since the available experimental support for actual conformational transition networks in native proteins is still very poor, we restrict our attention to a model network only. However, the paper can be also treated as an invitation to experimentalists to perform similar analysis on real systems. In the context of specific research focused on the biological molecular machines, we come to a general conclusion that the notion of information creation can be given a well-defined physical meaning of change in the value of a certain thermodynamic state function that we suggest to call the organization of the system under consideration.
Clarification of concepts and results

Biological molecular machines as nanoscopic chemo-chemical machines

For many historical reasons, the word “machine” has acquired several different meanings. In our context, a machine is understood to be any physical system that enables two other systems to perform work on each other. Work may be done by mechanical, thermal, electrical, chemical or still some other forces. From a theoretical point of view, it is convenient to treat all biological molecular machines as chemo-chemical machines\textsuperscript{33}. We assume that every molecular process can be considered as a certain effectively unimolecular chemical reaction. So, the protein-chemo-chemical machines are enzymes, that simultaneously catalyze two reactions: the energy-donating input reaction \( R_1 \rightarrow P_1 \) and the energy-accepting output reaction \( R_2 \rightarrow P_2 \) (Fig. 1a). Also, pumps and molecular motors can be treated in this way. Indeed, the molecules present on either side of a biological membrane can be considered to occupy different chemical states (Fig. 1b), whereas the external load influences the energy of binding the motor to its track (Fig. 1c), which can be expressed as a change in the effective concentration of this track\textsuperscript{33,34}.

![Figure 1](image1.png)

**Figure 1.** A schematic picture of the three types of the biological molecular machines: (a) enzymes that simultaneously catalyze two reactions, (b) pumps placed in a biological membrane, and (c) motors moving along a track. Constraints, symbolized by a frame of a broken line, are imagined by the observer to control the mean number of the externally incoming and outgoing molecules involved in the catalysis. It should be stressed that the entire system inside the frame represents the machine: it is these constraints that determine the value of the thermodynamic variables and the conjugate forces.

The machine that we consider consists of a single enzyme macromolecule, surrounded by a solution of its substrates and products, possibly involving the track (Fig. 1). The whole is an open non-equilibrium system with constraints controlling the mean number of incoming and outgoing reacting molecules per machine or, in particular, the motor load. Under specified relations between the concentrations of reactants\textsuperscript{6}, two independent non-equilibrium molar concentrations of products \([P_1]\) and \([P_2]\), related to enzyme total concentration \([E]\), are to be treated as dimensionless thermodynamic variables \(X_1\) and \(X_2\). These, together with the conjugate thermodynamic forces, chemical affinities \(A_1\) and \(A_2\), determine work performed on and by the machine, respectively. Assuming that the molecule solution is perfect, the formal definitions have the form\textsuperscript{33,35}:

\[
X_i := \frac{[P_i]}{[E]}, \quad \beta A_i := \ln \left( \frac{[P_i]^{\text{eq}}}{[R_i]^{\text{eq}}} \right) \frac{[R_i]}{[P_i]}.
\]

\( \beta \) is the reciprocal of the thermal energy \( k_B T \) and the superscript eq denotes the equilibrium concentrations.

In the language of conventional chemical kinetics, the action of the chemo-chemical machine can be presented as in Fig. 2a. The energy-donating reaction \( R_1 \rightarrow P_1 \) forces the direction of the energy-accepting reaction \( R_2 \rightarrow P_2 \), though the nonequilibrium concentration values \([R_2]\) and \([P_2]\) would prefer the opposite direction\textsuperscript{33,35}. In Fig. 2a, the enzyme is present in only two states \( E_1 \) and \( E_2 \) and its internal dynamics is modeled by a possible single non-productive transition. However, as already mentioned, it is now well established that most if not all enzymatic proteins display the slow stochastic dynamics of the transitions between a variety of conformational substates composing their native state. It follows that, on the nanoscopic level, the microscopic dynamics of a specific biological chemo-chemical machine is the Markov process described by a system of master equations, determining a network of conformational transitions that obey the detailed balance condition. In this network, a system of pairs of nodes (the “gates”) is distinguished, between which the input and output chemical reactions force transitions that break the detailed balance\textsuperscript{6,31} (Fig. 2b). The rule is the presence of many output gates allowing a choice between them (“dynamic disorder” or the “fluctuating reaction rate”)\textsuperscript{8,9,36}.

In Ref.\textsuperscript{6}, we hypothesized that the protein conformational networks, like networks of the systems biology, have evolved in a self-organized criticality process, as a result of which they are scale-free and display a transition from the fractal to small-world organization\textsuperscript{37,38}. Recently, the real conformational network that seems to possess similar properties was obtained in molecular dynamics simulation\textsuperscript{14}. On such networks, modeled by scale-free fractal trees\textsuperscript{39} extended by long-range shortcuts,
many different output gates can be realized in a natural way. A sample network of 100 nodes, constructed following the algorithm described in Methods, is depicted in Fig. 2c. Let us note two hubs, the states of the lowest free energy, that can be identified with the two main conformations of the protein machine, e.g., “open” and “closed”, or “bent” and “straight”, usually the only ones occupied sufficiently high to be notable for the observer under equilibrium conditions. So, the proposed kind of stochastic dynamics combines indeed a chemical description with a mechanical description.

The constraints control the mean concentrations $X_1$ and $X_2$ of product molecules $P_1$ and $P_2$, respectively. However, the actual concentrations $x_1$ and $x_2$ in the single biological nanomachine fluctuate over time, increasing or decreasing by one as a result of consecutive passes forward or backward through the input and output gates. The thermodynamic description of the nanosystem still requires averaging over a finite observation time $t$, not too short and not too long (further on, we explain what it means). The result is a pair of fluxes $x_i(t)/t =: j_i(t)$, $i = 1, 2$, the statistics of which is studied in detail in the rest of the paper. An exemplary distribution of fluctuating thermodynamic concentrations $x_1(t)$ and $x_2(t)$ is depicted in Fig. 2d. The time $t$ has been chosen so that the concentration values are statistically independent. Consequently, the concentrations distribution is Gaussian, and thus the stochastic thermodynamic behavior of a single biological nanomachine under stationary conditions is represented by diffusion in the two-dimensional parabolic potential on the plane $(x_1, x_2)$ with the minimum at point $(X_1, X_2)^{33}$.

**Free energy transduction in stationary isothermal machines**

As opposed to thermal engines, the biological molecular machines operate at a constant temperature. Under isothermal conditions, the internal energy is uniquely divided into free energy, the component that can be turned into work, and bound energy (entropy multiplied by temperature), the component that can be turned into heat$^{30}$. Both thermodynamic quantities can make sense in the non-equilibrium state if the latter is treated as a partial equilibrium state$^{33}$. Free energy can be irreversibly turned into bound energy in the process of energy dissipation (the internal entropy production). In view of such internal energy division, the protein molecular machines are referred to as free energy transducers$^{35}$.
During the stationary isothermal processes, both free and bound energy remains constant. The energy processing pathways in any stationary isothermal machine are shown in Fig. 3a, where the role of all the physical quantities being in use is also indicated. \( X_i \) denotes the input \((i = 1)\) and the output \((i = 2)\) thermodynamic variable, \( A_i \) is the conjugate thermodynamic force and the time derivative, \( J_i = dX_i/dt \), is the corresponding flux. \( T \) is the temperature and \( S \) is the entropy. To clearly specify the degree of coupling of the fluxes, \( \varepsilon := J_2/J_1 \), it is important that variables \( X_1 \) and \( X_2 \) be dimensionless as in Eq. (1). Corresponding forces \( A_1 \) and \( A_2 \) are then also dimensionless, if only multiplied by the reversal of the thermal energy \( \beta = (k_B T)^{-1} \). By convention, fluxes \( J_1 \) and \( J_2 \) are assumed to be of the same sign. Then, one system performs work on the other when forces \( A_1 \) and \( A_2 \) are of the opposite sign. We assume \( J_1, J_2, A_1 > 0 \) and \( A_2 < 0 \) throughout this paper. Dimensionless variables \( X_1 \) and \( X_2 \) are defined in such a way that always \( A_1 + A_2 \geq 0 \), i.e., the machine does not work as a gear.

