Maximizing power and velocity of an information engine

Tushar K. Saha[a], Joseph N. E. Lucero[a], Jannik Ehrich[a], David A. Sivak[a,1], and John Bechhoefer[a,1]

[a]Department of Physics, Simon Fraser University, Burnaby, BC, V5A 1S6, Canada

Edited by Felix Ritort, University of Barcelona, Barcelona, Spain, and accepted by Editorial Board Member Mehran Kardar April 11, 2021 (received for review November 10, 2020)

Information-driven engines that rectify thermal fluctuations are a modern realization of the Maxwell-demon thought experiment. We introduce a simple design based on a heavy colloidal particle, held by an optical trap and immersed in water. Using a carefully designed feedback loop, our experimental realization of an "information ratchet" takes advantage of favorable "up" fluctuations to lift a weight against gravity, storing potential energy without doing external work. By optimizing the ratchet design for performance via a simple theory, we find that the rate of work storage and velocity of directed motion are limited only by the physical parameters of the engine: the size of the particle, stiffness of the ratchet spring, friction produced by the motion, and temperature of the surrounding medium. Notably, because performance saturates with increasing frequency of observations, the measurement process is not a limiting factor. The extracted power and velocity are at least an order of magnitude higher than in previously reported engines.

information engine | stochastic thermodynamics | feedback trap

Over 150 y ago, Maxwell proposed a thought experiment to sharpen understanding of the second law of thermodynamics (1). He envisioned a "neat-fingered being" that could sort fast and slow molecules to create a temperature difference between two chambers, thereby converting the energy of a heat bath into a form that could be used to do work. In modern terms, Maxwell's thought experiment was the first example of an information engine. In 1929, Leo Szilard proposed a simpler variant consisting of a single gas molecule in a chamber, partitioned by a wall (2, 3). If the particle is observed in the left half, the demon attaches a mass raised by motion to the right, and vice versa. Then, an isothermal expansion of the chamber raises the mass and stores potential energy. When run cyclically, the engine converts information about the state of the molecule into gravitational potential of a raised mass, seemingly without doing any work to lift the mass. This apparent violation of the second law of thermodynamics was later resolved by considering costs associated with the processing of information (4, 5), leading to a clearer understanding of the thermodynamics of information (6–8).

Recent advances in technology and theoretical developments in stochastic thermodynamics (9–12) have made it possible to experimentally realize information engines, based on the ideas of Maxwell and Szilard (13–19). They have been used to evaluate the Landauer cost of manipulating the associated measurement-memory device (20–23) and to explore the efficiency of information-to-work conversion (24–27).

Here, we create and study the performance of a useful information engine that not only extracts energy from heat but also stores energy by raising a weight, as initially imagined by Szilard. The "fuel" for the motor is the information gathered from favorable system fluctuations.

Our information engine consists of an optically trapped, micron-scale bead in water. The laser beam of the trap is horizontal, perpendicular to the vertical gravitational axis. The optical tweezers create a harmonic potential, where the bead fluctuates about an equilibrium that is lower because of the bead's weight (Fig. 1*B*). The motion of the heavy bead can be modeled by a simple spring-mass system (Fig. 1*A*). The demon monitors the position of the mass and, when the mass fluctuates beyond a predefined threshold, raises the position of the spring anchor (top bar). Repeating the process, the mass is raised by exploiting favorable "up" fluctuations arising from thermal noise in the medium.

The experimental setup is similar to ref. 28, but here we store the extracted work in a reservoir. The ability to "spend" stored work on demand and for varying purposes greatly increases the utility of the engine. A previous experimental system introduced by Admon et al. (25) also stored work, but its design was based on a repulsive potential, which meant that the motor was always powered by a combination of external mechanical work and information. Here, with a design based on a trap potential having a local minimum, we ensure that no external work is done on the bead, which simplifies the physical picture.

In our study of this information engine we focus on understanding and then optimizing its performance: How much can it lift? How fast can it go? More precisely, what is the upper bound to the rate of gravitational energy storage and to the directed velocity? We reason that the value of the function of a motor can greatly exceed the cost of running it. For example, in biological applications such as chemotaxis, the metabolic costs of running cellular machinery (including information-processing costs) are usually unimportant compared to the benefit gained by the ability to move toward a new food source or away from a predator (29).

Significance

Around 100 y ago, Szilard imagined how to raise a weight without doing any work, just using the information gained by "looking" at a single gas molecule bouncing inside a box. Here, we designed an engine that stores energy by raising a bead against gravity, driven purely by information about the bead position. No work is done directly on the bead; instead, all dissipation occurs in the measuring apparatus. By choosing the right size of bead and through careful design, we improved information-engine performance by more than a factor of 10 over the best previous designs. We achieve a maximum velocity of 190 μm/s and maximum power of $10^3\ k_BT/s$, numbers comparable to those measured in microorganisms.

Author contributions: D.A.S. and J.B. designed research; T.K.S. performed experiments; T.K.S., J.N.E.L., and J.E. analyzed data; J.N.E.L. and J.E. performed theoretical analysis; and T.K.S., J.N.E.L., J.E., D.A.S., and J.B. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission. F.R. is a guest editor invited by the Editorial Board.

See online for related content such as Commentaries.

[1]To whom correspondence may be addressed. Email: dsivak@sfu.ca or johnb@sfu.ca.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2023356118/-/DCSupplemental.

Published May 10, 2021.

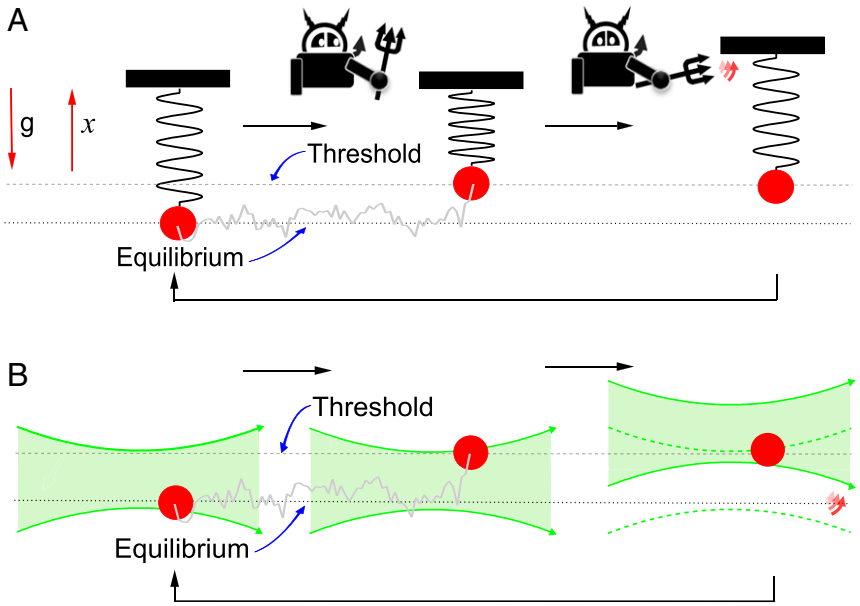

Fig. 1. Schematic of the information engine. (*A*) Ratcheted spring-mass system under gravity. (*B*) Experimental realization using horizontal optical tweezers in a vertical gravitational field. Feedback operations on the right side in *A* and *B* are indicated by the small red "swoosh" arrows.

We thus seek to maximize performance, independent of the energy required. As we will show, there is a maximum achievable energy-storage rate and a maximum achievable directed velocity, even when the signal-to-noise ratio of the measuring system is arbitrarily high (with correspondingly high costs for information processing); knowing the maximum level of performance independent of information costs can provide a benchmark to evaluate trade-offs between performance and operational costs. We will also show that the performance of an information engine is limited by its material parameters. In our case, these parameters include trap stiffness and bead size, and we provide a systematic method of choosing their values to maximize the desired performance measure.

Theory

Equation of Motion. The dynamics of an optically trapped bead are well described by an overdamped Langevin equation,

$$\gamma\dot{x}(t) = \underbrace{-\kappa\,(x(t)-\lambda(t))}_{\text{restoring force}} - \underbrace{mg}_{\text{grav. force}} + \underbrace{\sqrt{2k_{\mathrm{B}}T\gamma}\,\nu(t)}_{\text{thermal noise}}\,, \qquad [1]$$

where $x(t)$ denotes the position of a bead of radius r at time t, $\lambda(t)$ the center of the trap, κ the trap stiffness, γ the friction coefficient, and g the gravitational acceleration, and $\nu(t)$ represents Gaussian white noise with zero mean and $\langle\nu(t)\,\nu(t')\rangle = \delta(t-t')$. The effective mass $m = (4/3)\pi r^3\Delta\rho$ of the bead depends on the relative density $\Delta\rho = \rho_{\mathrm{bead}} - \rho_{\mathrm{medium}}$ (= 1.0 g/mL, for all bead diameters) and accounts for buoyancy. Scaling lengths by the equilibrium standard deviation $\sigma = \sqrt{k_{\mathrm{B}}T/\kappa}$ of the bead position and time by the bead relaxation time $\tau_r = \gamma/\kappa$, the overdamped Langevin equation becomes

$$\dot{x}(t) = -\,[x(t)-\lambda(t)] - \delta_{\mathrm{g}} + \sqrt{2}\,\nu(t)\,, \qquad [2]$$

where $\delta_{\mathrm{g}} \equiv mg/\kappa\sigma$ is a scaled effective mass that measures the sag of the bead due to gravity, relative to the scale of equilibrium fluctuations in the trap. The bead position is measured at discrete time intervals of $t_{\mathrm{s}}' = 20$ μs, and the feedback on the trap position is applied after a delay of one time step. Integrating Eq. **2** over one time step gives discrete-time dynamics (30),

$$x_{n+1} = \mathrm{e}^{-t_{\mathrm{s}}}x_n + \left(1-\mathrm{e}^{-t_{\mathrm{s}}}\right)(\lambda_n - \delta_{\mathrm{g}}) + \sqrt{1-\mathrm{e}^{-2t_{\mathrm{s}}}}\,\xi_n, \qquad [3]$$

where $t_{\mathrm{s}} \equiv t_{\mathrm{s}}'/\tau_{\mathrm{r}}$, $x_n \equiv x(nt_{\mathrm{s}})$ denotes the position at time step number n, and ξ_n is a Gaussian random variable, with zero mean and unit variance, satisfying $\langle\xi_m\,\xi_n\rangle = \delta_{mn}$.

The trap position λ_n is updated according to a feedback algorithm,

$$\lambda_{n+1} = \begin{cases} \lambda_n + \alpha(x_n - \lambda_n), & x_n - \lambda_n > X_{\mathrm{T}} \\ \lambda_n, & \text{otherwise.} \end{cases} \qquad [4]$$

Here, X_{T} is the threshold, and α is the feedback gain. Fig. 2, *Bottom-Right Inset* shows example time series of the upward motion $x(t)$ of the mass and $\lambda(t)$ of the trap.

For an instantaneous measurement and shift of the trap center, Fig. 2, *Top-Left Inset* shows that choosing $\alpha = 2$ would impose a zero-work condition (cf. *SI Appendix*, section B), where the stored potential energy results solely from the conversion of the information about the bead position; i.e., the work done by the trap is set to zero. In our experimental apparatus, there is a delay of 20 μs (one time step) arising (mostly) from the acousto-optic deflector (AOD) that controls the position of the trap (31). If uncompensated, such a delay can lead to a significant amount of input-trap work. During the delay, the bead tends to move back toward the equilibrium, reducing the value of α needed to impose zero work. To compensate, α is set empirically to implement the zero-work condition, which occurs at $\alpha \approx 1.5$ in Fig. 2. Note that the reset position X_{R} illustrated in Fig. 2 is related to α by $X_{\mathrm{R}} = (\alpha-1)X_{\mathrm{T}}$.

Energy Storage and Directed Motion. The input work is the change in energy of the bead that occurs when the position λ of the trap center is moved. Since the trap center is moved only at the sampling times $\{t_{n+1}\}$ and since the shift happens at a faster time

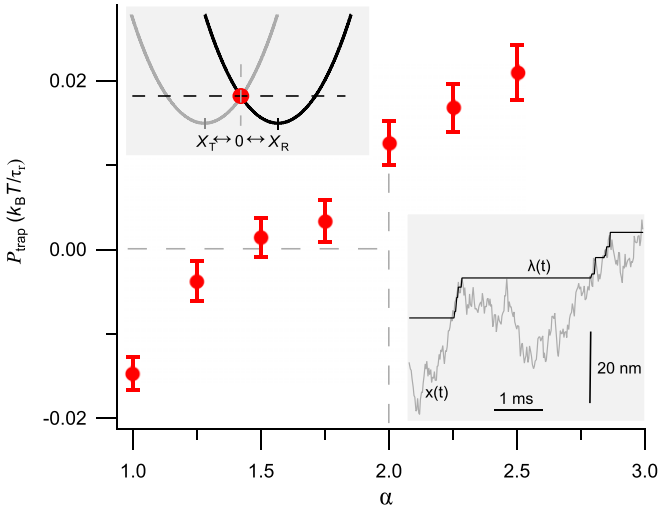

Fig. 2. Zero-work condition defining a pure information engine. Trap power P_{trap} as a function of feedback gain α for fixed threshold $X_T = 0$, scaled relaxation frequency $\tau_r/t'_s = 180$, scaled effective mass $\delta_g = 0.8$, relaxation time τ_r = 3.6 ms, diffusion constant $D = 0.16$ μm^2/s, trap stiffness κ = 7.0 pN/μm, and bead diameter of 3 μm. (*Bottom-Right Inset*) Experimental trajectories of the bead [$x(t)$, gray] and trap [$\lambda(t)$, black] during continuous ratcheting. (*Top-Left Inset*) Naive zero-work condition for a harmonic potential is $\alpha = 2$, equivalent to $X_R = X_T$. The black curve denotes the trap potential in the current step and gray curve in the previous step. Error bars here and in other figures are the standard error of the mean (*SI Appendix*, section M).

scale (< 1 μs, set by the response of the AOD) than bead motion, the work done at each update t_{n+1} is

$$W_{n+1} = \tfrac{1}{2}\left[(x_{n+1} - \lambda_{n+1})^2 - (x_{n+1} - \lambda_n)^2\right]. \qquad [5]$$

Similarly, the gain in gravitational potential is

$$\Delta U_{n+1} = \delta_g\,(x_{n+1} - x_n). \qquad [6]$$

By convention, the trap work is positive if energy flows into the system and negative if it flows out.

We quantify the performance of the information engine by the (long-time average) directed velocity and stored power, ideally for an infinitely long trajectory. Each trajectory can be viewed as a sequence of independent ratchet events, each starting with the particle at position $\lambda - X_R$ inside the trapping potential and ending when the particle fluctuates up and first reaches the position $\lambda + X_T$. The displacement $\Delta x = (\lambda + X_T) - (\lambda - X_R) = X_R + X_T$ is thus fixed for each event, but the time required for event m, the first-passage time τ_{FP}, is stochastic (32).

