Institute for Physical AI @ John Bailey Institute · The Charlot Lab

COMPUTE · SUBSTRATE

Technical Report TR–2026–40
Research / Position · Preprint v1
15 August 2026

COMPUTE · SUBSTRATE

The Physical AI Hardware Lottery

Energy-based models, post-von-Neumann architecture, and how a substrate shapes which ideas get developed.

The Charlot Lab · Institute for Physical AI @ John Bailey Institute

Builds on TR-2026-01, which surveys the substrates. This report asks which algorithm class they suit. Sweep conducted in thirteen vocabularies.

the inner loop was a Markov chain · diffusion models are energy-based models · two relationships to one device · three denominators, three arguments

In 2020 Sara Hooker described a mechanism she called the hardware lottery: a research idea advances partly because it suits the available hardware, and hardware can therefore delay a line of work by making it look unproductive [1]. This report traces that mechanism through energy-based models, and the tracing is mechanical rather than interpretive. Fitting an energy-based model by maximum likelihood requires the partition function, which is intractable in the dimension of a robot action, so the field estimates the gradient from samples drawn by Markov chain Monte Carlo. A Markov chain is sequential. A general-purpose graphics processor accelerates work that divides into independent pieces, so it accelerates backpropagation completely and a Markov chain not at all. That asymmetry is measurable and it is the subject of Section 3. What the record then shows is that the underlying idea continued to produce results in the parameterisation the hardware suited: a diffusion model is an energy-based model written in terms of the gradient of its energy rather than the energy itself, and that gradient costs one forward pass [5]. Physical AI is where this becomes concrete, because an embodied policy frequently needs to represent a set of correct actions rather than one, and the dominant visuomotor policy class does so by running **Langevin dynamics at inference**, at a published cost of 10 to 100 network evaluations per action that its own literature identifies as the constraint on real-time deployment [6]. Section 7 crosses that published range against a measured 100 millisecond end-to-end latency for an embedded sampler and finds the two costs meet at roughly 33 evaluations, inside the

range now in use [11]. Section 8 describes the second workload arriving on the same silicon, since transformer attention is being mapped into memory arrays [9], and identifies the physical quantity that distinguishes the two cases: one uses the device's fluctuations as its source of randomness and the other requires them suppressed. Section 10 separates the three denominators in which efficiency can be stated, which is what allows the datacentre argument and the embodied argument to be read without conflict.

1. What mechanism does this report trace?

Hooker's essay describes how hardware and software have repeatedly shaped which research directions advanced, and observes that domain-specialised hardware raises the cost of working away from the established path [1]. Her formulation is that a hardware lottery can delay progress by making a productive idea look like a failed one.

The subject here is specific and narrow. Energy-based models require an operation that the dominant hardware generation did not accelerate. Section 3 establishes that this is a property of the operation and can be measured; Sections 4 through 7 follow what happened to the idea afterwards. The report describes a relationship between an algorithm and a machine, and it does not make a claim about what any research community concluded or intended.

2. What is in scope, and what does this continue?

TR-2026-01 already surveys the substrates. It organises roughly forty methods into nine families across two axes, from energy-harvesting microcontrollers to orbital datacentres, and establishes that the roots of nearly every method are decades old and openly published: balanced ternary hardware from 1958, the memristor from 1971, the Ising model from 1925, reversible computing from 1961 [15]. Its closing observation is that the near-term advance is a convergence of existing paradigms rather than a single new one, and specifically the pairing of a multiply-free deterministic substrate with a physics-based stochastic one.

> That survey establishes which machines can be built. This report takes the adjacent question: which algorithm class the stochastic substrate suits, why that class is currently represented at small scale, and what it would offer an embodied system. The contribution is a description of *fit between an operation and a machine*.

What follows contributes no new measurement. It reads measurements made by others, together with two computations from the Institute's own benches which are marked where they appear. The literature sweep covered thirteen vocabularies: English, Chinese, Japanese, Korean, German, French, Spanish, Russian, Portuguese, Swedish, Dutch, Italian and Indian-institution English. Each figure is given with what was measured to obtain it, so that a reader can see which quantity a number describes. Where the sweep did not return a treatment of something, the text says that it did not locate one, which describes the search rather than the literature.

3. What is the operation, and how did the hardware meet it?

An energy-based model defines a distribution in which probability falls exponentially with an energy. Fitting it by maximum likelihood requires the partition function, the integral of that exponential over the whole answer space, which grows as grid resolution raised to the number of dimensions and is out of reach for a robot action. The approach developed from 1985 onward estimates the gradient using negative samples drawn by Markov chain Monte Carlo [3].

Three properties of that approach interact with parallel hardware in a way that can be stated exactly.

Property of the method	Behaviour on massively parallel hardware
A Markov chain is **sequential**: each state depends on the last	Parallel lanes accelerate work that divides into independent pieces. A chain does not divide, so lane count does not enter its wall-clock time.
A short chain carries bias; a long chain costs proportionally more	The trade between bias and cost is set by chain length, which lane count does not change.

Contrastive divergence is not the gradient of an objective function	Its convergence behaviour is studied empirically rather than derived, so tuning it is slower work.

Backpropagation has a different structure: a forward pass and a reverse pass, both dense linear algebra, both dividing perfectly across lanes, with an objective whose gradient is exact. When general-purpose graphics processors arrived, the cost of backpropagation fell by orders of magnitude and the cost of a Markov chain did not, because the chain's limit is its dependency structure rather than arithmetic throughput.

