Exceptional service in the national interest

Sandia National Laboratories

Physical Foundations of Landauer's Principle

Michael P. Frank
Center for Computing Research
Sandia National Laboratories

10th Conference on Reversible Computation
University of Leicester, Leicester, UK
Thu., Sep. 13th, 2018

Approved for Public Release
SAND2018-9935 C

Introduction

- **Context:** Landauer's Principle connecting information loss with entropy increase is a key motivation for reversible computing, but continues to be frequently misinterpreted/misunderstood.
  - As a result of these misunderstandings, the validity of Landauer's Principle has often been (misguidedly) challenged by engineers & physicists…
    - This has been a substantial barrier to R&D investment in reversible computing
  - We should all be well prepared to answer in-depth questions about the Principle, so we are better equipped to help defend & promote our field.
- There are a number of important subtleties that must be appreciated in order to have an understanding of Landauer's principle that allows one to answer questions about it properly:
  - Transformations of complex states – Focus of a paper at ICRC'16
  - Role of conditional reversibility – Focus of my RC'17 paper
  - **Treatment of stochastic operations** – Addressed in the current paper
    - I also first mentioned this issue many years ago, *e.g.*, at ISMVL'05
  - **Importance of correlations** – Addressed in the current paper
- Landauer's Principle follows as a rigorous theorem of fundamental physics, but *only* given a proper treatment of these issues.

Talk Outline

- Review of historical development of the concept of entropy
  - Shows how information theory emerged from, and is intimately connected with, statistical physics
- Review of some basic information theory concepts:
  - Entropy, known information, conditional entropy, mutual information
- Fundamental connections between computation and physics:
  - Bijective time evolution and the second law of thermodynamics
  - Relationship between computational and physical states
  - Types of computational operations, and thermodynamic implications
    - Many-to-one operations, and implications for entropy ejection
    - Stochastic (one-to-many) operations, and implications for entropy intake
      - Some physical examples
  - Essential role of mutual information in proving Landauer's Principle
- Review of empirical demonstrations of Landauer's Principle
  - Not needed if you know the physics, but helpful in countering skepticism
- Conclusion

A brief history of entropy… (1/5)

- Clausius (1862) identified a quantity $\Delta Q/T$ (with ΔQ = change in heat, T = temperature) that is always ≥ 0 when summed over all of the systems involved in any given thermodynamic process…
  - An early version of the Second Law of Thermodynamics
- In 1865 he proposed to call this quantity *entropy*, from Greek τροπή (tropḗ, transformation),
  - Connoting, that which gives an inherent (en-) direction to a physical transformation (tropḗ)…
  - He also introduced the use of the symbol S for it
    - Possibly to honor Sadi Carnot? (The inventor of the concept of a thermodynamically reversible heat engine)
  - **Note:** Entropy S has *physical* units of *heat/temperature.*
- Nowadays, we actually *define* temperature T in terms of the marginal change in the (maximum) entropy $\hat{S}$ of the system as heat is added to it…

Rudolph Clausius
Discoverer of entropy

$$\frac{1}{T} = \frac{\partial \hat{S}}{\partial Q}$$

T = Thermodynamic temperature
$\hat{S}$ = Maximum entropy (S at equilibrium)
Q = Quantity of heat

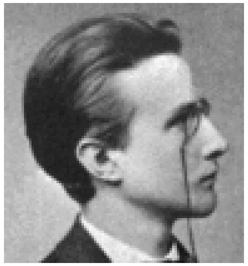
(continued on next slides…)

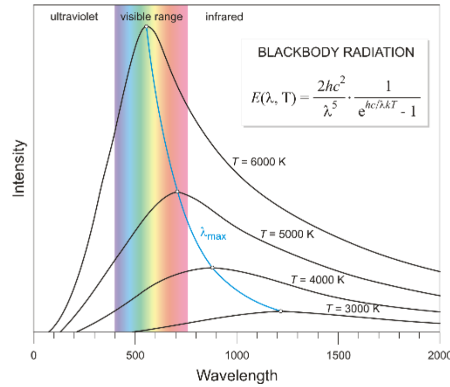
History of Entropy, cont. (2/5)

- Ludwig Boltzmann (1872) proved his *H-theorem* suggesting that entropy might have a *statistical* basis:
  - This paper was essentially an early exploration into what we now call *chaos theory*…
    - Boltzmann showed that statistical uncertainty about the states of common (classical) physical systems (*e.g.* gas molecules) tends to become amplified when those systems interact (*e.g.*, collide)…
  - This paper defined an abstract quantity $H = \int f \log f$ that, in essence, quantified the *degree of certainty* of (or *amount of knowledge* implicit in) any (continuous) probability density function f.
    - **Note:** This paper deserves significant credit as the historical antecedent (70 years ahead!) of Shannon's entropy H, and the entire field of *information theory!*
    - Boltzmann showed, in his theorem, that H tends to decrease over time as gas molecules collide, and suggested that physical entropy S was related (negatively) to H,
      - but he did not yet know how to derive an exact relation between these quantities…

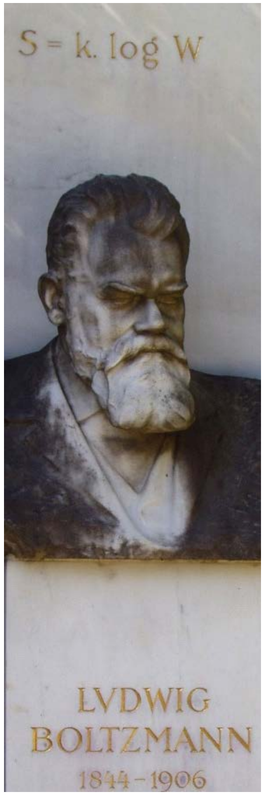
Ludwig Boltzmann
Sire of statistical mechanics, "great-grandfather" of chaos theory & information theory!

