Learning to erase quantum states

thermodynamic implications of quantum learning theory

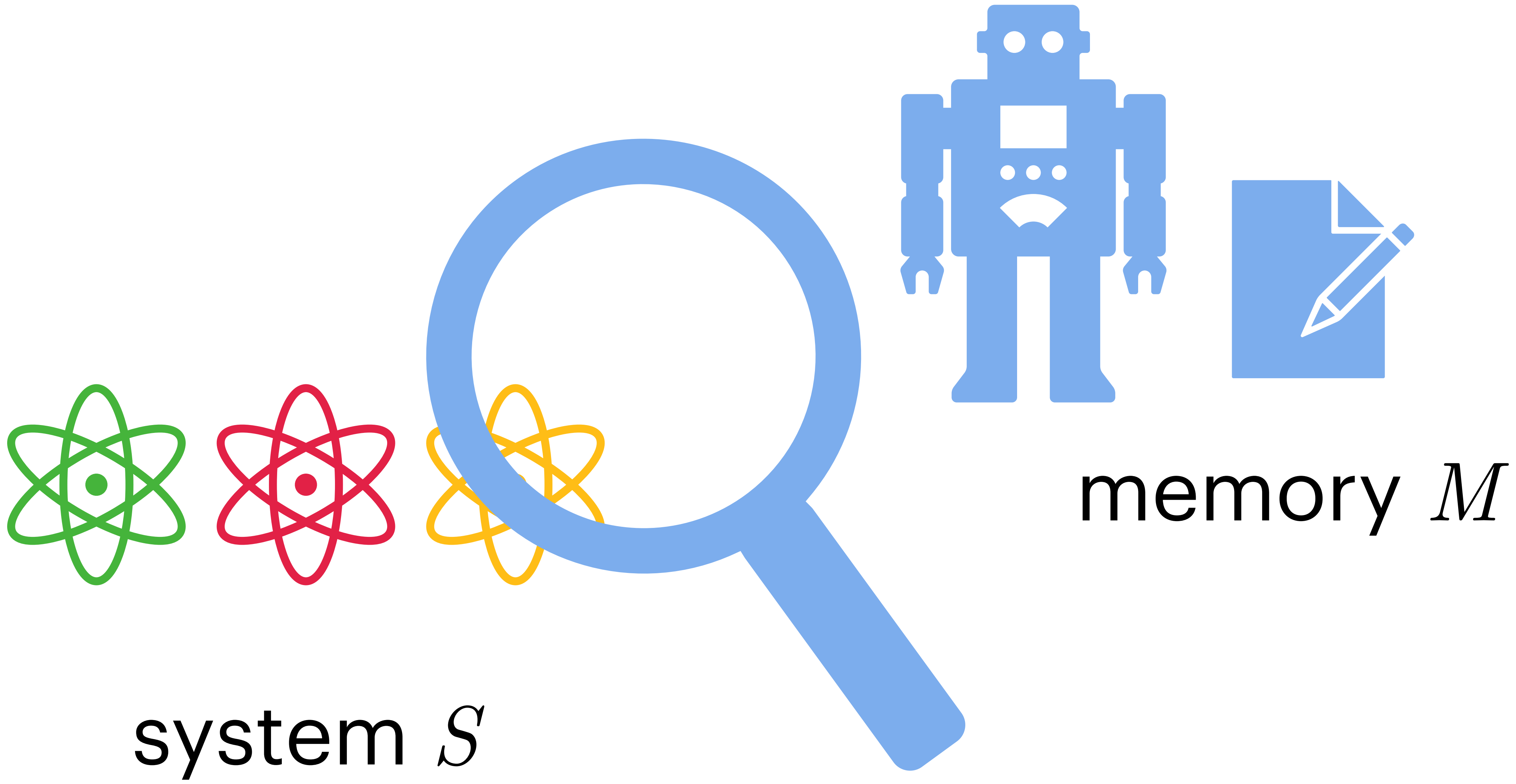
Haimeng Zhao
Caltech
haimeng@caltech.edu

Joint work with Yuzhen Zhang, John Preskill

arXiv:2504.07341

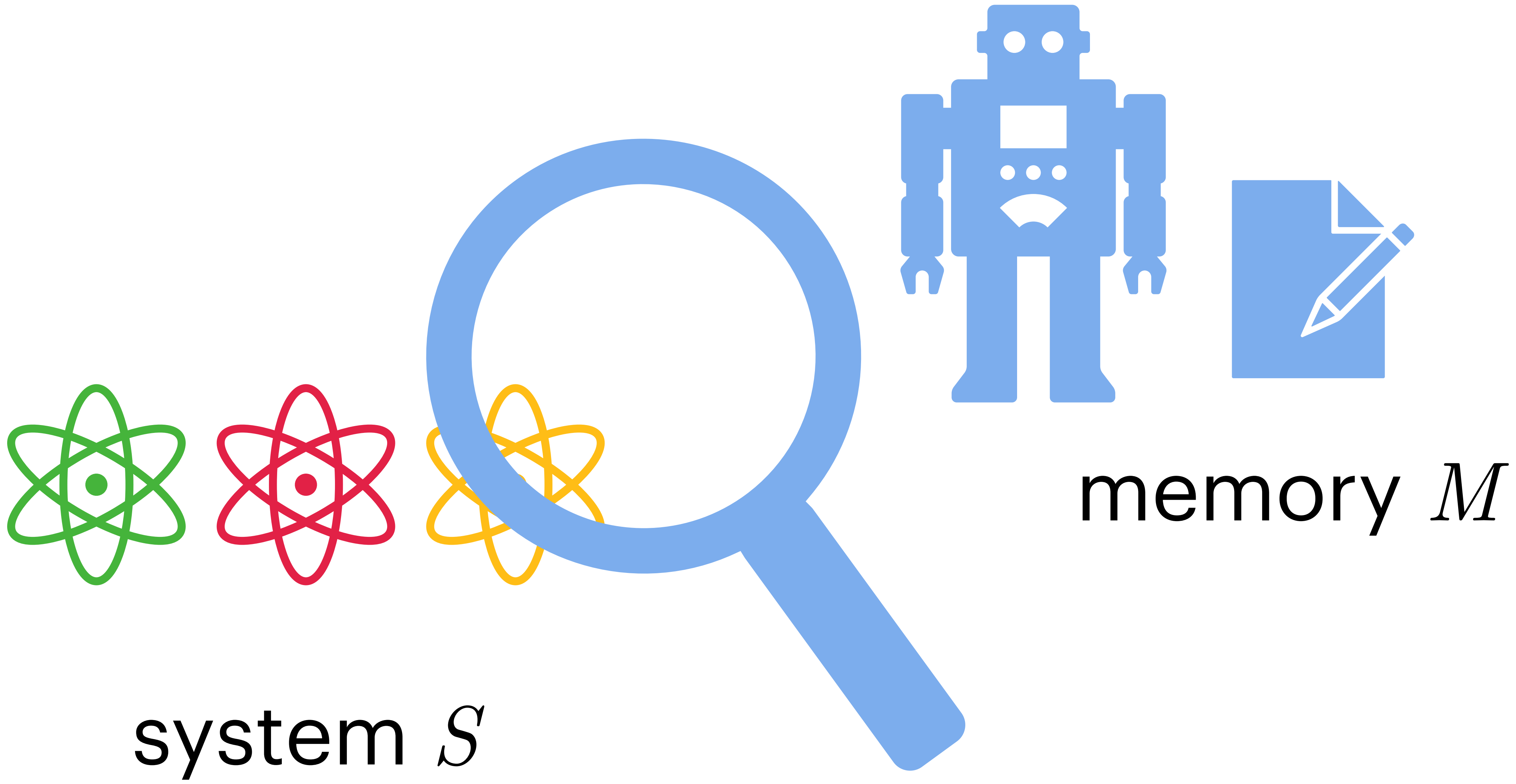
IQIM Caltech UC SANTA BARBARA aws

Motivation



Learning is a physical process.

Motivation

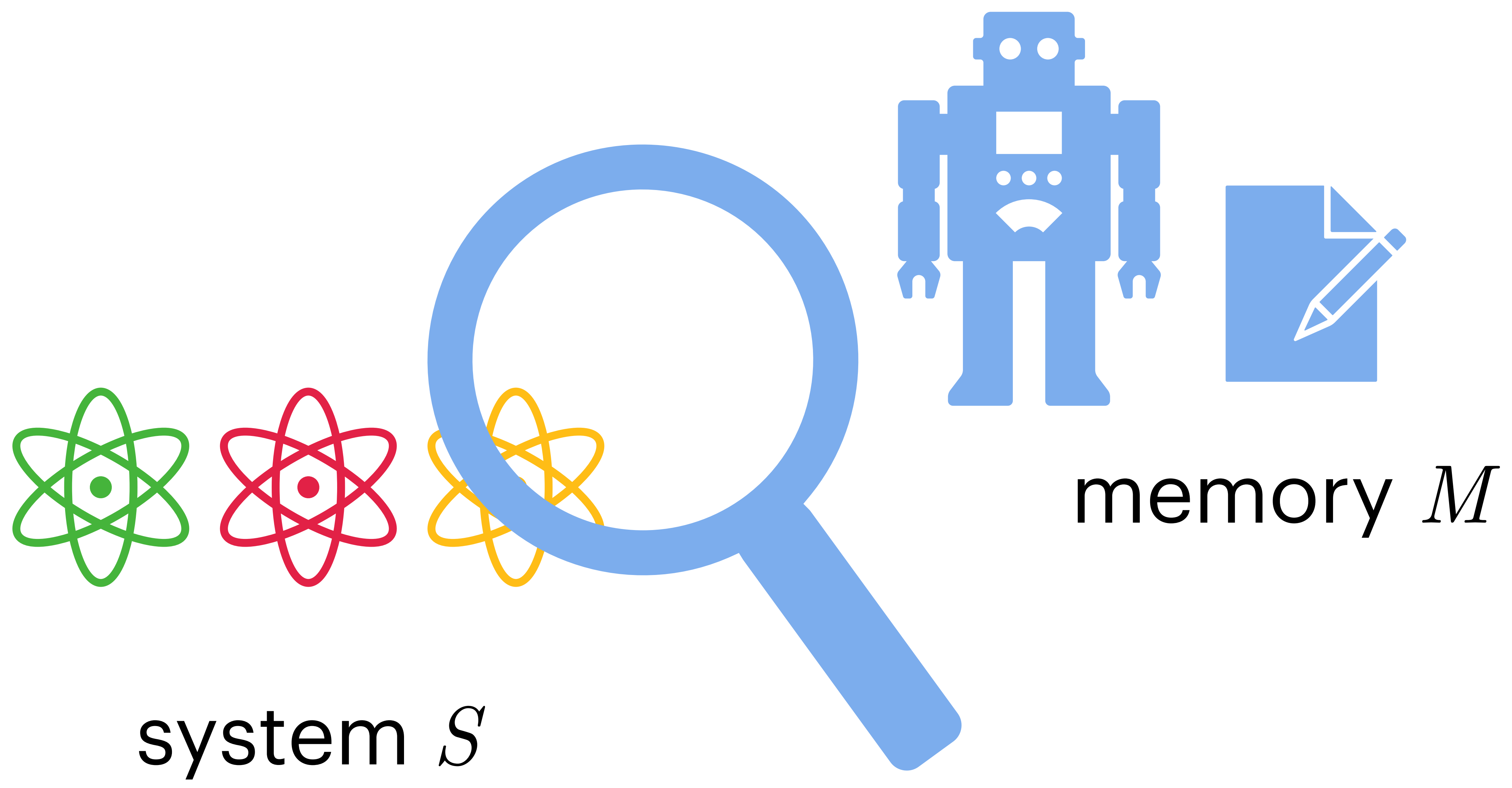


Learning is a physical process.

Questions:

1. *What physical properties does learning itself have?*

Motivation



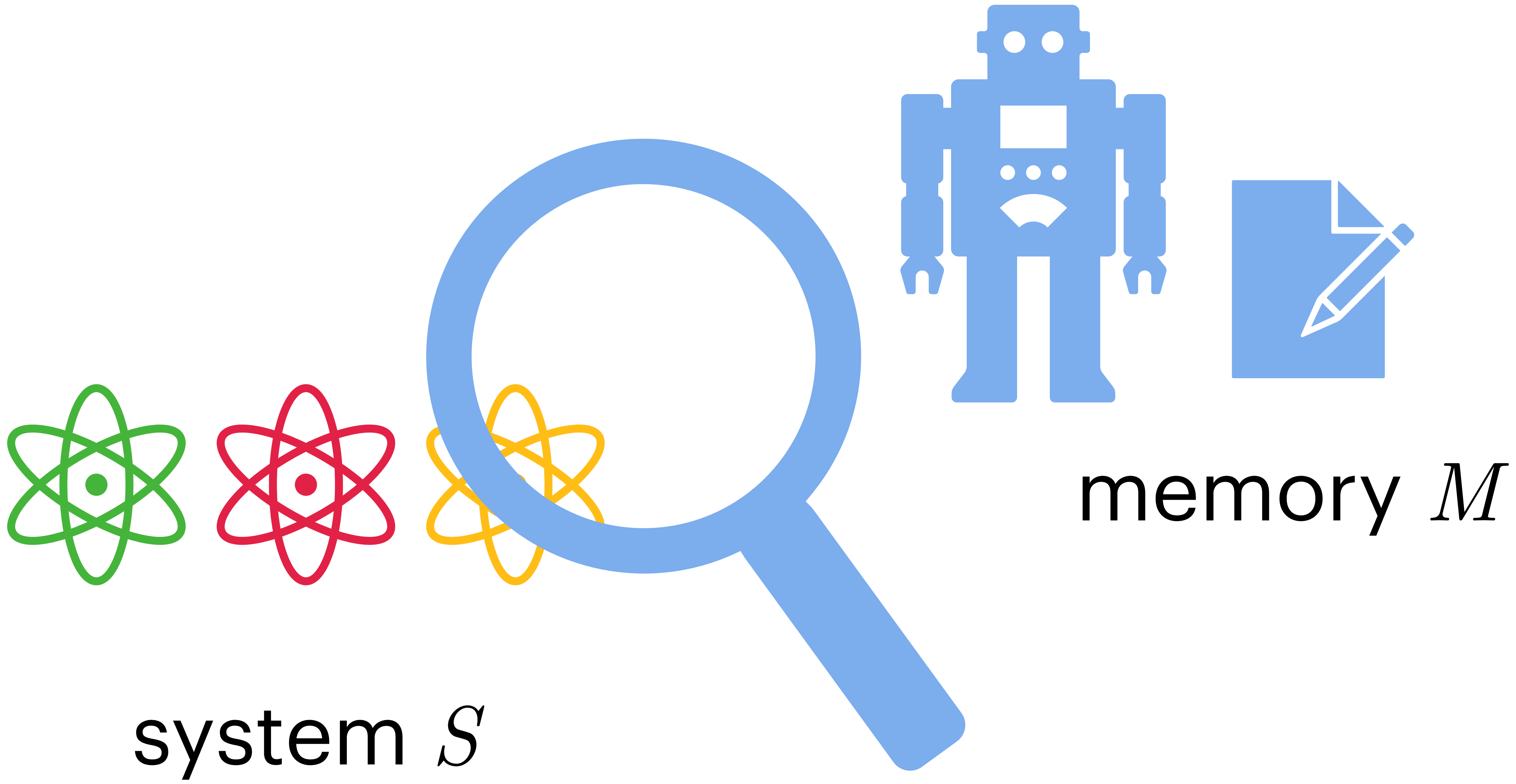
Learning is a physical process.

Questions:

1. *What physical properties does learning itself have?*

2. *What tangible physical consequences do abstract learning processes have?*

Motivation



Learning is a physical process.