> **The timeline follows the hardware.** Restricted Boltzmann machines powered deep belief networks in 2006 and were a principal route into deep learning. From 2012 onward the combination of graphics processors and backpropagation carried most of the field's activity, and work on the energy-parameterised form moved to smaller scales. Section 4 follows what happened to the idea itself over the same period.

4. What does the record show about the idea?

4.1 The physics lineage

The 2024 Nobel Prize in Physics was awarded jointly to John Hopfield and Geoffrey Hinton for foundational discoveries enabling machine learning with artificial neural networks. The citation identifies the work specifically: Hinton, between 1983 and 1985, used tools from statistical physics to create the **Boltzmann machine** [2]. The line originates in statistical mechanics, and that origin is what makes a physical substrate a natural question to ask about it.

4.2 Diffusion models are energy-based models

In a diffusion model the noise-prediction network is the gradient of an energy with respect to its input, and when data is perturbed with Gaussian noise the denoising score-matching loss coincides with the diffusion training loss. The relationship is stated directly in the literature: an energy-based model and a plain diffusion model **differ in parameterisation**, energy against score [5]. Lesson 6 of the accompanying course recovers one form from the other numerically and

shows the residual falling with the integration step, which is the behaviour of a discretisation error rather than a difference between the objects.

The same sources record that score parameterisation is dominant in practice, citing flexibility and efficiency, and that pre-trained energy-parameterised models are not publicly available at scale. Both forms describe the same distribution. They differ in what it costs to obtain a gradient: the score form returns it from a single network evaluation, and the energy form obtains it through sampling.

> **This is the mechanism of Section 3 visible in a current field rather than a historical one.** The observation is about which parameterisation carries the field's scale, and it supports a narrow statement: the cost of obtaining a gradient differs between the two forms by the structure of the computation, and the form with the cheaper gradient is the one with public models at scale.

4.3 Results on embodied tasks

Implicit Behavioral Cloning reported that energy-based policies often outperform explicit mean-squared-error and mixture-density policies on robot learning tasks, including high-dimensional action spaces and visual inputs, with physical robots learning contact-rich behaviours at one millimetre precision [4]. Its theoretical argument concerns representing functions that are **discontinuous and multi-valued**, which is the property Section 5 develops.

4.4 What is established, and at what scale?

The Institute's own bench series places the energy-based advantage on specific properties: out-of-distribution generalisation, composition by summing energies, verification by scoring candidates, and adaptive computation at inference. Every result in that series is at or below 800 million parameters [14]. Section 3 concerns the cost of an operation and the scale at which a parameterisation is represented; the question of capability at frontier scale is a separate one, and this report treats it as open rather than settled in either direction.

5. Why does this become concrete in Physical AI?

A robot's correct action is frequently a set rather than a point. There is an obstacle; going left is correct and going right is correct. A policy trained by minimising squared error against demonstrations of both returns their **mean**, and the mean of left and right passes through the obstacle. Additional demonstrations move the mean closer to the centre, because the mean is what the loss is defined to return.

The Institute's bench reproduces this on a multivalued algebraic system, where a fairly supervised feed-forward network scored zero at every model size tried [14]. An energy over the action space holds each valid answer as a separate minimum, and selection is a separate step. Lessons 1 and 2 of the course build both objects and compare them.

5.1 The dominant policy class samples an energy at inference

Diffusion Policy is described in its literature as the dominant paradigm for representing multimodal action distributions in robot learning, and its inference procedure iteratively optimises with respect to a learned gradient field through **a series of stochastic Langevin dynamics steps** [6]. The same literature states the cost and its consequence: these methods require typically ten to one hundred network function evaluations, and that requirement is the constraint on real-time deployment.

> **So the sampling operation of Section 3 is running today, on every action, in the highest-value robot policy class in the field, on general-purpose hardware.** Its cost appears directly as control latency, which makes the substrate question quantitative rather than architectural. Section 7 puts a number on it.

5.2 The controller and the contact model share the same form

Model Predictive Path Integral control, the sampling controller used on contact-rich manipulators and off-road vehicles, weights sampled rollouts by the exponential of negative cost over a temperature, which is a Boltzmann distribution over trajectories. The control literature derives this directly, obtaining MPPI as a preconditioned gradient step on a Kullback-Leibler-regularised **free energy** objective and as expectation

maximisation with Boltzmann reweighting [7]. A mapping of MPPI onto Ising hardware has been published, so far without hardware results or energy measurements [8].

Below the policy, contact dynamics is formulated as a complementarity problem, and modern treatments express contact constraints as **energy terms in an objective** and integrate variationally. The policy, the controller and the contact model are therefore all stated as energy objects, and all three are currently evaluated on hardware that represents an energy landscape numerically rather than physically.

6. What does a settling substrate change?

A Boltzmann machine, an Ising model and a probabilistic-bit fabric describe one object. A fabric held at a given inverse temperature occupies configurations with probability falling exponentially in their energy, so its resting distribution is the distribution a sampler is written to produce. Lesson 7 of the course verifies this against an exact enumeration and reports the sampling noise floor alongside the residual.

The same applies to fitting. Equilibrium propagation estimates gradients by comparing a freely settled state with one settled under a small nudge toward the target, using only local quantities, with no separate reverse pass and no stored activations [10]. Hardware demonstrations exist in analog, Ising, oscillator and memristive devices. Lesson 8 derives the estimate on a case where the exact gradient is available and shows the two converging as the nudge shrinks.