History of Entropy, cont. (3/5)





- Max Planck (1901) inferred, via his study of blackbody spectra, that physical states must be *quantized*,
  - As opposed to the *continuous* state variables used in classical mechanical models…
- *I.e.*, you can *count* a (finite) number of distinct states!
  - This let Planck finally give Boltzmann's intuition (that entropy is a statistical quantity) a precise quantitative foundation.
- Planck was first to calculate the magnitude, in physical units, of a natural logarithmic quantum of entropy k:
  - $[\log \mathrm{e}] = k = k_{\mathrm{B}} \cong 1.38 \times 10^{-23}\ \mathrm{J/K}$
    - Called "Boltzmann's constant" to honor Boltzmann's earlier role
- Defining k physically lets us write $\hat{S} = \log W = k \ln W$.
  - Rendered as $S = k \log W$ on Boltzmann's tombstone… →
    - The separate quantity H was, at this point, no longer needed…



History of entropy, cont. (4/5)

- John von Neumann developed not only the ENIAC's "von Neumann" architecture, but also (among many other things) the mathematical formulation of quantum mechanics…
  - Unifying the Schrödinger and Heisenberg formalisms
- In a 1927 paper, he showed how to translate the Boltzmann-Planck entropy concept (for the general case of nonuniform probabilities) to the language of quantum states: $S = -k\,\mathrm{Tr}\,(\rho \ln \rho)$.
  - ρ = density matrix; Tr = matrix trace operator
- This is (exactly) just $S = -k \sum p \ln p$, where the p are the probabilities of the pure quantum states that the ρ matrix represents a statistical mix of.
  - If you already know about Shannon entropy, this formula should be starting to look awfully familiar…

John von Neumann
Formulated the modern concept of entropy

History of entropy, cont. (5/5)

- Shannon (1948) is usually credited with "inventing the information-theoretic concept of entropy,"

$$H = -\sum p \log_2 p \ \text{ bits},$$

- However, in that formula, Shannon was really just *reformulating and reapplying* already-existing concepts that were already very well-established in physics:
  - **Note:** Shannon explicitly *cites* Boltzmann's contribution!
    - The symbol H, and the use of the expected log-probability quantity come straight out of Boltzmann's 1872 H-theorem!
      - The sign of H here is just changed to match that of S
    - Also, the transition from Boltzmann's continuous $\int$ to the discrete Σ case had <u>already been completed</u> over the period 1901-1927 (21 years prior!) by Planck and von Neumann.
  - Further, the change of the entropy unit from k to bit only reflects a shift in the conventional choice of the logarithmic unit from log e to log 2, nothing more.
- So really, the only true innovation in Shannon's entropy concept was:
  - The states that Shannon was explicitly concerned with in his work were not microscopic *physical* states, but macroscopic *digital* or symbolic states.
    - Yet, Shannon's entropy connects fundamentally to physical entropy, as we'll see…

Claude Shannon
Applied entropy in communication theory

Entropy in a Nutshell

Basic review + coining some useful terminology

- Define the "surprisingness" or *surprise* $s(x)$ of any event x that has a 1 in m chance of occurring as $s = s(x) = s(m) = \log m$.
  - Call the $m \geq 1$ "improbability;" it can be a non-integer.
    - s is logarithmic b/c the improbabilities of independent surprises multiply.
  - *Indefinite* logarithm; dimensioned in *arbitrary* logarithmic units.
    - Some example units: $\log 2 = 1$ bit; $\log \mathrm{e} = 1$ nat $= k_{\mathrm{B}}$; $\log 10 = 1$ bel.
- In terms of event's *probability* $p = p(x) = p(m) = 1/m$,

$$s(p) = \log\frac{1}{p} = -\log p\,.$$

- Define event's psychological "*heaviness*" $h = h(x) = h(p)$ as its surprise, weighted by its probability:

$$h(p) = s/m = p \cdot s = p\log m = -p\log p.$$

- Then for any probability distribution $p(x)$ over any mutually exclusive and exhaustive set of events $\mathbf{X} = \{x_1, \ldots, x_n\}$, we have that the ***expected surprise*** $S(X) = \mathrm{E}_p[s(x)]$ and the ***total heaviness*** $H(X) = \sum_{x\in X} h(x)$ associated with that particular set of possible events are the same, and are given by:

$$S(X) = \sum_{x\in X} p(x)\cdot s(x) = H(X) = -\sum_{x\in X} p(x)\cdot \log p(x)\,.$$

- We call this quantity $H = S$ the *entropy* of the given epistemic situation.
  - By convention, we'll prefer H for "computational" entropy, S for "physical" entropy.

Improbability:

$m = 6\times 6 = 36$

Surprise:

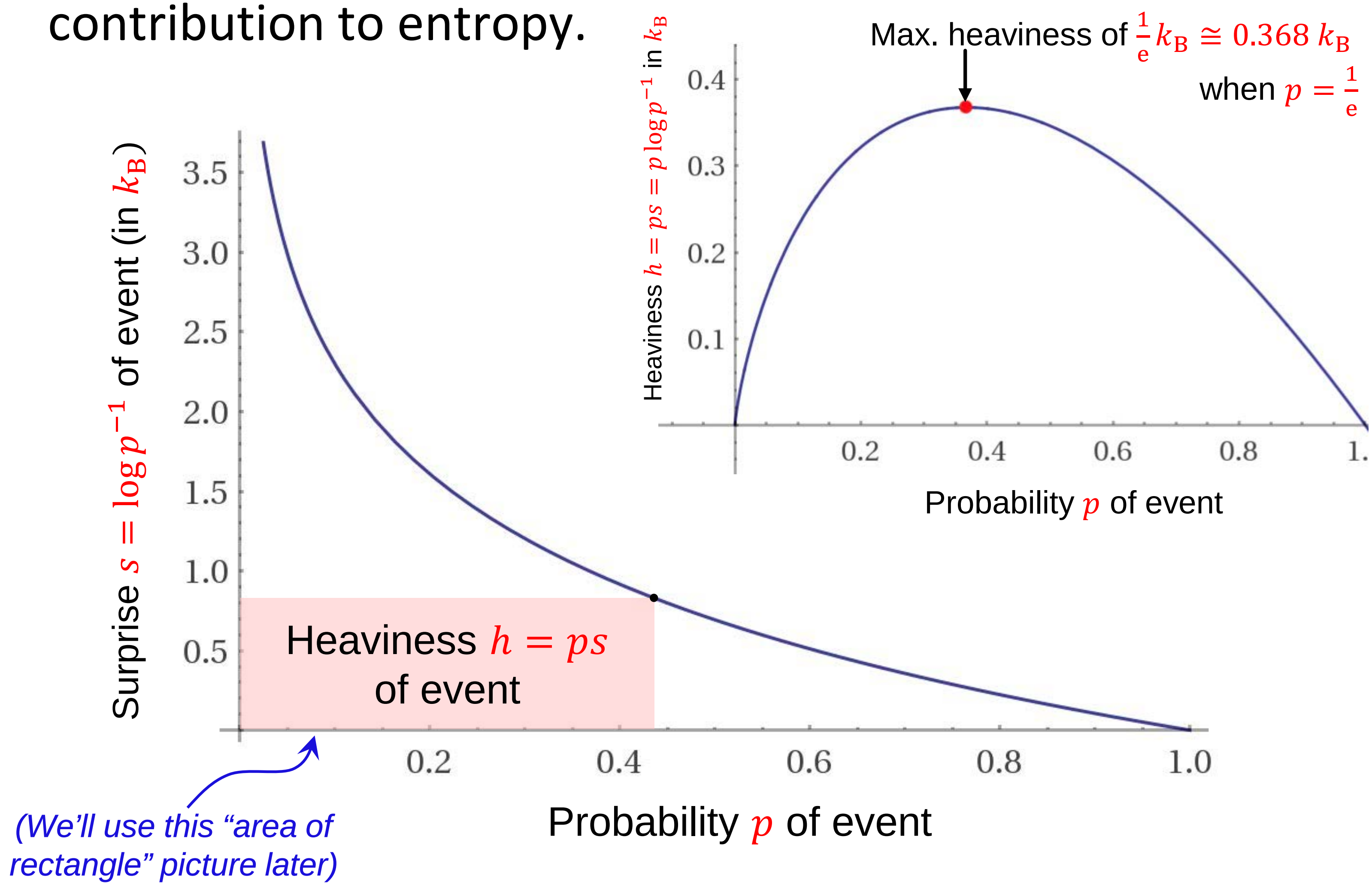
$s = 2(\log 6)$

Heaviness:

$h = \frac{s}{m} = \frac{2}{36}\log 6$

Surprise and Heaviness Functions

- For an individual state's contribution to entropy.



Known Information & Information Capacity

- Given a discrete variable V with a state set $\mathbf{V}$, and a probability distribution $P(V)$,
  - The *amount of known information* $K(V)$ about V is given by the maximum entropy $\widehat{H}$ minus the entropy H:

$$\begin{aligned} K(V) &= \widehat{H}(V) - H(V) \\ &= \log|\mathbf{V}| - H(V) \end{aligned}$$

  - The *information capacity* $\mathbf{I}(V)$ of V is the maximum known information or entropy, and is also the sum of known information and entropy.

$$\begin{aligned} \mathbf{I}(V) &= \widehat{K}(V) = \widehat{H}(V) = \log|\mathbf{V}| \\ &= K(V) + H(V) \end{aligned}$$

Conditional Entropy & Mutual Information

- Given a discrete variable V with a state set $\mathbf{V}$ expressible as a cartesian product $\mathbf{X} \times \mathbf{Y}$ of two state sets $\mathbf{X}, \mathbf{Y}$ for variables X, Y, and a joint probability distribution $P(X,Y)$,
  - The *conditional entropy* $H(X \mid Y) = H(X,Y) - H(Y)$.
    - Expected value of $H(X)$ that would result from learning the value of Y.
  - The *mutual information* is a symmetric function given by:

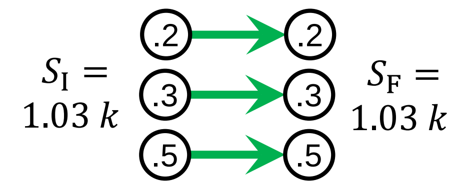
$$\begin{aligned} I(X;Y) = I(Y;X) &= K(X,Y) - K(X) - K(Y) \\ &= H(X) + H(Y) - H(X,Y) \\ &= H(X) - H(X \mid Y) \\ &= H(Y) - H(Y \mid X). \end{aligned}$$

    - The amount of shared/redundant information between X and Y.
      - The degree of information-theoretic *correlation* between the variables.

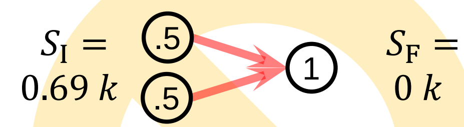
Entropy and Time Evolution

- Microscopic dynamics is one-to-one (injective).
  - A consequence of unitary quantum time evolution.
  - If we could track fully-detailed physical time evolution perfectly, we would see no entropy increase!
    - Probability distribution unchanged, just on new states
- In fact, this *reversibility* of microphysics <u>underlies</u> the Second Law of Thermodynamics.
  - If physics was not injective, entropy could decrease!
- But, entropy can be seen to *increase*, from our subjective perspective as modelers if we have any uncertainty about the microscopic dynamics, or cannot keep track of it in detail…
  - Thus, entropy increase only exists as a subjective epistemological phenomenon…
    - It is always fundamentally just a reflection of our degree of *ignorance & incompetence* in modeling the world…

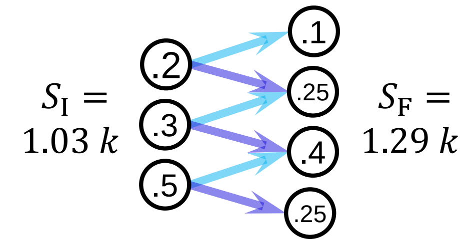
$$S[p] = \mathrm{E}_p[\log p^{-1}]$$



Bijective microphysics → No “true” entropy change



Irreversible microphysics → Entropy would decrease (Second Law of Thermo. would be violated)



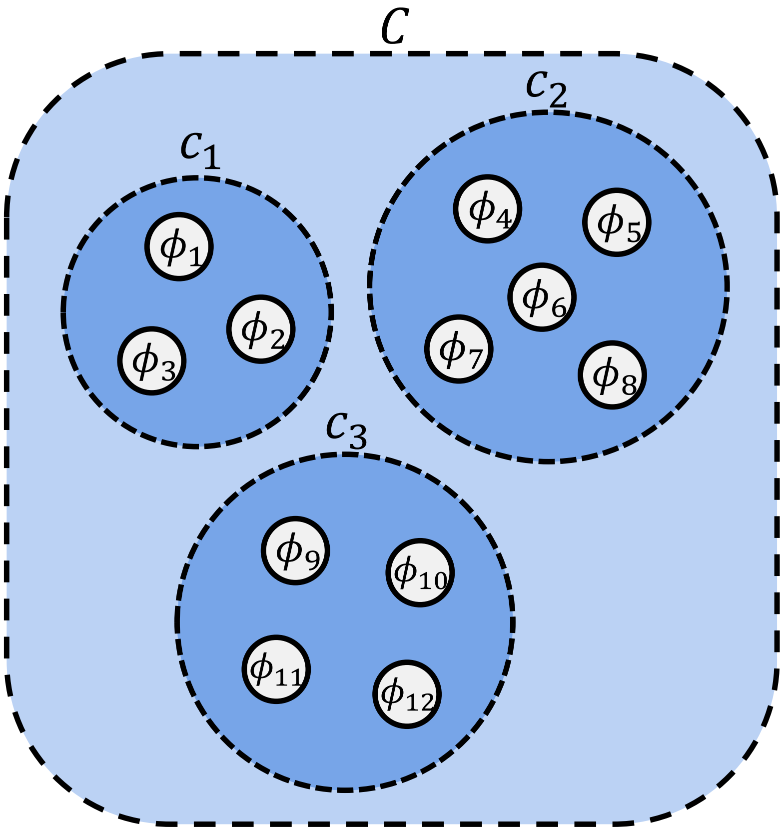
True dynamics uncertain (or not tracked in detail) → Entropy increases