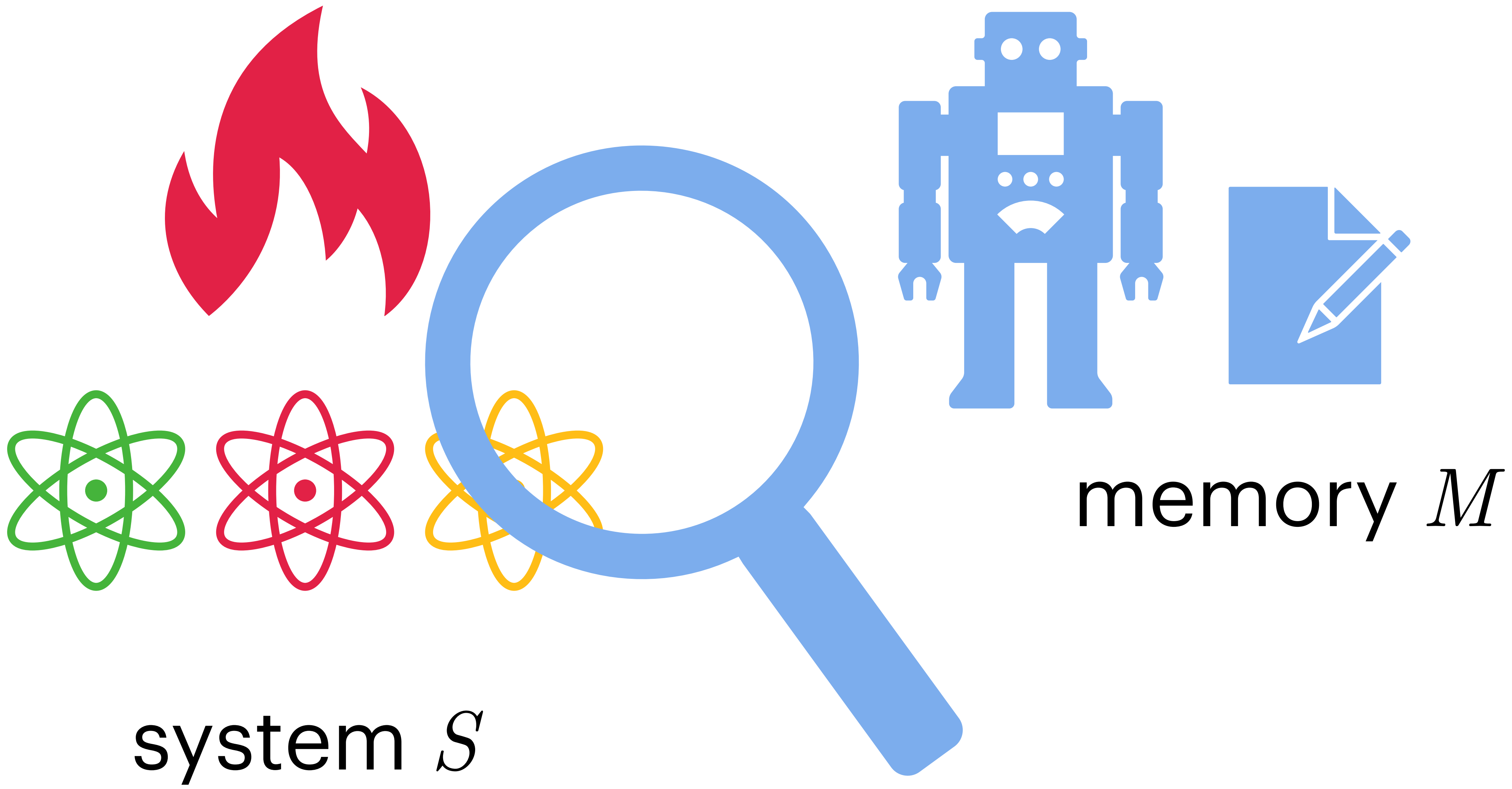
Questions:

1. *What physical properties does learning itself have?*

2. *What tangible physical consequences do abstract learning processes have?*

   Does our (in)ability to learn impact the amount of physical resources needed for certain tasks?

Motivation



Learning is a thermodynamic process.

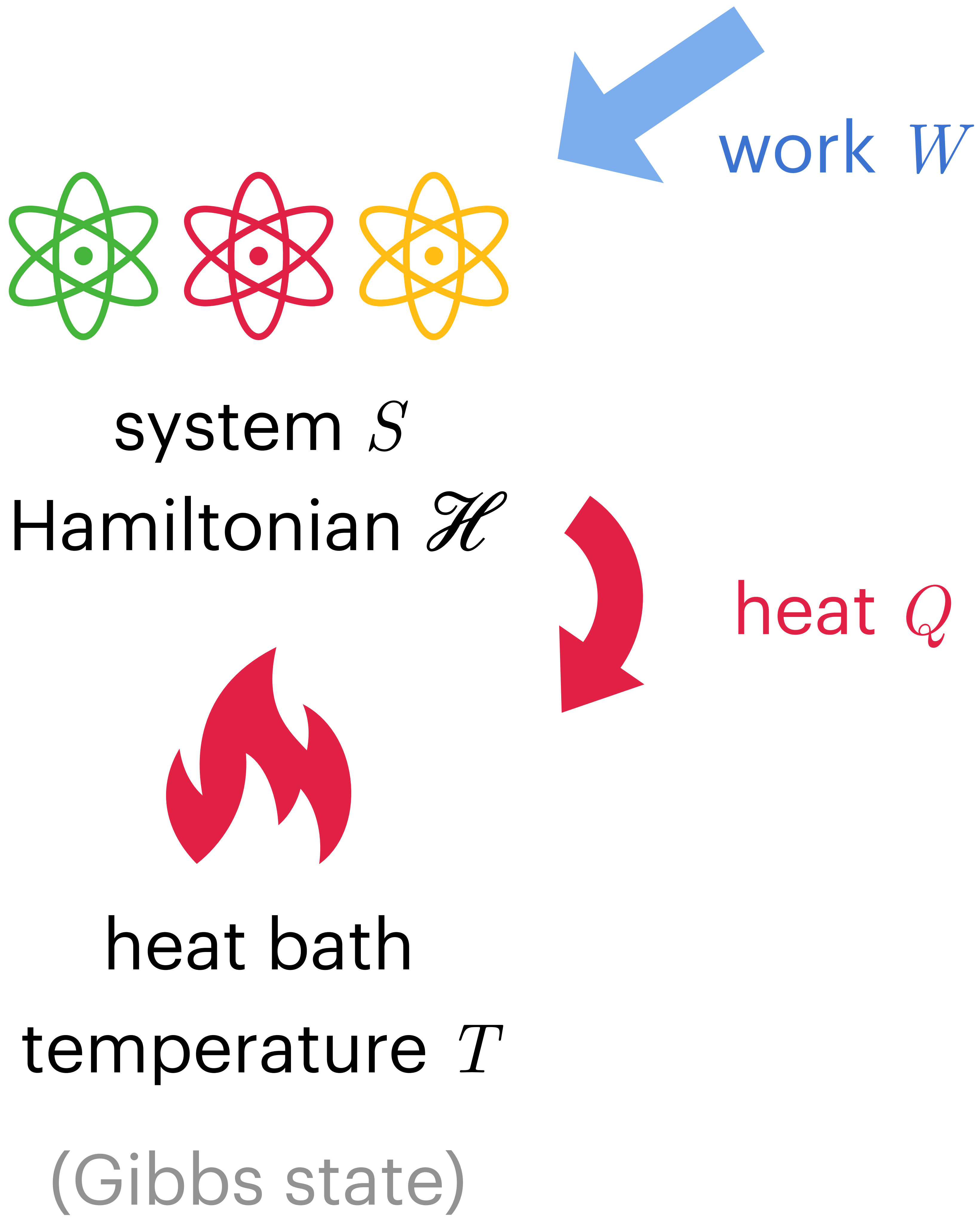
Questions:

1. *What physical properties does learning itself have?*
   No fundamental energy cost.
2. *What tangible physical consequences do abstract learning processes have?*

   Does our (in)ability to learn impact the amount of physical resources needed for certain tasks?

   Yes! Energy cost/gain in erasure/work extraction.

Thermodynamics

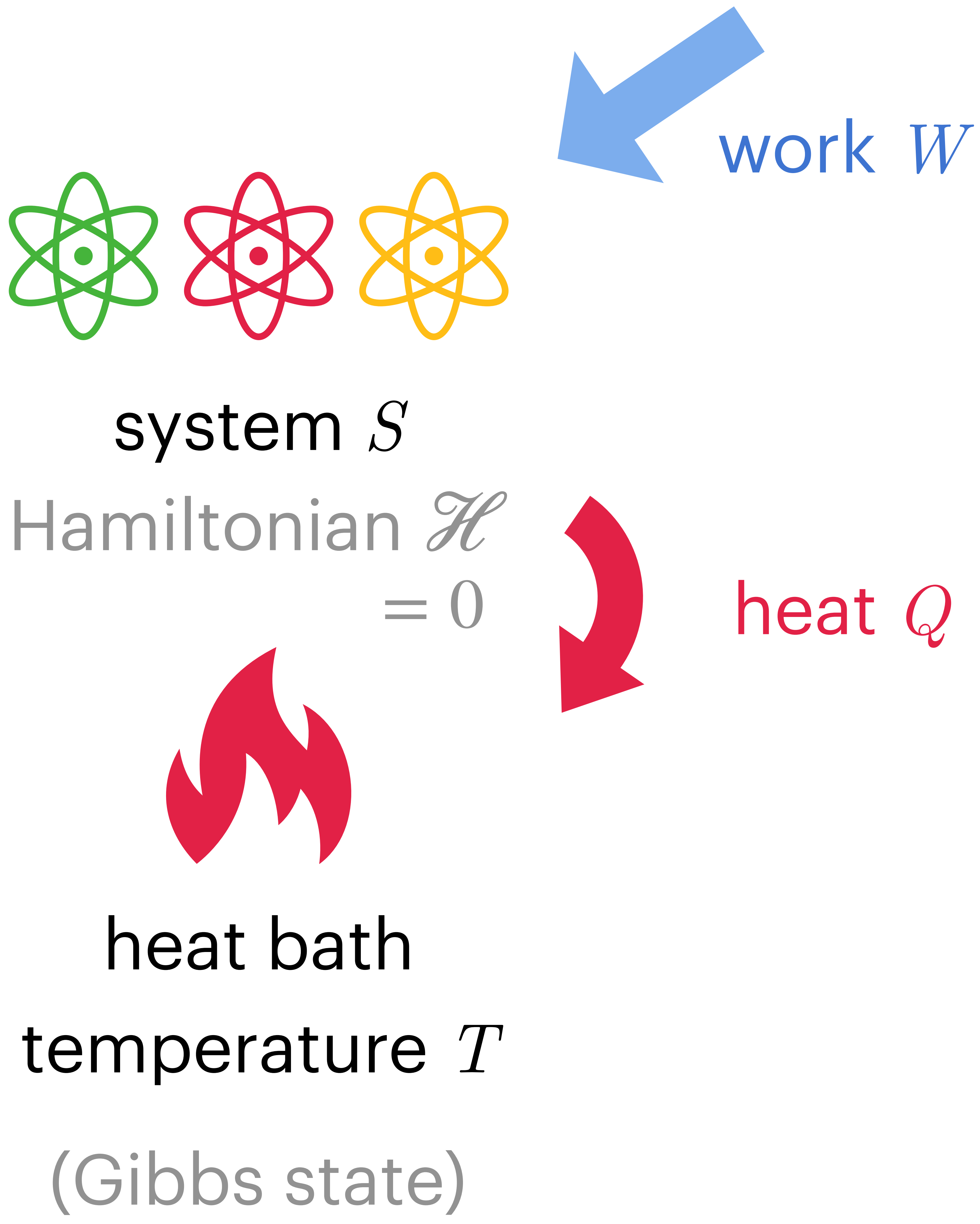


$$W = \Delta E + Q$$

Hamiltonian dependent

irreversibility

Thermodynamics

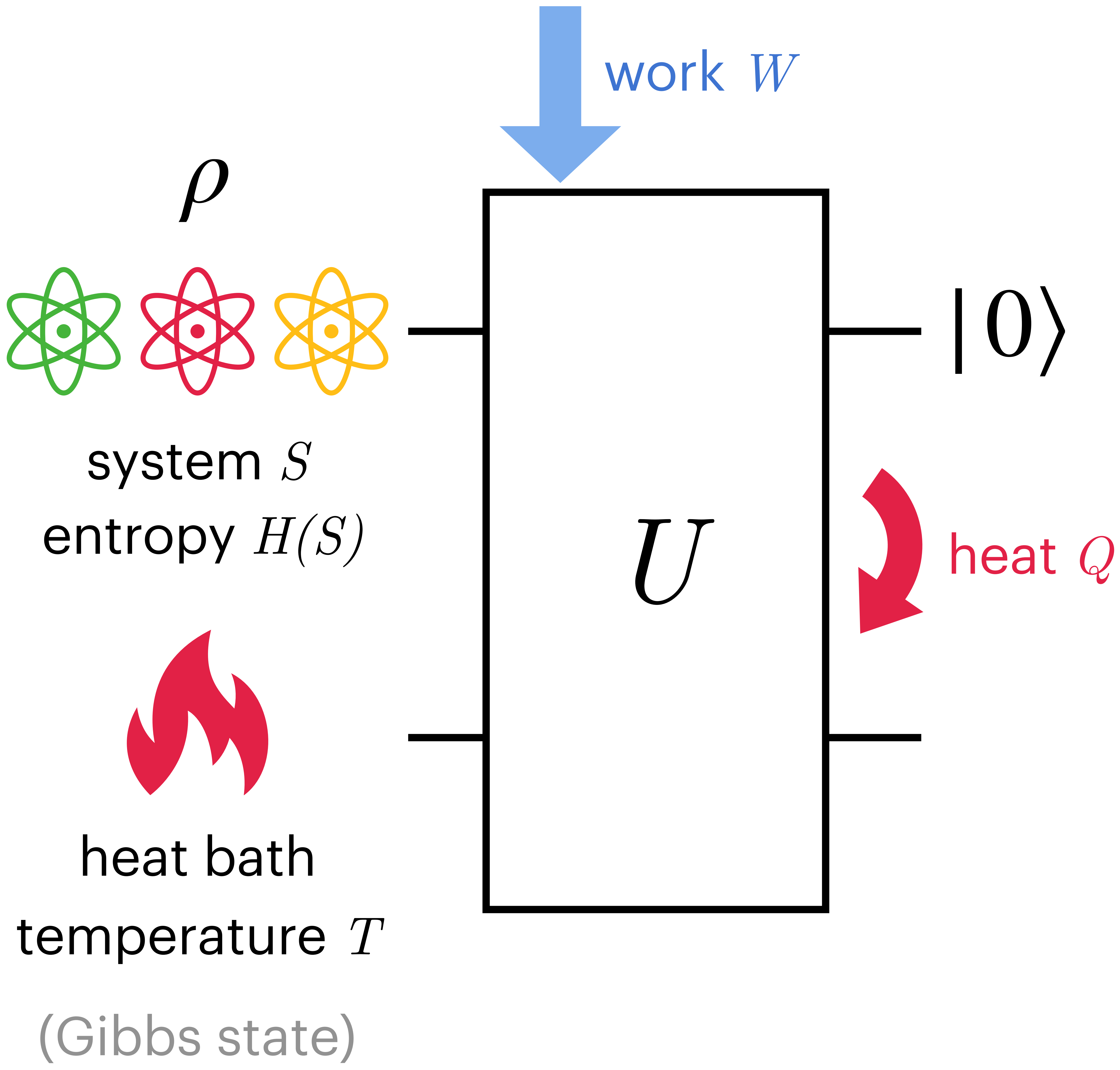


$$W = \cancel{\Delta E} + Q$$

Hamiltonian dependent

irreversibility

Erasure

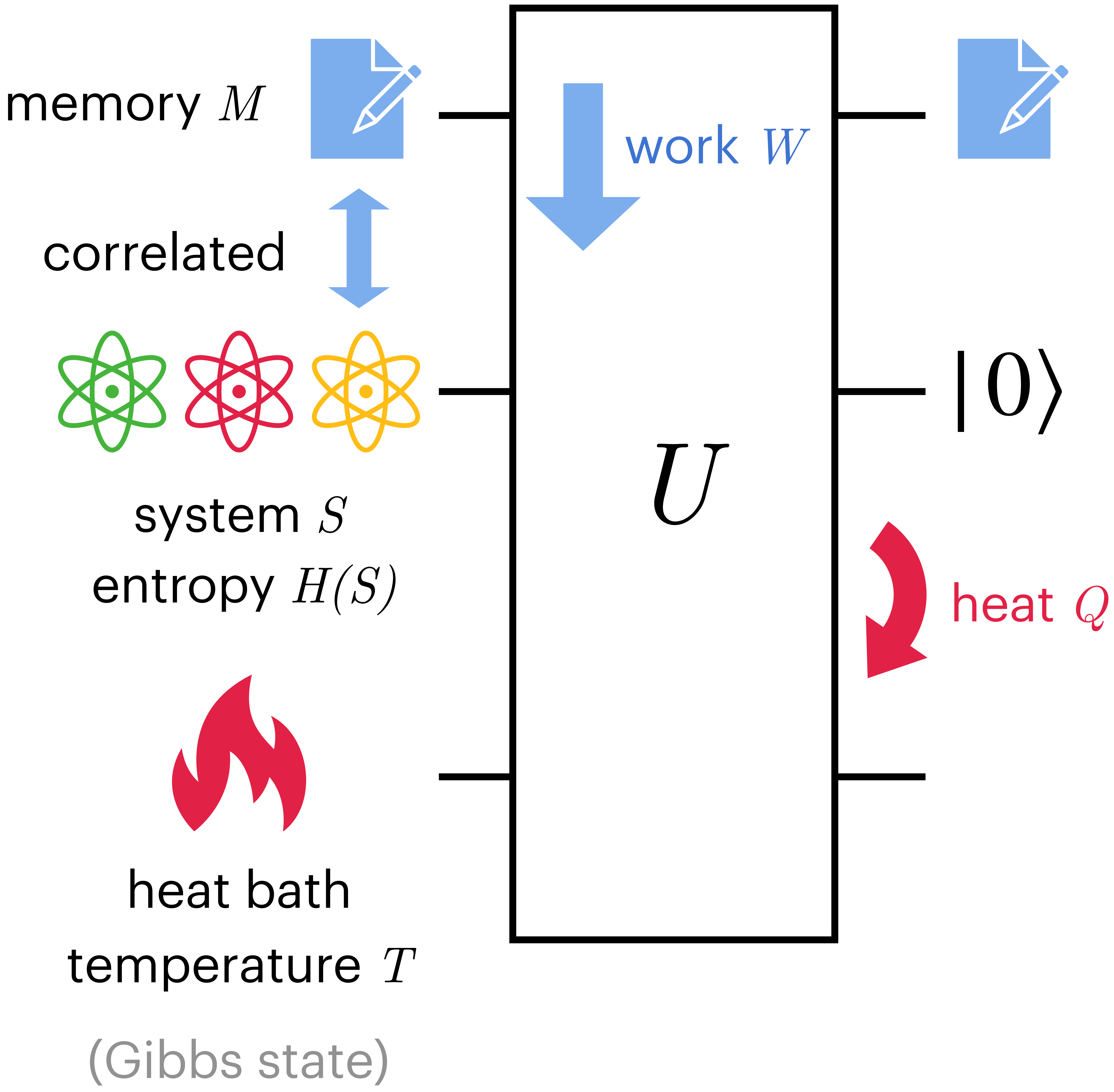


k: Boltzmann constant

$$W = Q = H(S)kT\ln 2$$

Landauer's principle

Reeb & Wolf New Journal of Physics 2014



Erasure with side info

An extreme case: $\rho_{SM} = \sum_x p_x |\psi_x\rangle\langle\psi_x|_S \otimes |x\rangle\langle x|_M$