Both loops therefore have a physical realisation on this class of device. What such a device requires in exchange is the subject of Section 7.

7. Where do the two costs meet?

Both sides of the comparison now carry a number. The published operating range for diffusion policies is ten to one hundred network evaluations per action, evaluated in sequence [6]. An embedded field-programmable gate array running simulated bifurcation over 2,048 spins has a measured end-to-end latency of approximately **100**

milliseconds, independent of problem size, obtained with 32-fold lossless compression of the coupling matrix and a learned parameter estimator in place of per-problem tuning [11].

One cost grows with the number of steps because the steps are sequential. The other does not vary with step count, because the settling substrate has no steps. The figure below crosses them.

Conditioning and solve latency **100 ms**

At **100 ms** the crossover is **33 NFE**, which is **inside** the published ten to one hundred range. Above that step count the fabric is already the faster path.

Figure 1. One published range and one measured latency. The 3 millisecond network forward pass is this report's assumption and is the quantity a reader is most likely to have better information about; changing it moves the crossing point proportionally. The shaded band marks the published ten to one hundred evaluation range.

At the measured latency and a 3 millisecond forward pass the costs meet near **33 evaluations**, inside the published range. The crossing point is therefore within the operating region already in use rather than beyond it.

7.1 What is the sequential part buying?

The comparison above prices two costs against each other. It does not say why the sequential part exists, and that turns out to be measurable on a system with an exact answer.

A sampling controller has two knobs. **Rollouts** are independent, so parallel lanes divide them; that is the cost hardware has spent two decades reducing. **Refinement passes** perturb the mean the previous pass produced, so they cannot overlap; that is the cost lane count does not touch. On a linear-quadratic plant the discrete algebraic Riccati equation gives the optimal cost in closed form, so each configuration can be scored against the true optimum rather than against a rival method. Median of seven seeds, spread reported alongside [16].

Configuration	Rollouts	Refinement passes	Excess over optimal	Spread
width alone	800	1	20.97%	5.21%
width alone, four times as much	3,200	1	21.04%	4.06%
the same width, with depth	3,200	4	2.00%	0.29%
the same width, more depth	3,200	8	0.64%	0.05%

Four times the parallel work bought nothing. Going from 800 to 3,200 rollouts at a single refinement pass moves the result from 20.97 percent to 21.04 percent, a difference far inside the 4 to 5 percent seed spread, so the curve has reached a floor rather than moved slowly. Holding that same width and adding refinement passes takes the error to **0.64 percent**, a reduction of roughly thirty-three times, and the spread falls with it.

> **So the sequential part is carrying the accuracy, and it is the part that lane count cannot reduce.** This is a measured reason for the ten to one hundred evaluations of Section 5.1 rather than a restatement of them: those passes are not an inefficiency waiting for a better implementation, they are buying precision that width has a floor on. A substrate that reaches the distribution by settling rather than by iterating toward it addresses exactly the cost that parallel hardware has never been able to shorten.

7.2 What happens when the action space grows?

A scalar plant is a model system, and dimension is what decides whether the result reaches robots: a seven-jointed arm samples in a seven-dimensional action space. The measurement was repeated on a coupled plant at dimensions one through eight, with the matrix Riccati solution supplying the exact optimum at each. The prediction was recorded in the source before the run: sampling covers a volume, volume grows exponentially in dimension, so a fixed rollout budget should explore a larger space worse and the floor should rise [16].

Comparing dimensions fairly requires one care. The exploration noise is set per dimension, so holding it fixed makes the total perturbation grow as the square root of the dimension, and a floor that rose under that setting could be over-perturbation rather than geometry. The sweep below therefore scales the per-dimension noise by one over the square root of the dimension, holding total perturbation constant, which removes the question by construction. All rows at 3,200 rollouts, median of five seeds.

Action dimension	1 pass	2 passes	4 passes	8 passes
n = 1	0.57%	0.44%	0.37%	0.35%
n = 2	21.29%	7.33%	2.01%	0.55%
n = 4	44.20%	18.05%	6.75%	1.72%
n = 8	78.74%	40.25%	15.85%	5.48%

Read down the first column: the accuracy reachable without any sequential refinement falls from 0.57 percent to **78.74 percent**, roughly **138 times worse** across three doublings of the action space. Read across a row: refinement recovers it at every dimension. Read the two together and the result is the one that matters for a body.

The number of sequential passes needed for a fixed accuracy grows with the number of joints. To stay within about six percent of optimal takes one pass at n=1, four at n=2, and eight at n=4 and n=8. To stay within two percent takes one pass at n=1 and more than eight by n=8. Over the range measured, on a fully actuated plant, the requirement grows roughly in step with dimension, so the part of the computation

> that cannot be divided across lanes is the part that scales with the body. Section 7.3 tests how far that survives a change of plant.

The alternative explanation was tested directly rather than assumed away. At eight dimensions and one pass, reducing the noise below the constant-perturbation setting makes width steadily worse, from 78.74 percent to 118.09 and then 281.61 percent, and raising it to the uncorrected 0.400 gives 69.50 percent with a far wider seed spread. There is no noise scale at which width alone recovers, and the setting that works is the one holding total perturbation constant.

7.3 Is the result about the plant?

The sweep above uses one structure: a sparse ring where each state feeds the next, fully actuated. Replicating across seeds tests the sampler; only replicating across structure tests the claim. The measurement was therefore repeated on four couplings and on an underactuated system, at dimensions two and eight, at 3,200 rollouts [16].