From Physics to Computation

- Thermodynamics and quantum mechanics show that any bounded physical system admits only a finite set $\Phi = \{\phi_1, \ldots, \phi_n\}$ of measurably distinguishable detailed physical states (*microstates*).
  - *E.g.*, Φ could be any orthogonal set of basis vectors for the system's Hilbert space.
- We can *group* these microstates, that is, partition them into subsets c_j of micro-states that we consider as *equivalent* to each other for some designated purpose…
  - *e.g.*, for purposes of representing some specific *computational* information
- Any probability distribution $p(\phi_i)$ over the physical state space Φ induces a probability distribution P over the computational state space (subsystem) $C = \{c_j\}$ as well…

$$P(c_j) = \sum_{\phi_i \in c_j} p(\phi_i).$$

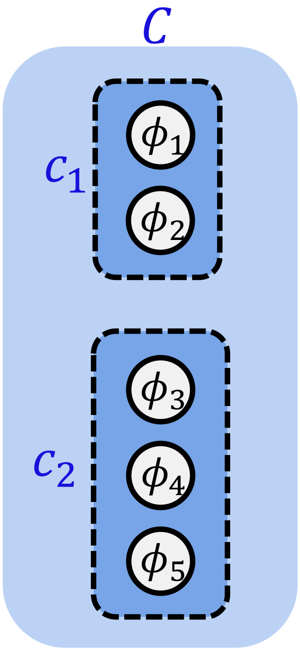
- This implies a *computational entropy* $H(C)$.

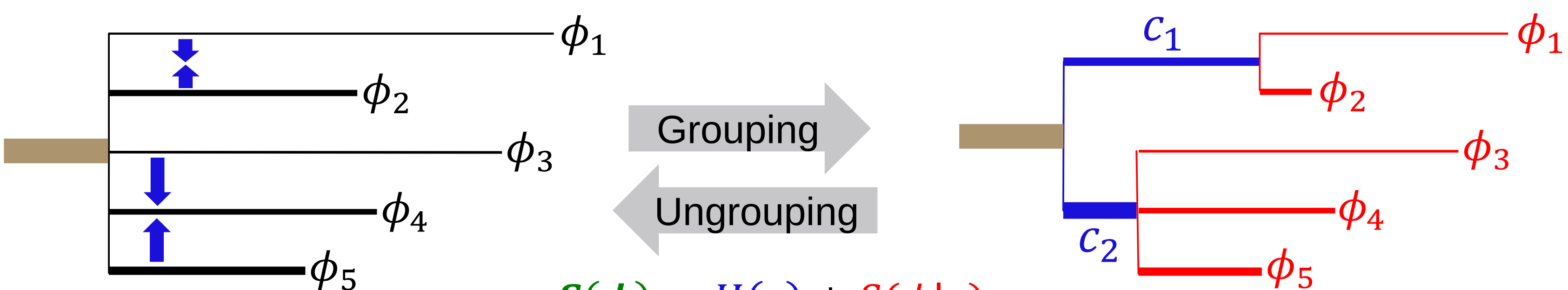


Example of a computational state space C consisting of 3 distinct computational states c_1, c_2, c_3, each defined as a set of equivalent physical states.

Visualizing Entropy of Grouped States

- Can represent a hierarchy of events in a tree structure…
  - Branch ***thickness*** = event probability p.
  - Branch ***length*** = *incremental surprise* Δs associated w. event,
    - relative to whatever base event it's branching off from.
  - Branch ***area*** = event's *incremental heaviness* $\Delta h = p\Delta s$, *i.e.*,
    - its contribution to total entropy, in addition to its base event's.
- *Grouping* events into larger events has these effects:
  - Thicknesses (probs.) of branches combine in parent branch
  - A corresponding part of the total length (surprise) of each branch is reassociated to parent (stem) branch.
  - Note: The total heaviness $\boldsymbol{H}$ of all branches and stems (total entropy ***S***) is not changed at all by any grouping/ungrouping!!