Using the above definitions, we write the velocity

$$v = \lim_{t_{traj}\to\infty} \frac{X_{traj}}{t_{traj}} = \lim_{N_{ratch}\to\infty} \frac{\sum_{m=1}^{N_{ratch}} \Delta x}{\sum_{m=1}^{N_{ratch}} (\tau_{FP})_m} = \frac{X_R + X_T}{\tau_{MFP}}, \qquad [7]$$

where X_{traj} and t_{traj} are the total trajectory length and time and N_{ratch} is the number of ratcheting events. We used the law of large numbers to write $\sum_m (\tau_{FP})_m \to N_{ratch}\tau_{MFP}$, with τ_{MFP} the mean first-passage time (MFPT), the average of τ_{FP} (cf. *SI Appendix*, section J).

The corresponding rate of energy extraction (power) is $v\,mg$, or, in scaled units,

$$P = v\,\delta_g. \qquad [8]$$

For each data point, typically 100 repeated trajectories are measured over a fixed distance of 340 nm. The velocity and power are estimated by replacing X_{traj} and t_{traj} in Eqs. **7** and **8** with their trajectory averages (*Materials and Methods*).

Predicted Maximum Output Power and Velocity. To predict the maximum output power and velocity, we first calculate the MFPT (Fig. 2, *Top-Left Inset*). A standard calculation (32, 33) (*SI Appendix*, section C) gives, in scaled units,

$$\tau_{MFP}(X_T) = \int_{-X_T}^{X_T} dx'\, e^{V(x')} \int_{-\infty}^{x'} dx''\, e^{-V(x'')}, \qquad [9]$$

for total potential $V(x) \equiv \frac{1}{2}x^2 + \delta_g\, x$. Although Eq. **9** in general must be solved numerically, a Taylor expansion for small threshold X_T gives

$$\tau_{MFP}(X_T) = \sqrt{2\pi}\, e^{\delta_g^2/2}\left[1 + \mathrm{erf}\left(\frac{\delta_g}{\sqrt{2}}\right)\right] X_T + \mathcal{O}\left(X_T^3\right), \qquad [10]$$

with positive higher-order corrections.

The velocity is then maximized by taking $X_T \to 0$:

$$v(X_T) = \frac{2X_T}{\tau_{MFP}(X_T)} \qquad [11a]$$

$$\xrightarrow{X_T\to 0} \sqrt{\frac{2}{\pi}}\, e^{-\delta_g^2/2}\left[1 + \mathrm{erf}\left(\frac{\delta_g}{\sqrt{2}}\right)\right]^{-1}. \qquad [11b]$$

Eq. **11b** was derived previously using a different method and in a slightly different context (34).

In physical units and for large force constants ($\kappa \to \infty$), the velocity and power are

$$v' = \left(\frac{\sigma}{\tau_r}\right) v \overset{\kappa\to\infty}{\sim} \sqrt{\frac{2k_B T}{\pi}}\frac{\sqrt{\kappa}}{\gamma}, \qquad [12a]$$

$$P' = \left(\frac{k_B T}{\tau_r}\right) P \overset{\kappa\to\infty}{\sim} \sqrt{\frac{2k_B T}{\pi}}\frac{\sqrt{\kappa}}{\gamma} mg. \qquad [12b]$$

Results

To maximize the rate of gravitational-energy extraction (the power), we first studied its dependence on the sampling frequency. Fixing the trap stiffness κ and hence the relaxation time τ_r, we varied the sampling time t'_s. Fig. 3*A* shows that the power saturates at large sampling frequencies ($f_s = \tau_r/t'_s \gg 1$). Thus, making more measurements may not increase the extracted power. Indeed, measurements faster than the relaxation time τ_r of the bead are correlated and thus provide less information than a single, isolated measurement (25). Nonetheless, sampling faster than τ_r reduces the chance of missing a favorable fluctuation that reaches X_T.

At low frequencies, the number of ratchet events is linearly proportional to the sampling frequency. The gray dotted line in Fig. 3*A* has slope ≈ 0.19, which is consistent with a calculation assuming the particle position distribution equilibrates during each interval (*SI Appendix*, section D). The solid curve in Fig. 3*A* is based on semianalytic calculations (*SI Appendix*, section E) that use the measured material parameters and agree well with experiments, with no free parameters. Thus, sampling more slowly than the fluctuation time scale τ_r of the dynamics misses possibly useful fluctuations; sampling more quickly eventually yields diminishing returns.

Having established that the extracted power is maximized for infinite sampling frequency, we henceforth use the fastest feedback time of 20 μs, which typically corresponds to $f_s \gtrsim 100$. Such a sampling frequency is high enough that analytic calculations

PHYSICS

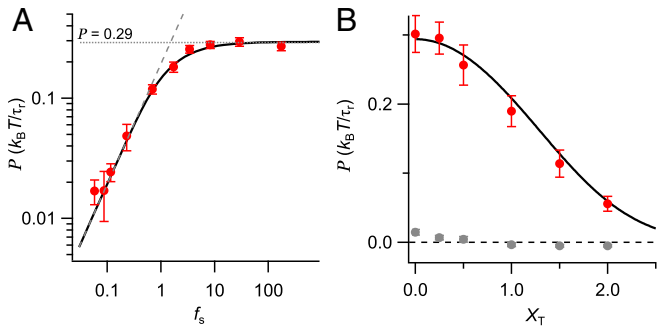

Fig. 3. Optimization of ratcheting power. (*A*) Power as a function of sampling frequency $f_s = \tau_r / t'_s$. The black solid curve denotes the semianalytic results (*SI Appendix*, section E) for the same material parameters, feedback gain $\alpha = 1.8$, and threshold $X_T = 0$. The horizontal dotted line indicates the infinite-frequency limit (Eq. **11b** multiplied by δ_g), and the dashed line denotes the low-frequency limit (*SI Appendix*, section D). (*B*) Power as a function of threshold X_T for fixed $\alpha = 1.9$ and sampling frequency of 50 kHz. The gray markers show that the input trap power is small. The black curve follows from Eq. **11a**. Red markers denote experimental values. For all data, scaled effective mass $\delta_g = 0.8$, relaxation time τ_r = 3.5 ms, diffusion constant = 0.16 $\mu m^2/s$, trap stiffness κ = 7.3 pN/μm, and bead diameter = 3 μm.

based on the continuous-sampling limit ($f_s \to \infty$) describe the data well.

We next explored how to set the position threshold X_T. This parameter controls the magnitude of the fluctuation that is captured during each ratchet event. The experiments were performed for $\delta_g = 0.8$. The feedback gain $\alpha = 1.9$ ensured that the input power was zero for the chosen threshold values, as confirmed by the gray solid markers in Fig. 3*B*.

Fig. 3*B* shows that the output power, under the constraint of zero input power, is maximized for $X_T \to 0$ ("continuous ratcheting"). The trap position $\lambda(t)$ then either ratchets to accommodate up fluctuations or pauses when the bead fluctuates down, before reaching the threshold again (Fig. 2, *Bottom-Right Inset*). As X_T increases, the fluctuations that take the bead to the threshold become increasingly rare (exponentially in X_T), leading to longer wait times between ratchet events; hence, the power tends to zero. The solid black curve is calculated by numerically integrating Eq. **9** to find $\tau_{MFP}(X_T)$.

Having determined that continuous sampling and continuous ratcheting maximize the extracted power, we explored the role of bead mass in experiments using nominal bead diameters of 0.5, 1.5, 3, and 5 μm. For each trap strength κ (set by the trapping laser power) and for each trapped bead (whose size varies slightly from the nominal size listed by the manufacturer) we determine the value of feedback gain α that makes $P_{trap} \approx 0$. The gray markers in Fig. 4*D* show that the trap power can be kept small, even though the required value of α is different for each value of δ_g.

We then measured the extracted velocity (Fig. 4*A*) and power (Fig. 4*C*) for the four nominal bead sizes. For a fixed bead size, the power and velocity increase monotonically with trap stiffness, as increasing the trap stiffness reduces the MFPT. For fixed trap stiffness and bead density, the velocity decreases with bead size. By contrast, the power is maximized at an optimal intermediate bead size. Scaling the length, time, and energy by the trap standard deviation σ, trap relaxation time τ_r, and $k_B T$, respectively, collapses the data onto single scaled power and velocity curves as a function of the scaled effective mass δ_g, Eq. **11b** (Fig. 4 *B* and *D*). The power is maximized at $P^* \approx 0.295\ k_B T/\tau_r$ for $\delta_g \approx 0.845$ and the velocity at $v^* \approx 0.8\ \sigma/\tau_r$ for $\delta_g \to 0$ (34). The maximum in extracted power at finite δ_g results from a competition between two effects: The potential energy of a raised object increases with mass, but so does the time to fluctuate beyond a threshold.

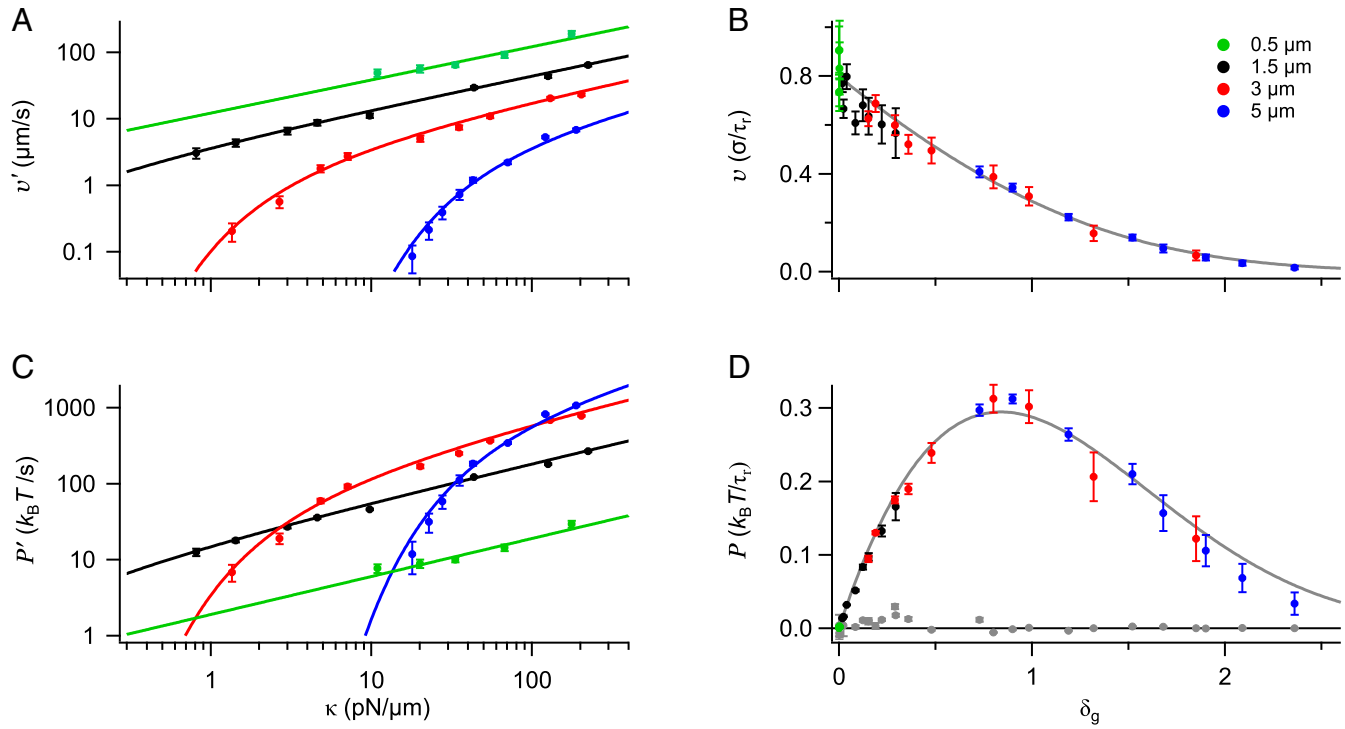

Fig. 4. Power and velocity for bead diameters 0.5 (green), 1.5 (black), 3 (red), and 5 (blue) μm for threshold $X_T = 0$. Markers denote experimental data (see *SI Appendix*, Table S2 for experimental parameters). (*A*) Velocity as a function of trap stiffness κ. (*B*) Scaled velocity as a function of scaled effective mass δ_g. (*C*) Power as a function of κ. (*D*) Scaled power as a function of δ_g. Gray markers show that the corresponding P_{trap} values remain small. Solid curves in *A* and *C* are calculated from Eqs. **12a** and **12b**. Solid curves in *B* and *D* are calculated from Eq. **11b**.

Finally, we explored the influence of gravity on directed motion. Our analysis suggests that gravitational effects should be quantified by the scaled effective mass, δ_g. To test this idea, we compared the measured directed velocity achieved for horizontal motion with that achieved for vertical motion (*SI Appendix*, section G). Horizontal velocities are consistent with predictions based on Eq. **12a** or, equivalently, Eq. **11b** for $\delta_g \to 0$. Thus, when the bead is sufficiently light or small (1.5 μm in this case), gravity becomes irrelevant: Particle speed is independent of direction (*SI Appendix*, Fig. S4). For heavier beads and smaller spring constants the motion is slower in the vertical direction.

Discussion

We have designed a simple information-fueled engine that can convert the heat of a surrounding bath into directed motion and hence store gravitational potential energy. A systematic study of conditions that optimize the performance limits of the engine shows that continuous measurements and continuous ratcheting are best. Fortunately, the analysis of the continuous-feedback limit is simpler than that for the corresponding discrete-time dynamics and can draw on well-known results from the analytic theory of MFPTs. From the optimization, we find simple expressions for extracted power and velocity establishing that the performance limits of the engine are set by material parameters such as the stiffness of the spring created by the optical tweezers.

Fig. 4 *A* and *C* and Eq. **12** show that smaller beads maximize directed motion, but larger beads maximize power extraction. That varying goals call for varying design principles is familiar in macroscopic applications. For example, the diesel engines used in trucks are optimized for power, whereas the turbocharged engines used in race cars are optimized for speed. More generally, systematic connections between material parameters and performance limits are common features of motors. Indeed, motors ranging from proteins to jet engines follow scaling laws whose form is determined by the failure modes of the materials used in the motor construction (35, 36).

By following optimal design principles, we have markedly improved performance relative to previous efforts, which focused instead on information-processing costs and the associated "information-to-work" efficiency of the engine (13, 24, 25, 28). The maximum extracted power is 10^4 times higher than that reported in ref. 25), although comparable laser powers are used. Most of the improvement in extracted power is achieved through the trap design. In the present case, power is applied where needed, via a single trap; an array of traps was used in ref. 25. Our design may also be compared with ref. 28, which uses a single trap, as here, but does not store work. The power levels achieved here exceed those in ref. 26 by an order of magnitude. The improvement relative to ref. 26 arises from careful optimization of parameters (bead size, X_T, etc.). Similarly, we increase the directed velocity by a factor of 30 compared to ref. 37 by choosing a smaller bead.