$H(S|M)=0$

$$W = Q = H(S|M)kT\ln 2$$

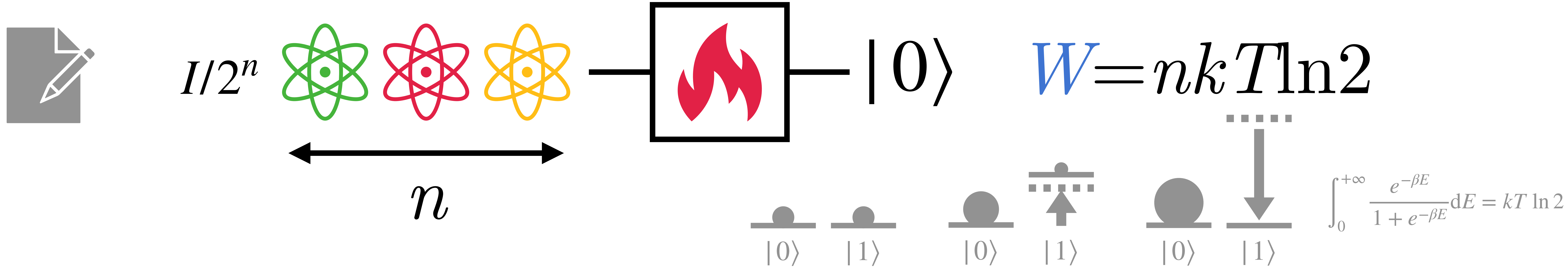
Landauer's principle

For quantum memory, $H(S|M)$ may be negative.

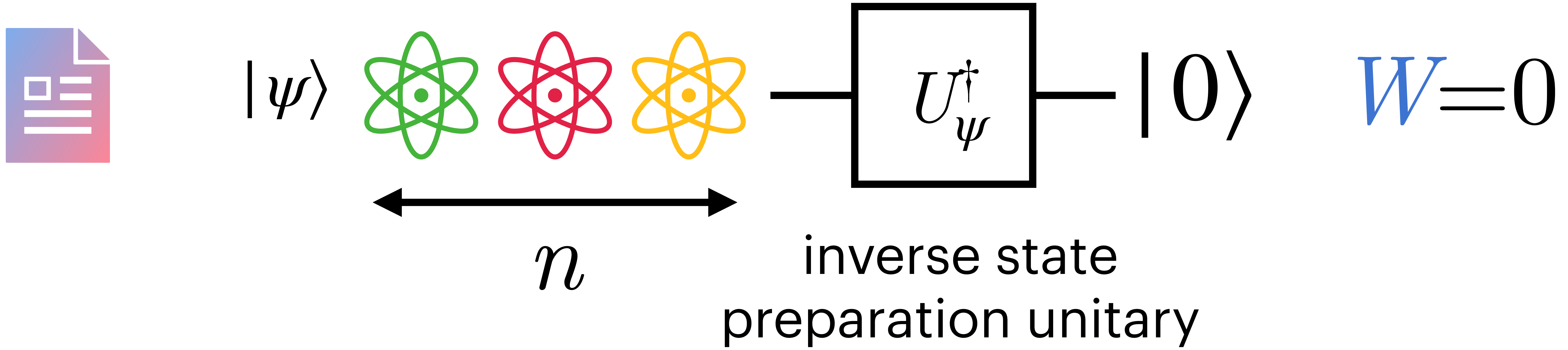
del Rio et al. Nature 2011

Erasure with side information

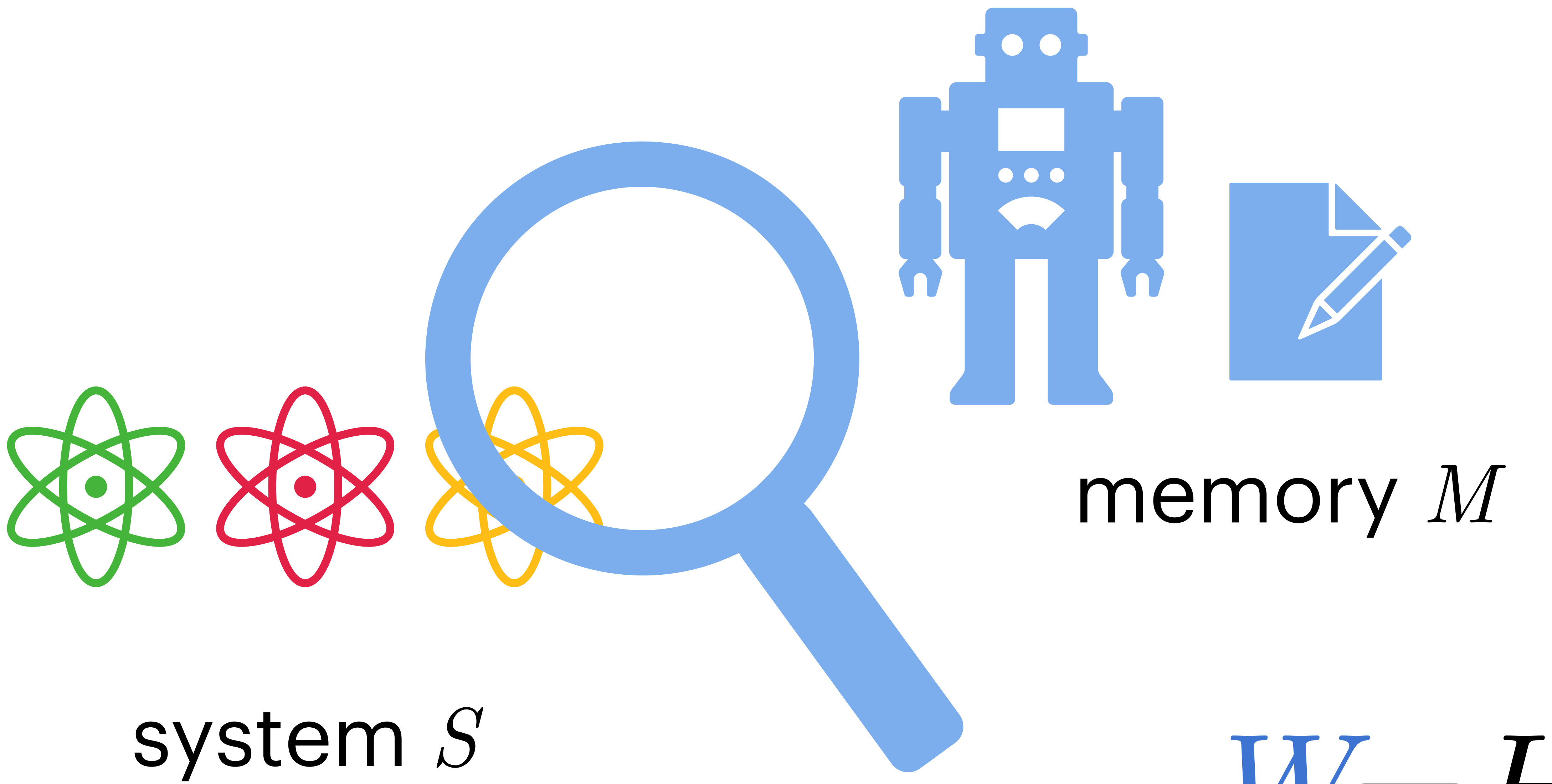
Complete ignorance: $H(S|M)=n$



Complete knowledge: $H(S|M)=0$



Question

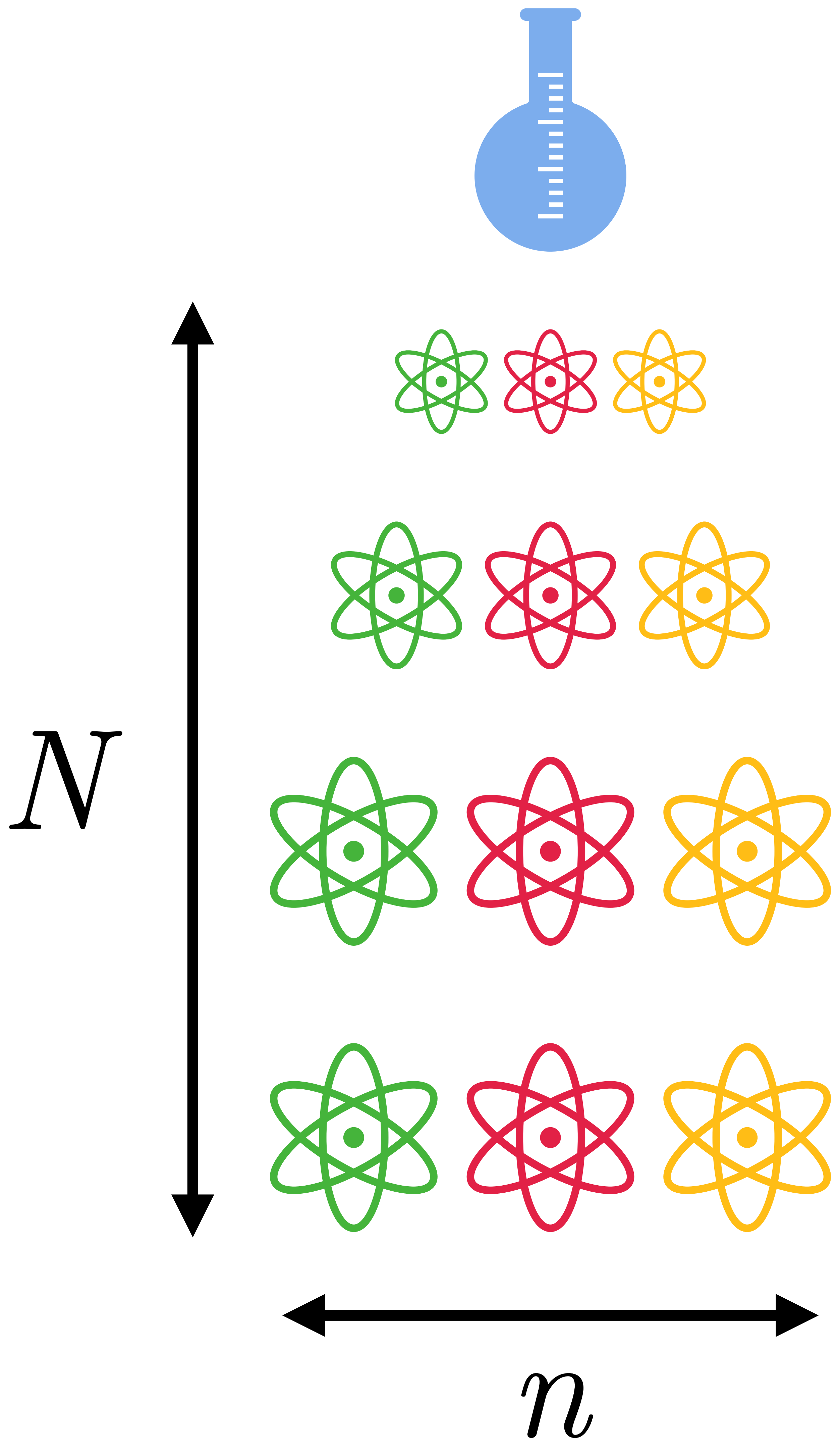


$$W = H(S|M)kT\ln 2$$

How to acquire such knowledge to reduce work cost?

What is the work cost of acquiring such knowledge?

Learning to erase



N copies of an unknown pure state
from m possibilities: $|\psi_x\rangle, x = 1,\ldots,m$ pairwise $\Theta(1)$ apart

$$\rho = \sum_x p_x (|\psi_x\rangle\langle\psi_x|)^{\otimes N}$$

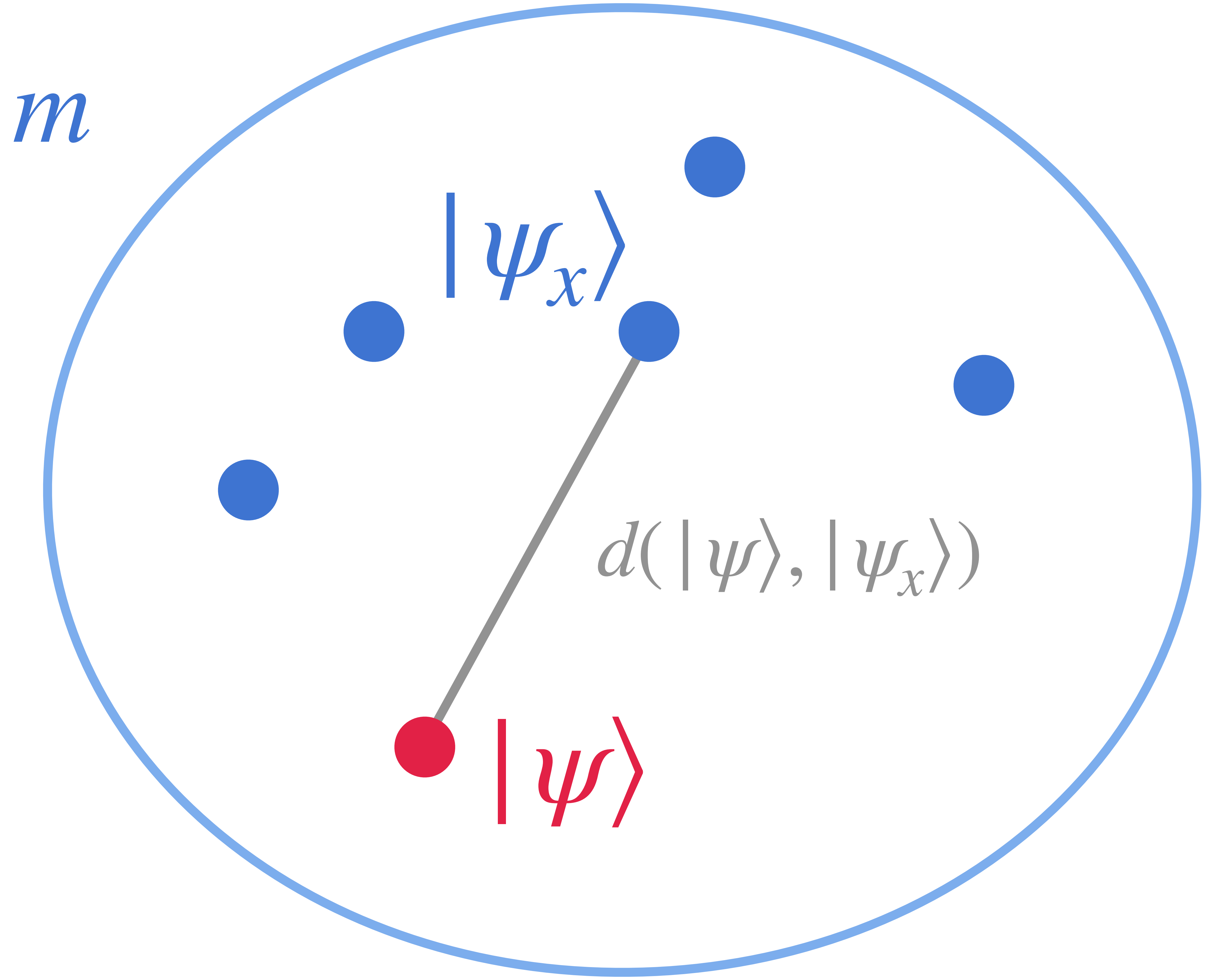
Intuition:

as we collect more copies

ignorance is reduced

=> less work cost per erasure $\frac{W}{N} \to 0$

Learning algorithm



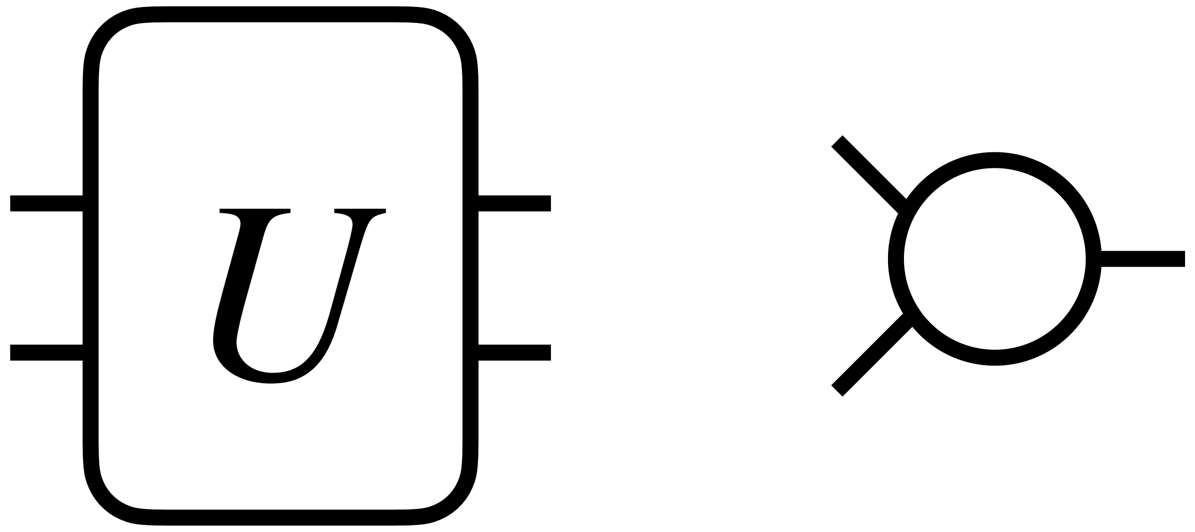
- Estimate the trace distance between target state $|\psi\rangle$ and every candidate $|\psi_x\rangle$ using Clifford classical shadow
- Output the closest one
- Optimal sample complexity $s = O(\log m)$

What is the energy cost?

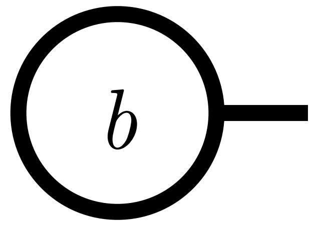
Zhao et al. PRX Quantum 2024, Huang et al. Nature Physics 2020, Badescu et al. TheoretiCS 2024

Reversible learning

Building blocks of a learning algorithm:

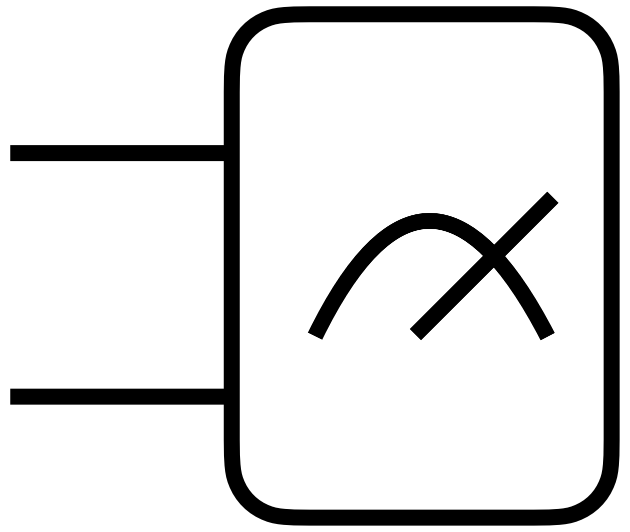


quantum/classical gates



$b \sim \mathrm{Bern}(1/2)$

random number



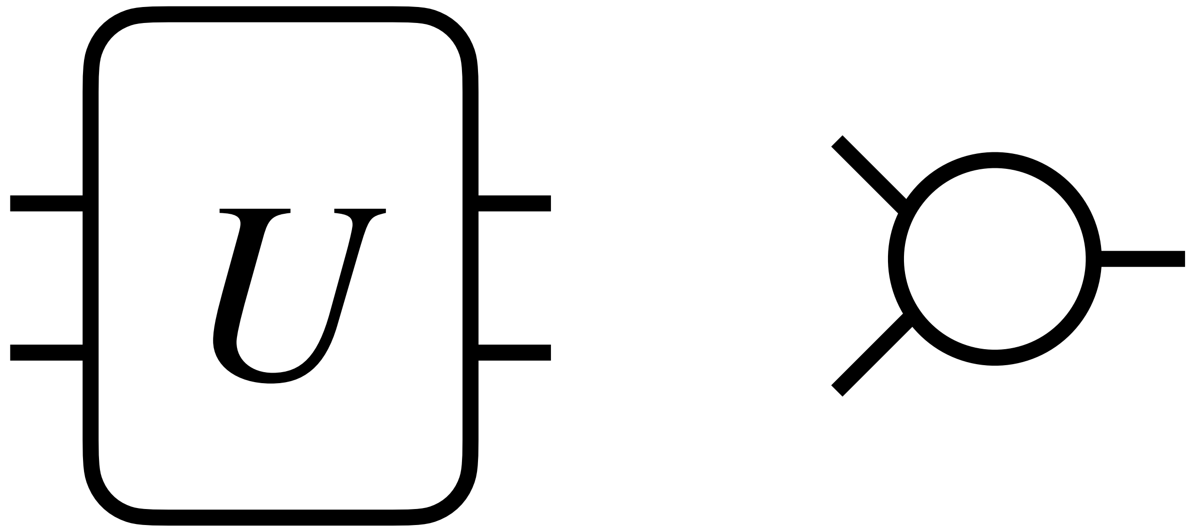
$\{P_a\}$

projectors

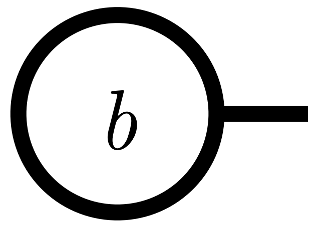
measurement

Reversible learning

Building blocks of a learning algorithm:

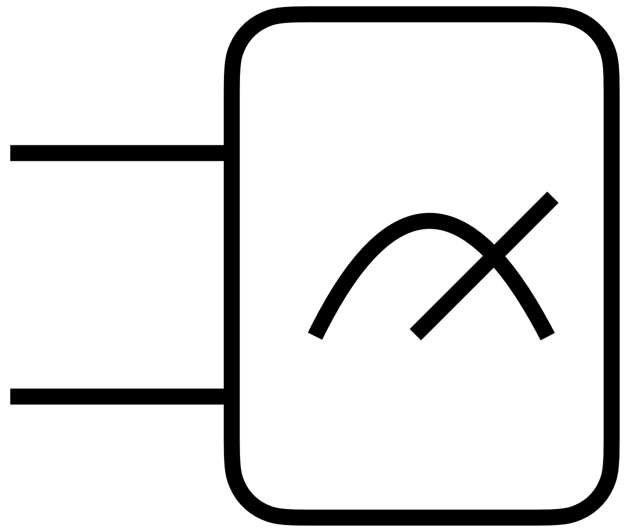


quantum/classical gates



$b \sim \mathrm{Bern}(1/2)$

random number

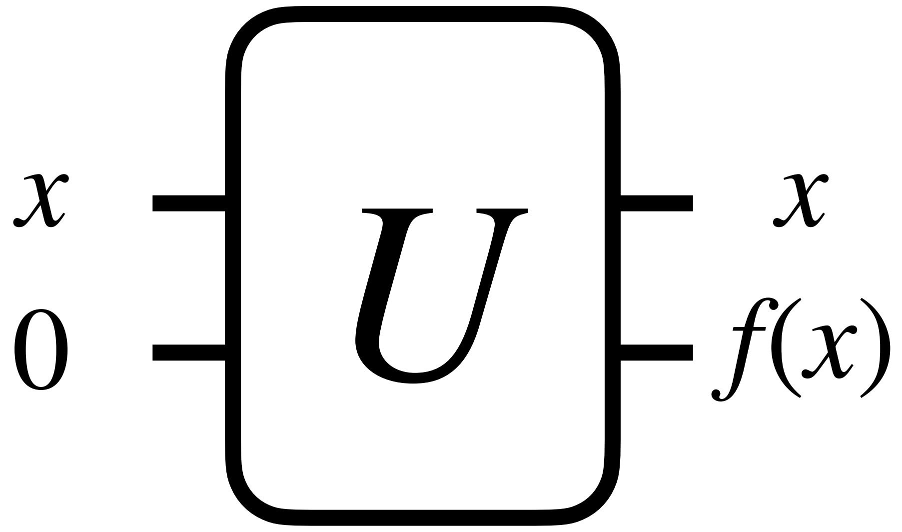


$\{P_a\}$

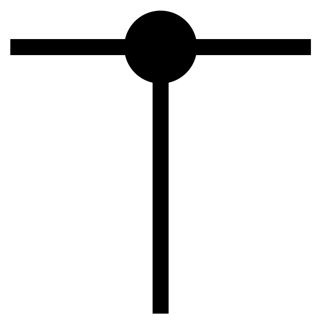
projectors

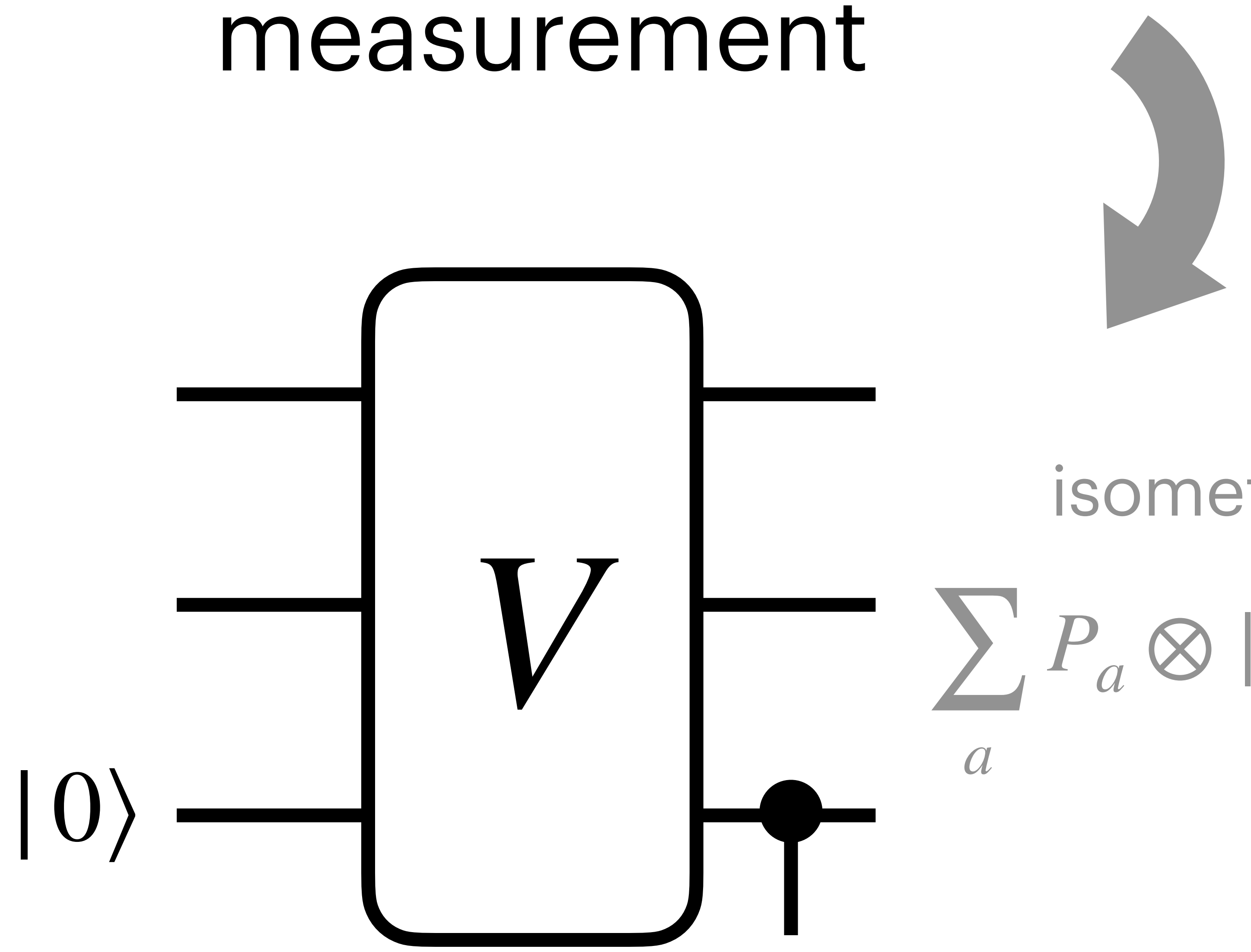
measurement

Make it reversible: fully coherent processing



$|+\rangle$ $\frac{|0\rangle + |1\rangle}{\sqrt{2}}$



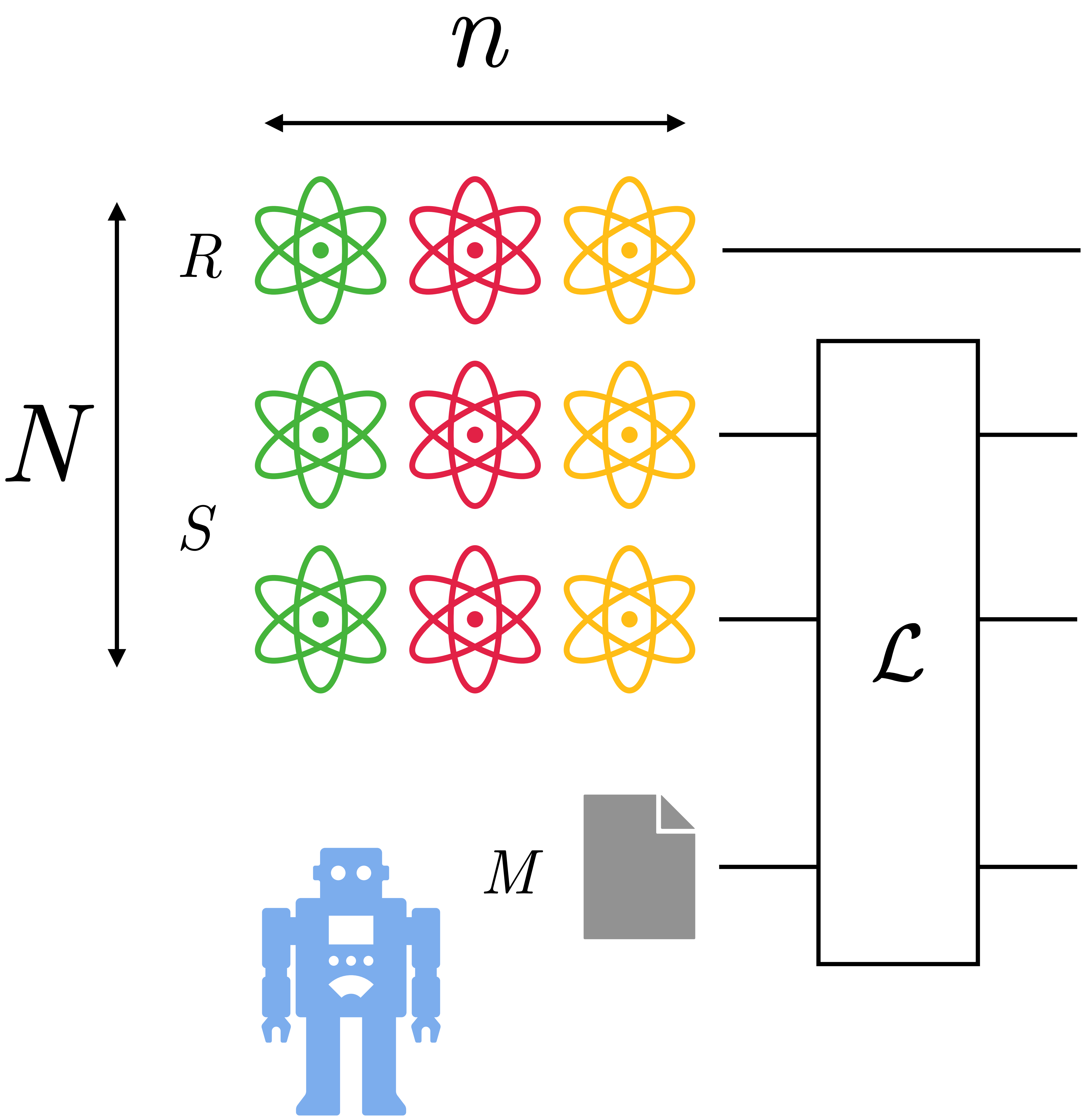


isometry

$\sum_a P_a \otimes |a\rangle\langle 0|$

$$V|\psi\rangle|0\rangle = \sum_a P_a|\psi\rangle|a\rangle$$

Reversible learning

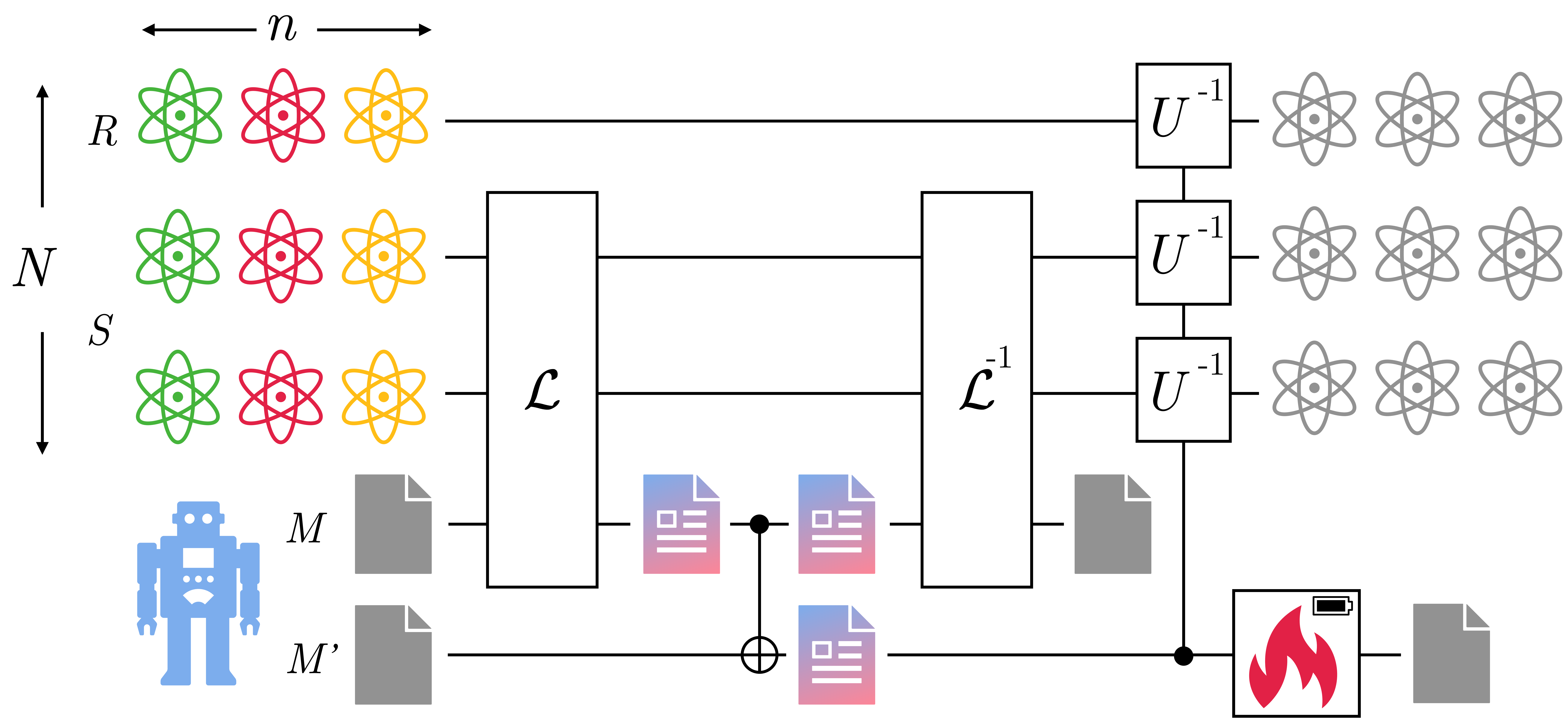


Reversible learning algorithm:

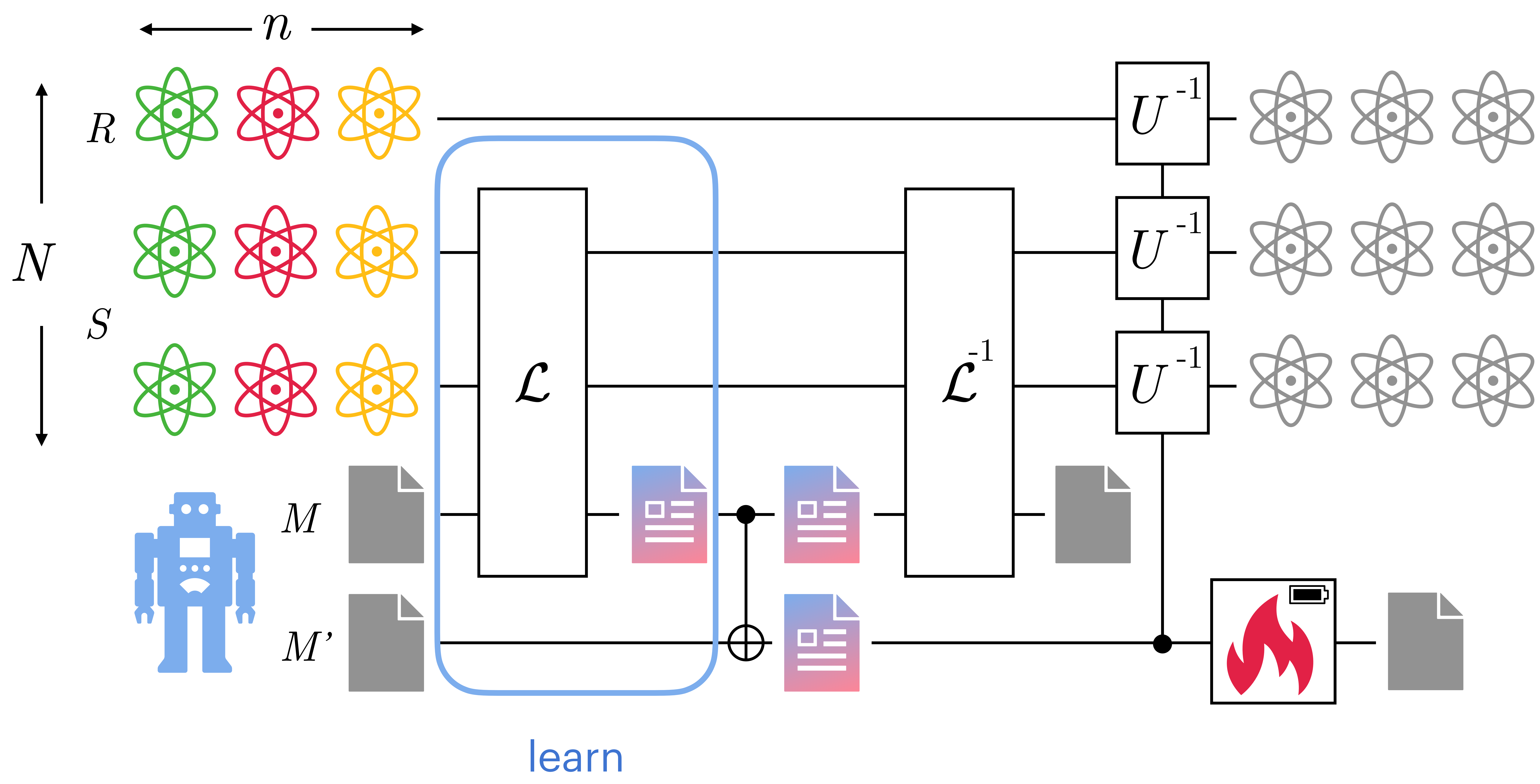
$$\mathcal{L}\,|\psi_x\rangle_S^{\otimes s}\,|0\rangle_M\,|0\rangle_A = \sum_{x'=1}^{m} c_{x'|x}\,|x'\rangle_M\,|\mathrm{junk}_{x'}\rangle_{S,A}$$