Plant	n = 2, one pass	n = 8, one pass	n = 8, eight passes
sparse ring (the one above)	21.29%	78.74%	5.48%
dense, every state coupled	21.29%	78.51%	5.46%
stronger coupling, weaker self-term	21.07%	77.64%	5.44%
deterministic random coupling	21.29%	79.03%	5.47%
underactuated, half the states uncontrolled	12.57%	18.47%	3.09%

Across the four couplings the numbers are almost indistinguishable: the one-pass result at eight dimensions spans 77.64 to 79.03 percent, and eight passes recover it by a factor of 14.3 to 14.4 in every case. **The coupling structure is close to irrelevant; the dimension is what sets the result.**

Underactuation holds the direction and shrinks the effect, which narrows the claim. With half the states uncontrolled the one-pass result still worsens with dimension, from 12.57 percent at two to 18.47 percent at eight, and refinement still recovers it. But the rise is about 1.5 times rather than the 3.7 times the fully actuated ring shows. Part of that is a measurement effect rather than a physical one: an underactuated plant has a larger optimal cost, so the same absolute error reads as a smaller percentage. **The defensible statement is therefore the direction, which holds in all five plants, and a magnitude that is largest when the body is fully actuated.**

Getting that setting wrong understates what refinement buys by nearly an order of magnitude. At eight dimensions the uncorrected noise reports eight passes as a factor of 1.7 improvement; at the corrected setting the same eight passes are a factor of fourteen. A hyperparameter held fixed across a swept variable is a confound rather than a control, and this one was only visible because the check was run before the reading was published.

7.4 What happens on a body, at matched compute?

Linear plants supply truth and no physics. The same question was therefore put to a three-joint arm with gravity torque, a first-order actuator, torque limits, and an identified electrical power model whose copper, viscous and Coulomb coefficients are calibrated to a measured 71.5 joule reach on a Unitree G1 arm [16]. There is no exact optimum here and none is claimed. The anchor is a **reference controller**, the analytic Jacobian-transpose expert used throughout this Institute's arm work, which reaches all 24 targets at a median **26.4 joules** of actuation. That is a measurement of that controller, not a statement about the best achievable.

Because rollouts and refinement passes both cost dynamics evaluations, the two can be compared at a **matched compute budget**, which removes the objection that one simply spent more. Each row below is the same total evaluations per action, split differently.

Evaluations per action	Spent on width	Spent on depth

2,400	200 x 1 pass: **9 of 24** reached, 38.2 J	50 x 4 passes: **24 of 24**, 36.9 J
4,800	200 x 2: **20 of 24**, 37.2 J	50 x 8: **24 of 24**, 24.8 J
9,600	800 x 1: **9 of 24**, 43.6 J	200 x 4: **24 of 24**, 27.6 J
19,200	800 x 2: **19 of 24**, 30.9 J	200 x 8: **24 of 24**, 17.8 J

Depth wins every pairing, on both the task and the joules. And the floor of Section 7.1 appears here in its plainest form: at one pass, **200 rollouts and 800 rollouts both reach 9 of 24**. Four times the parallel work, the same result.

With enough refinement the sampling controller uses less energy than the reference it is measured against. At 200 rollouts and eight passes it reaches every target at a median **17.8 joules** against the expert's 26.4, which is 33 percent less actuation for the same task. The compute that bought that is sequential, so it is the compute a parallel machine cannot shorten, and on this arm it is also the compute that pays for itself in the actuation term.

The controller was not optimising energy. Its objective is distance to target plus control effort; the power model appears nowhere in the rollout cost and is only evaluated on the true body, as a measurement. The 33 percent came from reaching more directly rather than from being asked to save joules, which is worth separating: the result is what better planning is worth in the actuation term, not what an energy-aware objective could achieve. An objective that does include the power model is the subject of Section 7.5.

A selection effect, and which way it cuts. Median energy is taken over the targets a configuration actually reached, so a row reaching 9 of 24 is reporting the nine easiest. Its energy therefore looks better than it is, and the rows that reach everything are being compared against a flattered opponent. The comparison is conservative in the direction of the conclusion rather than favourable to it.

The two published rates are the same measurement once stated as a budget. A policy running at H hertz with N evaluations per action has 1/(H x N) seconds for each evaluation, whatever machine it runs on, and that is the quantity neither source states. The distilled one-step policy at 62 Hz gets **16.13 ms** per evaluation [18]; the one-step flow

policy at 71 Hz gets **14.08 ms** [19]. They differ by 1.145 times as budgets against 14.5 percent as rates, so quoting one against the other compares model sizes rather than approaches. The undistilled row is the one that changes the reading: 1.5 Hz across 100 evaluations leaves each evaluation only **6.67 ms**. **Distillation made every evaluation 2.42 times more expensive and ran a hundred times fewer of them**, for a net 41-fold gain in rate. That literature is buying fewer sequential steps rather than cheaper arithmetic, which is the same quantity Section 6 says a settling substrate removes rather than reduces.

7.5 What if you price the joules while choosing?

Section 7.4 measures energy but does not optimise it: the controller scores rollouts on distance and effort, and the power model is evaluated only on the body. A sampling controller can do better than that without any new machinery, because its cost function has to be evaluable rather than differentiable, so the identified power model can go straight into the rollout score. The published work on this arm class keeps the power model out of the reward and uses it as a post-hoc meter, so what follows is the measurement that puts it in [16]. All rows at 200 rollouts and eight passes, the best configuration from Section 7.4.