For our setup, the "best" values achieved for power and velocity are 1,066 $k_B T$/s and 190 μm/s, respectively. These values are significant: They are roughly 10 times faster than *Escherichia coli* and are comparable to the speeds of faster motile bacteria such as those found in marine environments (which need to outswim their algae prey) (38) and are also comparable to the power used to drive molecular motors such as kinesin (39).

For setups similar to the one used here, the laser power can in principle be increased significantly, which would increase the trap constant κ; however, in many applications, heating will limit the power that can be applied. Another route to increasing performance is to optimize the response properties of the trapped particle. Here, we limited our particle choice to dielectric spheres; more sophisticated core-shell particle designs can reduce beam reflection and scattering forces, thereby increasing the trap stiffness at fixed laser power by a factor of approximately 10 (40).

In our experiments, the optical-tweezer setup imposed a harmonic potential. Could more power or higher velocities be possible using a different potential shape? We numerically studied a potential with controllable asymmetry and found no improvement, given a fixed maximum stiffness. Additionally, we can show that, for symmetric traps, the harmonic shape is optimal (*SI Appendix*, section K).

Another route to higher rates of energy extraction and storage is via systems with intrinsically shorter dynamical time scales. Experimental setups using electronic circuits (15, 22) and superconducting qubits (16, 17) have been used to extract power. With modification, they could also store power and then be optimized by following the techniques introduced here.

Beyond technological limits set by the stiffness of the material used to build the motor, the dynamical model used in our optimization can break down. Naively, decreasing dynamical time scales (e.g., by increasing the trap stiffness κ) always improves information-engine performance. However, our analysis assumes Eq. **1**, which describes a simple overdamped Langevin model with instantaneous damping and is characterized by the relaxation time τ_r. For the range of κ and bead sizes that we explore, this assumption holds; however, as κ increases, the time scale τ_r of the trap dynamics decreases.

If short enough, other dynamical time scales, linked to inertial and memory effects in the surrounding fluid, can act to filter high-frequency fluctuations, thereby limiting the ratcheting achievable through feedback that is based solely on the most recent measurement. To capture inertial effects, the term $m\ddot{x}$ should be included in Eq. **1**, which introduces the velocity relaxation time scale $\tau_v = m/\gamma$. To capture memory effects, the viscous friction term $\gamma\dot{x}$ generalizes to a convolution with a kernel that captures the effects of fluid rearrangements in response to bead motion. This introduces a time scale $\tau_f = r^2/\nu$, the time it takes the fluid to diffuse one particle radius r, where ν is the kinematic viscosity. The combined effects of inertia and hydrodynamic memory are captured by the Basset-Boussinesq–Oseen equation (41).

We have made informal numerical studies of these two effects. On the one hand, we find that our proposed feedback algorithm (Eq. **4**) leads to worse performance than implied by estimates based on the overdamped limit. The performance begins to degrade at trap dynamics time scales ≈ 3 μs, obtained by equating the overdamped relaxation time to the fluid memory time scale $\tau_r = \tau_f$. This regime is achieved by the 5-μm bead at a trap stiffness of $\kappa \approx 6{,}000$ pN/μm, which is about an order of magnitude greater than our current setup is capable of; however, deviations are empirically already seen for $\kappa \approx 200$ pN/μm (*SI Appendix*, Fig. S6). Nevertheless, "naively" extrapolating the overdamped theory to this time scale implies work extraction of $\approx 10^4$ $k_B T$/s (for a 5-μm bead) and speeds of $\approx 3{,}000$ μm/s (for a 0.5-μm bead). See *SI Appendix*, section L.

On the other hand, these new physical effects are characterized by new dynamical variables that can be used to further optimize the feedback algorithm. When inertial effects are important, measuring the velocity $\dot{x}$ can improve feedback; likewise, when the hydrodynamic memory kernel is important, the history of positions can help. In principle, one could modify the feedback rule to incorporate the recent history rather than just the most recent measurement. Operating an information engine in a gas of reduced pressure (42) would make these scales more accessible experimentally, and it would be interesting to explore whether improved algorithms can capture some of the performance that would otherwise be lost in these regimes.

PHYSICS

Although our focus here has been on maximizing either the rate of extracted work or the directed velocity, the energetic costs of running the engine are also of interest. For our macroscopic experimental apparatus, the operating costs dwarf the engine output; however, several studies have shown that the fundamental, unavoidable information costs in similar contexts can be much lower (43–47). In the limits of infinite sampling frequency and signal-to-noise ratio (SNR), different ways of estimating costs (24–26, 48) all lead to vanishing efficiency when applied to our setup. We have not here specified the details of our feedback controller, but one reasonable choice is to allow each new measurement to overwrite the previous one. For such an architecture, the appropriate measure of minimum costs is the information flow (46, 47). For our "typical" experiment with amplitude $\mathrm{SNR} \approx 10$ and scaled sampling frequency $f_s \approx 180$, this measure implies an input–output efficiency of about 8% (*SI Appendix*, section N). Even at this relatively low efficiency, the power and velocity are indistinguishable experimentally from those found in the limit of infinite sampling frequency and SNR. In future work, it would be interesting to explore more systematically the trade-off between performance and information-processing costs, using the tools of multiobjective optimization (49).

Finally, our information-engine design exploits only the "up" fluctuations. In Szilard's original proposal, the ability to change the connection between mass and partition as a function of the measurement outcome (the side on which the particle is found) allowed exploitation of all measurement outcomes. However, in our design "down" fluctuations lead to no feedback response. The information gathered in measuring those fluctuations cannot be exploited, reflecting a structural limitation of the engine (50). A design that could convert and store energy from all measurements would further enhance information-engine performance.

Materials and Methods

The experiments were performed using an optical-tweezer setup that can rapidly shift the beam position under feedback control (31, 51). For setup details, see *SI Appendix*, section A. To estimate the power and velocity from empirical data, we record trajectories over a fixed distance of 340 nm, a range set by the quadrant photodiode sensor, which records beam deflections due to bead movement. The measurement noise σ'_m ranged from 0.6 to 4.8 nm (see *SI Appendix*, Table S2); the amplitude SNR ranged from 1 to 60 and was typically about 10. Every time the bead reaches the upper bound, it is returned to the lower bound, and the ratchet protocol is repeated. Each 340-nm trajectory contains $N_{\mathrm{ratch}} \approx 80$ ratchet events when the threshold $X_T = 0$. Typically, the first relaxation time τ_r of the trajectory is not included when estimating power and velocity, to allow the system to reach steady state. The total displacement and time for each trajectory j is recorded, and then the procedure is repeated $N_{\mathrm{traj}} \approx 100$ times. The velocity and power are calculated from the average over the N_{traj} trajectories as $\overline{v} = \sum (x_n - x_{n-1})/t_s$ and $\overline{P} = \delta_g \overline{v}$, where the sum is over time steps within a trajectory and over the N_{traj} multiple trials. Because the total number of ratcheting events is large ($N_{\mathrm{tot}} = N_{\mathrm{ratch}} \times N_{\mathrm{traj}} \approx 8 \times 10^3$) and each first-passage time τ_{FP} is an independent random variable, we can aggregate the first-passage times from all N_{tot} events. As $N_{\mathrm{tot}} \gg 1$, the law of large numbers can be used to estimate the mean velocity and power. The approximation becomes exact when $N_{\mathrm{ratch}}, N_{\mathrm{traj}} \to \infty$ (*SI Appendix*, section J).

Sample Preparation. Four sizes of silica bead were used, with nominal diameters specified by the manufacturer of 1.49 ± 0.22 μm (Bangs Laboratories), 0.50 ± 0.05 μm, 3.00 ± 0.25 μm, and 5.00 ± 0.35 μm (Sigma-Aldrich). The properties (diffusion constant and force constant) associated with each bead were measured individually before each set of experiments done with the particular bead. The sphere solution from the manufacturer was diluted using deionized water. The sample chamber was prepared from a glass slide and a coverslip, which were separated by 100-μm spacer wires and sealed by nail polish. For the 0.5-μm-bead experiment, the sample chamber consisted of two coverslips, separated by 50-μm spacer wires.

Data Availability. Code and experimental data have been deposited in Zenodo (DOI: 10.5281/zenodo.4697882) (52).

ACKNOWLEDGMENTS. We thank Avinash Kumar and Luis Reinalter (Simon Fraser University Department of Physics) for contributions to the experimental setup and Susanne Still (University of Hawaii) for fruitful discussions. This research was supported by grant FQXi-IAF19-02 from the Foundational Questions Institute Fund, a donor-advised fund of the Silicon Valley Community Foundation. Additional support was from Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grants (D.A.S. and J.B.), a Tier-II Canada Research Chair (D.A.S.), an NSERC Undergraduate Summer Research Award, a BC Graduate Scholarship, and an NSERC Canadian Graduate Scholarship–Masters (J.N.E.L.). Computational support was provided by WestGrid and Compute Canada Calcul Canada.

1. C. G. Knott, *Life and Scientific Work of Peter Guthrie Tait* (Cambridge University Press, London, 1911).
2. L. Szilard, Über die Entropieverminderung in einem thermodynamischen System bei Eingriffen intelligenter Wesen. *Z. Phys.* **53**, 840–856 (1929).
3. L. Szilard, "On the decrease in entropy in a thermodynamic system by the intervention of intelligent beings" in *Maxwell's Demon 2*, H. S. Leff, A. F. Rex, Eds. (IOP Publishing, 2003), pp. 110–119.
4. R. Landauer, Irreversibility and heat generation in the computing process. *IBM J. Res. Dev.* **5**, 183–191 (1961).
5. C. H. Bennett, The thermodynamics of computation—A review. *Int. J. Theor. Phys.* **21**, 905–940 (1982).
6. D. Mandal, C. Jarzynski, Work and information processing in a solvable model of Maxwell's demon. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 11641–11645 (2012).
7. R. K. Schmitt, J. M. R. Parrondo, H. Linke, J. Johansson, Molecular motor efficiency is maximized in the presence of both power-stroke and rectification through feedback. *New J. Phys.* **17**, 065011 (2015).
8. J. M. R. Parrondo, J. M. Horowitz, T. Sagawa, Thermodynamics of information. *Nat. Phys.* **11**, 131–139 (2015).
9. K. Sekimoto, Kinetic characterization of heat bath and the energetics of thermal ratchet models. *J. Phys. Soc. Jpn.* **66**, 1234–1237 (1997).
10. K. Sekimoto, *Stochastic Energetics* (Springer, 2010), vol. 799.
11. U. Seifert, Stochastic thermodynamics, fluctuation theorems and molecular machines. *Rep. Prog. Phys.* **75**, 126001 (2012).
12. C. Van den Broeck, M. Esposito, Ensemble and trajectory thermodynamics: A brief introduction. *Physica A* **418**, 6–16 (2015).
13. S. Toyabe, T. Sagawa, M. Ueda, E. Muneyuki, M. Sano, Experimental demonstration of information-to-energy conversion and validation of the generalized Jarzynski equality. *Nat. Phys.* **6**, 988–992 (2010).
14. P. A. Camati *et al.*, Experimental rectification of entropy production by Maxwell's demon in a quantum system. *Phys. Rev. Lett.* **117**, 240502 (2016).
15. J. V. Koski, A. Kutvonen, I. M. Khaymovich, T. Ala-Nissila, J. P. Pekola, On-chip Maxwell's demon as an information-powered refrigerator. *Phys. Rev. Lett.* **115**, 260602 (2015).
16. N. Cottet *et al.*, Observing a quantum Maxwell demon at work. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 7561–7564 (2017).
17. Y. Masuyama *et al.*, Information-to-work conversion by Maxwell's demon in a superconducting circuit quantum electrodynamical system. *Nat. Commun.* **9**, 1291 (2018).
18. J. V. Koski, V. F. Maisi, J. P. Pekola, D. V. Averin, Experimental realization of a Szilard engine with a single electron. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 13786–13789 (2014).
19. K. Chida, S. Desai, K. Nishiguchi, A. Fujiwara, Power generator driven by Maxwell's demon. *Nat. Commun.* **8**, 15310 (2017).
20. A. Bérut *et al.*, Experimental verification of Landauer's principle linking information and thermodynamics. *Nature* **483**, 187–189 (2012).
21. Y. Jun, M. Gavrilov, J. Bechhoefer, High-precision test of Landauer's principle in a feedback trap. *Phys. Rev. Lett.* **113**, 190601 (2014).
22. J. V. Koski, V. F. Maisi, T. Sagawa, J. P. Pekola, Experimental observation of the role of mutual information in the nonequilibrium dynamics of a Maxwell demon. *Phys. Rev. Lett.* **113**, 030601 (2014).
23. J. Hong, B. Lambson, D. Scott, J. Bokor, Experimental test of Landauer's principle in single-bit operations on nanomagnetic memory bits. *Sci. Adv.* **2**, e1501492 (2016).
24. M. Ribezzi-Crivellari, F. Ritort, Large work extraction and the Landauer limit in a continuous Maxwell demon. *Nat. Phys.* **15**, 660–664 (2019).
25. T. Admon, S. Rahav, Y. Roichman, Experimental realization of an information machine with tunable temporal correlations. *Phys. Rev. Lett.* **121**, 180601 (2018).
26. G. Paneru *et al.*, Optimal tuning of a Brownian information engine operating in a nonequilibrium steady state. *Phys. Rev. E* **98**, 052119 (2018).
27. G. Paneru, S. Dutta, T. Sagawa, T. Tlusty, H. Kyu Pak, Efficiency fluctuations and noise induced refrigerator-to-heater transition in information engines. *Nat. Commun.* **11**, 1012 (2020).
28. G. Paneru, D. Yun Lee, T. Tlusty, H. Kyu Pak, Lossless Brownian information engine. *Phys. Rev. Lett.* **120**, 020601 (2018).
29. H. C. Berg, *E. Coli in Motion* (Springer-Verlag, New York, 2004).