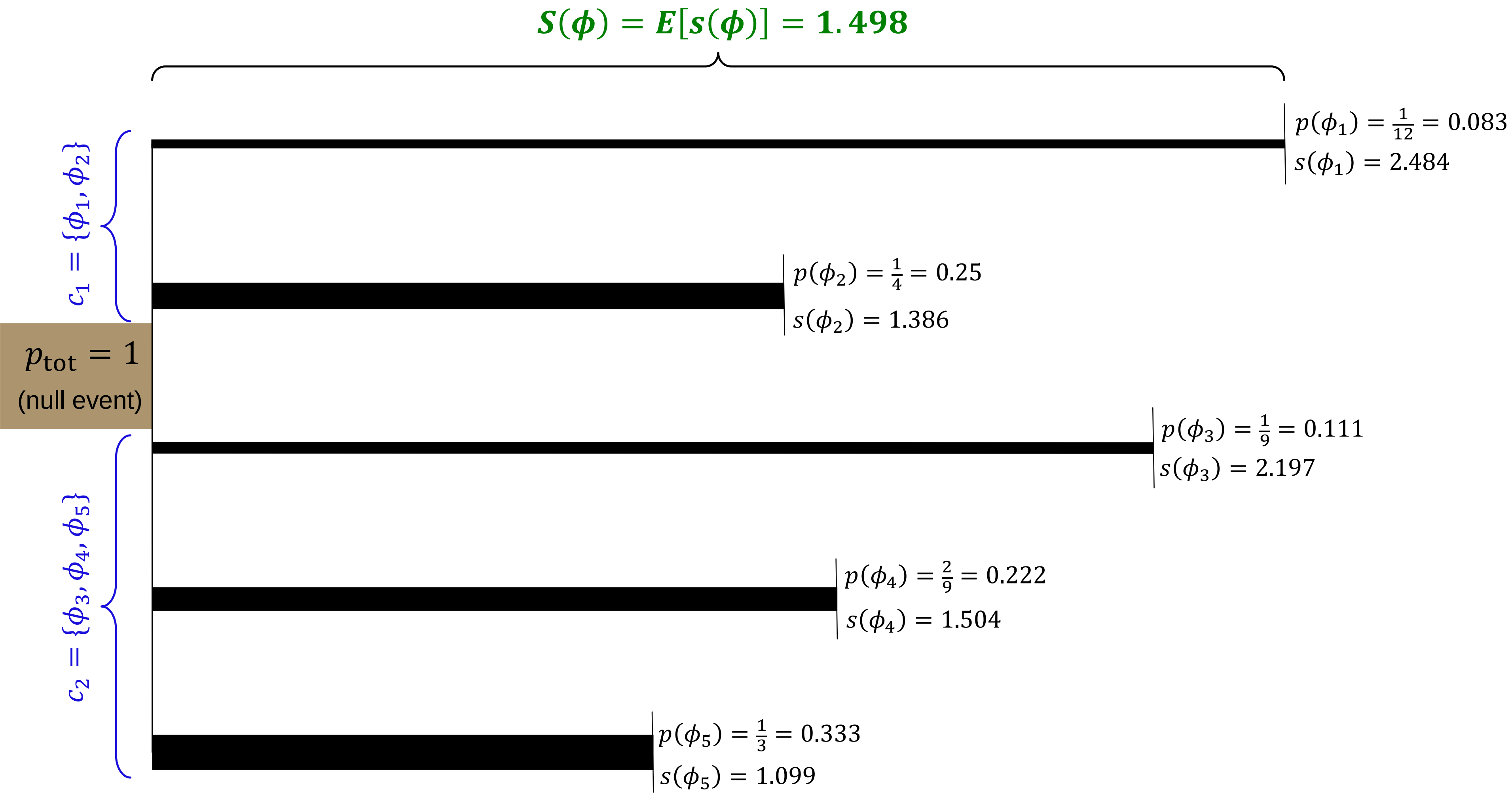




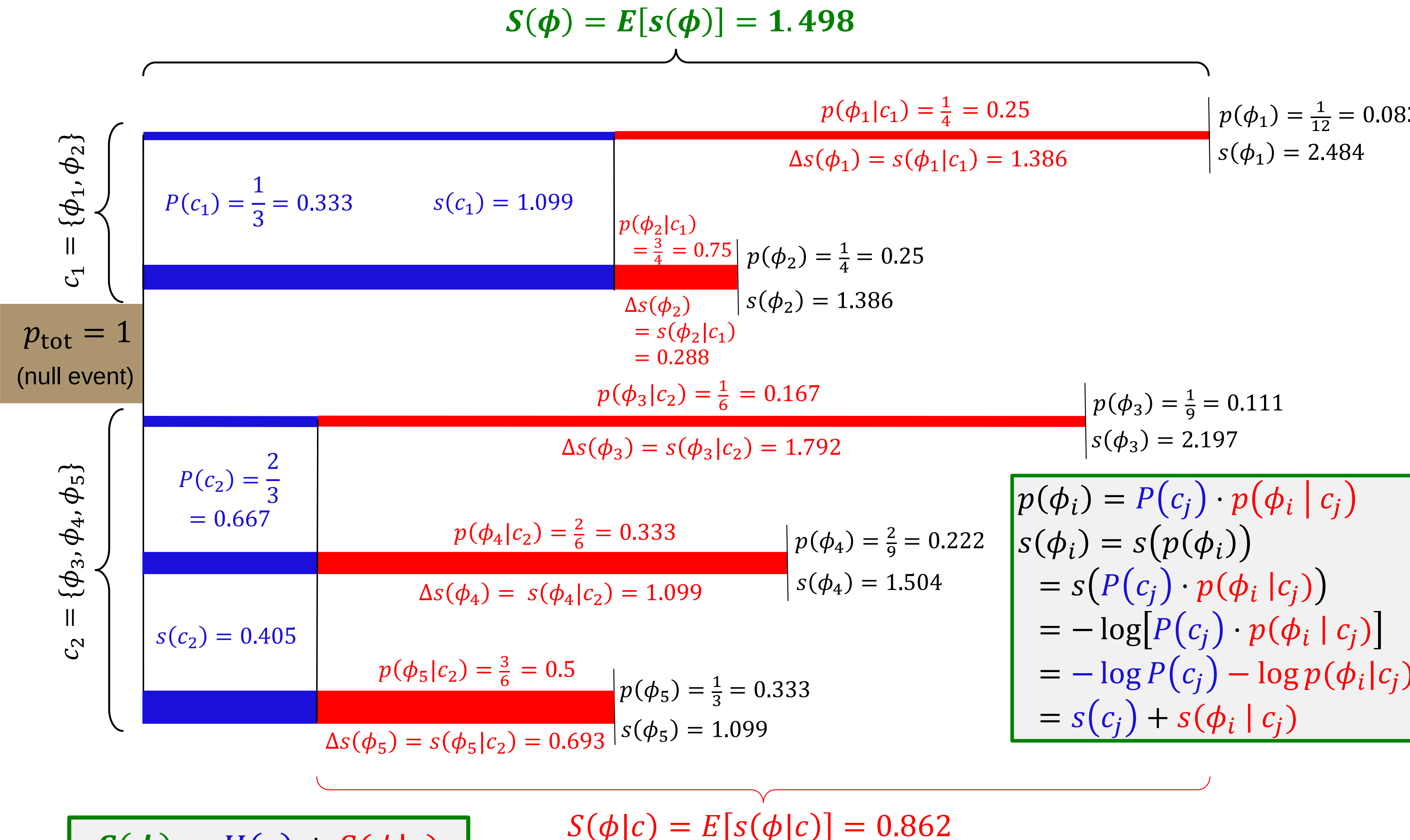
$$\boldsymbol{S}(\boldsymbol{\phi}) = H(c) + S(\phi|c)$$

Total system entropy = computational entropy + non-computational entropy

Grouping of States (slide 1 of 3)