Weight on joules	Reached	Median J	J per completed task	Worst J	Worst time
0, the Section 7.4 controller	24 of 24	17.8	17.4	37.3	2.82 s
0.2	24 of 24	17.7	17.5	39.2	2.98 s
0.8	24 of 24	17.4	18.1	50.8	3.96 s
3	24 of 24	17.1	19.7	56.3	4.75 s

Pricing joules per tick made the delivered cost **worse, monotonically**. Nothing was gained: the best joules per completed task is at a weight of zero, and the heaviest weighting is 13 percent worse than not pricing energy at all.

The mechanism is the idle draw, and it is visible only in the tail. Median time is flat across the whole sweep at 1.33 to 1.35 seconds, so the typical reach is untouched. The worst case is not: it goes from 2.82 to 4.75 seconds, and its energy from 37.3 to 56.3 joules. The arithmetic closes on the standing 9 watt draw. Of the 13.5 joules of extra worst-case energy at a weight of 0.8, **10.3 joules, or 76 percent, is idle draw over the extra 1.14 seconds**, and idle accounts for 68 to 70 percent of worst-case energy at every weight. A controller told to spend less per tick becomes reluctant on the targets that need decisive motion, and standing still on a body is not free.

Two statistics, opposite conclusions, same runs. Read the median column and this sweep is a 4 percent improvement. Read joules per completed task and it is a 13 percent regression. The median is measuring the targets that were never the problem, and a median cannot report a cost that lives entirely in a tail. Where a mean and a median move in opposite directions, the mean is naming where to look.

Two things follow for an energy-aware objective, and neither is that the idea fails. The quantity to price is the **completed task** rather than the instantaneous rate, which means a term for time or for predicted total energy to completion rather than for power in the next tick. And any such objective has to be scored on the tail, because that is the only place its cost appears.

Scope. Sections 7.1 to 7.3 use linear plants with an exact oracle, which is what makes the comparison against truth possible and also what bounds it: those are model systems. Section 7.4 uses a physical arm model with an identified power model and no exact optimum, so it is anchored on a reference controller and says so. The two probes use different noise and temperature settings, so their absolute numbers are not comparable with each other; each comparison above is within a single probe. Section 7.2's figures are all at 3,200 rollouts with noise scaled to hold total perturbation constant. A refinement pass in this controller and a denoising step in a diffusion policy are both sequential refinements of a sampled estimate, and they are not the same operation. What transfers is the structural result about which knob carries the accuracy.

Two operations are easy to read as one, and the distinction sets the number above. *Restreaming* loads a whole model onto a fabric and is slow, on the order of a problem per second. *Clamping* imposes new boundary conditions on a model already resident and is much faster. A controller does the second on every tick, so the clamp rate is the quantity that governs closed-loop use, and the 100 millisecond figure above is the measured end-to-end latency of a system performing it. The two rates differ by roughly two orders of magnitude, so which one a comparison uses decides whether the gap to control rates reads as one to two orders or as three to four.

8. Can two workloads share one substrate?

Transformer attention is being mapped onto the same class of device. An analog in-memory architecture built on gain cells reports up to a 70,000-fold reduction in energy and a hundred-fold speed-up against graphics processors for a 1.5 billion parameter model, reaching text-processing performance comparable to GPT-2 through an initialisation algorithm rather than retraining [9]. Attention accounts for roughly seventy to eighty percent of large language model inference energy.

The two workloads use the device differently, and the difference is physical.

	Attention in an analog memory array	**An energy-based model on a sampling fabric**
What the device performs	analog multiply-accumulate	occupation of states with Boltzmann probability
Relation to the algorithm	**acceleration**: the arithmetic is faster and the algorithm is unchanged	**realisation**: the resting distribution is the sampled distribution
What it asks of the device	linearity, precision, low drift	statistics that match the intended distribution
Role of device fluctuation	a source of error, to be suppressed	the source of randomness, to be characterised

That last row is a statement about what each workload needs from the same silicon. Analog computing has a long record of difficulty in production, and the reported causes are consistent across two decades of reviews: conductance variability, non-linear and asymmetric switching, drift, mismatch, and the integration of emerging non-volatile memory into standard flows. Digital neuromorphic parts are closer to commercial availability; analog crossbar in-memory computing is at an earlier stage.

> **Those device properties enter the two cases differently.** An arithmetic result is displaced by fluctuation, so the acceleration case improves as fluctuation falls. A sampled distribution is defined by fluctuation, so the realisation case improves as

> fluctuation is characterised. Both are engineering programmes; they are not the same programme, and a device that is early for one may be usable for the other.

9. Is sampling a general-compute workload?

Inference-time sampling has become routine across the field: generate candidates, score them, select. Coverage rises log-linearly with the number of samples, and verifier-based selection scales more robustly than verifier-free selection. Purpose-built inference silicon has begun to include dedicated acceleration for token sampling.

The Institute's benches show the same shape from the other side. Using an energy to **select** among candidates rather than to **descend** a landscape reached complete accuracy on conjunctions where gradient descent on the same landscape was sensitive to step size [14]. Generate-and-verify is an energy-based procedure in structure.