In the steady state, the total work flux (the power) \( A_1 J_1 + A_2 J_2 \) equals the dissipation flux, and that equals the heat flux. According to the second law of thermodynamics, the total dissipation flux (the internal entropy production rate, multiplied by the temperature) must be nonnegative\textsuperscript{35}. However, it consists of two components. The first component, \( (A_1 + A_2)J_1 \), achieved when the fluxes are tightly coupled, \( J_2 = J_1 \), must also be, according to the same law, nonnegative. The sign of the complement \( A_2 (J_2 - J_1) \) of \( (A_1 + A_2)J_1 \) to the total dissipation flux \( A_1 J_1 + A_2 J_2 \) is open to discussion.

In the macroscopic machines, entropy \( S \) is additive and can always be divided into the two parts \( S_1 \) and \( S_2 \), relating to input and output thermodynamic variables \( X_1 \) and \( X_2 \), respectively. As a consequence, flux \( A_2 (J_2 - J_1) \), which corresponds to \( S_2 \), must also be, under isothermal conditions, nonnegative. This means that, for the assumed negative \( A_2 \), the value of the degree of coupling between the adequately defined fluxes cannot be higher than unity, \( \varepsilon \leq 1 \). Macroscopically, the second component of the dissipation flux has the obvious interpretation of a slippage in the case of the mechanical machines, a short-circuit in the case of the electrical machines, or a leakage in the case of pumps.

However, because of non-vanishing correlations, entropy \( S \) in the nanoscopic machines is not additive\textsuperscript{19–23,25,27} and cannot be divided into two parts like free energy. This allows the transfer of information within the bound energy subsystem, which could result in a partial reduction of energy dissipation. In fact, for the biological molecular machines, output flux \( J_2 \) can surpass input flux \( J_1 \). Such a case was observed for single-headed biological motors: myosin II\textsuperscript{41}, myosin V\textsuperscript{42}, kinesin 3\textsuperscript{43}, and both flagellar\textsuperscript{44} and cytoplasmic\textsuperscript{45} dynein, which can take several steps along the track per ATP or GTP molecule hydrolyzed.

From the point of view of output force \( A_2 \), subsystem 1 carries out work on subsystem 2 while subsystem 2 carries out work on the environment. Jointly, the flux of the resultant work (the resultant power) \( A_2 (J_2 - J_1) \) is driven by force \( A_2 \). The complement to \( A_1 J_1 + A_2 J_2 \) is flux \( (A_1 + A_2)J_1 \) driven by force \( A_1 + A_2 \). Consequently, the free energy processing from Fig. 3a can be alternatively presented as in Fig. 3b, with the free energy transduction path absent. Here, the subsystem described by variable \( X_1 \) and constituting the perfect, tightly coupled machine is energetically independent from the subsystem described by variable \( X_2 - X_1 \) representing correction for a possible loose coupling. However, like the subsystems described by the two variables \( X_1 \) and \( X_2 \), the newly defined subsystems are still statistically correlated. Note that in Fig. 3b, only free energy is specified. Bound energy (thermodynamic entropy) and the heat bath are considered jointly as the energy reservoir.
**Generalized fluctuation theorem**

In the nanoscopic systems, work, dissipation and heat are fluctuating random variables and their nonequilibrium behavior is described by the fluctuation theorem\textsuperscript{15-18}. The total dynamics of the biological nanomachine is the Markov stochastic process on a network in which the macroscopic dynamics is distinguished by detailed balance breaking external transitions through a number of selected gates (Fig. 2b and c). In general, this network has no simple bipartite structure unlike the networks for which the concept of Shannon’s mutual information was considered\textsuperscript{21,24,25}. However, the partial entropy production in the distinguished set of external transitions satisfies the fluctuation theorem in the form of the Jarzynski equality\textsuperscript{26,27}:

\[ \langle \exp(-\sigma) \rangle = 1. \] (2)

The stochastic dimensionless entropy production \( \sigma \) for a period of time \( t \) is equal to energy dissipation in the stationary isothermal machine presented in Fig. 3, multiplied by the reciprocal thermal energy \( \beta \):

\[ \sigma = \beta [A_1 J_1(t) + A_2 J_2(t)] t = \beta [(A_1 + A_2) J_1(t) + A_2 J_0(t)] t, \quad J_0(t) := J_2(t) - J_1(t). \] (3)

Above, we wrote two alternative divisions of the dissipation flux discussed in Fig. 3.

The detailed fluctuation theorem corresponding to the integral fluctuation theorem (2) with the second division (3) is, in the logarithmic form,

\[ \ln \frac{p(j_1(t), j_0(t))}{p(-j_1(t), -j_0(t))} = \beta [(A_1 + A_2) j_1(t) + A_2 j_0(t)] t. \] (4)

\( \mathscr{J}_i(t) \) in (3), \( i = 1, 2 \) or \( i = 1, 0 \), denotes the random variable of the mean net flux over the time period \( t \), whereas \( j_i(t) \) in (4) denotes its particular value. \( p \) is the joint probability distribution function of these fluxes. The average in (2) is performed over this distribution. The chosen time \( t \) must be long enough for the considered ensemble to comprise only stationary fluxes, but finite for observing any fluctuations. All stationary fluxes in the ensemble are statistically independent, hence the probability distribution function \( p(j_1(t), j_2(t)) \) is Gaussian. The stationary averages of the random fluxes are time independent, \( \langle \mathscr{J}_i(t) \rangle = J_i \) for any \( t \). This time independence is not true for the higher moments of probability distribution and in Supplementary Information we prove that the variances (the squares of the standard deviations) of this distribution, \( \Delta_i^2 := \langle (\mathscr{J}_i(t) - J_i)^2 \rangle \), are inversely proportional to \( t \). Note that Eq. (4) with the well determined condition for \( t \) is stronger than the fluctuation theorem for the stationary fluxes earlier introduced by Gaspard\textsuperscript{46}.

The convexity of the exponential function provides the second law of thermodynamics

\[ \langle \sigma \rangle = \beta [(A_1 + A_2) J_1 + A_2 J_0] t \geq 0 \] (5)
to be a consequence of the Jarzynski equality (2).

In further discussion, for brevity, we will omit argument \( t \) specifying all the fluxes. One can calculate the marginal probability distributions \( p(j_1) \) and \( p(j_0) \) for the Gaussian probability distribution \( p(j_1, j_0) \) that satisfies fluctuation theorem (4). With their help we obtain the detailed and integral fluctuation theorems for the separate flux \( \mathscr{J}_1 \):

\[ \ln \frac{p(j_1)}{p(-j_1)} = \beta (A_1 + A_2) j_1 t + \rho \frac{\Delta_1}{\Delta_i} \beta A_2 j_1 t =: \sigma_1(j_1) + t_1(j_1), \quad \langle \exp(-\sigma_1 - t_1) \rangle = 1, \quad \langle \sigma_1 \rangle + \langle t_1 \rangle \geq 0, \] (6)

and the separate flux \( \mathscr{J}_0 \):

\[ \ln \frac{p(j_0)}{p(-j_0)} = \beta A_2 j_0 t + \rho \frac{\Delta_1}{\Delta_0} \beta (A_1 + A_2) j_0 t =: \sigma_0(j_0) + t_0(j_0), \quad \langle \exp(-\sigma_0 - t_0) \rangle = 1, \quad \langle \sigma_0 \rangle + \langle t_0 \rangle \geq 0. \] (7)

The generalized forms of the second law of thermodynamics are also listed. The averages are performed over the one-dimensional distributions, either \( p(j_1) \) or \( p(j_0) \). The relationships of standard deviations \( \Delta_1 \) and \( \Delta_0 \) with forces \( \beta A_1 \) and \( \beta A_2 \) are discussed in Supplementary Information and the dimensionless correlation coefficient \( \rho := \langle (\mathscr{J}_1 - J_1)(\mathscr{J}_0 - J_0) \rangle / \Delta_1 \Delta_0 \) is to be determined numerically for a particular total stochastic dynamics. As \( \Delta_1 \) and \( \Delta_0 \) are positive and forces \( \beta A_1 + A_2 \) and \( \beta A_1 \) are of the opposite signs, components \( \langle \sigma_i \rangle \) and \( \langle t_i \rangle \) add up if \( \rho \) is negative and substrate if \( \rho \) is positive.