30. P. E. Kloeden, E. Platen, *Numerical Solution of Stochastic Differential Equations: Stochastic Modelling and Applied Probability* (Springer, Berlin, 2013).
31. A. Kumar, J. Bechhoefer, Nanoscale virtual potentials using optical tweezers. *Appl. Phys. Lett.* **113**, 183702 (2018).
32. P. Hänggi, P. Talkner, Reaction-rate theory: Fifty years after Kramers. *Rev. Mod. Phys.* **62**, 251–341 (1990).
33. M. Chupeau, J. Gladrow, A. Chepelianskii, U. F. Keyser, E. Trizac, Optimizing Brownian escape rates by potential shaping. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 138–1388 (2020).
34. J.-M. Park, J. S. Lee, J. D. Noh, Optimal tuning of a confined Brownian information engine. *Phys. Rev. E* **93**, 032146 (2016).
35. J. H. Marden, L. R. Allen, Molecules, muscles, and machines: Universal performance characteristics of motors. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 4161–4166 (2002).
36. H. Hess, G. Saper, Engineering with biomolecular motors. *Acc. Chem. Res.* **51**, 3015–3022 (2018).
37. D. Y. Lee, J. Um, G. Paneru, H. K. Pak, An experimentally-achieved information-driven Brownian motor shows maximum power at the relaxation time. *Sci. Rep.* **8**, 12121 (2018).
38. G. M. Barbara, J. G. Mitchell, Bacterial tracking of motile algae. *FEMS Microbiol. Ecol.* **44**, 79–87 (2003).
39. T. Ariga, M. Tomishige, D. Mizuno, Nonequilibrium energetics of molecular motor kinesin. *Phys. Rev. Lett.* **121**, 218101 (2018).
40. A. Jannasch *et al.*, Nanonewton optical force trap employing anti-reflection coated, high-refractive-index titania microspheres. *Nat. Photonics* **6**, 469–473 (2012).
41. S. L. Seyler, S. Pressé, Long-time persistence of hydrodynamic memory boosts microparticle transport. *Phys. Rev. Res.* **1**, 032003(R) (2019).
42. F. Tebbenjohanns, M. Frimmer, V. Jain, D. Windey, L. Novotny, Motional sideband asymmetry of a nanoparticle optically levitated in free space. *Phys. Rev. Lett.* **124**, 013603 (2020).
43. J. M. Horowitz, S. Vaikuntanathan, Nonequilibrium detailed fluctuation theorem for repeated discrete feedback. *Phys. Rev. E* **82**, 061120 (2010).
44. T. Sagawa, M. Ueda, Nonequilibrium thermodynamics of feedback control. *Phys. Rev. E* **85**, 021104 (2012).
45. S. Still, D. A. Sivak, A. J. Bell, G. E. Crooks, Thermodynamics of prediction. *Phys. Rev. Lett.* **109**, 120604, (2012).
46. J. M. Horowitz, M. Esposito, Thermodynamics with continuous information flow. *Phys. Rev. X* **4**, 031015 (2014).
47. J. M. Horowitz, H. Sandberg, Second-law-like inequalities with information and their interpretations. *New J. Phys.* **16**, 125007 (2014).
48. G. Paneru, H. K. Pak, Colloidal engines for innovative tests of information thermodynamics. *Adv. Phys. X* **5**, 1823880 (2020).
49. A. P. Solon, J. M. Horowitz, Phase transition in protocols minimizing work fluctuations. *Phys. Rev. Lett.* **120**, 180605 (2018).
50. S. Still, Thermodynamic cost and benefit of memory. *Phys. Rev. Lett.* **124**, 050601 (2020).
51. J. A. C. Albay, G. Paneru, H. K. Pak, Y. Jun, Optical tweezers as a mathematically driven spatio-temporal potential generator. *Opt. Express* **26**, 334542 (2018).
52. J. Lucero, T. Saha, J. Ehrich, tk5/maximum_power_info_ratchet: InfoRatchet2021. Zenodo. https://doi.org/10.5281/zenodo.4697882. Deposited 30 March 2021.

PNAS
www.pnas.org

Supplementary Information for

Maximizing power and velocity of an information engine

Tushar K. Saha, Joseph N. E. Lucero, Jannik Ehrich, David A. Sivak[1], and John Bechhoefer[1]

[1]To whom correspondence should be addressed. E-mail: johnb@sfu.ca, dsivak@sfu.ca

This PDF file includes:

Supporting Information Text

A. Experimental apparatus. Figure S1 shows the schematic of the experimental apparatus, whose basic configuration has been described previously (1). The optical trap is made using a 532-nm, green, ring-cavity laser (HÜBNER Photonics, Cobolt Samba, 1.5 W, 532 nm). Two acousto-optic deflectors (AODs, DTSXY-250-532, AA Opto Electronic) can deflect the beam in two directions, to shift the trap's position along both x- and y-axes. The laser propagates along the z-axis; the x-axis is parallel to the gravitational force; and the y-axis is perpendicular to both gravity and the laser beam. The AOD's plane is imaged onto the back aperture of the trapping objective (MO1, water-immersion 60x Olympus objective 1.2 NA) using a two-lens relay system. The two lenses also magnify the green laser to overfill the back aperture of the trapping objective. The green laser is focused inside the sample chamber (SC), which is filled with a solution of micron-sized beads. Objective MO1 creates the optical trap. The AODs also change the power of the trapping laser, to control trap stiffness.

A 660-nm fiber pig-tailed red laser (Thorlabs, LP660-SF20, 20 mW, 660 nm) was used for detection. The diameter of the laser beam was kept smaller than the back aperture of the second objective (MO2, 40x Nikon). The reduced beam size leads to a larger focus spot size in the trapping plane, which increases the linear range of the position measurements. The detection plane was adjusted using the relay lens (L) system between the red laser and MO2. The detection beam is collected by MO1 and focused on the quadrant photo-diode (QPD, First Sensor, QP50-6-18u-SD2). The detection laser is reflected on the QPD using a red dichroic filter (F2). This signal is used to detect the position of the trapped bead.

A blue LED was introduced in the laser's path using a high-pass filter (F1), in order to visualize the trapped bead. The collected blue light is focused onto a USB-3 digital camera (BASLER, ace acA800-510µm) to image the trapped bead. Finally, the green laser that is collected by MO2 is reflected by the green dichroic filter (F3) and then directed to a beam dump (BD).

The voltage from the QPD is sent to the Analog/Digital input of the Field Programmable Gate Arrays (FPGA, National Instruments, NI PCIe-7857), which converts it to a discretized position signal. The voltage-to-position conversion factor is obtained by calibrating the QPD-AOD-CAM system (details can be found in (2)). The FPGA then makes the feedback decision for the trap reset based on the feedback rule given in the main text. The feedback loop time was 20 µs. In addition to the loop time, the *feedback latency* (time between acquisition of a measurement and subsequent shift of the trap) is also important (3). Here, it is always set equal to the sampling time. The feedback latency and scaled effective mass $\delta_{\rm g}$ both affect the value of feedback gain α needed to meet the zero-work condition. We typically find α in the range 1.3–1.9.

Our theoretical model implies that the directed velocity of the information engine can be increased by reducing the bead size. To test this prediction, we performed experiments with 0.5-µm beads, for which the apparatus had to be modified. A 1.2 NA water-immersion objective was used to focus a higher-power detection laser (HÜBNER Photonics, Cobolt 06-MLD, 50 mW) to increase the signal from the bead. The maximum velocity and power were obtained for a trap stiffness $\kappa \approx 200$ pN/µm and a (bead-dependent) dynamical time scale as short as $\tau_{\rm r} \approx 23$ µs, achieved with the maximum trapping laser power of 0.4 W at the trapping plane.

B. Zero-work condition. The work done by the trap is defined, in scaled units, as

$$W_{n+1} = \frac{1}{2}\left[(x_{n+1}-\lambda_{n+1})^2 - (x_{n+1}-\lambda_n)^2\right]. \qquad \text{[S1]}$$

As the trap position is updated after a delay of one time step, the trap update for time step $n+1$ is based on the bead and trap position at time step n. Under the assumption that the bead is stationary during the trap update, the trap update that sets the work to zero is given by

$$\begin{aligned} 0 &= \frac{1}{2}\left[(x_n-\lambda_{n+1})^2 - (x_n-\lambda_n)^2\right] \\ (x_n-\lambda_{n+1}) &= \pm\,(x_n-\lambda_n) \\ \lambda_{n+1} &= \lambda_n + 2\,(x_n-\lambda_n)\,, \end{aligned} \qquad \text{[S2]}$$

which corresponds to a feedback gain $\alpha = 2$. Figure 2 shows that, experimentally, the zero-work condition is observed for $\alpha \approx 1.5$. This small discrepancy can be explained by time delays in the experiment: There is a one-sample lag ($t_{\rm s} = 20$ µs) in the experiment that is not accounted for in the simple model, which reduces the value of α required for zero work.

C. Mean first-passage time. The particle's position propagator $p(x,t|x_0,0)$ obeys the (forward) Fokker-Planck equation:

$$\frac{\partial}{\partial t}p(x,t|x_0,0) = \frac{\partial}{\partial x}\left(p(x,t|x_0,0)\frac{\partial}{\partial x}V(x)\right) + \frac{\partial^2}{\partial x^2}p(x,t|x_0,0)\ , \qquad \text{[S3]}$$

for total potential $V(x) \equiv V_{\rm t}(x) + \delta_{\rm g}\,x$ consisting of the sum of the trap potential $V_{\rm t}(x) = \frac{1}{2}(x-\lambda)^2$ and the gravitational potential, as shown in Fig. S2.

In general, the mean first-passage time $\tau_{\rm MFP}(x;b)$ gives the mean time until a particle started at $x < b$ first reaches b. It is given by (4):

$$\tau_{\rm MFP}(x;b) = \int_x^b {\rm d}y\ \ {\rm e}^{V(y)} \int_{-\infty}^{y} {\rm d}z\ \ {\rm e}^{-V(z)}\ . \qquad \text{[S4]}$$

In our case, we are interested in the MFPT $\tau_{\rm MFP}(X_{\rm T})$ from $\lambda - X_{\rm T}$ to $\lambda + X_{\rm T}$. Substituting $x = \lambda - X_{\rm T}$ and $b = \lambda + X_{\rm T}$ into Eq. (S4) gives

$$\tau_{\rm MFP}(X_{\rm T}) = \int_{\lambda - X_{\rm T}}^{\lambda + X_{\rm T}} {\rm d}y \, \exp\left\{V_{\rm t}(y-\lambda) + \delta_{\rm g} y\right\} \int_{-\infty}^{y} {\rm d}z \, \exp\left\{-V_{\rm t}(z-\lambda) - \delta_{\rm g} z\right\} \tag{S5}$$

$$= \int_{\lambda - X_{\rm T}}^{\lambda + X_{\rm T}} {\rm d}y \, \exp\left\{V_{\rm t}(y-\lambda) + \delta_{\rm g}(y-\lambda)\right\} \int_{-\infty}^{y} {\rm d}z \, \exp\left\{-V_{\rm t}(z-\lambda) - \delta_{\rm g}(z-\lambda)\right\} \tag{S6}$$

$$= \int_{-X_{\rm T}}^{X_{\rm T}} {\rm d}x' \, \exp\left\{V_{\rm t}\left(x'\right) + \delta_{\rm g} x'\right\} \int_{-\infty}^{x'} {\rm d}x'' \, \exp\left\{-V_{\rm t}\left(x''\right) - \delta_{\rm g} x''\right\} , \tag{S7}$$

where in the last line we substituted $x'' \equiv z - \lambda$ and $x' \equiv y - \lambda$.

D. Low sampling-frequency limit. When the (scaled) sampling frequency $f_{\rm s} = \tau_{\rm r}/t_{\rm s}$ is small, the position distribution equilibrates between successive feedback steps. It is then given by the Boltzmann distribution (in scaled units)

$$p_{\rm eq}(x;\lambda) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\tfrac{1}{2}(x - \lambda + \delta_{\rm g})^2\right\} . \tag{S8}$$

In that limit, the average work $W_{\rm eq}$ extracted per feedback step is determined from the feedback rules given in Eq. 4 in the main text:

$$W_{\rm eq} = \delta_{\rm g} \Bigg[\underbrace{0}_{\text{positions left of threshold}} + \underbrace{\int_{\lambda}^{\infty} {\rm d}x \, 2(x-\lambda) p_{\rm eq}(x;\lambda)}_{\text{positions right of threshold}} \Bigg] \tag{S9}$$

$$= \sqrt{\frac{2}{\pi}} \, \delta_{\rm g} \int_{\delta_{\rm g}}^{\infty} {\rm d}x' \, (x' - \delta_{\rm g}) \, e^{-(x')^2/2} \tag{S10}$$

$$= \delta_{\rm g} \left\{ \sqrt{\frac{2}{\pi}} \, e^{-\delta_{\rm g}^2/2} + \delta_{\rm g} \left[{\rm erf}\left(\frac{\delta_{\rm g}}{\sqrt{2}}\right) - 1 \right] \right\} , \tag{S11}$$

where in the second line we substituted $x' \equiv x - \lambda + \delta_{\rm g}$.

Therefore, in the limit of low sampling frequency, the power in scaled units is

$$P = f_{\rm s} \, W_{\rm eq} \tag{S12}$$

$$= f_{\rm s} \left\{ \sqrt{\frac{2}{\pi}} \, \delta_{\rm g} \, e^{-\delta_{\rm g}^2/2} + \delta_{\rm g}^2 \left[{\rm erf}\left(\frac{\delta_{\rm g}}{\sqrt{2}}\right) - 1 \right] \right\} , \tag{S13}$$

or $P \approx 0.19 f_{\rm s}$ for $\delta_{\rm g} = 0.8$, the value used in Fig. 3*A*.

E. Arbitrary sampling frequency. To calculate velocity and power for arbitrary sampling frequencies, we first derive a self-consistency equation for the steady-state position distribution as a function of sampling frequency, which we evaluate numerically. We begin by noting the causal structure given by Fig. S3A. From this causal structure we derive the following decomposition for the joint transition probability to go from bead position x_n and trap center λ_n at sampling time t_n to x_{n+1} and λ_{n+1} at sampling time t_{n+1}:

$$p\left(x_{n+1}, \lambda_{n+1} | x_n, \lambda_n\right) = p_x\left(x_{n+1} | x_n, \lambda_n\right) p_\lambda\left(\lambda_{n+1} | x_{n+1}, \lambda_n\right) . \tag{S14}$$

The bead-position propagator $p_x\left(x_{n+1} | x_n, \lambda_n\right)$, in scaled units, is given by the infinitesimal generator for an Ornstein-Uhlenbeck process (5, Section 5.3),

$$p_x\left(x_{n+1} | x_n, \lambda_n\right) = \mathcal{N}\left(x_{n+1};\; x_n \, {\rm e}^{-t_{\rm s}} + \left(1 - {\rm e}^{-t_{\rm s}}\right)(\lambda_n - \delta_{\rm g}), 1 - {\rm e}^{-2t_{\rm s}}\right) , \tag{S15}$$

where $\mathcal{N}(x;\mu,\sigma^2)$ denotes a normal (Gaussian) distribution over x with mean μ and variance σ^2. The propagator $p_\lambda\left(\lambda_{n+1} | x_{n+1}, \lambda_n\right)$ for the trap center is given by

$$p_\lambda\left(\lambda_{n+1} | x_{n+1}, \lambda_n\right) = \delta\left(\lambda_{n+1} - \lambda_n\right) \Theta[-(x_{n+1} - \lambda_n)] + \delta[\lambda_{n+1} - (2x_{n+1} - \lambda_n)] \, \Theta\left(x_{n+1} - \lambda_n\right) . \tag{S16}$$

Here, $\delta(\cdot)$ denotes the Dirac-delta function, and $\Theta(\cdot)$ denotes the Heaviside function. The first term on the right-hand side reflects fluctuations that are not sufficient to trigger a ratcheting, while the second term corresponds to ratcheting events.