Beyond machine learning the sampling class is long established: Bayesian inference, Monte Carlo, uncertainty quantification, combinatorial optimisation. And the economics of the substrate transition are measurable at the level of the memory hierarchy. A dynamic-memory access costs on the order of three thousand times a sixteen-bit multiply at a 45 nanometre node, data movement runs roughly a hundred times the arithmetic, and a large majority of energy in a conventional machine is spent moving operands rather than operating on them. These figures describe **machine-level energy per operation**, which is one of the three denominators separated in Section 10.

10. Which three denominators, and what does each support?

Efficiency in this field is stated in at least three different denominators, and a figure is only comparable to another figure in the same one. Naming them is what allows the datacentre results of Section 9 and the embodied results below to be read together.

Denominator	What a figure in it describes	What it supports
Energy per operation	the cost of an arithmetic result or a sample on a given device	device and architecture comparisons, including the memory-wall figures of Section 9
Energy per completed task on a body	compute and actuation together, integrated over a closed loop until the task succeeds	deployment planning for a robot, and bounded as below
Time per action	latency against a control period	whether a policy can run at a given control rate, which is Section 7

In the second denominator, compute is one term of a sum. On a reaching task the compute share runs from five percent at one watt of compute to fifty-nine percent at thirty watts, and on a mobile robot under autonomous navigation the graphics processor alone draws 37.3 percent while the motors draw 16.6 percent [13]. Improving compute efficiency by a factor leaves the non-compute term unchanged, so total task energy falls to the non-compute share plus the compute share divided by that factor, and the limit is the reciprocal of the non-compute share.

The share is a property of the body and the workload, and it moves further than this report first assumed. Measured since publication on the same three-joint arm as Section 7, one dynamics evaluation of the controller's rollout loop costs about two ten-millionths of a joule at the wall, and a completed reach costs 2,592,000 of them, so the planning costs roughly 0.6 joules against 17.8 joules of actuation [17]. That is a three percent compute share, and the motors outweigh the processor about thirty to one. Set beside the mobile robot's 37.3 percent it spans more than a factor of ten across two bodies, which is the useful form of the result: **a whole-robot energy ceiling has to name the machine it was computed for.**

On a humanoid the same arithmetic gives the policy budget directly, from the vendor's own published numbers. Unitree publishes a 9 Ah pack, a 54 V charger and about two hours of battery

life for the G1 [20]. A 13-series lithium pack sits near 48 V nominal, so the pack holds 432 to 486 Wh and the whole machine averages **216 to 243 W** across the stated runtime. At a 50 Hz whole-body control rate, sitting at the arm's three percent means a policy drawing 6.5 to 7.3 W, and sitting at the mobile robot's 37.3 percent means 81 to 91 W. **The two published shares are 12.4 times apart in policy power on one unchanged body**, and that ratio does not depend on which body figure is used. The share is therefore something a designer chooses, and the measurement of which model size lands where has now been taken [22]. The standing 2026 survey of how artificial intelligence changes energy use across robotic platforms is the work to position it against [21]. These are derived from a specification rather than metered, and are stated as such.

Measured at batch one across seven policy sizes, every one of them lands below the band's floor. On a laptop-class SoC against an idle baseline whose wander stayed within eight percent of every signal, a policy evaluation costs from 5.0×10^{-5} joules at 49 thousand parameters to 5.9×10^{-2} joules at 75.5 million [22]. At a 50 Hz whole-body rate even the 75-million-parameter policy draws 2.9 W, which is **1.2 to 1.4 percent** of the G1's average body power: below the arm's three percent, and a twenty-fifth of the mobile robot's 37.3. The prediction recorded in the bench expected the tested range to reach into the band and was falsified in the informative direction. **A control policy is not where a humanoid's compute share comes from: the 37.3 percent figure is a full autonomy stack, perception and planning and infrastructure, not a policy.** The joules per evaluation sit flat near 7×10^{-10} per parameter across the three largest sizes, the bandwidth-bound regime, and extrapolating that trend, stated as an extrapolation rather than a measurement, a policy would need to reach roughly **two billion parameters** before it claimed the mobile robot's share of this body at this rate. Both published one-step budgets of Section 7.5 admit every size tested, the largest with 12.2 ms to spare.

Compute share of a completed task	Limit on whole-robot saving
3 percent, three-joint arm, planning only, measured	1.03x

37.3 percent, mobile robot, GPU alone	1.60x
42 percent, reaching at 15 W	1.72x
59 percent, reaching at 30 W	2.44x
74 percent, reaching at 60 W	3.85x

At the 59 percent share a thirty-fold compute improvement delivers 95.4 percent of what an unbounded improvement could deliver in this denominator. A figure of 70,000-fold stated in energy per operation and a limit of 2.44x stated in energy per completed task are therefore both correct and describe different quantities. At the arm's three percent share the same unbounded improvement is worth 1.03x, which is the clearest statement available of why this denominator has to be quoted with its body attached: the identical substrate advance is worth a factor of two on one machine and three percent on another.

> **This is why the embodied case in this report is developed in the third denominator.** Time per action has no bound of this kind: it is set by the control period, and Section 7 shows the two substrate costs meeting inside the operating range. Two further quantities behave the same way. **Duty cycle:** event-driven perception idles near hundredths of a watt against tens of watts for continuous inference, and a deployed machine spends most of its time between tasks rather than in them, so the completed-task denominator does not govern that interval. **Thermal envelope:** a body dissipates through its own surface area, which is a physical limit on sustained power rather than an economic one.