Functions \( \sigma_i \) and \( \sigma_0 \) describe separate entropy productions but the interpretation of functions \( t_1 \) and \( t_0 \) is not so obvious. Because microdynamics of the system is hidden for the observer, the expressions for \( t_1 \) and \( t_0 \) provided by Eqs. (6) and (7) cannot be rewritten in the language of mutual information. However, these functions actually represent information produced by the microdynamics acting as an information reservoir\textsuperscript{22,23,27}. Indeed, the produced information has the form of a sequence of bits, or rather signs, e.g. \( \ldots, +, +, -, +, +, -, +, +, -, +, \ldots \) that describe successive passages forward or back through
As an application of the theory, we performed computer simulations of random walk on the network constructed following the gates, associated with the creation or annihilation of the appropriate product molecules. Only the sum of both functions \( t_1 \) and \( t_0 \) is related to mutual information between macrovariables \( J_1 \) and \( J_2 \) as from Eqs. (4), (6) and (7), the relationship

\[
\langle t_1 \rangle + \langle t_0 \rangle = -\left( \ln \frac{p(J_1, J_0)}{p(J_1)p(J_0)} - \ln \frac{p(-J_1, -J_0)}{p(-J_1)p(-J_0)} \right) \tag{8}
\]

results. Let us emphasize that because \( t_1 \) is the function of only \( j_1 \), and \( t_0 \) is the function of only \( j_0 \), the value of the difference of the logarithms in (8) remains the same when averaged over the correlated distribution \( p(j_1, j_0) \) or the product \( p(j_1)p(j_0) \).

Of course, this is not true for the separate components of the logarithm difference in (8).

Because Eq. (8) is symmetric with respect to the replacement of \( J_1 \) with \( J_0 \), it does not describe any directed information transfer between the input and output macrodynamics. Rather, Eq. (8) represents information transfer between the total macrodynamics and the environment. In fact, Eqs. (6) and (7) can be also obtained directly as the fluctuation theorems for each of the macrodynamics separately, with the signal of random passages of the product molecules through the constraints playing the role of an additional chemical force\(^{27} \). Then, the complementary macrodynamics together with the microdynamics are jointly treated as hidden\(^{47} \).

**Application to the model system**

As an application of the theory, we performed computer simulations of random walk on the network constructed following the algorithm described in Methods and shown in Fig. 2c. We assumed \( \beta A_1 = 1 \) and a few smaller, negative values of \( \beta A_2 \). To get statistical ensembles of the stationary fluxes, we have divided long stochastic trajectories of random walk into segments of equal lengths \( t \). In such divisions, the initial microstate in each segment is random. However, the long stochastic trajectory of random walk may also be divided into the successive free energy-transduction cycles of varying lengths, which start and end at the same microstate right after the free energy-donating transit through the single input gate. For the fluxes \( j_i \) to be statistically independent, the selected time \( t \) must have been a lot longer than the mean duration of the transient stage.

**Figure 4.** The generalized fluctuation theorem dependences found in the random walk simulations on the network shown in Fig. 2c with the single output gate (the squares) and the fourfold output gate (the circles). We assumed \( \beta A_1 = 1 \) and a few smaller, negative values of \( \beta A_2 \) determining the ratio of averaged fluxes \( \varepsilon = J_2/J_1 \) noted in the graphs. (a) The case of marginal \( p(j_1) \), compare Eq. (6). Most of simulation points practically cover each other. (b) The case of marginal \( p(j_0) \), compare Eq. (7). In order to distinguish clearly the entropic contribution from the informational one, the results are divided by the values of forces \( \beta (A_1 + A_2) \) and \( \beta A_2 \), respectively. The thick lines in both the graphs correspond then to the first, entropic components in Eqs. (6) and (7). Let us note that \( \beta (A_1 + A_2) \) is positive while \( \beta A_2 \) is negative.

Detailed results are discussed in Complementary Information. The two-dimensional distribution functions of fluxes obtained are well approximated by Gaussian (see Fig. 2d) and actually satisfy fluctuation theorem (4). From the two-dimensional distributions, we calculated the marginal distributions \( p(j_1) \) and \( p(j_0) \). The logarithms of the ratio of marginals \( p(j_i)/p(-j_i) \) are presented in Fig. 4a and b as the functions of \( j_i t \) for \( i = 1 \) and 0, respectively. All the dependences are actually linear.
as predicted by generalized fluctuation theorems (6) and (7). In order to clearly distinguish the entropic from informational contributions, the results are divided by the values of dimensionless forces $\beta(A_1 + A_2)$ and $\beta A_2$ for $i = 1$ and $0$, respectively, which makes the entropic contribution to be represented by a straight line with the unit tangent of inclination angle.

For the single output gate (the squares in Fig. 4), we got both $t_1$ and $t_0$ positive, of the same sign as $\sigma_1$ and $\sigma_0$. This is consistent with the determined negative values of correlation coefficient $\rho$. Both informational contributions differ from zero only for $\varepsilon$ close to unity, i.e., for very small values of force $A_2$, which does not brake practically the detailed balance condition for the transitions through the only output gate. In consequence, information $t_1$ is unnoticeable in Fig. 4a, and the signal of passages through the output gate, determining information $t_2$ in Fig. 4b, is a usual white noise. For the fourfold output gate (the circles in Fig. 4), however, the situation is much more interesting.

The stationary values of $J_1$ and $J_2$ are unambiguously related to forces $A_1$ and $A_2$, so we can directly determine the averaged entropy productions (dissipations) $\beta D_i := \langle \sigma_i \rangle$ and information productions $I_i := \langle \varepsilon_i \rangle$ resulting both from Eq. (6) and (7). For the fourfold output gate, these averages are presented in Fig. 5a as the functions of $J_2/J_1$. Note that information productions $I_1$ and $I_0$ are of the opposite sign to entropy productions $\beta D_1$ and $\beta D_0$, respectively, which is consistent with the now positive determined values of correlation coefficient $\rho$. For $J_2/J_1 < 1$, which corresponds to large values of output force $\beta A_2$, information $I_0$ is negative, i.e. taken from microdynamics, to be next subtracted from the positive production of entropy $\beta D_0$. For $J_2/J_1 > 1$, which corresponds to small values of output force $\beta A_2$, information $I_0$ is positive, i.e. sent to microdynamics and production of entropy $\beta D_0$ becomes negative. The biological molecular machines, for which this is the case, may be said to act as Maxwell’s demons, although they do not use the information creation directly for work, but only for reduction of energy losses.

![Image](image-url)

**Figure 5.** (a) The dependence of information productions $I_i$ and dissipation (entropy productions) $\beta D_i$, $i = 1, 0$, on the ratio of fluxes $J_2/J_1$ or, equivalently, forces $A_2/A_1$ for the fourfold output gate. The points represent the data obtained from Eqs. (6) and (7) for the same values of $\varepsilon = J_2/J_1$ as shown in Fig. 4. All quantities are counted in nats (natural logarithm is used instead of binary logarithm) per random walking step. (b) Transduction of useful energy versus transduction of organization in the nanoscopic machine which participates in a stationary isothermal process seen from a finite time perspective. In addition to energy processing (top), it is also necessary to consider the information processing (bottom). The transfer directions shown are for the model system under study with the fourfold output gate and $J_2/J_1 > 1$.