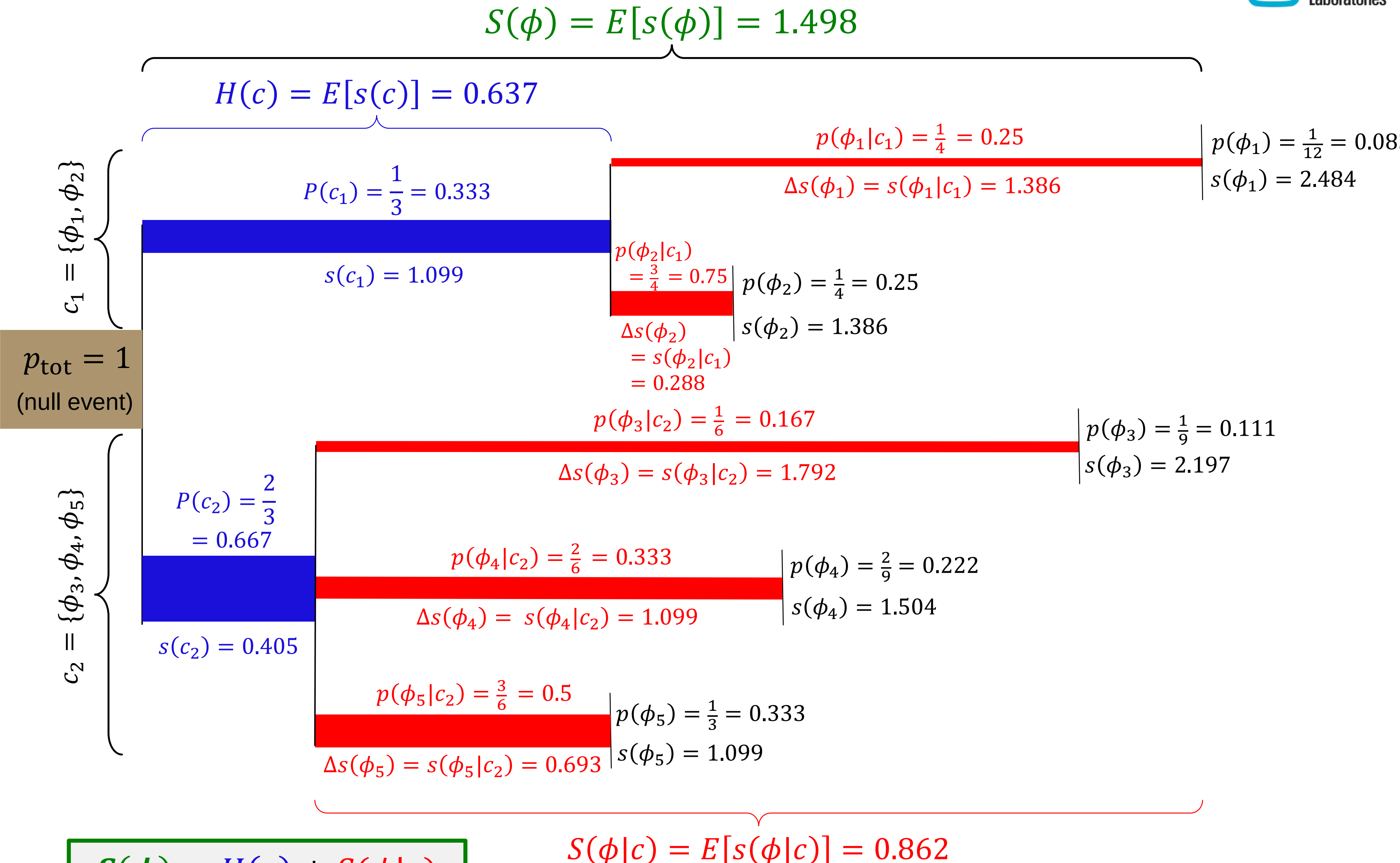
Grouping of States (slide 2 of 3)



$$
\begin{aligned}
p(\phi_i) &= P(c_j) \cdot p(\phi_i \mid c_j) \\
s(\phi_i) &= s(p(\phi_i)) \\
&= s\big(P(c_j) \cdot p(\phi_i \mid c_j)\big) \\
&= -\log\big[P(c_j) \cdot p(\phi_i \mid c_j)\big] \\
&= -\log P(c_j) - \log p(\phi_i|c_j) \\
&= s(c_j) + s(\phi_i \mid c_j)
\end{aligned}
$$

$$\boxed{S(\phi) = H(c) + S(\phi|c)}$$

Grouping of States (slide 3 of 3)