A change of variables to relative coordinates

$$x^{\rm r}_{n^+} \equiv x_{n+1} - \lambda_n \qquad \text{and} \qquad x^{\rm r}_n \equiv x_n - \lambda_n \tag{S17}$$

yields

$$p\left(x^{\mathrm{r}}_{n+}|x_n,\lambda_n\right) = \int \mathrm{d}x_{n+1}\ \delta[x_{n+1}-(x^{\mathrm{r}}_{n+}+\lambda_n)]\, p_x\left(x_{n+1}|x_n,\lambda_n\right) \tag{S18a}$$

$$= \mathcal{N}\left(x^{\mathrm{r}}_{n+};\ x^{\mathrm{r}}_n\,\mathrm{e}^{-t_{\mathrm{s}}} - \left(1-\mathrm{e}^{-t_{\mathrm{s}}}\right)\delta_{\mathrm{g}}, 1-\mathrm{e}^{-2t_{\mathrm{s}}}\right) \tag{S18b}$$

$$\equiv p_1\left(x^{\mathrm{r}}_{n+}|x^{\mathrm{r}}_n\right)\ . \tag{S18c}$$

Similarly, using $x^{\mathrm{r}}_{n+1} \equiv x_{n+1} - \lambda_{n+1}$, we obtain

$$p\left(x^{\mathrm{r}}_{n+1}|x_{n+1},\lambda_n\right) = \int \mathrm{d}\lambda_{n+1}\ \delta[x^{\mathrm{r}}_{n+1}-(x_{n+1}-\lambda_{n+1})]\, p_\lambda\left(\lambda_{n+1}|x_{n+1},\lambda_n\right) \tag{S19a}$$

$$= \delta\left(x^{\mathrm{r}}_{n+1}-x^{\mathrm{r}}_{n+}\right)\Theta\left(-x^{\mathrm{r}}_{n+}\right) + \delta\left(x^{\mathrm{r}}_{n+1}+x^{\mathrm{r}}_{n+}\right)\Theta\left(x^{\mathrm{r}}_{n+}\right) \tag{S19b}$$

$$\equiv p_2\left(x^{\mathrm{r}}_{n+1}|x^{\mathrm{r}}_{n+}\right)\ . \tag{S19c}$$

Importantly, this change of variables simplifies the causal structure, which is shown in Fig. S3B.

The steady-state solutions $\pi_+\left(x^{\mathrm{r}}_{n+}\right)$ and $\pi\left(x^{\mathrm{r}}_n\right)$ for these two variables x^{r}_{n+} and x^{r}_n, respectively, are given by the self-consistent integral equations

$$\pi_+(x^{\mathrm{r}}_{n+}) = \int \mathrm{d}u \underbrace{\left[\int \mathrm{d}v\, p_1\left(x^{\mathrm{r}}_{n+}|v\right) p_2\left(v|u\right)\right]}_{\equiv T\left(x^{\mathrm{r}}_{n+}|u\right)} \pi_+(u). \tag{S20}$$

$$\pi(x^{\mathrm{r}}_n) = \int \mathrm{d}v \underbrace{\left[\int \mathrm{d}u\, p_2\left(x^{\mathrm{r}}_n|u\right) p_1\left(u|v\right)\right]}_{\equiv \tilde{T}(x^{\mathrm{r}}_n|v)} \pi(v), \tag{S21}$$

where u and v are dummy variables of integration.

The propagator $T\left(x^{\mathrm{r}}_{n+}|u\right)$ is given by

$$T\left(x^{\mathrm{r}}_{n+}|u\right) = \Theta(-u)\mathcal{N}\left(x^{\mathrm{r}}_n; (u+\delta_{\mathrm{g}})\,\mathrm{e}^{-t_{\mathrm{s}}} - \delta_{\mathrm{g}}, 1-\mathrm{e}^{-2t_{\mathrm{s}}}\right) + \Theta\left(u\right)\mathcal{N}\left(x^{\mathrm{r}}_n; -\delta_{\mathrm{g}} - (u-\delta_{\mathrm{g}})\,\mathrm{e}^{-t_{\mathrm{s}}}, 1-\mathrm{e}^{-2t_{\mathrm{s}}}\right)\ . \tag{S22}$$

Similarly, the propagator $\tilde{T}\left(x^{\mathrm{r}}_n|v\right)$ is given by

$$\tilde{T}\left(x^{\mathrm{r}}_n|v\right) = \Theta\left(-x^{\mathrm{r}}_n\right)\left\{\mathcal{N}\left(x^{\mathrm{r}}_n; \delta_{\mathrm{g}} - (v+\delta_{\mathrm{g}})\,\mathrm{e}^{-t_{\mathrm{s}}}, 1-\mathrm{e}^{-2t_{\mathrm{s}}}\right) + \mathcal{N}\left(x^{\mathrm{r}}_n; (v+\delta_{\mathrm{g}})\,\mathrm{e}^{-t_{\mathrm{s}}} - \delta_{\mathrm{g}}, 1-\mathrm{e}^{-2t_{\mathrm{s}}}\right)\right\}\ . \tag{S23}$$

We numerically solve for the steady-state distributions, Eqs. (S20) and (S21), by discretizing the propagator Eq. (S22) in x^{r}_{n+} and u and the propagator Eq. (S23) in x^{r}_n and v. The associated eigenvectors of $T\left(x^{\mathrm{r}}_{n+}|u\right)$ and $\tilde{T}\left(x^{\mathrm{r}}_n|v\right)$ with eigenvalue 1 give the distributions $\pi_+\left(x^{\mathrm{r}}_{n+}\right)$ and $\pi\left(x^{\mathrm{r}}_n\right)$, respectively. Here, we use 2000 uniformly spaced grid points in the domain $x^{\mathrm{r}}_n, x^{\mathrm{r}}_{n+}, u, v \in [-20, 20]$.

The steady-state output power P in a given cycle can be recast in terms of the new variables x^{r}_n and x^{r}_{n+} as

$$P = \langle \Delta W_{\mathrm{g}} \rangle\, f_{\mathrm{s}} = \delta_{\mathrm{g}}\left(\langle x^{\mathrm{r}}_{n+}\rangle - \langle x^{\mathrm{r}}_n\rangle\right) f_{\mathrm{s}}\ , \tag{S24}$$

where $\langle \Delta W_{\mathrm{g}} \rangle$ denotes the work done per cycle, and the averages are taken over the appropriate steady-state distributions for each of the variables. Figure 3A of the main text compares numerical results using this approach and the experiment.

F. Large-stiffness asymptotics for power and velocity. As shown in the main text (Eq. 12a,b), power and velocity are given (in physical units) by

$$P' = \frac{k_{\mathrm{B}}T}{\tau_r}\sqrt{\frac{2}{\pi}}\,\delta_{\mathrm{g}} e^{-\delta_{\mathrm{g}}^2/2}\left[1+\mathrm{erf}\left(\frac{\delta_{\mathrm{g}}}{\sqrt{2}}\right)\right]^{-1} \tag{S25}$$

$$v' = \frac{\sigma}{\tau_r}\sqrt{\frac{2}{\pi}}\,e^{-\delta_{\mathrm{g}}^2/2}\left[1+\mathrm{erf}\left(\frac{\delta_{\mathrm{g}}}{\sqrt{2}}\right)\right]^{-1}\ , \tag{S26}$$

where $\tau_{\mathrm{r}} = \gamma/\kappa$, $\delta_{\mathrm{g}} = mg/\kappa\sigma$, and $\sigma = \sqrt{k_{\mathrm{B}}T/\kappa}$.

An expansion for large trap stiffness κ leads to the following asymptotic relations:

$$P' \sim \sqrt{\frac{2k_{\mathrm{B}}T}{\pi}}\frac{mg\sqrt{\kappa}}{\gamma}\ , \tag{S27}$$

$$v' \sim \sqrt{\frac{2k_{\mathrm{B}}T}{\pi}}\frac{\sqrt{\kappa}}{\gamma}\ . \tag{S28}$$

G. Horizontal vs. vertical velocity. In the horizontal direction, there is no gravitational effect. The bead fluctuates subject solely to the harmonic potential of the trap. Setting $\delta_g = 0$ in Eq. (S26) gives the velocity

$$v' = \sqrt{\frac{2k_BT}{\pi}}\frac{\sqrt{\kappa}}{\gamma}\,. \tag{S29}$$

Figure S4 shows the velocity along the horizontal and vertical directions as a function of trap stiffness. For the 1.5-µm bead, the horizontal and vertical velocities are indistinguishable. The system is in the high-κ limit, so the vertical velocity is well approximated by Eq. S28. For larger beads, the vertical velocity asymptotes to the horizontal velocity as the stiffness increases.

H. Trap stiffness. The trap stiffness κ is measured by fitting the power spectrum of the position data to an *aliased Lorentzian* (6) with Gaussian noise floor. We report the standard deviation of the power, integrated over the measured bandwidth. The fit parameters, the diffusion constant D and corner frequency f_c, are used to evaluate the trap stiffness. The quantities are related by

$$\kappa = \frac{2\pi f_c\, k_BT}{D}. \tag{S30}$$

Figure S5 shows the power spectrum and the fit.

I. Optimal bead size that maximizes power. Figure 4*C* shows that there is an optimal δ_g that maximizes the power. The optimization can be understood in physical units as well, where we change the bead size subject to fixed trap stiffness and bead density. To study the size dependence, Eq. S25 can be written as

$$P' = \frac{\sqrt{\kappa k_BT}}{6\eta}\sqrt{\frac{2}{\pi}}\frac{4r^2\Delta\rho g}{3}\exp\left\{-\frac{1}{2}\left(C\frac{r^3}{\sqrt{\kappa}}\right)^2\right\}\left[1+\mathrm{erf}\left(\frac{1}{\sqrt{2}}\left(C\frac{r^3}{\sqrt{\kappa}}\right)\right)\right]^{-1}\,, \tag{S31}$$

where $C = \frac{4\pi\Delta\rho}{3}\frac{g}{\sqrt{k_BT}}$, $\Delta\rho$ is the effective density of the bead in water, η the viscosity of water, and we have used the relations $\tau_r = \gamma/\kappa$, $\gamma = 6\pi\eta r$ and $\sigma = \sqrt{k_BT/\kappa}$. Figure S6 shows the extracted power as a function of bead diameter. For $\kappa = 30$ pN/µm, the optimal bead diameter is ≈ 3.5 µm; when the trap stiffness is increased, the optimum shifts towards larger beads.

J. Velocity estimator. We calculate the velocity by averaging over N_{traj} trajectories indexed by j. Each trajectory has a different duration; consequently, the number $N(j)$ of position measurements varies for each trajectory:

$$v = \frac{\sum_{j=1}^{N_{\mathrm{traj}}}\sum_{n=1}^{N(j)}(x_n - x_{n-1})}{\sum_{j=1}^{N_{\mathrm{traj}}}\sum_{n=1}^{N(j)}t_s}\,, \tag{S32}$$

where t_s is the sampling time, and x_n denotes the particle position at sampling time t_n.

In the following, we derive Eq. 7 from this estimator in the limit of long trajectories. For $N_{\mathrm{traj}} \gg 1$, the law of large numbers permits us to write

$$\langle v\rangle = \frac{N_{\mathrm{traj}}\left\langle\sum_{n=1}^{N}(x_n - x_{n-1})\right\rangle_{p(N)}}{N_{\mathrm{traj}}\left\langle\sum_{n=1}^{N}t_s\right\rangle_{p(N)}}\,, \tag{S33}$$

where we average over the random number N of time samples in a trajectory.

We rewrite this expression in terms of the average over the number N_{ratch} of ratchet events during a trajectory,

$$\langle v\rangle = \frac{\left\langle\sum_{m=1}^{N_{\mathrm{ratch}}}\Delta x_m\right\rangle_{p(N_{\mathrm{ratch}})}}{\left\langle\sum_{m=1}^{N_{\mathrm{ratch}}}\tau_{\mathrm{FP}}\right\rangle_{p(N_{\mathrm{ratch}})}}\,, \tag{S34}$$

where Δx_m denotes the bead displacement of ratchet event m, and τ_{FP} is the time interval between successive ratchet events (a first-passage time).

The system quickly reaches a steady state, which implies that all ratchet events are independent and identically distributed events. Steady state is ensured by neglecting the first relaxation time of each trajectory. Then, with $N_{\mathrm{ratch}} \gg 1$, the central limit theorem implies

$$\langle v\rangle = \frac{\langle N_{\mathrm{ratch}}\langle\Delta x\rangle\rangle_{p(N_{\mathrm{ratch}})}}{\langle N_{\mathrm{ratch}}\langle\tau\rangle\rangle_{p(N_{\mathrm{ratch}})}}\,, \tag{S35}$$

where the inner averages $\langle\cdot\rangle$ are now taken over a single ratchet event.

Now, $\langle\Delta x\rangle = X_T + X_R$ is an exact relation in the fast-sampling limit, which also implies that the average time to ratchet (i.e., reach X_T from $-X_R$) is given by the mean first-passage time: $\langle\tau\rangle \to \tau_{\mathrm{MFP}}$. Thus, the velocity relation in Eq. S32 leads to Eq. 7:

$$\langle v\rangle = \frac{X_T + X_R}{\tau_{\mathrm{MFP}}}\,. \tag{S36}$$

K. Non-harmonic and asymmetric potentials. Here, we discuss the influence of the shape of the trapping potential on the maximum extracted power and provide support for the claim made in discussion section of the main text that choosing a potential whose shape differs from a quadratic cannot improve the performance of an information engine. To restrict the discussion to physically relevant settings, we assume that the material used to construct the trap has a maximum stiffness. Otherwise, it would be possible to extract work at infinite rate with an infinitely stiff (*wall-like*) potential.

We first investigate whether asymmetric trapping potentials can increase the power above that achieved by symmetric traps.

K.1. Numerical exploration of asymmetric trapping potentials. To study whether an asymmetric trapping potential can increase the power even further, beyond an equivalent symmetric trap, we consider the following *quadratic-to-linear* trap:

$$V_{\mathrm{t}}(x) = \begin{cases} -f_1 x - \frac{f_1^2}{2}, & x < -f_1 \\ \frac{1}{2}x^2, & -f_1 \le x < f_2 \\ f_2 x - \frac{f_2^2}{2}, & f_2 \le x\,, \end{cases} \tag{S37}$$

which describes an asymmetric continuous and differentiable trapping potential that is linear at large displacements with a quadratic minimum (see Fig. S7A). Importantly, this potential can never exceed the harmonic trap (which corresponds to $f_1, f_2 \to \infty$), and its curvature is always less than or equal to that of a harmonic trap with potential $\frac{1}{2}x^2$.