Flux difference $J_0 = J_2 - J_1$ is the time derivative of thermodynamic variable $X_0 = X_1$ that characterizes the organization of the system. In the case of the macroscopic mechanical machines, it is the mutual linear or angular slip of two coupled wheels and in the case of the macroscopic battery (the electrochemical machine), it is the voltage drop on internal resistance. Contrary to the macroscopic machines, the organization of the nanomachines can grow in a non-stationary state, which results from the conversion of entropy into information. Flux $J_1$ itself is the time derivative of thermodynamic variable $X_1$ that characterizes this component of free energy, which is responsible for the most optimal work of the machine in the tight coupling mode. It is justifiable to call it the useful energy. In the case of nanomachines, useful energy is applied also to erase information from microdynamics that functions as memory. The main result of our study is that useful energy transduction must be distinguished from organization transduction, which is schematically illustrated in Fig. 5b.
Under stationary isothermal conditions, the free energy conservation, expressed by the first law of thermodynamics, implies equality of the unused inserted work $W$ with dissipation $D$, which according to the second law of thermodynamics must be non-negative:

$$W = D \geq 0.$$  

Our interpretation of Eq. (8) indicates an additional conservation law, which implies equality of information $I^{\text{int}}$, transferred from the macrodynamics to the internal microdynamics and then to the environment, with information $I^{\text{ext}}$, coming back from the environment through the constraints to the macrodynamics. We have noted, that the signal of random passages of the product molecules through the constraints plays the role of the additional force\textsuperscript{37}. Thus, external information $k_B T I^{\text{ext}}$ can actually be considered as an additional work. Taking into account that the total sum of dissipation $D$ and internal information $I^{\text{int}}$ in $k_B T$ units must be nonnegative, it follows that both the first and second law of thermodynamics for the nanoscopic systems, stationary operating under isothermal conditions, should be generalized to the form

$$W + k_B T I^{\text{ext}} = D + k_B T I^{\text{int}} \geq 0.$$  

Like work $W$ and dissipation $D$, external and internal informations $I^{\text{ext}}$ and $I^{\text{int}}$ are functions of the process, not the state of the system. Both generalized laws of thermodynamics (10) apply separately to useful energy (subscript 1) and organization (subscript 0), see Fig. 5b. If the nanomachine has no choice possibility, both $D_1$ and $k_B T I^{\text{int}}_1$, as well as $D_0$ and $k_B T I^{\text{int}}_0$, are positive and information or, rather, noise processing results in an additional energy loss. Only in the nanomachines that choose randomly the way of doing work and exchange the information about it with the microdynamics, $D_0$ and $k_B T I^{\text{int}}_0$ are of the opposite sign, which makes $D_0$ to be negative, when $k_B T I^{\text{int}}_0$ is positive.

**Discussion**

Under physiological conditions, the protein molecular machines fluctuate constantly between lots of conformational substates composing their native state. The probabilities of visiting individual substates are far from the equilibrium and determined by the concentration of the surrounding molecules involved in the process. During the full free energy-transduction cycle, the possibility to choose different realizations of the free energy-accepting reaction results in the transient limitation of the dynamics to different regions of the conformational network, that is, to breaking of the ergodicity\textsuperscript{9,48}. The transient ergodicity breaking transforms the machine’s internal dynamics into a memory for storing and manipulating information. Information is erased each time the energy-donating reaction starts the next cycle from the input gate.

Information is exchanged between memory and the fluctuating thermodynamic variables. Two components of free energy determine the thermodynamic state of the machine. The first is the useful energy, proportional to the concentration of the input reaction product molecules $X_1$. And the second is organization, proportional to the difference in the concentrations of the output and input reaction product molecules $X_2 - X_1$. Similarly as work and dissipation, information is also a change in these well defined functions of state. In the protein machines with no possibility of choice, changes in organization are related, like in the macroscopic machines, only to energy losses. Otherwise, in the protein machines with such a possibility, changes in organization can essentially reduce these losses.

All of our conclusions are based on the distribution of input and output fluxes, which in turn results from the computer generated strings of signs at random moments of time (see Supplementary Information). Since similar strings can now be registered in the experiments\textsuperscript{31,43-45,46}, all the theses of the paper are open for experimental verification.

The storage capacity of memory is higher the larger and more complex is the network of conformational substates. This network is particularly large in the case of protein motors, since it then contains the substates of the motor together with the whole track on which it moves\textsuperscript{33}. In Fig. 5a, which presents the most important result of the paper, two cases of zero dissipation $D_0$ stand out. The first one is realized for tight coupling, $\varepsilon = J_2 / J_1 = 1$, and occurs simultaneously with the zero information $I_0$ transfer. The necessary condition for it is a special, critical value of output force $A_2$. There are some arguments, that such an effective force value is realized for processive motors with a feedback control between their two identical components\textsuperscript{49,50}. Because the corresponding informations and forces are transfered within microdynamics and not between micro and macrodynamics, this topic is out of scope of the present paper, although it is certainly worth detailed research. Achieving tight coupling may also prove to be the reason for most protein machines to operate as dimers or higher organized assemblies.

The second case of zero dissipation $D_0$ happens for $\beta A_2$ tending to 0, when the system is no longer working as a machine, but dissipation $D_1$ is used for collecting information $I_0$. Such conditions are met by the transcription factors in their one-dimensional search for the target on DNA, intermittent by three-dimensional flights\textsuperscript{51}. The flights correspond to our randomly chosen external transitions and the sliding between them is not simple diffusion but the active continuous time random walk\textsuperscript{48}.
(compare Complementary Information). In fact, the external transitions, assumed in the present model to be very fast, are also continuous time random walk\textsuperscript{52}. During information collection, ergodicity is broken and no work is performed by the system, but the initial state for searching is prepared at the expense of some free energy donation.

At the end, let us take the liberty for a couple of speculations. The generalized first and second laws of thermodynamics (10) were justified only for the stationary isothermal processes. Their universality remains an open problem. One thing is certain. In order for the information and entropy productions to be of opposite signs, the two necessary conditions must be met. These are the presence of an intermediate level of stochastic dynamics and the possibility of a choice. Besides the nanoscopic machines, we know two other systems sharing such properties. The first are the systems with thermodynamic fluctuations associated with phase transitions, whose organization is determined by the extra thermodynamic variables that survive stochastization\textsuperscript{1}, referred to as order parameters\textsuperscript{40} or, in various contexts, emergent\textsuperscript{39} or structural\textsuperscript{33} variables. Here, the long-living transient stage can be identified, e.g., with the nonergodic condensation of gas through the state of fog or the nonergodic solidification of liquid (the glass transition), and the local order parameter is a random variable. The second example is the systems displaying quantum fluctuations entangled with the environment, the organization of which is determined by extra variables that survive decoherence, identified with the classical variables\textsuperscript{54,55}. The value of this variable found in a local measurement is also random.

\textbf{Methods: Specification of the computer model}

The algorithm of constructing the stochastic scale-free fractal trees is from Goh et al.\textsuperscript{39}. Shortcuts, though more widely distributed, were considered by Rozenfeld, Song and Makse\textsuperscript{37}. Here, we randomly chose three shortcuts from the set of all the pairs of nodes distanced by six links. The network of 100 nodes in Fig. 2c is too small to determine its scaling properties, but a similar procedure of construction applied to $10^5$ nodes results in a scale-free network, which is fractal on a small length-scale and a small world on a large length-scale.