Learning guarantee: $|c_{x|x}|^2 \geq p_{\mathrm{succ}} \to 1$

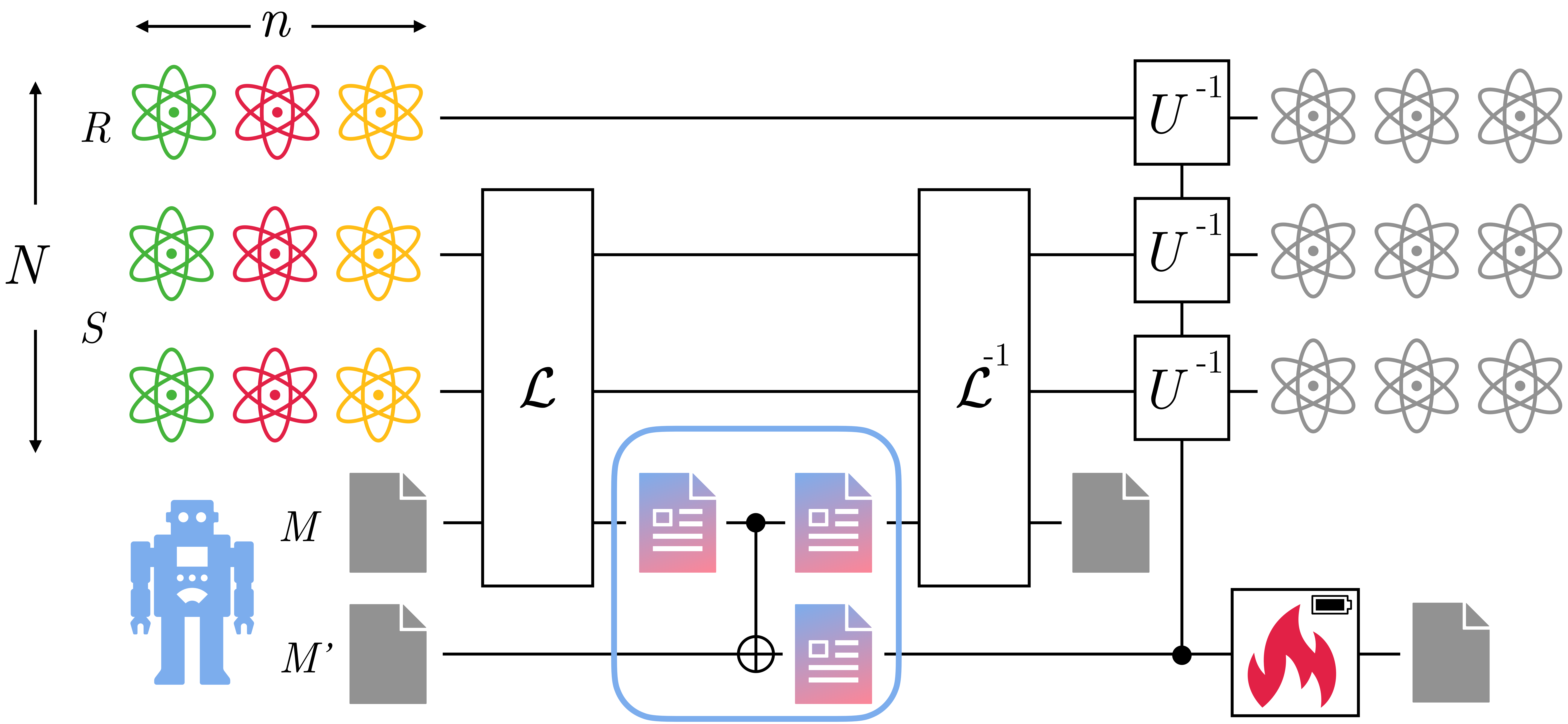
Erasure protocol



Erasure protocol

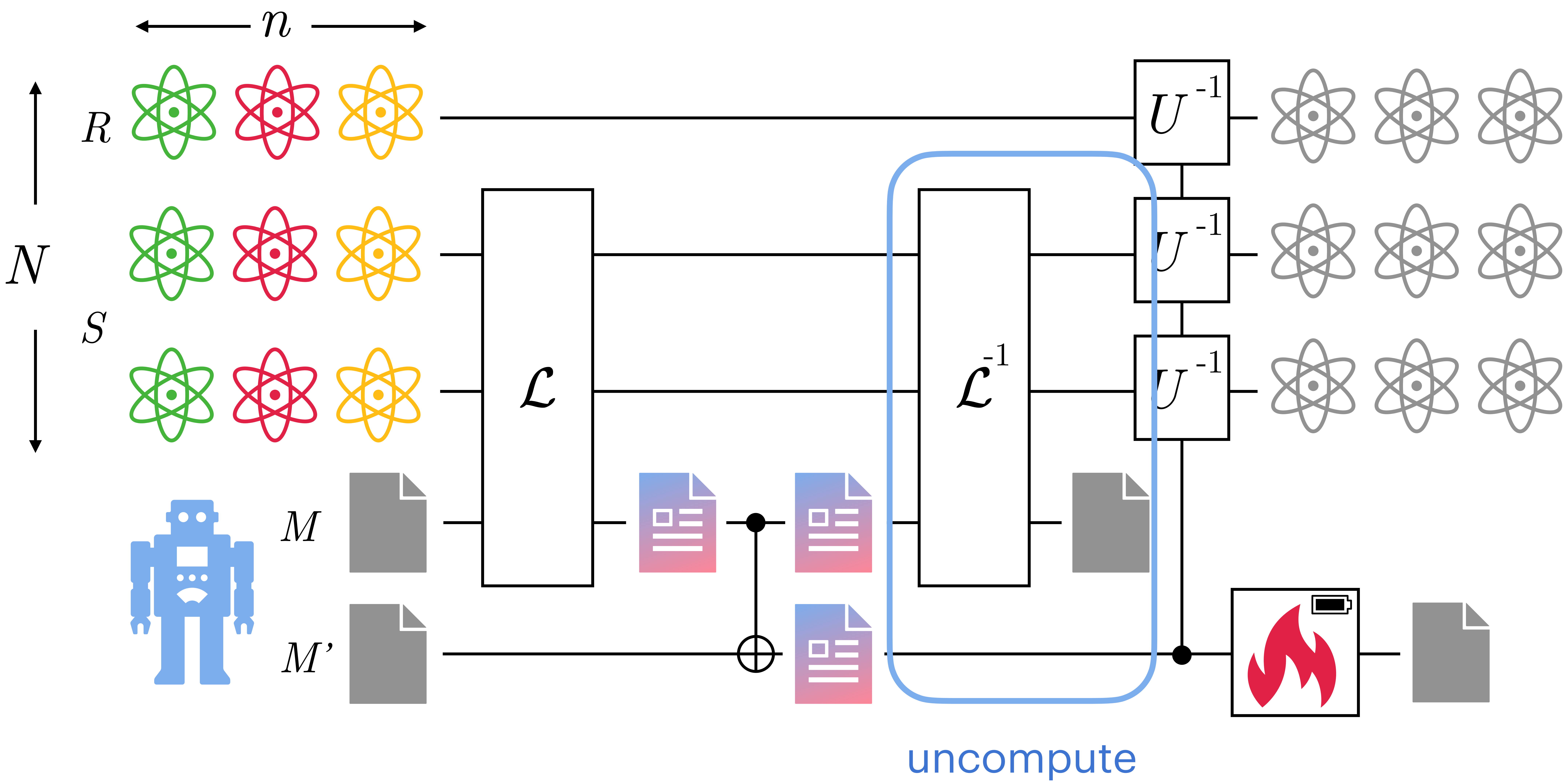


Erasure protocol

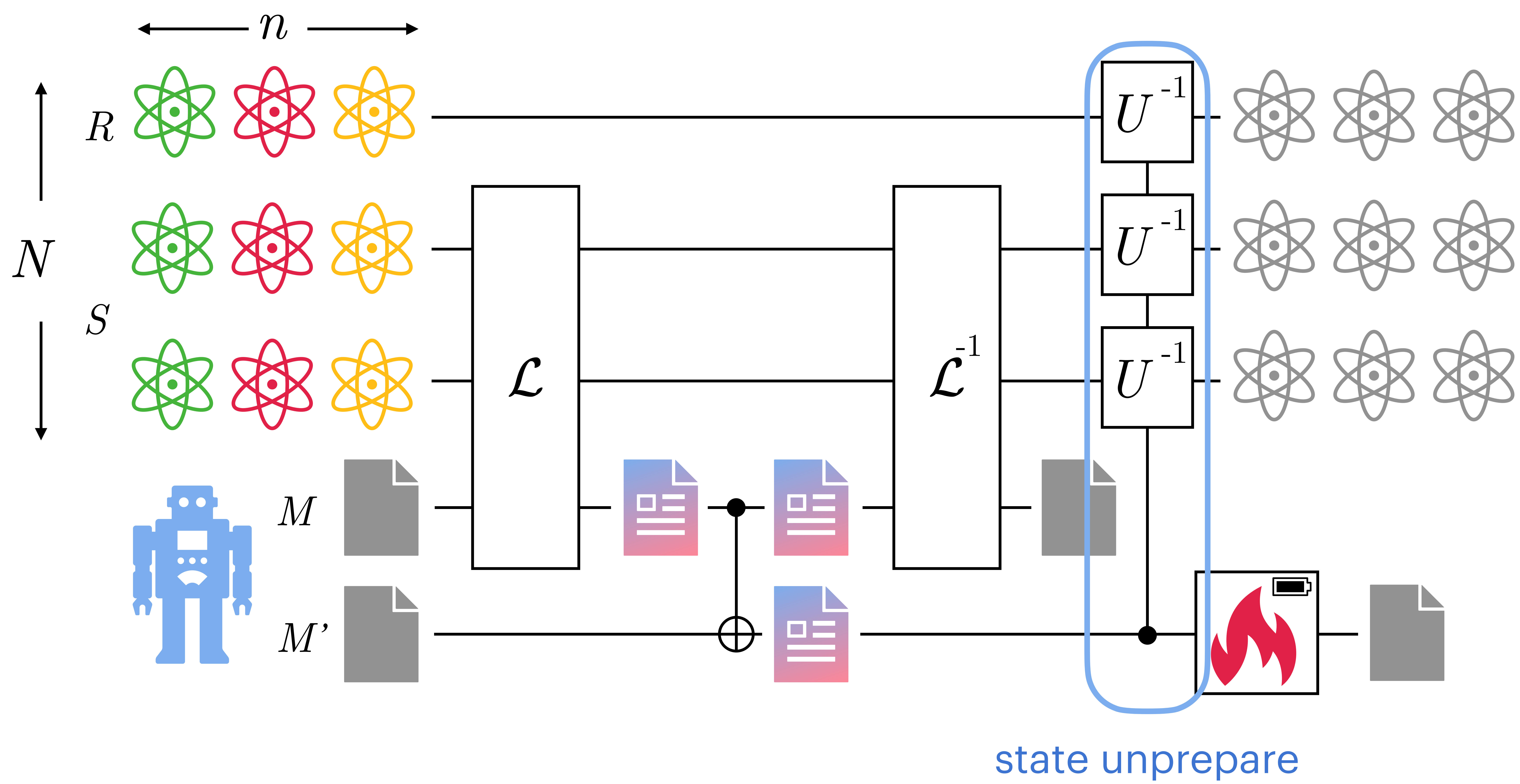


copy out learning outcome

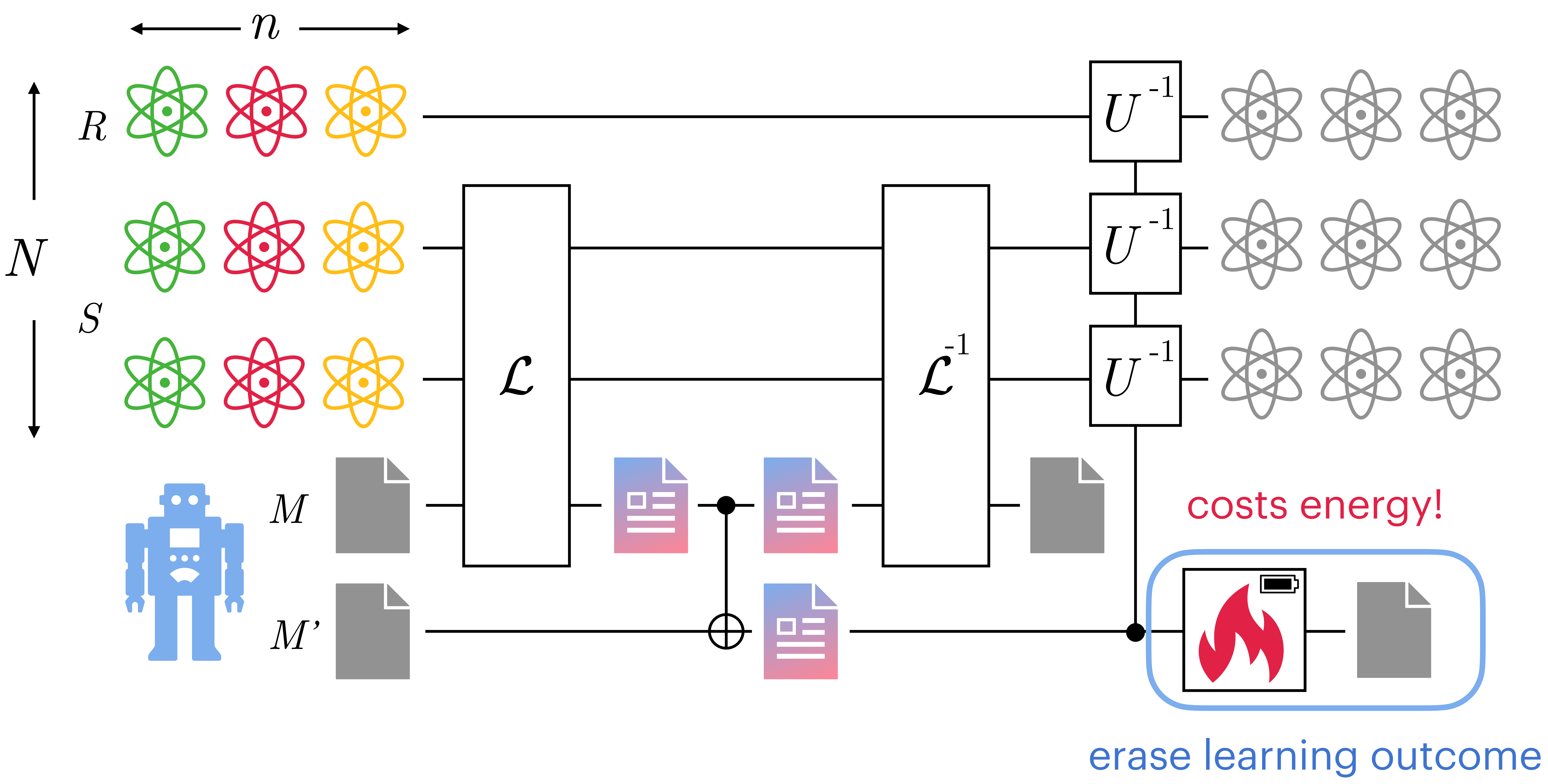
Erasure protocol



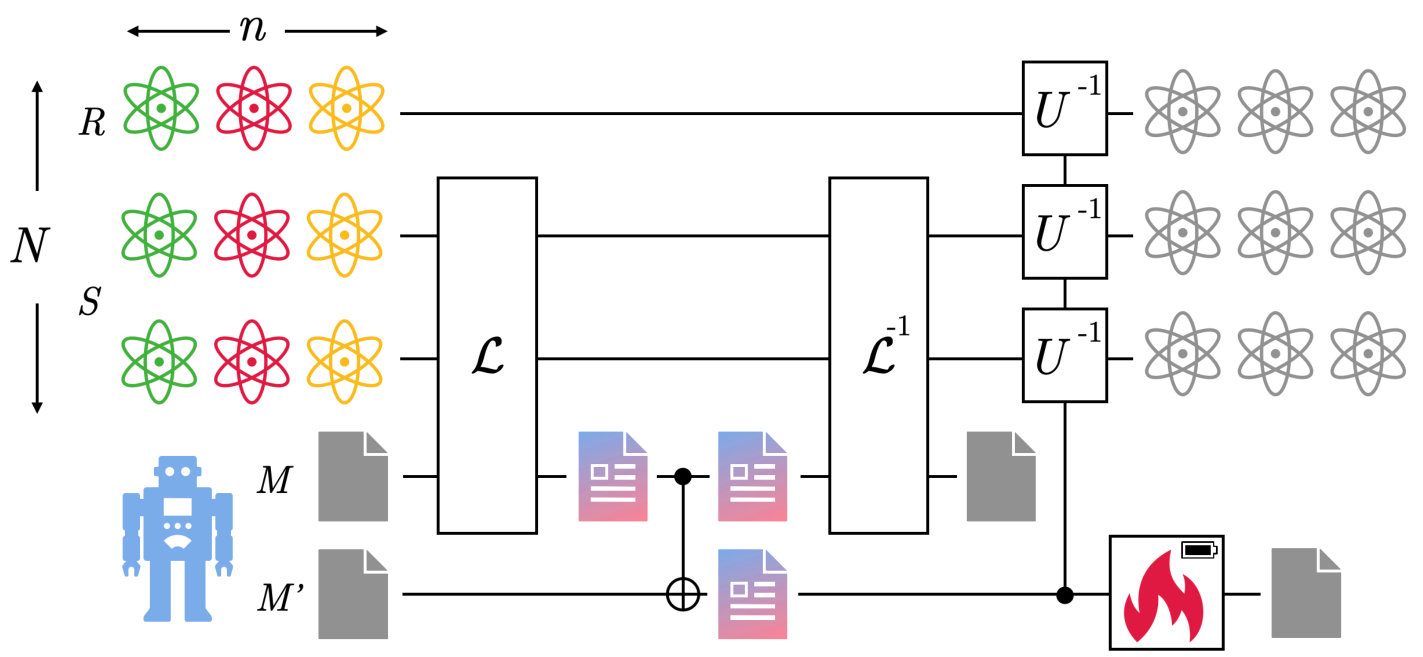
Erasure protocol



Erasure protocol



Erasure protocol



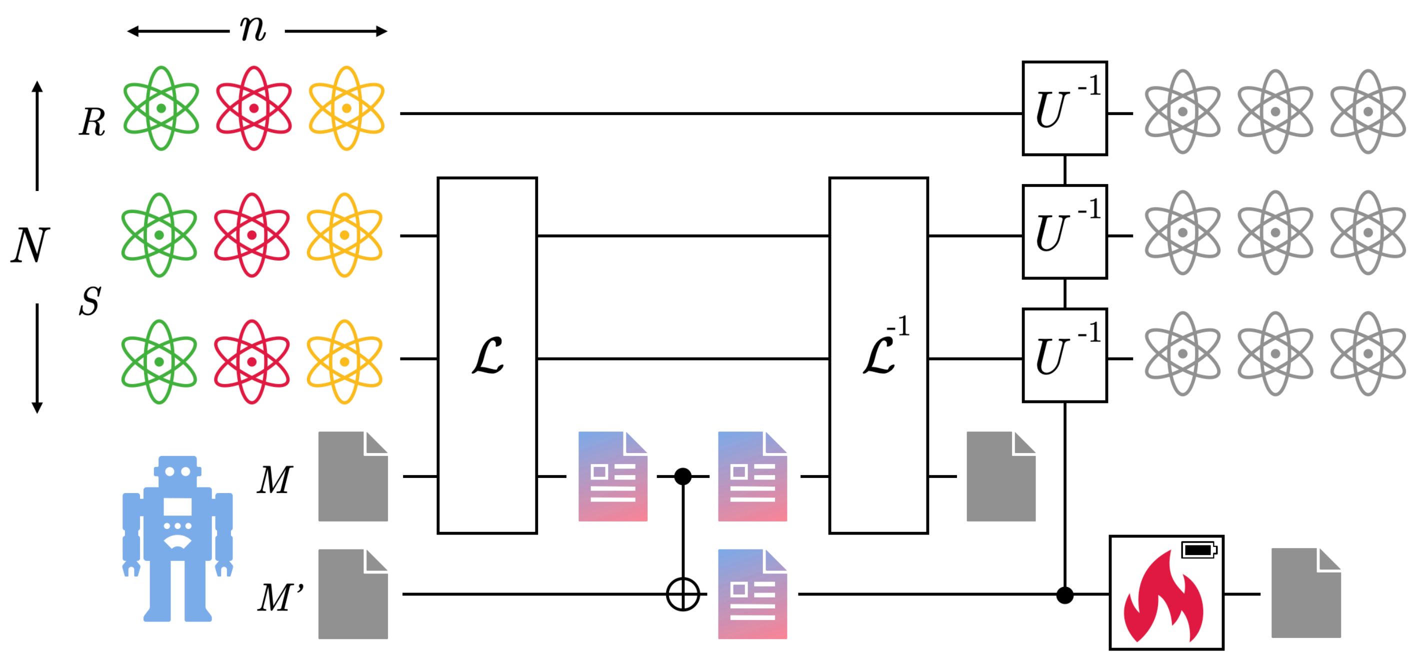
Remarks:

1. Correctness: trace distance from $|0\rangle$ bounded by $\sqrt{1-p_{\text{succ}}^2} \to 0$

#bits to store learning outcome

2. Energy cost: $W = (\log_2 m)kT\ln 2$, independent of N

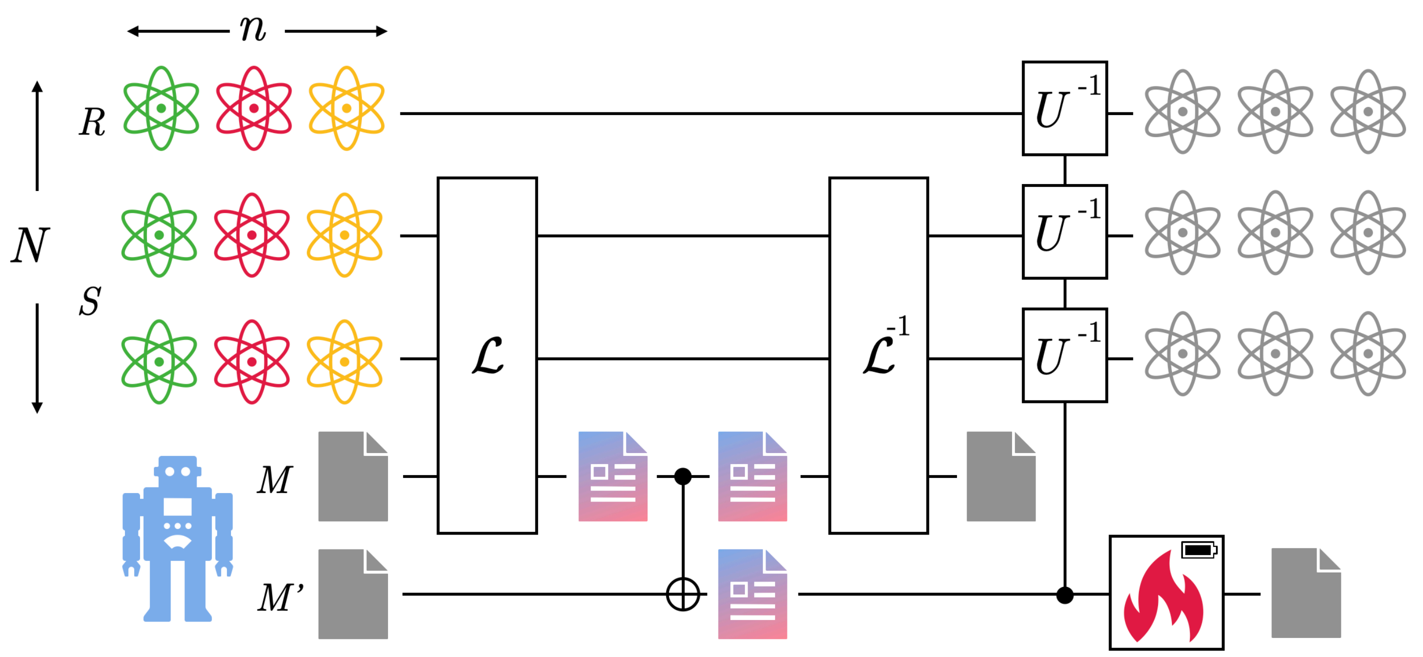
Erasure protocol



Remarks:

1. Correctness: trace distance from $|0\rangle$ bounded by $\sqrt{1 - p_{\text{succ}}^2} \to 0$
2. Energy cost: $W = (\log_2 m)kT \ln 2$, independent of N (#bits to store learning outcome)
3. Learning can be reversible and has no fundamental energy cost itself!
4. The energy cost occurs when we erase the learning outcome.