The condition of zero work done by the trap implies

$$X_{\mathrm{R}}(X_{\mathrm{T}}) := \begin{cases} X_{\mathrm{T}}, & X_{\mathrm{T}} < f_2 \\ \sqrt{2f_2 X_{\mathrm{T}} - f_2^2}, & f_2 \le X_{\mathrm{T}} < \frac{f_1^2+f_2^2}{2f_2} \\ \frac{f_1^2-f_2^2+2f_2X_{\mathrm{T}}}{2f_1}, & \frac{f_1^2+f_2^2}{2f_2} \le X_{\mathrm{T}} \end{cases} \tag{S38}$$

when $f_1 \ge f_2$ and

$$X_{\mathrm{R}}(X_{\mathrm{T}}) := \begin{cases} X_{\mathrm{T}}, & X_{\mathrm{T}} < f_1 \\ \frac{f_1^2+X_{\mathrm{T}}^2}{2f_1}, & f_1 \le X_{\mathrm{T}} < \frac{f_2^2}{2} \\ \frac{f_1^2-f_2^2+2f_2X_{\mathrm{T}}}{2f_1}, & \frac{f_2^2}{2} \le X_{\mathrm{T}} \end{cases} \tag{S39}$$

when $f_2 > f_1$.

The mean first-passage time can be computed by numerically evaluating the integral

$$\tau_{\mathrm{MFP}}(X_{\mathrm{T}}) = \int_{-X_{\mathrm{R}}(X_{\mathrm{T}})}^{X_{\mathrm{T}}} \mathrm{d}x\, e^{V_{\mathrm{t}}(x)+\delta_{\mathrm{g}}x} \int_{-\infty}^{x} \mathrm{d}y\, e^{-V_{\mathrm{t}}(y)-\delta_{\mathrm{g}}y}\,. \tag{S40}$$

Finally, the power as a function of the threshold reads

$$P(X_{\mathrm{T}}) = \frac{\delta_g\,[X_{\mathrm{R}}(X_{\mathrm{T}}) + X_{\mathrm{T}}]}{\tau_{\mathrm{MFP}}(X_{\mathrm{T}})}\,. \tag{S41}$$

Figure S7B shows the power as a function of the threshold, for different slopes f_1 and f_2. We see that an asymmetric trapping potential does not outperform the harmonic potential. Furthermore, maximum power is still achieved at vanishing threshold and is determined by the maximum slope of the left side of the potential, which prevents downward fluctuations. This conclusion is expected since, at vanishing threshold, the mean first-passage time is independent of the right side of the potential.

An intuitive lesson from this example is that, perhaps unsurprisingly, one should design a trap to minimize unproductive downward fluctuations.

K.2. For symmetric trapping potentials, power is maximized at vanishing threshold. Since symmetric trapping potentials do as well as asymmetric ones, we investigate whether a vanishing threshold is always optimal. Here, we consider a generic symmetric trapping potential $V_{\mathrm{t}}(x) = V_{\mathrm{t}}(-x)$ with a well-defined minimum $V(0) = 0$ at $x = 0$.

The mean first-passage time is given by [see Eq. (S7)]:

$$\tau_{\mathrm{MFP}}(X_{\mathrm{T}}) = \int_{-X_{\mathrm{T}}}^{X_{\mathrm{T}}} \mathrm{d}x\, e^{V_{\mathrm{t}}(x)+\delta_{\mathrm{g}}x} \int_{-\infty}^{x} \mathrm{d}y\, e^{-V_{\mathrm{t}}(y)-\delta_{\mathrm{g}}y}\,. \tag{S42}$$

We rewrite this expression by splitting the integral first at $x = 0$ and then at $y = 0$,

$$\tau_{\mathrm{MFP}}(X_{\mathrm{T}}) = \int_{-X_{\mathrm{T}}}^{0} \mathrm{d}x\, e^{V_{\mathrm{t}}(x)+\delta_{\mathrm{g}}x} \int_{-\infty}^{x} \mathrm{d}y\, e^{-V_{\mathrm{t}}(y)-\delta_{\mathrm{g}}y} + \int_{0}^{X_{\mathrm{T}}} \mathrm{d}x\, e^{V_{\mathrm{t}}(x)+\delta_{\mathrm{g}}x} \int_{-\infty}^{x} \mathrm{d}y\, e^{-V_{\mathrm{t}}(y)-\delta_{\mathrm{g}}y} \tag{S43}$$

$$= \int_{-X_{\mathrm{T}}}^{0} \mathrm{d}x\, e^{V_{\mathrm{t}}(x)+\delta_{\mathrm{g}}x} \int_{-\infty}^{0} \mathrm{d}y\, e^{-V_{\mathrm{t}}(y)-\delta_{\mathrm{g}}y} - \int_{-X_{\mathrm{T}}}^{0} \mathrm{d}x\, e^{V_{\mathrm{t}}(x)+\delta_{\mathrm{g}}x} \int_{x}^{0} \mathrm{d}y\, e^{-V_{\mathrm{t}}(y)-\delta_{\mathrm{g}}y}$$
$$+ \int_{0}^{X_{\mathrm{T}}} \mathrm{d}x\, e^{V_{\mathrm{t}}(x)+\delta_{\mathrm{g}}x} \int_{-\infty}^{0} \mathrm{d}y\, e^{-V_{\mathrm{t}}(y)-\delta_{\mathrm{g}}y} + \int_{0}^{X_{\mathrm{T}}} \mathrm{d}x\, e^{V_{\mathrm{t}}(x)+\delta_{\mathrm{g}}x} \int_{0}^{x} \mathrm{d}y\, e^{-V_{\mathrm{t}}(y)-\delta_{\mathrm{g}}y}\,, \tag{S44}$$

and substitute $x' = -x$ and $y' = -y$ in the second line and use the symmetry of the trapping potential, to obtain

$$\tau_{\mathrm{MFP}}(X_{\mathrm{T}}) = \int_{-X_{\mathrm{T}}}^{0} \mathrm{d}x\, e^{V_{\mathrm{t}}(x)+\delta_{\mathrm{g}}x} \int_{-\infty}^{0} \mathrm{d}y\, e^{-V_{\mathrm{t}}(y)-\delta_{\mathrm{g}}y} - \int_{0}^{X_{\mathrm{T}}} \mathrm{d}x'\, e^{V_{\mathrm{t}}(x')-\delta_{\mathrm{g}}x'} \int_{0}^{x'} \mathrm{d}y'\, e^{-V_{\mathrm{t}}(y')+\delta_{\mathrm{g}}y'}$$
$$+ \int_{0}^{X_{\mathrm{T}}} \mathrm{d}x\, e^{V_{\mathrm{t}}(x)+\delta_{\mathrm{g}}x} \int_{-\infty}^{0} \mathrm{d}y\, e^{-V_{\mathrm{t}}(y)-\delta_{\mathrm{g}}y} + \int_{0}^{X_{\mathrm{T}}} \mathrm{d}x\, e^{V_{\mathrm{t}}(x)+\delta_{\mathrm{g}}x} \int_{0}^{x} \mathrm{d}y\, e^{-V_{\mathrm{t}}(y)-\delta_{\mathrm{g}}y} \quad [S45]$$
$$= \int_{-X_{\mathrm{T}}}^{X_{\mathrm{T}}} \mathrm{d}x\, e^{V_{\mathrm{t}}(x)+\delta_{\mathrm{g}}x} \int_{-\infty}^{0} \mathrm{d}y\, e^{-V_{\mathrm{t}}(y)-\delta_{\mathrm{g}}y} + \int_{0}^{X_{\mathrm{T}}} \mathrm{d}x\, e^{V_{\mathrm{t}}(x)} \int_{0}^{x} \mathrm{d}y\, e^{-V_{\mathrm{t}}(y)} \left[e^{\delta_{\mathrm{g}}(x-y)} - e^{-\delta_{\mathrm{g}}(x-y)}\right] , \quad [S46]$$

where we re-labelled x' and y' to x and y? In general, $x-y \geq 0$ in the second double integral making it non-negative. Therefore, neglecting it yields a lower bound on the mean first-passage time. Also, since $V_{\mathrm{t}}(x) \geq V_{\mathrm{t}}(0) = 0$, we find,

$$\tau_{\mathrm{MFP}}(X_{\mathrm{T}}) \geq \int_{-X_{\mathrm{T}}}^{X_{\mathrm{T}}} \mathrm{d}x\, e^{V_{\mathrm{t}}(x)+\delta_{\mathrm{g}}x} \int_{-\infty}^{0} \mathrm{d}y\, e^{-V_{\mathrm{t}}(y)-\delta_{\mathrm{g}}y} \quad [S47]$$
$$\geq \int_{-X_{\mathrm{T}}}^{X_{\mathrm{T}}} \mathrm{d}x\, e^{\delta_{\mathrm{g}}x} \int_{-\infty}^{0} \mathrm{d}y\, e^{-V_{\mathrm{t}}(y)-\delta_{\mathrm{g}}y} \quad [S48]$$
$$= \frac{2}{\delta_{\mathrm{g}}} \sinh\left(\delta_{\mathrm{g}} X_{\mathrm{T}}\right) \int_{-\infty}^{0} \mathrm{d}y\, e^{-V_{\mathrm{t}}(y)-\delta_{\mathrm{g}}y} \quad [S49]$$
$$\geq 2X_{\mathrm{T}} \int_{-\infty}^{0} \mathrm{d}y\, e^{-V_{\mathrm{t}}(y)-\delta_{\mathrm{g}}y} . \quad [S50]$$

That is, for symmetric trapping potentials, the mean first-passage time exceeds its linear expansion for small X_{T}. Consequently, the power is maximized at vanishing threshold:

$$P(X_{\mathrm{T}}) = \frac{2\delta_{\mathrm{g}} X_{\mathrm{T}}}{\tau_{\mathrm{MFP}}(X_{\mathrm{T}})} \leq P(X_{\mathrm{T}} \to 0) = \delta_{\mathrm{g}} \left[\int_{-\infty}^{0} \mathrm{d}y\, e^{-V_{\mathrm{t}}(y)-\delta_{\mathrm{g}}y}\right]^{-1} . \quad [S51]$$

K.3. For symmetric potentials having a maximum stiffness, the harmonic shape is optimal. In the preceding subsections, we have established that power is maximized for symmetric trapping potentials at vanishing ratcheting threshold. Here, we show that, when the stiffness of the potential is limited, a quadratic trap achieves maximum power.

We assume that the stiffness of the material is upper bounded by one. Since the trap has a single minimum fixed at $V(0) = 0$, we can conclude that

$$V(x) \leq \frac{1}{2}x^2 , \quad [S52]$$

and, consequently, the extractable power at vanishing threshold reads

$$P(X_{\mathrm{T}} \to 0) = \delta_{\mathrm{g}} \left[\int_{-\infty}^{0} \mathrm{d}y\, e^{-V_{\mathrm{t}}(y)-\delta_{\mathrm{g}}y}\right]^{-1} \quad [S53]$$
$$\leq \delta_{\mathrm{g}} \left[\int_{-\infty}^{0} \mathrm{d}y\, e^{-y^2/2-\delta_{\mathrm{g}}y}\right]^{-1} \quad [S54]$$
$$= \sqrt{\frac{2}{\pi}}\, \delta_{\mathrm{g}}\, \mathrm{e}^{-\delta_{\mathrm{g}}^2/2} \left[1 + \mathrm{erf}\left(\frac{\delta_{\mathrm{g}}}{\sqrt{2}}\right)\right]^{-1} , \quad [S55]$$

i.e., when the stiffness is limited, the harmonic potential extracts maximum power.

L. Limits on velocity and work extraction. We provide here more details on how to estimate the limits to the rate of work extraction and directed velocity. At sufficiently high trap stiffness, time scales in addition to the overdamped relaxation time τ_{r} become important. Fluid memory effects make relevant the time scale required for fluid vorticity to diffuse a single particle radius r,

$$\tau_{\mathrm{f}} = \frac{r^2 \rho}{\eta} , \quad [S56]$$

where η the dynamical viscosity, and ρ the density of the surrounding water. Solving for the trap stiffness where these two time scales are equivalent, $\tau_{\mathrm{r}} = \tau_{\mathrm{f}}$, we obtain a critical trap stiffness

$$\kappa_{\mathrm{f}}^{*} = \frac{6\pi\eta^2}{\rho r} , \quad [S57]$$

where we recall that the Stokes dissipation for a spherical particle in an unbounded medium is $\gamma = 6\pi\eta r$. Computing the velocity and power from Eq. (S26) and Eq. (S25), respectively, using the different values of κ_{f}^* for the various bead diameters, we obtain rough estimates for the maximum velocity and output power (Table S1). Inertial effects also make relevant the time scale associated with the relaxation of momentum (velocity),

$$\tau_v = \frac{m}{\gamma}\,. \tag{S58}$$

The ratio of these time scales is $\tau_v/\tau_{\mathrm{f}} = (2/9)\Delta\rho/\rho \approx 0.2$ for silica beads in water, independent of particle size. Empirically, we have observed deviations from the single-time-scale overdamped theory when $\tau_r \geq 10\tau_{\mathrm{f}}$: in Fig. S6, the red marker corresponding to a 5 µm bead at $\kappa = 200$ pN/µm falls well below the predicted curve.

M. Error analysis. The error bars in all experimental figures represent the standard error of the mean. The relative uncertainties for the mean rate of stored gravitational energy P differ from those for the trap P_{trap}, even though both quantities are calculated from the same position measurements. To understand why, we discuss how each quantity may be estimated from the underlying theoretical model.