11. What is established, and what would get us there faster?

Established. That the sampling operation is sequential and that lane count does not enter its wall-clock time, which Section 3 states and Lesson 5 measures. That the score and energy parameterisations describe the same object, which Lesson 6 verifies numerically. That the dominant visuomotor policy class samples an energy at inference at ten to one hundred evaluations per action, which its literature states [6]. That an embedded sampler has a measured end-to-end latency of about 100 milliseconds independent of problem size [11]. That total task energy on a

body bounds any compute improvement to the reciprocal of the non-compute share, which across the bodies measured so far runs from 1.03 times on a three-joint arm doing planning only [17] to 2.44 times on a reaching task at thirty watts of compute [13]. An earlier revision of this report gave that range as roughly 1.6 to 2.4 times, which was the span of the sources then in hand and understated it at the low end.

Open. Whether energy-parameterised models are competitive at frontier scale, which the Institute's own results at 800 million parameters and below do not settle in either direction [14]. Whether conditioning latency reaches control rates, where roughly one to two orders remain at 100 Hz and the published improvements are compression and estimation rather than new physics [11]. Whether analog device statistics can be characterised closely enough to serve as a sampling source at scale, which is an active materials and circuits programme.

Prior art. Boltzmann machines, Ising annealing, equilibrium propagation, MPPI, diffusion policies and in-memory computing are established fields with substantial literatures. What this report contributes is their joining: the mechanical reading of the operation in Section 3, the acceleration and realisation distinction in Section 8, conditioning bandwidth as the governing quantity for embodied use in Section 7, and the separation of denominators in Section 10. This review did not locate these three stated together in the literature it read, and the field's consensus vision document, which makes intelligence per joule its stated north star and treats embodied AI at length, did not return energy-based models or thermodynamic computing as a named workload direction on the seeds used [12].

12. Conclusions

Energy-based models require sampling, sampling is sequential, and the hardware generation that carried the field forward accelerates work that divides into independent pieces. That is a relationship between an operation and a machine, and it is measurable. Over the same period the underlying idea continued to develop in the parameterisation whose gradient costs one forward pass, and it is that parameterisation which carries the field's public models at scale. Physical AI makes the cost

visible, because an embodied policy needs to represent a set of correct actions and must produce one within a control period, and the dominant policy class pays ten to one hundred sequential network evaluations to do so. A substrate whose resting distribution is the sampled distribution removes the sequence rather than shortening it, and on one published range and one measured latency the two costs meet inside the region already in use. The quantity that distinguishes this case from the parallel case of analog attention is what each asks of device fluctuation: one characterises it, the other suppresses it.

13. The forcing function

What is bounded	The physics that sets it	The engineering change that moves it	What becomes possible
Ten to one hundred network evaluations per action	Sampling an implicit landscape by iterated denoising on a machine that represents the distribution numerically	A substrate whose resting distribution is the sampled one, so a single settling replaces a sequence of passes	Multimodal policies at full control rate, with the latency term removed rather than reduced
Conditioning bandwidth, about 100 ms per new problem	A fabric re-equilibrates after each new set of boundary conditions	Resident models with fast clamp ports; coupling compression, measured at 32-fold; estimated rather than tuned parameters	Roughly one to two orders remain to robot control rates, and each is an interface and memory-hierarchy problem
Fitting requires a sequential sampler	The intractable partition function	Equilibrium propagation, where the device supplies its own gradient by settling twice	Learning on the device at the edge, without a stored activation tape
Analog device statistics vary and drift	Nanoscale device physics	For arithmetic, tighter linearity and lower drift. For sampling, characterisation of	A device can become usable for the sampling case at a maturity earlier

		the distribution the device produces	than the arithmetic case requires
Whole-robot energy saving, limit 1.03x to 2.44x depending on the body	Total task energy is compute plus actuation, and the compute share is a property of the machine rather than of the field	Work on the actuation term: winding and gearing design, thermal management, trajectory shaping against an identified power model	The larger term of the sum becomes addressable, and on the bodies where the motors dominate thirty to one that is where nearly all the joules are. The embodied case is developed in latency, duty cycle and thermal envelope, where no such limit applies

Energy-based models were **out of reach at scale** when every gradient step required a sequential sampler on a machine built to divide work across lanes. They are **demonstrably productive today**, in score parameterisation, at the head of generative modelling and of robot policy learning, with a physics lineage recognised by a Nobel Prize and a sampling cost that the robotics literature identifies as its real-time constraint. They become **ordinary** when a substrate holds the distribution physically and can be conditioned at control rate, which requires no new physics: it requires conditioning bandwidth, device characterisation, and an interface, which is the most tractable class of obstacle there is.

> **A companion reading guide is published separately.** Eight questions, answered plainly, with the seventy-eight works that answer them. What is an energy-based model. What makes a landscape hard to learn. Why today's models predict slopes instead of landscapes. What settling costs a robot, and what it buys. What kind of machine fits, and how close it is. Written for a reader meeting this field for the first time, with each work placed on the road from a proof to a fabricated chip. Start with the questions.