To provide the network with stochastic dynamics, we assumed the probability of changing a node to any of its neighbors to be the same in each random walking step\textsuperscript{4,6}. Then, following the detailed balance condition, the free energy of a given node is proportional to the number of its links (the node degree). The most stable nodes are the hubs, which are the only practically observed conformational substates under equilibrium conditions. For given node $l$, the transition probability to any of the neighboring nodes per one random walking step is one over the number of links and equals $(p_{eq}^{l} \tau_{int})^{-1}$, where $p_{eq}^{l}$ is the equilibrium occupation probability of the given node and $\tau_{int}$ is the mean time to repeat a chosen internal transition, counted in the random walking steps. This time is determined by the doubled number of links minus one\textsuperscript{4}, $\tau_{int} = 2(100 + 3 - 1) = 204$ random walking steps for the 100 node tree network with 3 shortcuts assumed.

The forward external transition probability, related to stationary concentration $[P_i]$, is determined by the mean time of external transition $\tau_{ext}$, and equals $(p_{eq}^{l} \tau_{ext})^{-1}$ per random walking step, $p_{eq}^{l}$ denoting the equilibrium occupation probability of the initial input or output node ($i = 1, 2$, respectively). The corresponding backward external transition probability is modified by detailed balance breaking factors $exp(-\beta A_{i})$. The assumed mean time of forward external transition $\tau_{ext} = 20$ was one order of the magnitude shorter than the mean time of internal transition $\tau_{int} = 204$, so that the whole process is controlled by the internal dynamics of the system. The opposite situation, when the process is controlled by the external spatial diffusion of the involved molecules, is not considered in the present paper.

\textbf{References}

1. Penrose, O. Foundations of statistical mechanics. \textit{Rep. Prog. Phys.} \textbf{42}, 1937–2006 (1979).
2. Volkenstein, M. V. \textit{Entropy and Information} (Birkhäuser, Basel, 2009).
3. Eigen, M. \textit{From Strange Simplicity to Complex Familiarity} (Oxford University Press, 2013).
4. Karplus, M. & McCammon, J. A. Molecular dynamics simulations of biomolecules. \textit{Nat. Struct. Molec. Biol.} \textbf{9}, 646–652 (2002).
5. Frauenfelder, H. \textit{The Physics of Proteins} (Springer, Berlin, 2010).
6. Kitao, A., Hayward, S. & Go, N. Energy landscape of a native protein: jumping-among-minima model. \textit{Proteins} \textbf{33}, 496–517 (1998).
7. Kurzynski, M. A synthetic picture of intramolecular dynamics of proteins. Towards a contemporary statistical theory of biochemical processes. \textit{Prog. Biophys. Molec. Biol.} \textbf{69}, 23–82 (1998).
8. Lu, H. P., Xun, L. & Xie, X. S. Single-molecule enzymatic dynamics. \textit{Science} \textbf{282}, 1877–1882 (1998).
9. English, B. P. \textit{et al.} Ever-fluctuating single enzyme molecules: Michaelis-menten equation revisited. \textit{Nature Chem. Biol.} \textbf{2}, 87–94 (2006).
10. Henzler-Wildman, K. & Kern, D. Dynamic personalities of proteins. *Nature** 450, 964–972 (2007).
11. Uversky, V. N. & Dunker, A. K. Understanding protein non-folding. *Biochim. Biophys. Acta** 1804, 1231–1263 (2010).
12. Chodera, J. D. & Noe, F. Markov state models of biomolecular conformational dynamics. *Curr. Opin. Struct. Biol.* 25, 135–144 (2014).
13. Shukla, D., Hernandez, C. X., Weber, J. K. & Pande, V. S. Markov state models provide insight into dynamic modulation of protein function. *Acc. Chem. Res.* 48, 414–422 (2015).
14. Hu, X. *et al.* The dynamics of single protein molecules is non-equilibrium and self-similar over thirteen decades in time. *Nature Phys.* 12, 171–174 (2016).
15. Evans, D. J. & Searles, D. J. The fluctuation theorem. *Adv. Phys.* 11, 1529–1585 (2002).
16. Jarzynski, C. Nonequilibrium equality for free energy differences. *Phys. Rev. Lett.* 78, 2690–2693 (1997).
17. Crooks, G. E. Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences. *Phys. Rev. E* 60, 2721–2726 (1999).
18. Seifert, U. Stochastic thermodynamics, fluctuation theorems and molecular machines. *Rep. Prog. Phys.* 75, 126001 (2012).
19. Sagawa, T. & Ueda, M. Generalized Jarzynski equality under nonequilibrium feedback control. *Phys. Rev. Lett.* 104, 090602 (2010).
20. Ponmurugan, M. Generalized detailed fluctuation theorem under nonequilibrium feedback control. *Phys. Rev. E* 82, 031129 (2010).
21. Sagawa, T. & Ueda, M. Fluctuation theorem with information exchange: Role of correlations in stochastic thermodynamics. *Phys. Rev. Lett.* 109, 180602 (2012).
22. Mandal, D. & Jarzynski, C. Work and information processing in a solvable model of Maxwell’s demon. *Proc. Natl. Acad. Sci. USA* 109, 11641–11645 (2012).
23. Barato, A. C. & Seifert, U. Unifying three perspectives on information processing in stochastic thermodynamics. *Phys. Rev. Lett.* 112, 090601 (2014).
24. Hartrich, D., Barato, A. C. & Seifert, U. Stochastic thermodynamics of bipartite systems: Transfer entropy inequalities and Maxwell’s demon interpretation. *J. Stat. Mech.* P02016, 2–19 (2014).
25. Horowitz, J. M. & Esposito, M. Thermodynamics with continuous information flow. *Phys. Rev. X* 4, 031015 (2014).
26. Shiraishi, N. & Sagawa, T. Fluctuation theorem for partially masked nonequilibrium dynamics. *Phys. Rev. E* 91, 012130 (2015).
27. Shiraishi, N., Matsumoto, T. & Sagawa, T. Measurement-feedback formalism meets information reservoirs. *New J. Phys.* 18, 013044 (2016).
28. Maruyama, K., Nori, F. & Verdal, V. Colloquium: The physics of Maxwell’s demon and information. *Rev. Mod. Phys.* 81, 1–23 (2009).
29. Sagawa, T. *Thermodynamics of Information Processing in Small Systems* (Springer, Berlin, 2013).
30. Parrondo, J. M. R., Horowitz, J. M. & Sagawa, T. Thermodynamics of information. *Nature Phys.* 11, 131–139 (2015).
31. Kurzynski, M. & Chelminiak, P. Mean first-passage time in stochastic theory of biochemical processes. Application to actomyosin molecular motor. *J. Stat. Phys.* 110, 137–181 (2003).
32. Kurzynski, M., Torchala, M. & Chelminiak, P. Output-input ratio in thermally fluctuating biomolecular machines. *Phys. Rev. E* 89, 012722 (2014).
33. Kurzynski, M. *The Thermodynamic Machinery of Life* (Springer, Berlin, 2006).
34. Kolomeisky, A. B. & Fisher, M. E. Molecular motors: A theorist’s perspective. *Annu. Rev. Phys. Chem.* 58, 675–695 (2007).
35. Hill, T. L. *Free Energy Transduction and Biochemical Cycle Kinetics* (Springer, New York, 1989).
36. Kurzynski, M. Statistical properties of the dichotomous noise generated in biochemical processes. *Cell. Mol. Biol. Lett.* 13, 502–513 (2008).
37. Rozenfeld, H. D., Song, C. & Makse, H. A. Small-world to fractal transition in complex networks: A renormalization group approach. *Phys. Rev. Lett.* 104, 025701 (2010).
38. Escolano, F., Handcock, E. R. & Lozano, M. A. Thermodynamic depth complexity of networks. Phys. Rev. E 85, 036206 (2012).
39. Goh, K. I., Salvi, G., Kahng, B. & Kim, D. Skeleton and fractal scaling in complex networks. Phys. Rev. Lett. 96, 018701 (2006).
40. Callen, H. B. Thermodynamics and an Introduction to Thermostatistics, 2nd edn. (Wiley, New York, 1985).
41. Kitamura, K., Tokunaga, M., Iwane, A. H. & Yanagida, T. A single myosin head moves along an actin filament with regular steps of 5.3 nanometers. Nature 397, 129–134 (1999).
42. Watanabe, T. M. et al. A one-headed class V myosin molecule develops multiple large (32-nm) steps successively. Proc. Natl. Acad. Sci. USA 101, 9630–9635 (2004).
43. Okada, Y. & Hirokawa, N. A processive single-headed motor: Kinesin superfamily protein kif1a. Science 283, 1152–1157 (1999).
44. Kojima, M., Kikumoto, M., Sakakibara, K. & Oiwa, K. Mechanical properties of a single-headed processive motor, inner-arm dynein subspecies-c of chlamydomonas studied at the single molecule level. J. Biol. Phys. 28, 335–345 (2002).
45. Malik, M., Carter, B. C., Lex, S. A., King, S. J. & Gross, S. P. Cytoplasmic dynein functions as a gear in response to load. Nature 427, 649–652 (2004).
46. Gaspard, P. Multivariate fluctuation relations for currents. New J. Phys. 15, 115014 (2013).
47. Mehl, J., Lander, B., Bechinger, C., Blickle, V. & Seifert, U. Role of hidden slow degrees of freedom in the fluctuation theorem. Phys. Rev. Lett. 108, 220601 (2012).
48. Metzler, R., Jeon, J.-H., Cherstvy, A. G. & Barkai, E. Anomalous diffusion models and their properties: Non-stationarity, non-ergodicity, and ageing at the centenary of single particle tracting. Phys. Chem. Chem. Phys. 16, 24128–24164 (2014).
49. Yildiz, A., Tomishige, M., Gennerich, A. & Vale, R. D. Intramolecular strain coordinates kinesin stepping behaviour along microtubules. Cell 134, 1030–1041 (2008).
50. Bier, M. & Cao, F. J. How occasional backstepping can speed up a processive motor protein. BioSystems 103, 355–359 (2011).
51. Kolomeisky, A. B. Physics of protein-dna interactions: mechanisms of facilitated target search. Phys. Chem. Chem. Phys 13, 2088–2095 (2011).
52. Liu, L., Cherstvy, A. G. & Metzler, R. Facilitated diffusion of transcription factor proteins with anomalous bulk diffusion. J. Phys. Chem, B 121, 1284–1289 (2017).
53. Anderson, P. W. More is different. Science 177, 393–396 (1972).
54. Zurek, W. H. Decoherence, einselection, and the quantum origins of the classical. Revs. Mod. Phys. 75, 715–765 (2003).
55. Zurek, W. H. Quantum darwinism. Nature Phys. 5, 181–188 (2009).
56. Chelminiak, P. & Kurzynski, M. Steady-state distributions of probability fluxes on complex networks. Physica A 468, 540–551 (2017).