The uncertainty in estimating P is calculated by propagating the errors in mean work extracted in N_{traj} independent trajectories and in the N_{traj} corresponding trajectory durations. The mean power is defined as $P \equiv W^{\mathrm{traj}}/t^{\mathrm{traj}}$, where W^{traj} and t^{traj} are the mean extracted work and trajectory times, respectively. Note that although the trajectory lengths are fixed, typically, we exclude the first relaxation time of each trajectory, so that the remainder of the trajectory has approximately steady-state statistics. This exclusion leads to fluctuations in the stored gravitational energy across the trajectories, because the distance traveled in a fixed time is stochastic. Including the contribution from the stochastic first-passage time and neglecting the (small) covariance between work and time, we have

$$\left(\frac{\delta P}{P}\right)^2 = \left(\frac{\delta W^{\mathrm{traj}}}{W^{\mathrm{traj}}}\right)^2 + \left(\frac{\delta t^{\mathrm{traj}}}{t^{\mathrm{traj}}}\right)^2\,, \tag{S59}$$

where $\delta W^{\mathrm{traj}} = \sqrt{\mathrm{Var}(W^{\mathrm{traj}})/N_{\mathrm{traj}}}$, and $\mathrm{Var}(W^{\mathrm{traj}})$ is the variance of the gravitational energy stored in the trajectories. The error in the power P_{trap} input by the trap is governed by a similar expression,

$$\left(\frac{\delta P_{\mathrm{trap}}}{P_{\mathrm{trap}}}\right)^2 = \left(\frac{\delta W_{\mathrm{trap}}^{\mathrm{traj}}}{W_{\mathrm{trap}}^{\mathrm{traj}}}\right)^2 + \left(\frac{\delta t^{\mathrm{traj}}}{t^{\mathrm{traj}}}\right)^2\,. \tag{S60}$$

In Eqs. S59 and S60, the uncertainties arising from trajectory-time fluctuations δt^{traj} are identical. As a result, we focus on estimating the work fluctuations in the first terms. Since the actual values of these terms vary with parameters such as the trap stiffness κ, we focus on a typical case, $\delta_g = 0.8$, which maximizes P.

For gravitational energy, the fluctuation arises mainly from excluding the first relaxation time of the trajectory. This leads to a position variance $\approx 2Dt_{\mathrm{r}}$; hence, the variance of W^{traj} is $\left(mg\sqrt{2Dt_{\mathrm{r}}}/k_{\mathrm{B}}T\right)^2$. The mean work is $mg\,X^{\mathrm{traj}}/k_{\mathrm{B}}T$, where X^{traj} is the mean trajectory length. For $\delta_g = 0.8$, the predicted error estimates (mean ± standard error), in units of $k_{\mathrm{B}}T$, are 11.4 ± 0.1, which is very close to the experimental error estimates $W^{\mathrm{traj}} = 11.3 \pm 0.1$.

For trap energy, the variance in W_{trap} arises in part from feedback latency, the delay between measurement and response: If the response were instantaneous, then the only uncertainty would arise from measurement errors; however, the feedback delay t_{s} allows the bead to relax. By empirically tuning the feedback gain α, we compensate for the mean shift, but fluctuations are still present. The variance in measured position at the reset time step is $2Dt_{\mathrm{s}} + (\sigma_m')^2$, where $(\sigma_m')^2$ is the variance of position measurements. As the trap work depends on the square of the position, the variance of the square of the position is given by $2\left(2Dt_{\mathrm{s}} + (\sigma_m')^2\right)^2$. Thus, for a trajectory with N ratchet events, the trap-work variance is $\approx \left(\sqrt{2N}(2Dt_{\mathrm{s}} + (\sigma_m')^2)\,\kappa/k_{\mathrm{B}}T\right)^2$. The predicted value of $\delta W_{\mathrm{trap}}^{\mathrm{traj}}$ for $N_{\mathrm{traj}} = 81$, $N = 83$ and $\sigma_m' = 1.7$ nm is 2.43×10^{-2}, which is very close to the experimental value, 2.47×10^{-2}.

The 0.5-µm-bead experiments show larger relative fluctuations than the experiments using larger beads (Fig. S8). Measuring the position of this small bead size required a higher-NA objective and thus a smaller detection range. The trajectories then had half the length (≈ 170 nm) relative to the other experiments, with correspondingly fewer ratchet events. Here we focus on errors in power in the limit of small δ_g, where the directed velocity is maximized. For $\delta_g = 7.5 \times 10^{-4}$, the experimental mean work $W^{\mathrm{traj}} = 3.4 \pm 0.3 \times 10^{-2}$, which is comparable to the expected values of $3.3 \pm 0.1 \times 10^{-2}$, but with higher uncertainty. For trap energies, the experimental error estimates are $\delta W_{\mathrm{trap}}^{\mathrm{traj}} = 3.8$, and the predicted values are of the order $\delta W_{\mathrm{trap}}^{\mathrm{traj}} = 0.95$. The position measurements were significantly affected by the measurement noise ($\sigma_m' \approx 4$ nm) as it is comparable to the standard deviation of the bead in the trap ($\sigma' \approx 5$ nm).

The experimental variances of the energies are larger than predicted, because of vibrations and low-frequency drift in the mechanical setup and because of high-frequency intensity fluctuations in the detection laser. From the above two case studies, we see that the uncertainties, both absolute and relative, change considerably with δ_g. Nonetheless, they follow the trend that $\delta P_{\mathrm{trap}}/P_{\mathrm{trap}} > \delta P/P$. The reason is that although both energies are calculated from the bead positions, the uncertainty in P arises from the "single" measurement uncertainty at the beginning of the trajectory, whereas P_{trap} arises from the accumulated errors in the N ratchet events in a trajectory.

N. Quantifying information and associated thermodynamic costs. In each measurement-feedback cycle, we acquire information about the position of the particle. The information gained about the particle position leads to a lower bound on the fundamental energetic requirements of the controller that achieves the ratcheting. Various "information costs" have been proposed in the literature. While all are based on Landauer's principle, i.e., that erasing information from a memory comes at an unavoidable thermodynamic cost, their appropriateness depends on the explicit "architecture" of the measurement-feedback process. A main conclusion from a detailed treatment (7) is that if the controller stores all past measurements, its dissipation is bounded by the *transfer entropy* (8–10). By contrast, when only the current measurement is stored, the controller's dissipation is bounded by *information flow* (11). For our discrete-time setup, the latter coincides with the *nonpredictive information* (12) that the controller has about the next position of the particle.

Since the exact details of the controller are irrelevant when maximizing the ratchet performance in the limit of fast feedback loops and precise measurements, there is some freedom in choosing between transfer entropy and nonpredictive information. Importantly, both information costs diverge for high signal-to-noise ratios and high sampling frequencies. In that limit, carrying out the feedback algorithm (Eq. 4 in the main text) has infinite cost, implying that the information-to-work conversion efficiency is zero.

Although the power output of our ratchet is well described by a simple theory that assumes infinite sampling frequency and no measurement error, the experiment is actually carried out using a finite sampling rate and with measurements that have finite error. Typical values for the amplitude signal-to-noise ratio are ≈ 10 (Table S2), and typical scaled sampling frequencies are ≈ 180 (Fig. 3A in the main text). Below, we show that these values lead to an efficiency of $\approx 8\%$.

Efficiency estimate – We estimate the efficiency by considering the stored gravitational work per feedback cycle and the costs of running the controller per measurement under the effect of finite measurement error σ_m, which, when scaled by the trap size σ, can be expressed via the amplitude signal-to-noise ratio: $\sigma_{\mathrm{m}} = 1/\mathrm{SNR}$.

Since the experimental power spectra (Fig. S5) are compatible with Gaussian noise, the observations y of position x are distributed as

$$p\left(y|x\right) = \mathcal{N}\left(y; x, \sigma_m^2\right) \ . \tag{S61}$$

Contrary to the derivations in Section E, the trap potential now reacts to a noisy measurement of the particle position. Therefore, the propagator of the trap (λ) dynamics changes from that given in Eq. (S19) to

$$p\left(\lambda_{n+1}|x_{n+1},\lambda_n\right) = \frac{1}{2}\delta\left(\lambda_{n+1}-\lambda_n\right)\left[1+\mathrm{erf}\left(-\frac{x_{n+1}-\lambda_n}{\sqrt{2\sigma_m^2}}\right)\right] + \Theta\left(\lambda_{n+1}-\lambda_n\right)\mathcal{N}\left(\lambda_{n+1}; \alpha^* x_{n+1} + \left[1-\alpha^*\right]\lambda_n, \left[\alpha^*\sigma_m\right]^2\right) , \tag{S62}$$

where α^* is determined numerically such that the input work is zero.

We carry out the same procedure as in Section E and compute the output gravitational work $\langle \Delta W_{\mathrm{g}} \rangle$ per cycle [Eq. (S24)]. A possible controller architecture is one in which the memory of the control variable is overwritten upon each new measurement. Then, following Ref. (7), the input work, in scaled units, is determined by the information flow from the system x to the control variable λ. At steady state, this is equivalent to the non-predictive information (12),

$$I_{\mathrm{np}} \equiv I\left[x_n;\lambda_n\right] - I\left[x_{n+1},\lambda_n\right] \tag{S63}$$

$$= H\left[x_n\right] - H\left[x_n|\lambda_n\right] - \left\{H\left[x_{n+1}\right] - H\left[x_{n+1}|\lambda_n\right]\right\} \ , \tag{S64}$$

where I denotes the mutual information and H the Shannon entropy (13). At steady state, the system entropy does not change, and performing a change of variables to relative coordinates [as in Eq. (S17)] gives

$$I_{\mathrm{np}}^{\mathrm{ss}} = H\left[x_{n^+}^{\mathrm{r}}\right] - H\left[x_n^{\mathrm{r}}\right] \ . \tag{S65}$$

These marginal entropies over the relative coordinates are computed using

$$H\left[x_{n^+}^{\mathrm{r}}\right] = -\int \mathrm{d}x_{n^+}^{\mathrm{r}}\, \pi\left(x_{n^+}^{\mathrm{r}}\right) \log \pi\left(x_{n^+}^{\mathrm{r}}\right) \tag{S66}$$

$$H\left[x_n^{\mathrm{r}}\right] = -\int \mathrm{d}x_n^{\mathrm{r}}\, \pi\left(x_n^{\mathrm{r}}\right) \log \pi\left(x_n^{\mathrm{r}}\right) \ , \tag{S67}$$

with steady-state distributions $\pi_+\left(x_{n^+}^{\mathrm{r}}\right)$ and $\pi\left(x_n^{\mathrm{r}}\right)$ obtained by Eqs. (S20) and (S21), respectively, with the modified propagator from Eq. (S62).

We numerically compute these distributions using a uniformly spaced grid with 15,000 points in the domain $x_n^{\mathrm{r}},\ x_{n^+}^{\mathrm{r}} \in [-60, 60]$. The efficiency is

$$\eta = \frac{\langle \Delta W_{\mathrm{g}} \rangle}{I_{\mathrm{np}}^{\mathrm{ss}}} . \tag{S68}$$

For the typical experimental values stated above and scaled effective mass $\delta_{\mathrm{g}} = 0.8$, we compute the non-predictive information to be $I_{\mathrm{np}}^{\mathrm{ss}} \approx 0.019$ and the output work to be $\langle \Delta W_{\mathrm{g}} \rangle \approx 0.0016$. This yields an efficiency of $\eta \approx 0.084$.

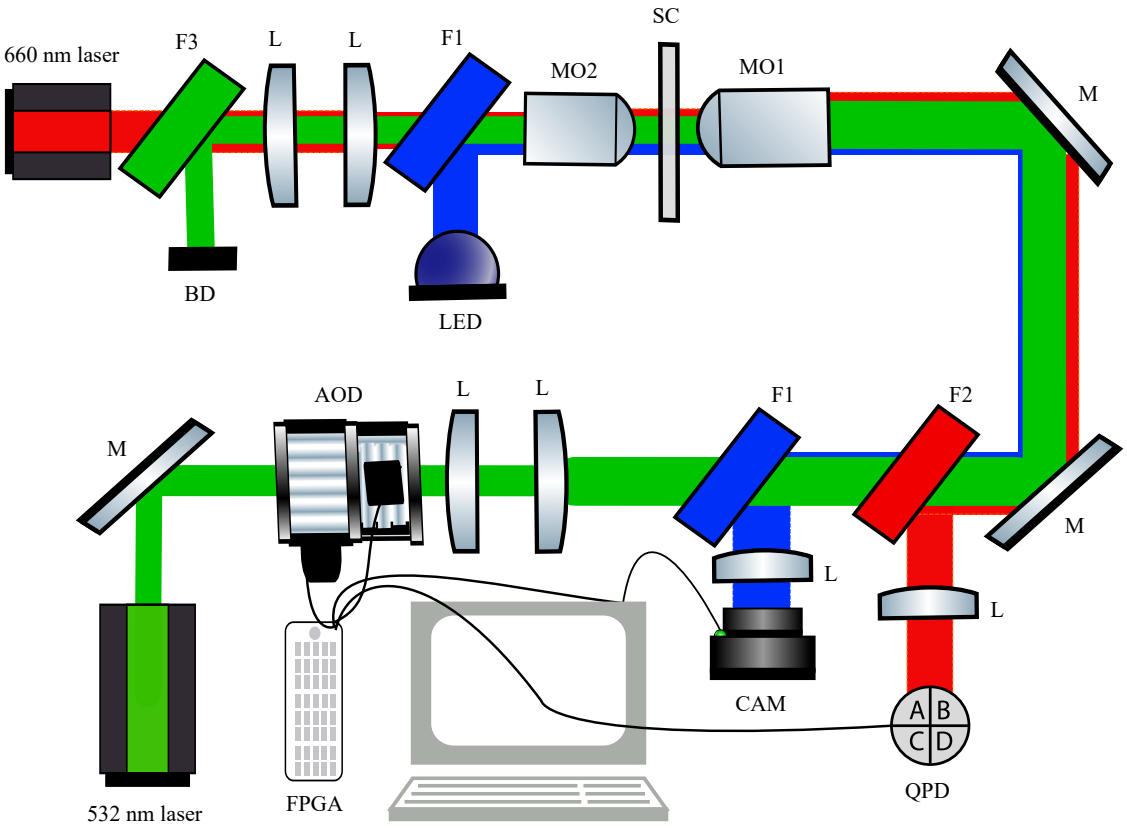

Fig. S1. Schematic diagram of the experimental apparatus. M = mirror, AOD = acousto-optic deflectoor, L = lens, F1 = blue dichroic filter, CAM = camera, F2 = red dichroic filter, MO1 = trapping microscope objective, MO2 = detection microscope objective, SC = sample chamber, F3 = green dichroic filter.

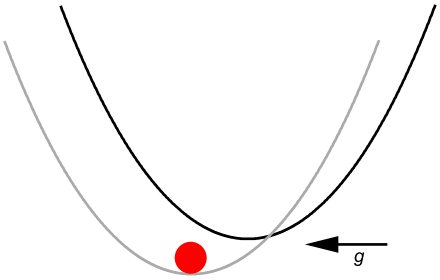

Fig. S2. Bead at the minimum-energy state of the total potential (gray curve) due to the trap and gravity. The black plot is the trap potential.

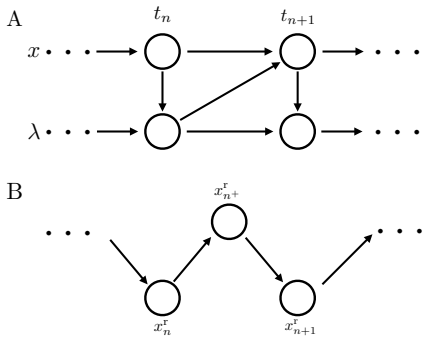

Fig. S3. Causal structure of bead (x) and trap (λ) dynamics. A) Structure of dynamics for true position x and trap center λ. B) Structure of dynamics for the relative coordinates x^{r}_{n} and x^{r}_{n+} defined in $\mathrm{Eq.}$ (S17).