$$\boldsymbol{S(\phi) = H(c) + S(\phi|c)}$$

Total system entropy = computational entropy + non-computational entropy

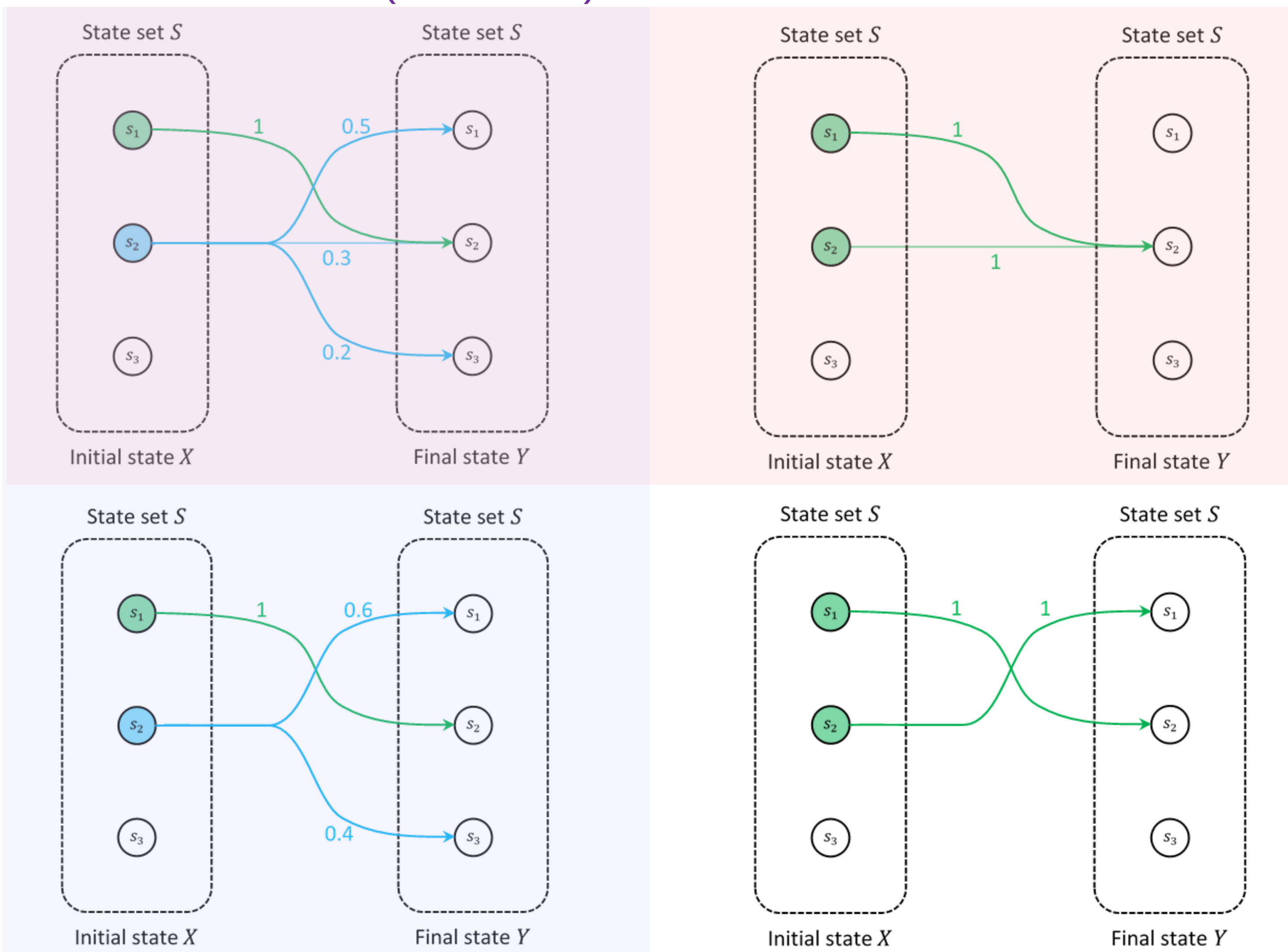
Types of Computational Operations

Define operations as (possibly partial) probabilistic transition relations

Nondeterministic (stochastic) | Deterministic

Irreversible (many-to-one)

(Unconditionally) Reversible (injective)

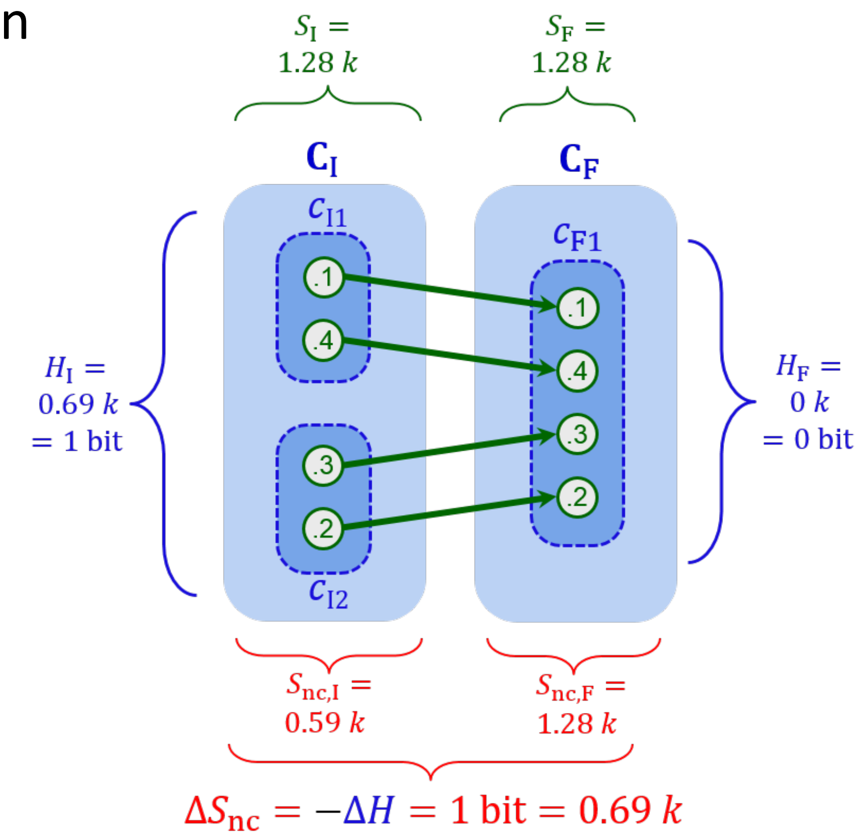


Entropy Ejection in Many-to-One Operations

- Again, a *computational state* c_j is just an equivalence class of physical states ϕ_i
  - On the left we see two computational states c_0, c_1, each with probability 0.5
- The computational subsystem has an induced information entropy $H(c)$.
  - Here, it is $H(c) = \log 2 = 1\text{ bit} = k \ln 2$.
- The *non*-computational subsystem (everything else) has expected entropy

$$S_{\mathrm{nc}} = S(\phi|c) = S(\phi) - H(c) = S - H$$

  - The conditional entropy of the physical state ϕ, given the computational state c.



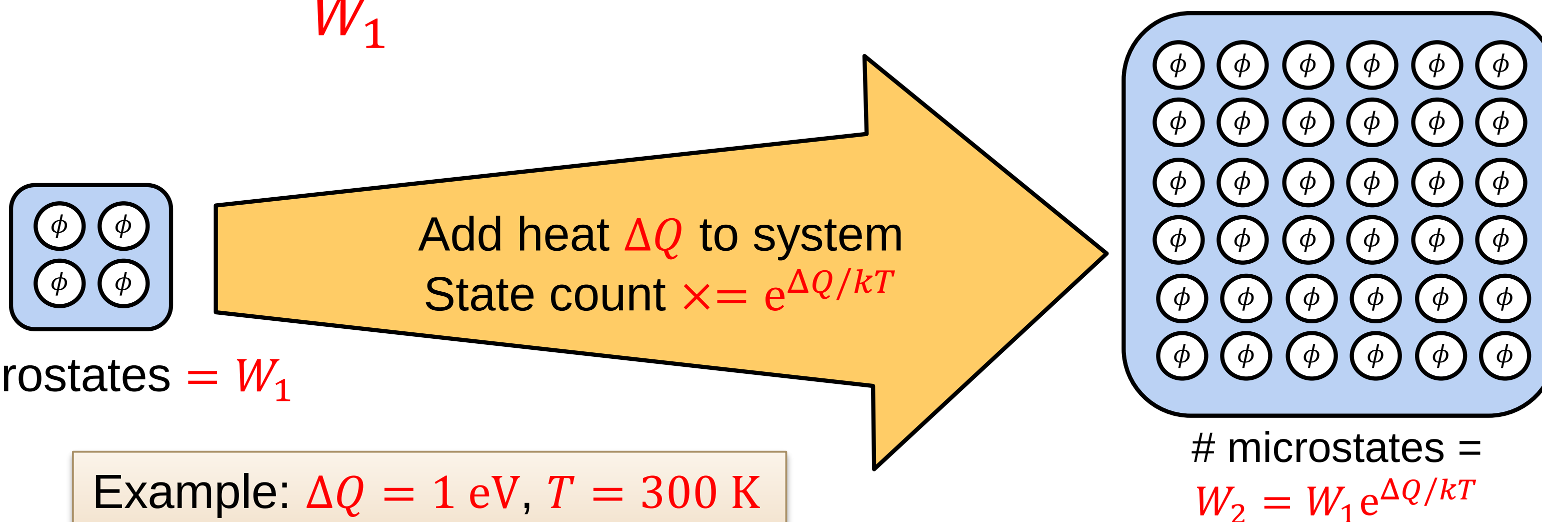
- Thus, if the computational entropy <u>*decreases*</u> (note here $\Delta H = -1\text{ bit}$),
  - The *non*-computational entropy *must <u>increase</u>* by $\Delta S_{\mathrm{nc}} = -\Delta H$ (here, $k \ln 2$).
- Thus, ejecting computational entropy $H = 1\text{ bit}$ implies we must add heat $\Delta Q = kT \ln 2$ to an environment at some temperature T.