Acknowledgements
M.K. thanks Yasar Demirel and Herve Cailleau for discussing the problem in the early stages of the investigation. Both authors are indebted to several anonymous referees of earlier versions of the paper for valuable remarks.

Author contributions statement
The general concept and the theory is mainly due to M.K., who also wrote the manuscript. The specification of the critical branching tree model and the numerical simulations are mainly due to P.C., who also wrote Supplementary Information.
Supplementary information

This supplement concerns a technical description of the methods we used to obtain all the results presented in the main paper. These methods include: (1) performance of computer simulations of random walks on the network extended with a system of the input and output gates; (2) determination of probability distributions of the dwell times between consecutive transitions through the gates; (3) analysis of the time courses of these transitions by means of a coarse-graining procedure and, consequently, construction of statistical ensembles of the fluctuating stationary fluxes proceeding through the gates; (4) evidence that the joint probability distributions of fluxes in the ensemble are Gaussian and (5) the proof that they satisfy the generalized fluctuation theorem. We demonstrate that (6) the appropriate fluctuation theorems also apply to the marginal distributions. Finally, (7) the dependences of the average fluxes through the gates on thermodynamic forces are briefly discussed, from which the degree of coupling can be determined.

Results of computer simulations

All the results presented in the main paper concern the statistical properties of stationary fluxes on a complex network shown in Fig. 2c with the dynamics described in Methods. The random walk simulations were performed in $t = 10^{10}$ computer steps for the assumed value of thermodynamic force $\beta A_1 = 1$ for the input gate and a few smaller, negative values of $\beta A_2$ for the output gate. In Fig. S.1, the sample time course of the net number $x_i(t)$ of external transitions through the single input gate ($i = 1$) and the fourfold output gate ($i = 2$) is shown for the first $t = 10^4$ random walking steps and the chosen value of $\beta A_2 = -0.05$. Both records reveal numerous steps up and down intermittent by shorter or longer periods of time, when the network is explored without passing through the gates. From this point of view the dynamics of external transitions resemble a continuous time random walk $^1$, according to which a system stays in a some state with a time drawn from the waiting time probability distribution before it appears in a subsequent state.

![Fig. S.1. Simulated time course of the net number of external transitions $x_i(t)$ through a single input gate ($i = 1$, below) and a fourfold output gate ($i = 2$, above). Time $t$ is counted in random walking steps on the network considered.](image)

Determination of the dwell time distributions $\psi_1(t)$ and $\psi_2(t)$

The flux is defined as the net number $x_i$ of external transitions through the gate $i = 1, 2$ in a certain time $t$. This time must be chosen to be long enough to ensure that the fluxes in the ensemble are both stationary and statistically independent. For this purpose we proceed as follows.
Fig. S.2. Log-log plot of distributions $\psi_i(t)$ between the successive external transitions through a single input gate ($i = 1$) and a fourfold output gate ($i = 2$). Continuous lines present the fits to Eq. S.1.

From the whole trajectory of a random walk, the probability distributions of the waiting, or rather dwell times $\psi_i(t)$ between the successive external transitions through the input and output gates can be determined. However, it needs to point out numerous vertical dashes of imperceptible width on the trajectories depicted in Fig. S.1. In the real experiments, such transitions are not observed due to inertia and a finite time resolution of the instrument. In order to avoid such effects and to smooth out the trajectories, we first determined the cumulative distribution functions $\Psi_i(t)$ and only then, by numerical differentiation, the very distribution functions $\psi_i(t) = \frac{d\Psi_i(t)}{dt}$ were established. The numerical differentiation interval (we chose it to be 10 random walking steps) determines the time resolution scale. The dwell time probability distributions $\psi_1(t)$ and $\psi_2(t)$ are shown in Fig. S.2. They fit the dependence

$$\psi_i(t) = \left[(1 - c_i)(\eta_i t)^{-\alpha_i}/\sqrt{\pi} + c_i\right] e^{-t/\tau_i}$$

with $i = 1, 2$ very well. It includes a certain power law range before the exponential completion, which indicates that the stochastic process displayed in Fig. S.1 is not the Markov process, but rather a complex continuous time random walk. From the fitting, we found $\eta_1 = 0.18$ and $\eta_2 = 0.16$, the reciprocal of which equals approximately the smoothing time 10 random walking steps, and $\alpha_1 = \alpha_2 = 1.8$, which is an exponent that specifies the fractal properties of the tree-like network. The contributions from the exponential decay are almost the same for the input and output transitions, $c_1 = 0.030$ and $c_2 = 0.038$, and only values of the disappearance time are evidently different, $\tau_1 = 233$ and $\tau_2 = 55$ random walking steps.