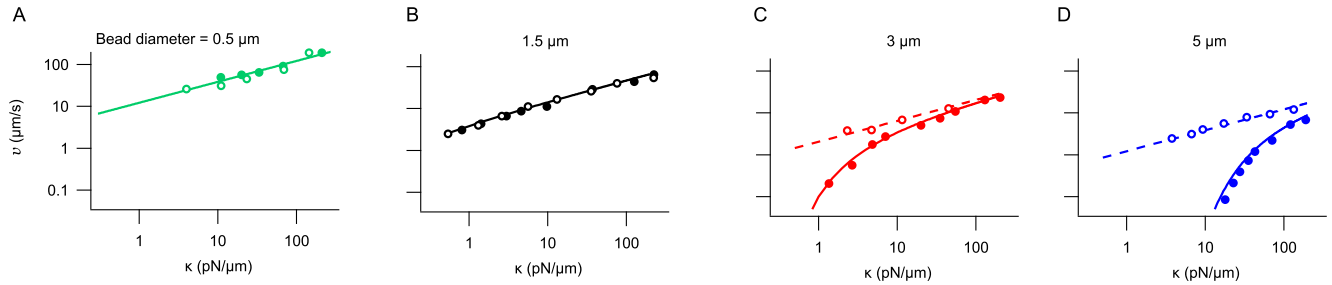

Fig. S4. Horizontal and vertical velocity (hollow and solid markers, respectively) as a function of trap stiffness κ for different bead diameters. The solid lines correspond to Eq. (S26) and reflect the influence of the gravitational force on upward fluctuations. The dashed lines are calculated from Eq. (S29).

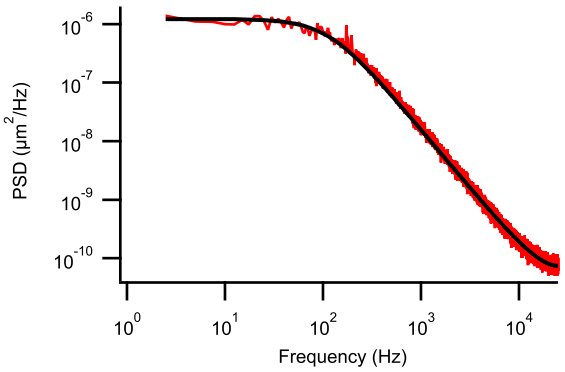

Fig. S5. Power spectral density for 3 μm bead. The red curve is the experimental data and solid black curve is the fit to an aliased Lorentzian (6) with Gaussian noise floor. The fit parameters are $f_c = 113.42 \pm 1.34$ Hz, $D = 0.167 \pm 0.001$ μm^2/s, and $\sigma'_m = 0.54 \pm 0.08$ nm.

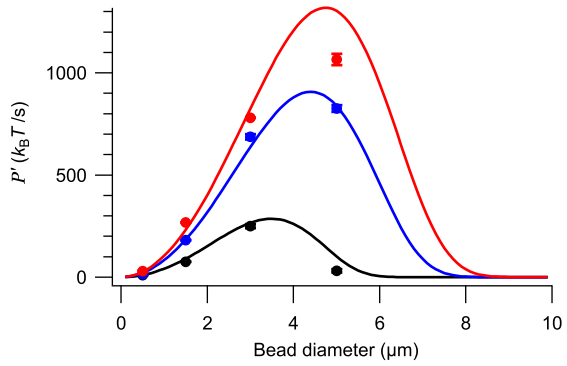

Fig. S6. Optimum bead size. Power as a function of bead size for trap stiffness $\kappa = 30$ pN/μm (black), $\kappa = 100$ pN/μm (blue) and $\kappa = 200$ pN/μm (red). The circles are experimental data. Curves show Eq. S31.

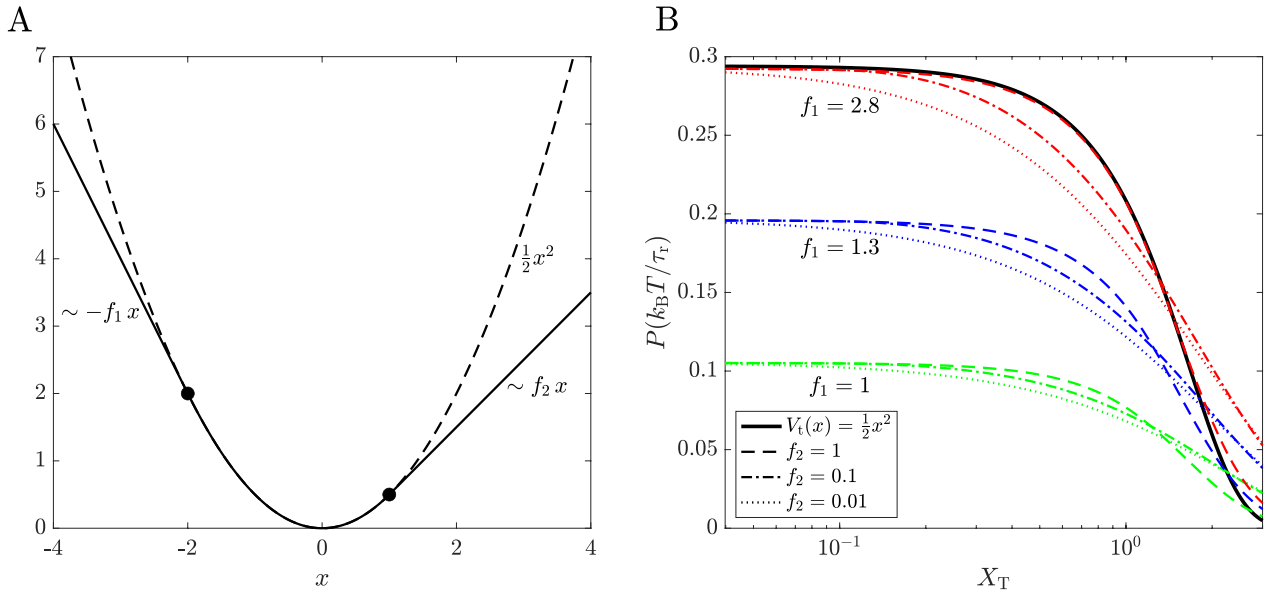

Fig. S7. The *quadratic-to-linear* trap. A) Sketch of the potential for maximum slope magnitudes $f_1 = 2$ for negative x and $f_2 = 1$ for positive x. B) Power as a function of threshold for different maximum slope magnitudes f_1 and f_2 and $\delta_{\rm g} = 0.8$.

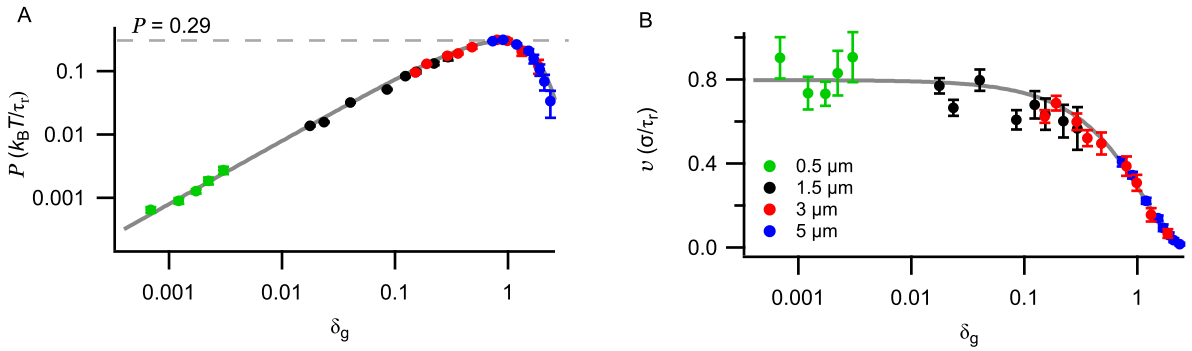

Fig. S8. A) Scaled power and B) velocity as a function of δ_g to show small δ_g dependence. The symbols shows the same data as Fig. 4 (B) and (D) on log scale for scaled velocity and power, respectively.

d [μm]	0.5	1.5	3	5
τ_{f} [μs]	0.07	0.6	2.5	7
τ_v [μs]	0.01	0.1	0.6	1.6
κ_{f}^* [μN/m]	60000	20000	10000	6000
v^* [μm/s]	3000	570	200	80
P^* $[k_{\mathrm{B}}T/s]$	470	2400	6800	10000

Table S1. Fluid memory time scale, velocity relaxation time, critical trap stiffness and the associated velocity and output power (rows) for the different bead diameters (columns) that we consider.

ϕ (μm)	σ'_m (nm)	σ' (nm)	SNR	κ (pN/μm)	τ_r (ms)	D (μm^2/s)	γ (pN·s/μm)	δ_g	α
0.5	4.8 ± 0.3	19.4	4.0	10.9	0.36	1.0	3.9×10^{-3}	3.0×10^{-3}	0.65
		14.3	3.0	20.1	0.21	0.98	4.2×10^{-3}	2.2×10^{-3}	0.66
		11.1	2.3	33.3	0.13	0.93	4.4×10^{-3}	1.7×10^{-3}	0.85
		7.8	1.6	67.6	0.06	0.98	4.2×10^{-3}	1.2×10^{-3}	1.0
		4.8	1.0	177	0.02	1.0	4.1×10^{-3}	6.9×10^{-4}	1.0
1.5	3.7 ± 0.2	71.7	19.4	0.8	13	0.38	1.1×10^{-2}	2.9×10^{-1}	1.0
		54.2	14.6	1.4	7.4	0.39	1.0×10^{-2}	2.2×10^{-1}	0.8
		37.0	10.0	3.0	3.6	0.38	1.1×10^{-2}	1.5×10^{-1}	1.0
		29.9	8.1	4.6	2.3	0.38	1.1×10^{-2}	1.2×10^{-1}	1.1
		20.5	5.5	9.8	1.1	0.38	1.1×10^{-2}	8.5×10^{-2}	1.1
		9.7	2.6	43.6	0.26	0.36	1.1×10^{-2}	4.0×10^{-2}	1.0
		5.7	1.5	126	0.09	0.38	1.1×10^{-2}	2.4×10^{-2}	1.0
		4.3	1.2	224	0.05	0.36	1.1×10^{-2}	1.8×10^{-2}	1.1
3	0.9 ± 0.1	54.2	60.2	1.4	17.8	0.17	2.4×10^{-2}	1.8	1.9
		39.0	43.4	2.7	10.8	0.14	2.9×10^{-2}	1.3	1.9
		29.3	32.5	4.8	5.1	0.17	2.4×10^{-2}	9.8×10^{-1}	1.8
		24.1	26.7	7.1	3.4	0.17	2.4×10^{-2}	8.0×10^{-1}	1.6
		14.3	15.9	20.1	1.4	0.14	2.8×10^{-2}	4.8×10^{-1}	1.5
		10.8	12.0	35.2	0.8	0.15	2.7×10^{-2}	3.6×10^{-1}	1.3
		8.6	9.6	54.9	0.48	0.16	2.6×10^{-2}	2.9×10^{-1}	1.5
		5.6	6.2	130	0.19	0.17	2.4×10^{-2}	1.9×10^{-1}	1.45
		4.5	5.0	203	0.12	0.17	2.4×10^{-2}	1.5×10^{-1}	1.4
5	0.6 ± 0.1	15.1	25.2	18	2.8	0.080	5.1×10^{-2}	2.4	1.7
		13.4	22.4	22.8	2.2	0.082	5.0×10^{-2}	2.1	1.65
		12.2	20.3	27.7	1.8	0.082	5.0×10^{-2}	1.9	1.63
		10.8	17.9	35.4	1.4	0.083	5.0×10^{-2}	1.7	1.6
		9.8	16.3	43.0	1.1	0.084	4.9×10^{-2}	1.5	1.45
		7.6	12.7	70.7	0.77	0.078	5.4×10^{-2}	1.2	1.6
		5.8	9.7	121	0.38	0.089	4.6×10^{-2}	0.9	1.5
		4.6	7.8	189	0.28	0.078	5.2×10^{-2}	0.7	1.7

Table S2. Experimental parameters corresponding to the data shown in Fig. 4. Table columns (left to right) list the nominal bead diameter ϕ reported by the manufacturer (Sigma-Aldrich), associated measurement noise σ'_m for the different bead diameters, measured trap length (σ'), signal-to-noise ratio (SNR), trap stiffness κ, relaxation time τ_r and diffusion constant D for each experiment. The last three columns list derived quantities: the corresponding friction coefficient γ, scaled effective mass δ_g, and applied feedback gain α.

References

1. A Kumar, J Bechhoefer, Nanoscale virtual potentials using optical tweezers. *Appl. Phys. Lett.* **113**, 183702 (2018).
2. A Kumar, J Bechhoefer, Optical feedback tweezers in *Optical Trapping and Optical Micromanipulation XV*. (International Society for Optics and Photonics), Vol. 10723, p. 107232J (2018).
3. Y Jun, J Bechhoefer, Virtual potentials for feedback traps. *Phys. Rev. E* **86**, 061106 (2012).
4. P Hänggi, P Talkner, M Borkovec, Reaction-rate theory: fifty years after Kramers. *Rev. Mod. Phys.* **62**, 251 (1990).
5. H Risken, *The Fokker-Planck equation*. (Springer, Berlin Heidelberg New York), 2nd edition, (1996).
6. K Berg-Sørensen, H Flyvbjerg, Power spectrum analysis for optical tweezers. *Rev. Sci. Instruments* **75**, 594–612 (2004).
7. JM Horowitz, H Sandberg, Second-law-like inequalities with information and their interpretations. *New J. Phys.* **16**, 125007 (2014).
8. JM Horowitz, S Vaikuntanathan, Nonequilibrium detailed fluctuation theorem for repeated discrete feedback. *Phys. Rev. E* **82**, 061120 (2010).
9. M Ponmurugan, Generalized detailed fluctuation theorem under nonequilibrium feedback control. *Phys. Rev. E* **82**, 031129 (2010).
10. T Sagawa, M Ueda, Nonequilibrium thermodynamics of feedback control. *Phys. Rev. E* **85**, 021104 (2012).
11. JM Horowitz, M Esposito, Thermodynamics with Continuous Information Flow. *Phys. Rev. X* **4**, 031015 (2014).
12. S Still, DA Sivak, AJ Bell, GE Crooks, Thermodynamics of Prediction. *Phys. Rev. Lett.* **109**, 120604 (2012).
13. TM Cover, JA Thomas, *Elements of Information Theory*. (Wiley-Interscience, Hoboken, NJ), 2nd edition, (2006).