Erasure protocol



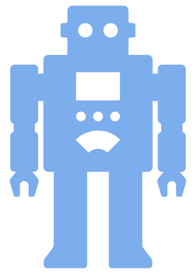
Remarks:

1. Correctness: trace distance from $|0\rangle$ bounded by $\sqrt{1-p_{\text{succ}}^2} \to 0$

#bits to store learning outcome

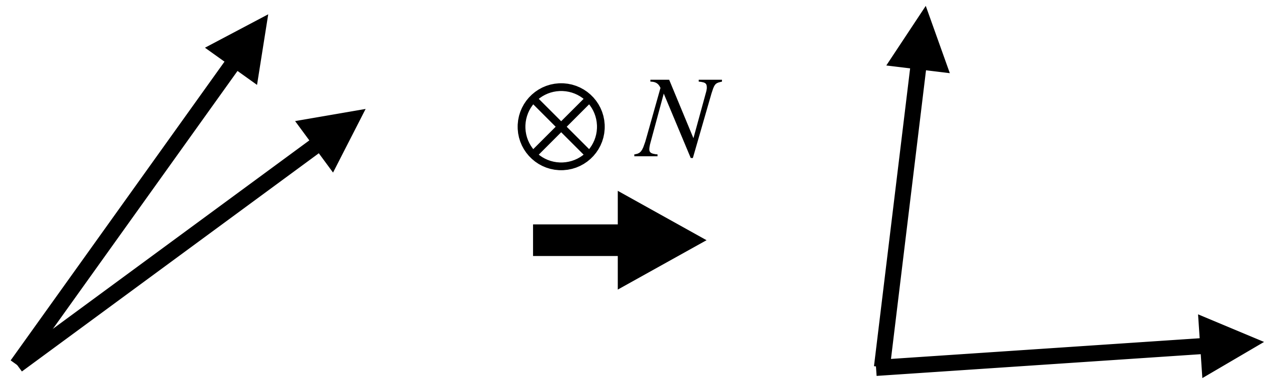
2. Energy cost: $W = (\log_2 m)kT \ln 2$, independent of N
3. Learning can be reversible and has no fundamental energy cost itself!
4. The energy cost occurs when we erase the learning outcome.
5. Sample complexity => minimal quantum memory requirement
6. Time complexity: $O(T_{\text{learn}} + \log m + NT_{\text{prep}})$
7. Efficient learning & state preparation => efficient erasure

Energy optimality

Energy cost of learning to erase: $W = (\log_2 m)kT\ln 2$ 

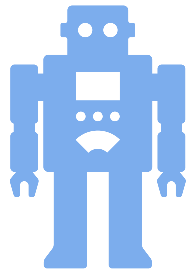
(One-shot) Landauer's limit: $W \geq H_{\max}(\rho)kT\ln 2$ $\quad H_{\max}(\rho) = \log_2 \mathrm{rank}(\rho)$

$$\rho = \sum_{x=1}^{m} p_x(|\psi_x\rangle\langle\psi_x|)^{\otimes N}$$



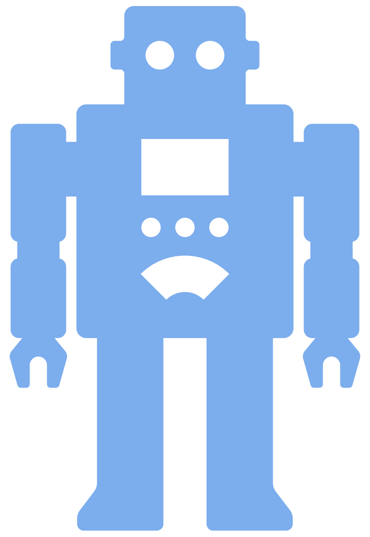
(sample complexity)

When $N \geq \Omega(\log m)$, Gram matrix $G_{x,x'} = \langle\psi_x|\psi_{x'}\rangle^N$ is diagonally dominant

=> $\mathrm{rank}(\rho) = m$ and Landauer's limit coincides with $W = (\log_2 m)kT\ln 2$ 

Dahlsten et al. New Journal of Physics 2011

Efficiently erasable states



Energy cost of learning to erase: $W = (\log_2 m) kT \ln 2$

Erasure is efficient when learning & state preparation is efficient.

efficient: time is polynomial in n, N

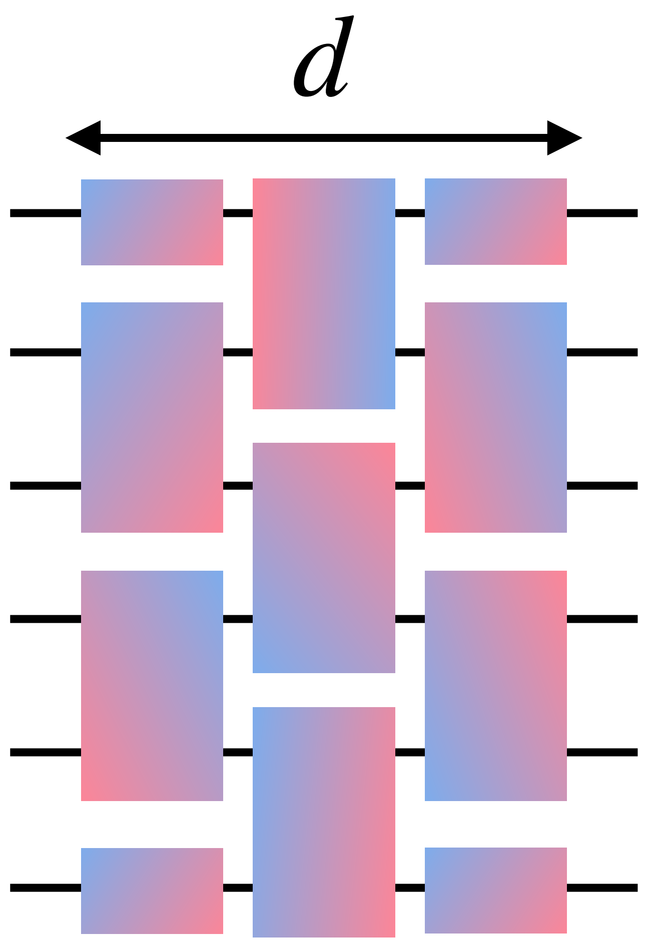
inefficient: for $n \sim 100$, may take longer than the age of the universe

Physically relevant examples:

short time dynamics

NISQ

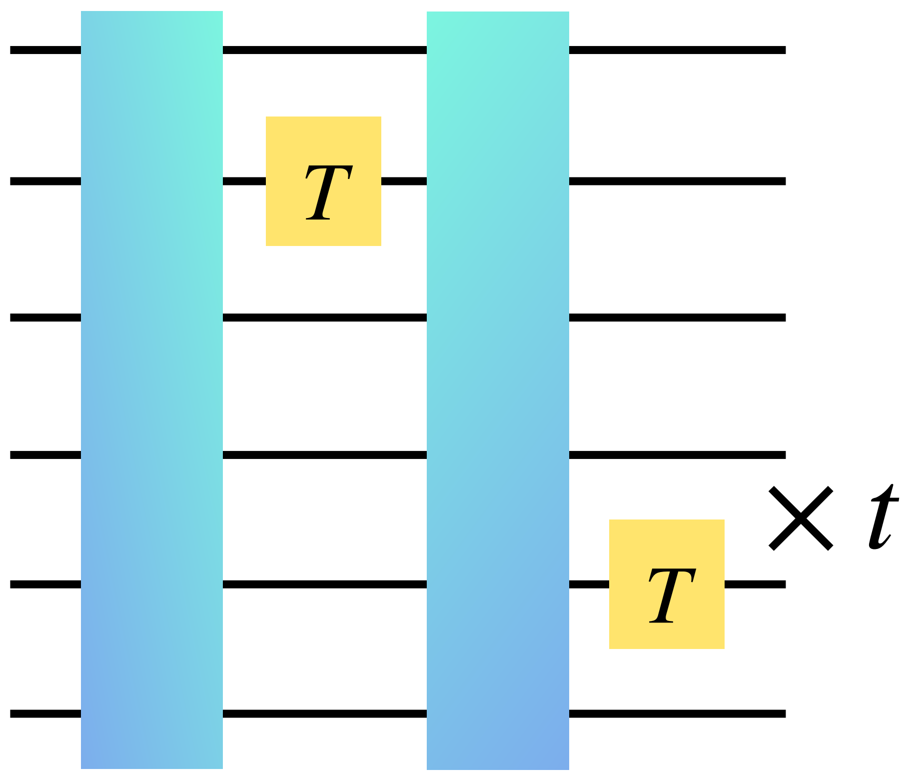
quantum advantage



shallow circuit states

low complexity

QEC



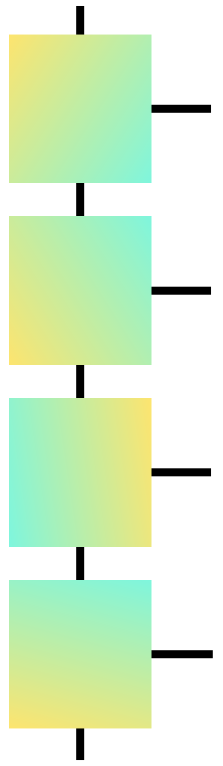
doped stabilizer states

low magic

ground state

SPT

MBQC



$\log_2 \chi \leq \mathcal{S}$

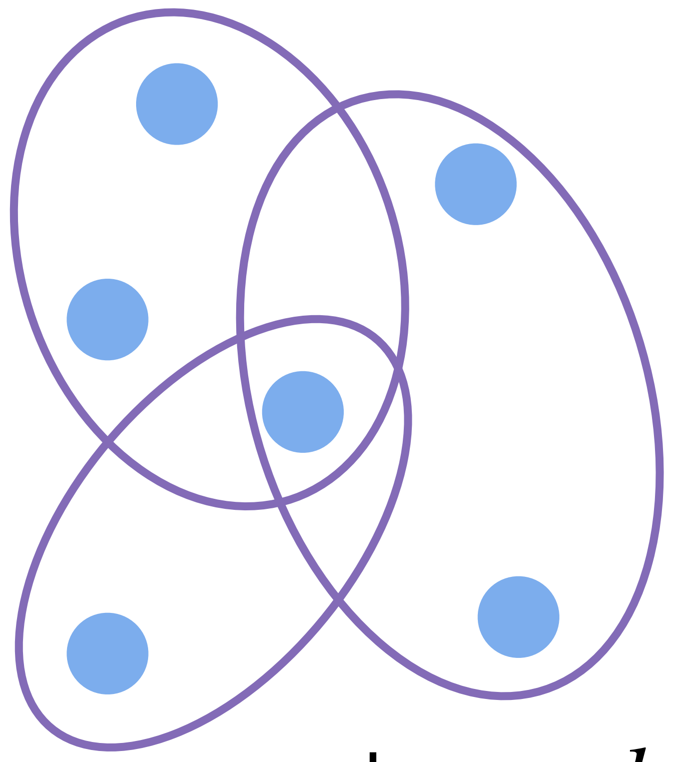
matrix product states

low entanglement

query algorithm

MBQC

IQP

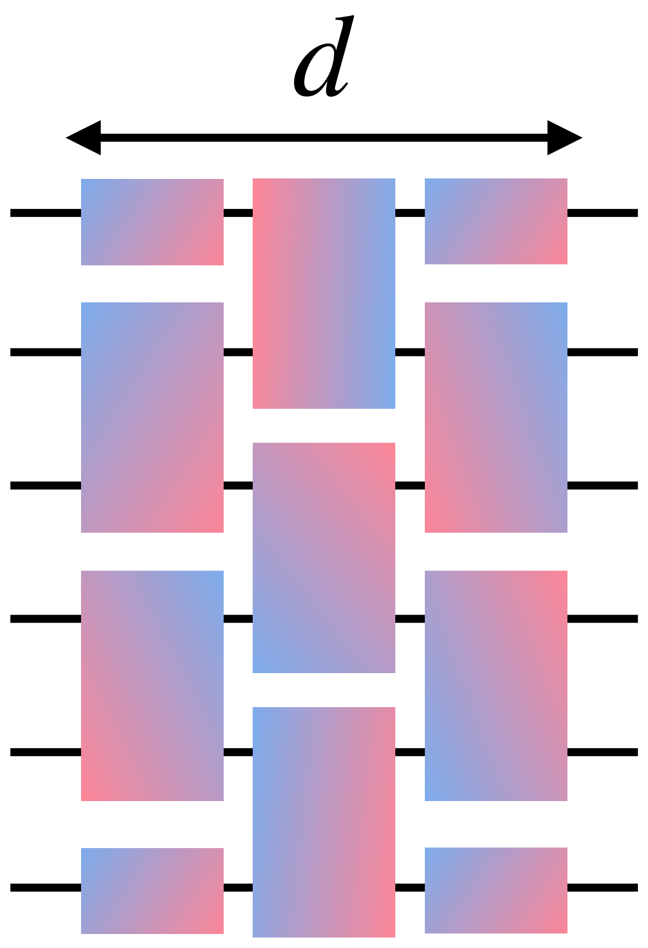


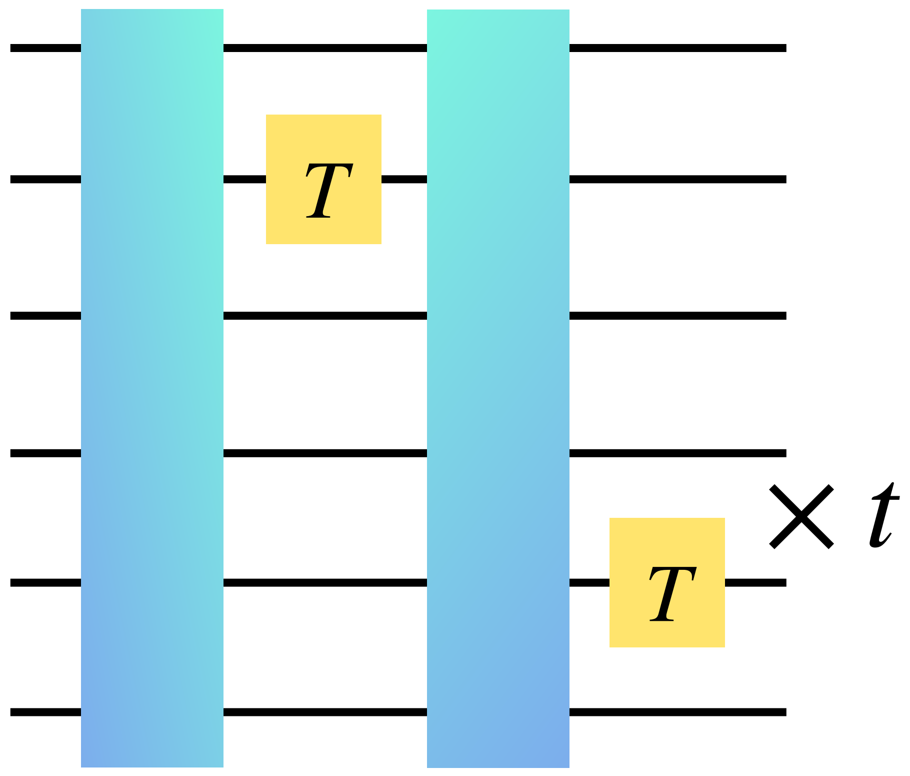
degree k

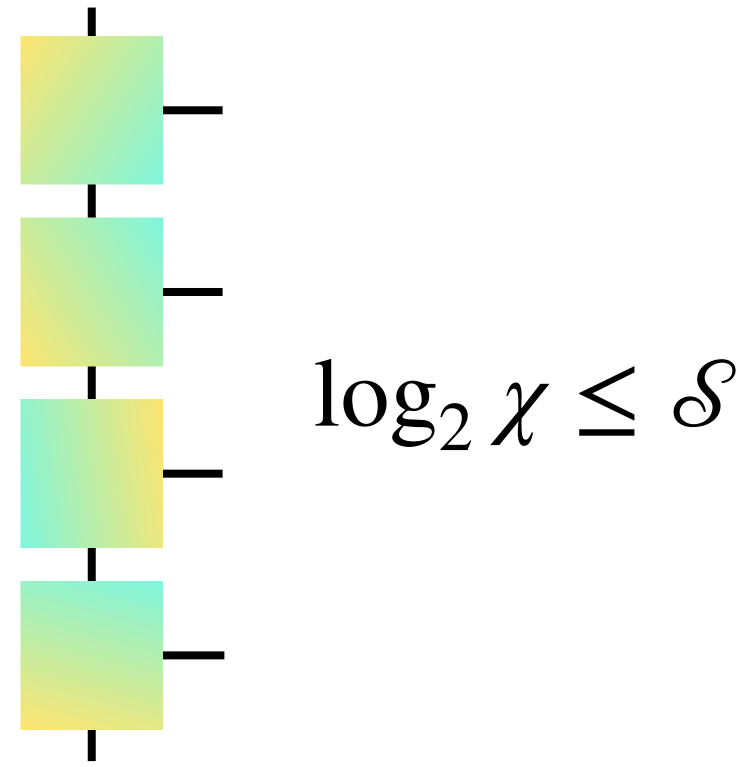
phase states

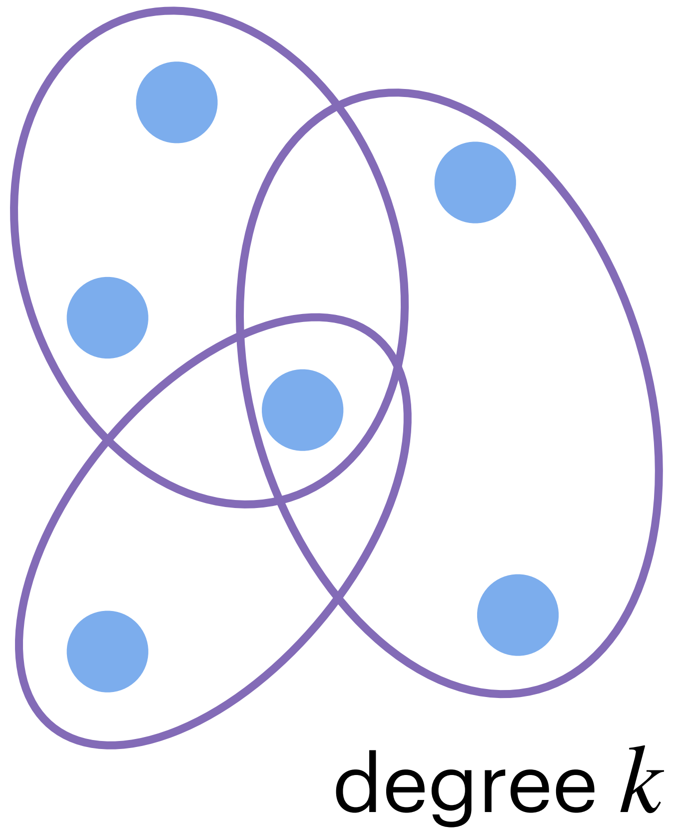
low degree

Work vs complexity





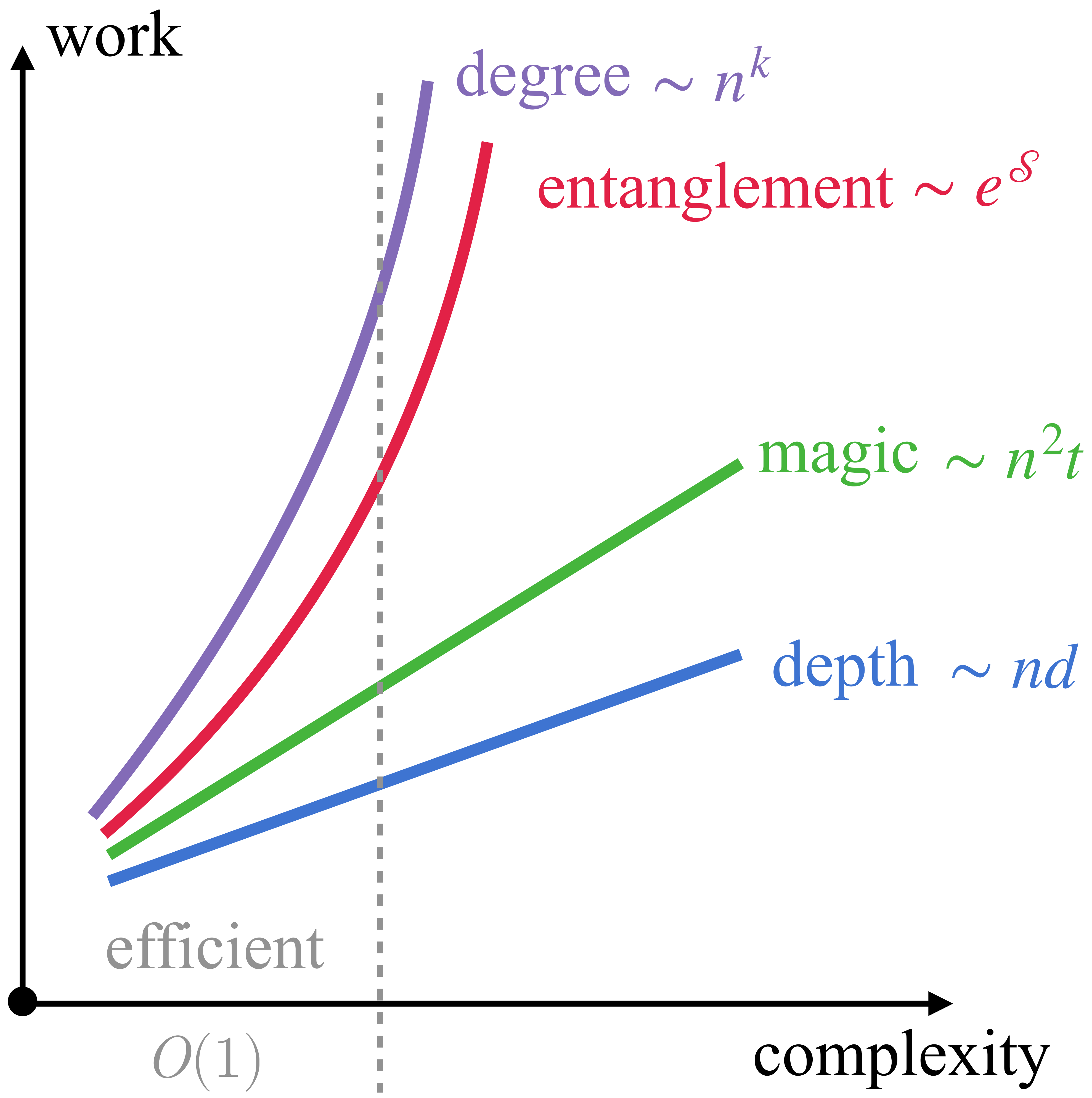




	shallow circuit states	doped stabilizer states	matrix product states	phase states
complexity	depth d	t T-gate	entanglement entropy $\mathcal{S}$	degree k
work cost	$\Theta(nd)kT\ln 2$ gate count	$\Theta(n^2 t)kT\ln 2$ Clifford has gate count $O(n^2)$	$\exp(\tilde{\Theta}(\mathcal{S}))kT\ln 2$ continuous class needs covering/packing	$\Theta(n^k)kT\ln 2$ #polynomial with degree k
time complexity	$O(\mathrm{poly}(n)2^{d^{O(1)}} + ndN)$ local inversion	$O(\mathrm{poly}(n,2^t) + n^2 tN)$ Bell sampling	$O(\mathrm{poly}(n,2^{\mathcal{S}}) + n4^{\mathcal{S}}N)$ sequential unentangling	$O(n^{3k-2} + kn^k N)$ directional gradient

Landau et al. arXiv:2410.23618, Leone et al. Quantum 2024, Cramer et al. Nat Comm 2010, Arunachalam et al. TQC 2023

Work vs complexity



For these special classes of states, we give provably-efficient energy-optimal erasure protocols based on *learning*.

What about more general states?

Can we achieve Landauer's limit in polynomial time?

Computational hardness

Can we achieve Landauer's limit in polynomial time for general states?

No!

There is a class of states for which the Landauer's limit is

$$\Theta(n\,\mathrm{polylog}(n))kT\ln 2,$$

independent of N

but any polynomial time quantum algorithm must pay

$$W_{\mathrm{Haar}} = \log_2 \binom{N+2^n-1}{N} kT \ln 2$$

#qubits

$$\sim nN(1-o(1))$$

omitting $1/\mathrm{poly}(n), \log(1/\epsilon)$

uncertainty principle

failure probability

$N = \mathrm{poly}(n)$

joules of work to erase them!

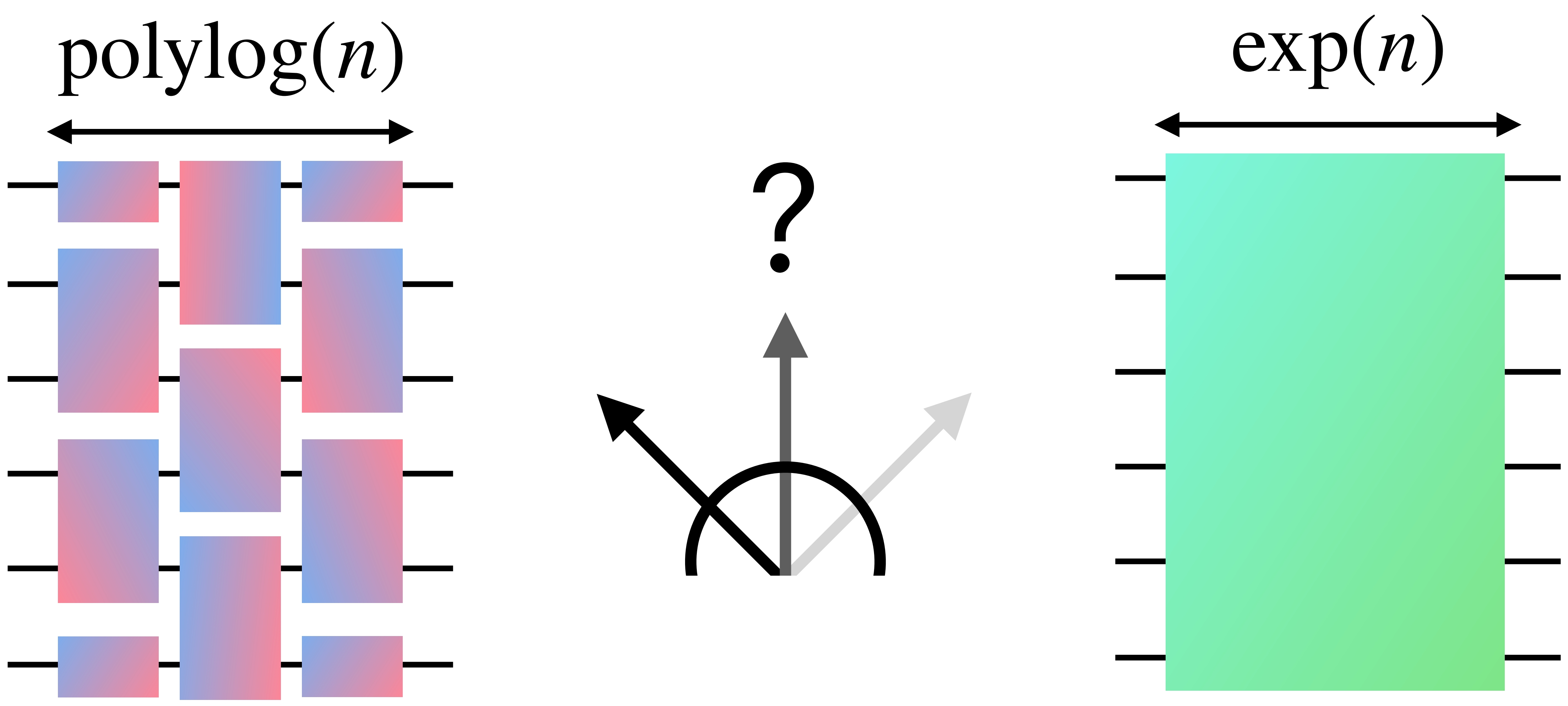
Pseudorandom states

Under standard cryptographic assumption,

existence of one-way functions secure against any sub-exponential time quantum adversary

pseudorandom states can be constructed in $d = \mathrm{polylog}(n)$ depth.

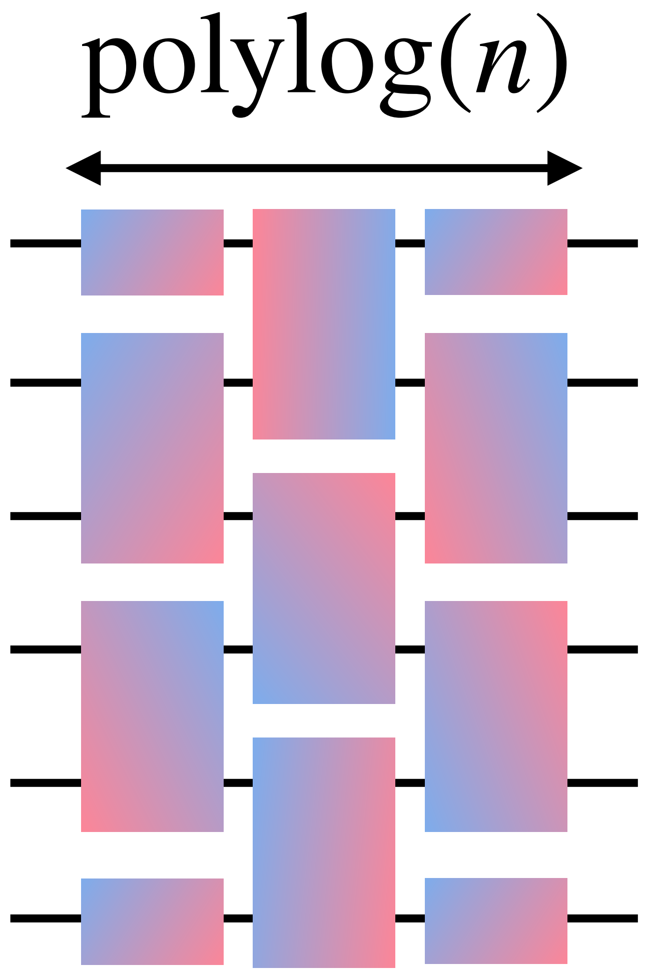
They cannot be efficiently distinguished from Haar random states with non-negligible probability.



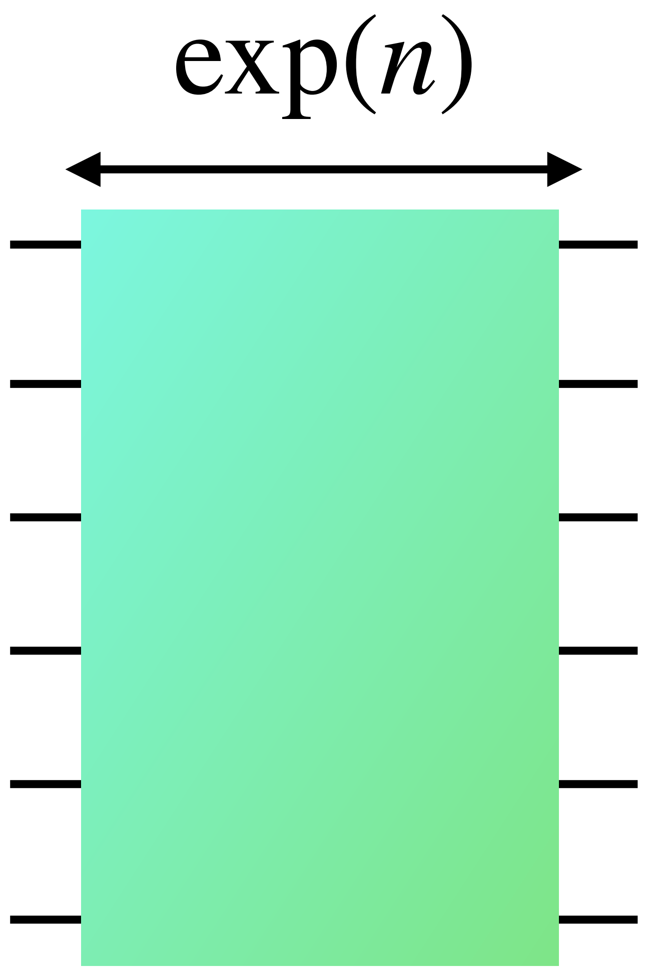
Ji et al. CRYPTO 2018, Schuster et al. arXiv:2407.07754

Pseudorandom states

They cannot be efficiently distinguished from Haar random states with non-negligible probability.



$$W = \Theta(n\mathrm{polylog}(n))kT\ln 2$$



$$W_{\mathrm{Haar}} = \log_2 \binom{N + 2^n - 1}{N} kT \ln 2$$

Full reduction:

1.erase

2.test if erase succeeded

3.measure work cost

Measuring the work cost of erasure gives a way to distinguish them!

=> no polynomial time quantum algorithm can achieve Landauer's limit!

a much stronger no-go result compared to the 3rd law of thermodynamics

(this is a genuine quantum many-body phenomenon!)

Summary

- A rigorous connection between thermodynamics and quantum learning theory.
- This allows us to answer several fundamental questions:

Learning:

Learning has no fundamental energy cost itself.

Our (in)ability to learn significantly impact the energy cost of thermodynamic tasks.

omitted: energy gain in work extraction

Thermodynamics:

Learning provides provably-efficient energy-optimal protocols.

The complexity of quantum many-body systems leads to drastic change of fundamental physical laws.

e.g., Landauer's principle

Open questions

- Extension to more realistic scenarios & continuous variable systems
- Other thermodynamic tasks; other physical properties of learning itself
- Consequences in high energy physics (pseudorandom models of black hole)