The mean dwell time for the transition through the input gate obtained from probability distribution $\psi_1(t)$, $\bar{\tau}_1 = 163$, is shorter than the exponential disappearance time $\tau_1 = 233$, because of the essential contribution of the fast power law decay to this distribution function. Both $\bar{\tau}_1$ and $\tau_1$, as well as the mean time of internal transition repeating $\tau_{int} = 204$ (see Methods in the main paper), may be considered as estimations of a mean duration of the free energy-transduction cycle. For comparison, we found the mean-first passage time between the most distant nodes of the network to be equal 710 random walking steps.

Taking all of it into account, we decided to choose a reasonable flux averaging time $t = 1000$ computer steps in the case of the network equipped with the fourfold output gate.

**Determination of fluctuating stationary fluxes $j_1(t)$ and $j_2(t)$**

In Fig. S.3, the successive steady-state fluxes $j_i(t) := \frac{x_i(t)}{t}$, $i = 1, 2$ averaged over $t = 1000$ random walking steps, are shown and compared with the mean fluxes $J_i$, averaged over the whole trajectory. This plot contains the same time course of the net number of external transitions $x_i$ as the one depicted in Fig. S.1, to which the coarse-graining procedure has been applied. The slopes of segments that make up the zigzag lines correspond to the values of fluctuating stationary fluxes $j_1$ proceeding through the input gate and $j_2$ passing through the output gate. In turn, the mean values of the fluxes $J_i$, averaged over $t = 10^{10}$...
Fig. S.3. Simulated time course of the net number of external transitions $x_i(t)$ through a single input gate ($i = 1$, below) and a fourfold output gate ($i = 2$, above). Time $t$ is counted in random walking steps on the network studied. The fluctuating fluxes $j_i$, averaged over successive $t = 1000$ steps, are shown and compared with the mean fluxes $J_i$, averaged over the whole trajectory. The origins of the reference systems for average fluxes (the straight lines) are established at $t = 5000$ to compare them with the fluctuating stationary fluxes.

steps of the complete computer simulation do not fluctuate in time and are determined in Fig. S.3 by the slopes of the straight lines.

**Flux probability distribution** $p(j_1(t), j_2(t))$

For the correctly estimated value of the time $t$, using the procedure illustrated in Fig. S.3, we can construct a statistical ensemble of the stationary and statistically independent fluxes $j_1(t)$ and $j_2(t)$ and determine their joint probability distribution $p(j_1(t), j_2(t))$. In the case of the network shown in Fig. 2c of the main paper we assumed $t = 10^3$ random walking steps for the fourfold output gate and, for comparison, $t = 10^4$ random walking steps for the single output gate. Sample results are depicted in Fig. S.4. They confirm that the averaging times $t$, we set up in our numerical analysis, were chosen properly, as the joint probability distribution of the fluctuating stationary fluxes has the form of the two-dimensional Gaussians:

$$p(j_1, j_2) = \frac{1}{2\pi \Delta_1 \Delta_2 \sqrt{1 - \rho^2}} \exp\left\{ -\frac{1}{2(1 - \rho^2)} \left[ \frac{(j_1 - J_1)^2}{\Delta_1^2} - 2\rho \frac{(j_1 - J_1)(j_2 - J_2)}{\Delta_1 \Delta_2} + \frac{(j_2 - J_2)^2}{\Delta_2^2} \right]\right\}. \tag{S.2}$$

Here and henceforth, for the sake of brevity, we omit argument $t$. This result indicates that the coarse-grained stochastic process which proceeds on the two-dimensional plane $(x_1, x_2)$ is the Markov process, though the time course of external transitions through the gates shown in Fig. S.1 is a non-Markovian continuous time random walk. The Gaussian distribution in Eq. S.2 is characterized by the correlation coefficient $\rho$, the two average values of the net fluxes $J_1$ and $J_2$ and the corresponding standard deviations $\Delta_1$ and $\Delta_2$. The values of these parameters are to be determined from the numerical data, shown in example in Fig. S.4, for a particular realization of stochastic dynamics on a network extended with the input and output gates. Four parameters can be directly calculated from the two marginals

$$p(j_i) = \frac{1}{\sqrt{2\pi \Delta_i}} \exp\left[ -\frac{(j_i - J_i)^2}{2\Delta_i^2} \right], \tag{S.3}$$
Fig. S.4. Exemplifying two-dimensional probability distribution functions $p(j_1(t), j_2(t))$ found in the model random walk simulations with a single output gate (left) and a fourfold output gate (right). For the single output gate, we assumed $t = 10^4$ random walking steps and obtained $J_2/J_1 = 0.95$. For the fourfold output gate, we assumed $t = 10^3$ random walking steps and obtained $J_2/J_1 = 1.59$. The mean values of the individual fluxes, multiplied by the time $t$ of the determination, are marked by the dashed lines.

for $i = 1, 2$, which arise from the Gaussian distribution in Eq. S.2. In this case we can use a standard fitting procedure limited to two parameters. However, the correlation coefficient

$$\rho = \frac{\langle (j_1 - J_1)(j_2 - J_2) \rangle}{\Delta_1 \Delta_2},$$

(S.4)

where $\rho$ is the random variable of the mean net flux over the time period $t$, can be determined only numerically. The average in Eq. S.4 is performed over the joint probability distribution $p(j_1, j_2)$.

**Generalized fluctuation theorem**

The important conclusion arising from our studies is that the Gaussian probability distributions of the stationary fluxes, shown for example in Fig. S.4, satisfy the detailed fluctuation theorem. We express it as in the main paper, upon transformation from $j_1$ and $j_2$ to the new pair of variables $j_1$ and $j_0 := j_2 - j_1$, thus

$$\ln \frac{p(j_1, j_0)}{p(-j_1, -j_0)} = \beta [(A_1 + A_2) j_1 + A_2 j_0] t.$$  

(S.5)

$\beta = 1/k_B T$ is the inverse of thermal energy, whereas $\beta A_1$ and $\beta A_2$ stand for the dimensionless thermodynamic forces. In general, the above formula describes the equation of a plane embedded in three-dimensional space. In Fig. S.5 only a bit of the plane is depicted due to the finite values of the fluxes $j_1$ and $j_0$ which constitute the restricted domain of the joint distribution function $p(j_1, j_0)$ in Eq. S.5. We applied an interpolation of the data points to make the plane a continuous manifold. The distortions visible on the edges of the plane result from an insufficient number of large values of the fluxes $j_1$ and $j_0$ in the statistical ensemble.

A direct combination of the detailed fluctuation theorem in Eq. S.5 with the Gaussian probability distribution in Eq. S.2, where instead of $j_2$ and $J_2$, the fluxes $j_0$ and $J_0$ appear, results in a system of two coupled equations

$$\frac{2}{1 - \rho^2} \left( \frac{J_1}{\Delta_1} - \rho \frac{J_0}{\Delta_1 \Delta_0} \right) = \beta (A_1 + A_2) t$$

$$\frac{2}{1 - \rho^2} \left( \frac{J_0}{\Delta_0} - \rho \frac{J_1}{\Delta_1 \Delta_0} \right) = \beta A_2 t.$$  

(S.6)
We use them in the next section to express the logarithms of the marginal distributions \( p(j_1) \) and \( p(j_0) \) in the form of two components, each of which is a function of time. Before that, however, we rewrite the above formulae in another form of the coupled quadratic equations for \( \Delta_i (i = 0, 1) \):

\[
\begin{align*}
\Delta_0^2 + \rho \left( 1 + \frac{A_1}{A_2} \right) \Delta_0 \Delta_1 &= \frac{2J_0}{\beta A_2 t} \\
\Delta_1^2 + \rho \left( 1 + \frac{A_1}{A_2} \right)^{-1} \Delta_0 \Delta_1 &= \frac{2J_1}{\beta (A_1 + A_2) t}.
\end{align*}
\]

These imply that the variances, or the squares of the standard deviations \( \Delta_0^2 \) and \( \Delta_1^2 \), which appear in the joint and marginal Gaussian probability distributions, are inversely proportional to time \( t \). Moreover, if the correlation coefficient \( \rho = 0 \), then the Eqs. \( S.7 \) become independent of each other, and we find that the standard deviations

\[
\Delta_0 = \sqrt{\frac{2J_0}{\beta A_2 t}}, \quad \Delta_1 = \sqrt{\frac{2J_1}{\beta (A_1 + A_2) t}}.
\]

This particular case was studied in\(^{4}\), where we have considered the stochastic dynamics on the complex network equipped with a single gate.