References

1. Hooker, S. *The Hardware Lottery*. arXiv:2009.06489 (2020). READ · FRAMING

2. The Nobel Prize in Physics 2024, press release and popular science background, Royal Swedish Academy of Sciences. READ · PRIMARY
3. Ackley, D., Hinton, G. and Sejnowski, T. *A Learning Algorithm for Boltzmann Machines*. Cognitive Science 9, 147 (1985). CITED VIA PRIOR INSTITUTE RECORD
4. *Implicit Behavioral Cloning*. Florence et al., Conference on Robot Learning 2021, arXiv:2109.00137. READ · ABSTRACT AND RESULT SUMMARY
5. Denoising score matching and energy parameterisation: the equivalence of diffusion training and EBM score matching, and the dominance of score parameterisation in practice. Composite of the review literature located in this sweep. READ · SECONDARY REVIEW
6. *Diffusion Policy: Visuomotor Policy Learning via Action Diffusion*. arXiv:2303.04137; International Journal of Robotics Research 44 (2025), with its subsequent survey literature for the ten to one hundred evaluation figure. READ · ABSTRACT AND SURVEY
7. *Generalized Model Predictive Path Integral Control as Expectation Maximization*, arXiv:2606.00317; *Model Predictive Path Integral Control as Preconditioned Gradient Descent*, arXiv:2603.24489; *Model Predictive Control via Probabilistic Inference: A Tutorial and Survey*, arXiv:2511.08019. READ · ABSTRACTS
8. Werthen-Brabants, L. and Simoens, P. *Ising Machines for Model Predictive Path Integral-Based Optimal Control*. arXiv:2512.15533 (2025). READ · FULL RECORD CHECKED FOR HARDWARE AND ENERGY CONTENT, BOTH ABSENT
9. *Analog in-memory computing attention mechanism for fast and energy-efficient large language models*. arXiv:2409.19315; Nature Computational Science (2025). REPORTED · VENDOR-INDEPENDENT JOURNAL, FIGURES NOT REPRODUCED HERE
10. Scellier, B. and Bengio, Y. *Equilibrium Propagation: Bridging the Gap Between Energy-Based Models and Backpropagation*. arXiv:1602.05179 (2016), with 2026 hardware implementations including arXiv:2606.13454 and arXiv:2606.09112. READ · ABSTRACTS
11. Hamakawa, Y., Kashimata, T., Yamasaki, M. and Tatsumura, K. (Toshiba). *Machine Learning-assisted High-speed Combinatorial Optimization with Ising Machines for Dynamically Changing Problems*. arXiv:2503.23966; Nature Communications (2026). READ · FULL RECORD, LATENCY AND COMPRESSION FIGURES VERIFIED
12. *AI+HW 2035: Shaping the Next Decade*. arXiv:2603.05225 (2026). READ · FULL RECORD, CHECKED FOR ENERGY-BASED MODEL CONTENT
13. Liu, Shi and Shin, arXiv:2511.20467, mobile robot power decomposition, via TR-2026-36. MEASURED · THIRD PARTY
14. Institute for Physical AI @ JBI, Charlot Lab, internal energy-based model bench series and total task energy bench. MEASURED · OURS
15. Institute for Physical AI @ JBI, *Post-von Neumann and Energy-Efficient Computing Paradigms for Physical AI at the Edge: A Survey*, Technical Report TR-2026-01. The substrate taxonomy this report builds on and does not repeat. OURS · PRIOR REPORT
16. Institute for Physical AI @ JBI, Charlot Lab, `ferrotherm` Section 7 bench series: `examples/sequential_depth.rs` (accuracy against the exact Riccati optimum, sweeping rollouts and refinement passes), `examples/depth_vs_dimension.rs` (the same question

against action-space dimension, four couplings and an underactuated plant), and `examples/arm_width_vs_depth.rs` (three-joint arm with gravity, actuator lag, torque limits and the identified power model). Median of seven seeds, spread reported alongside. MEASURED · OURS

17. Institute for Physical AI @ JBI, Charlot Lab, `ferrotherm gpu/examples/planner_joules.rs`: the compute half of a completed task, measured at the wall against an idle baseline on the SoC's own power counters. Two clean passes on separate runs; passes whose baseline wandered are printed and set aside rather than averaged in. MEASURED · OURS
18. One-Step Diffusion Policy: Fast Visuomotor Policies via Diffusion Distillation. Wang, Li, Mandlekar et al., arXiv:2410.21257 (2024). Source of the 1.5 Hz and 62 Hz figures. READ · ABSTRACT AND RESULT
19. ElasticFlow: One-Step Physics-Consistent Policy with Elastic Time Horizons for Language-Guided Manipulation. Chen, Long, Li & Long, arXiv:2605.08799 (2026). Source of the ~71 Hz figure at one function evaluation. READ · ABSTRACT AND RESULT
20. Unitree Robotics, G1 published specification: 9000 mAh smart battery, 54 V charger, "about 2h" battery life, about 35 kg, 23 degrees of freedom. The pack energy and average draw above are arithmetic on those figures. VENDOR SPECIFICATION · DERIVED, NOT METERED
21. Vodovozov, V. and Raud, Z. *Assessment of AI Impact on Energy Utilization in Robotics*. Energies 19(14), 16 July 2026, doi:10.3390/en19143364. The standing 2026 survey on this axis. Metadata verified via Crossref; the full text was not reachable from this review's tooling, and no measured compute-versus-actuation split for a humanoid was located in the 2026 sources searched. MAPPED · SURVEY
22. Institute for Physical AI @ JBI, Charlot Lab, `research/ebm_substrate/policy_budget.py`: joules per policy evaluation at batch one across seven sizes from 49 thousand to 75.5 million parameters, single-stream latency and all-cores marginal energy measured separately, each size against a fresh idle baseline and refused when the baseline wanders past a quarter of the signal. Run record preserved beside the bench. Every intermediate activation is asserted finite before the platform BLAS's spurious floating-point-flag warnings are silenced. MEASURED · OURS