**Fluctuation theorems for marginals** \( p(j_1) \) and \( p(j_0) \)

By virtue of the detailed fluctuation theorem in Eq. \( S.5 \) we can derive the detailed fluctuation theorems for the separate flux \( J_1 \):

\[
\ln \frac{p(j_1)}{p(-j_1)} = \beta (A_1 + A_2) j_1 t - \left\langle \ln \frac{p(J_0 | j_1)}{p(-J_0 | -j_1)} - \beta A_2 J_0 t \right\rangle,
\]

and the separate flux \( J_0 \):

\[
\ln \frac{p(j_0)}{p(-j_0)} = \beta A_2 j_0 t - \left\langle \ln \frac{p(J_1 | j_0)}{p(-J_1 | -j_0)} - \beta (A_1 + A_2) J_1 t \right\rangle.
\]

The averages in the above equations are performed over the one-dimensional marginals \( p(j_0) \) and \( p(j_1) \) respectively. The conditional probability distribution for the two-dimensional Gaussian is

\[
p(j_0 | j_1) = \frac{1}{\Delta_0 \sqrt{2\pi(1-\rho^2)}} \exp \left\{ -\frac{1}{2\Delta_0^2(1-\rho^2)} \left[ (j_0 - J_0) - \frac{\rho \Delta_0}{\Delta_1} (j_1 - J_1) \right]^2 \right\}.
\]

![Fig. S.5. Numerical verification of the detailed fluctuation theorem in Eq. S.5. The data have been obtained from the random walk simulation on the network shown in Fig. 2c of the main text, extended with a single input gate and a fourfold output gate for, respectively, \( \beta A_1 = 1 \) and \( \beta A_2 = -0.05 \).](image-url)
Upon substituting this formula into Eq. S.9, then performing the adequate averaging over the marginal probability distribution in Eq. S.3 for \(i = 0\), and finally utilizing Eq. S.6, we arrive at
\[
\ln \frac{p(j_1)}{p(-j_1)} = \beta(A_1 + A_2)j_1t + \frac{\rho \Delta_0}{\Delta_1} \beta A_2 j_1 t =: \sigma_1(j_1) + t_1(j_1).
\]
(S.12)

Similarly, repeating the same calculations for Eq. S.10, we find that
\[
\ln \frac{p(j_0)}{p(-j_0)} = \beta A_2 j_0 t + \frac{\rho \Delta_1}{\Delta_0} \beta (A_1 + A_2) j_0 t =: \sigma_0(j_0) + t_0(j_0).
\]
(S.13)

The last two equations are identical with Eqs. 6 and 7 in the main paper.

The average values of \(t_1(j_1)\) in Eq. S.12 and \(t_0(j_0)\) in Eq. S.13 have the meaning of information creations in the internal dynamics within a network (their sign is discussed in the main paper). In the first case, the averaging is performed over the marginal probability distribution \(p(j_1)\), and hence
\[
\langle t_1 \rangle = \frac{\rho \Delta_0}{\Delta_1} \beta A_2 J_1 t,
\]
(S.14)

whereas in the second case we use the probability distribution \(p(j_0)\) to obtain
\[
\langle t_0 \rangle = \frac{\rho \Delta_1}{\Delta_0} \beta (A_1 + A_2) J_0 t.
\]
(S.15)

The sum of \(\langle t_1 \rangle\) and \(\langle t_0 \rangle\) can be expressed in a more general form when we add the second components on the r.h.s of Eqs. S.9 and S.10 and apply the detailed fluctuation theorem in Eq. S.5. As the final result we get that
\[
\langle t_1 \rangle + \langle t_0 \rangle = -\left( \ln \frac{p(J_1, J_0)}{p(J_1)p(J_0)} - \ln \frac{p(-J_1, -J_0)}{p(-J_1)p(-J_0)} \right),
\]
(S.16)

what reconstructs Eq. 8 in the main paper. Here, the average can be calculated over the joint probability distribution \(p(j_1, j_0)\) or the product \(p(j_1)p(j_0)\) of the marginal probability distributions, because \(t_1\) is the function of only the flux \(j_1\), and \(t_0\) is the function of only the flux \(j_0\) (see Eqs. S.12 and S.13).

If instead the averages are performed over the Gaussian probability distribution in Eq. S.2 with \(j_2\) replaced with \(j_0\), then the mean value of the first logarithm function in Eq. S.16 is:
\[
\left\langle \ln \frac{p(J_1, J_0)}{p(J_1)p(J_0)} \right\rangle = \ln \frac{1}{\sqrt{1 - \rho^2}},
\]
(S.17)

whereas of the second one is:
\[
\left\langle \ln \frac{p(-J_1, -J_0)}{p(-J_1)p(-J_0)} \right\rangle = \ln \frac{1}{\sqrt{1 - \rho^2}} + \frac{\rho \Delta_0}{\Delta_1} \beta A_2 J_1 t + \frac{\rho \Delta_1}{\Delta_0} \beta (A_1 + A_2) J_0 t.
\]
(S.18)

Eq. S.17 defines the mutual information between the fluxes \(J_1\) and \(J_0\), while Eq. S.18 determines the mutual information between the opposite fluxes \(-J_1\) and \(-J_0\).

**Flux-force dependences**

Assuming \(\beta A_1 = 1\) and varying the value of \(\beta A_2\) in a broad range, we were able, after sufficiently long simulations, to determine the mean values of fluxes \(J_2\) and \(J_1\) as functions of \(\beta A_2\). Both functions are depicted in Fig. S.6. They are characterized by two asymptotes, an inflection point and a specific value of the stalling force, for which they become equal to zero. In Fig. S.6, the fitting to the formula with four parameters, considered in Ref. 6, is given. Only the small region of positive \(J_1\) and \(J_2\) and negative \(\beta A_2\) corresponds to the system that operates as a machine. This is presented as an insertion in Fig. S.6. In this region, the dependences \(J_2(A_2)\) and \(J_1(A_2)\) are linear, which is a consequence of small values of \(\beta A_2\), much less than unity. With the use of data given in Fig. S.6, it is easy to determine the degree of coupling value \(\varepsilon = J_2/J_1\).
Fig. S.6. The flux-force dependences $J_2(A_2)$ and $J_1(A_2)$ for $\beta A_1 = 1$. The points denote the results of the random walk simulations and the continuous lines represent the fit mentioned in the text. The unit of time for the fluxes is one random walking step. Only if $J_2$ is positive and $A_2$ is negative, does the system acts as a machine. This region is shown in the magnification as the insertion.

References

1. Metzler, M. & Klafter, J. The random walk’s guide to anomalous diffusion: a fractional dynamics approach. *Phys. Rep.* 339, 1–77 (2000).
2. Kurzynski, M. A synthetic picture of intramolecular dynamics of proteins. Towards a contemporary statistical theory of biochemical processes. *Prog. Biophys. Molec. Biol.* 69, 23–82 (1998).
3. Širca, S. Probability for Physicists. *Springer* (2016).
4. Chelminiak, P. & Kurzynski, M. Steady-state distributions of probability fluxes on complex networks. *Physica A* 468, 540–551 (2017).
5. Cover, T. M. & Thomas, J. A. Elements of Information Theory. *John Wiley & Sons* (2006).
6. Kurzynski, M., Torchala, M. & Chelminiak, P. Output-input ratio in thermally fluctuating biomolecular machines. *Phys. Rev. E* 89, 012722 (2014).