Heat Increase inflates State Count

- The Boltzmann relation $\hat{S} = k \ln W$, together with the definition of temperature $1/T = \partial\hat{S}/\partial Q$, immediately implies that whenever a quantity of heat ΔQ gets added to a thermal system, its total number of accessible microstates gets multiplied by a factor $\blacksquare W$ that is given by $\mathrm{e}^{\Delta Q/kT}$:

$$\hat{S} = k \ln W \quad \Rightarrow \quad W = \mathrm{e}^{\hat{S}/k}$$

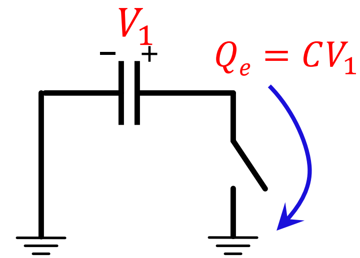
$$\blacksquare W = \frac{W_2}{W_1} = \mathrm{e}^{(\hat{S}_2 - \hat{S}_1)/k} = \mathrm{e}^{\Delta\hat{S}/k} = \mathrm{e}^{\Delta Q/kT}$$



Example: $\Delta Q = 1\ \mathrm{eV}$, $T = 300\ \mathrm{K}$ implies $\blacksquare W \cong 6.3 \times 10^{16}$

Energy Dissipation

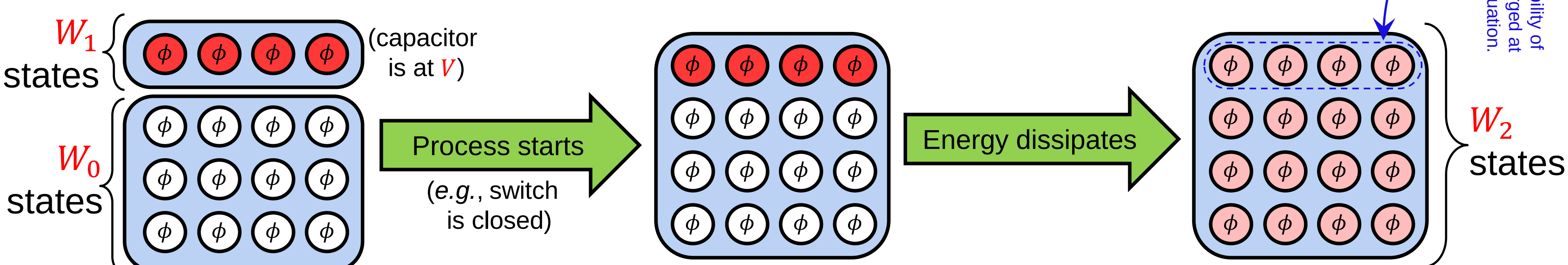
- For *any* means of dissipating some energy E_1 (*e.g.*, by discharging a capacitor), the dissipated energy $E_{\text{diss}} = E_1$ ends up as increased heat $\Delta Q = E_{\text{diss}}$ in a thermal environment at some temperature T…
  - Thus, any means of dissipating energy E_1 results in the same state count multiplier $\blacksquare W = \mathrm{e}^{E_{\text{diss}}/kT} = \mathrm{e}^{E_1/kT}$…



$$E_{\text{diss}} = E_1 = \frac{1}{2}CV_1^2$$

- Thus, we can represent any energy dissipation process as a merging of computational state sets, as follows:
  - Here, red shading indicates concentration of probability mass…

Note there is still a nonzero probability of still finding the capacitor to be charged at V, corresponding to a thermal fluctuation.



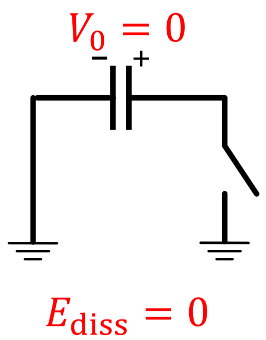
$W_0 = W_2 - W_1$
$= W_1(\mathrm{e}^{E_1/kT} - 1)$

(Before system thermalizes)

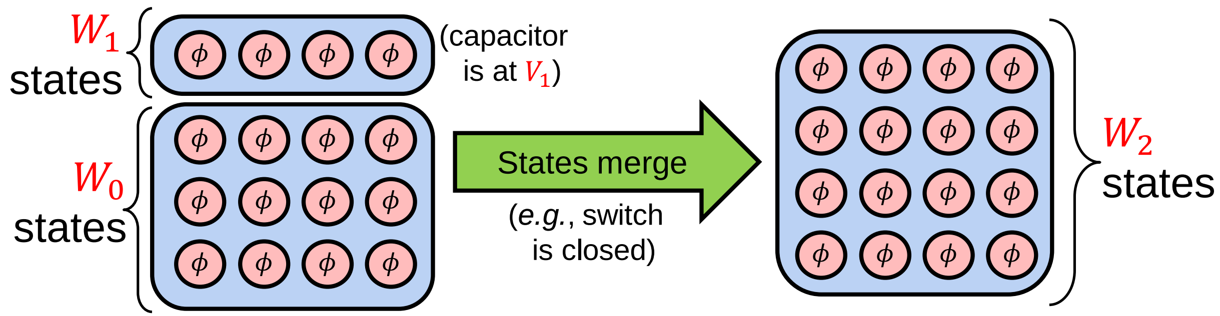
$W_2 = W_1 \cdot \blacksquare W$
$= W_1 \cdot \mathrm{e}^{E_1/kT}$

Entropy increase of $\Delta S = E_1/T$

Asymptotically conditionally reversible many-to-one operations



- In my RC'17 paper, I discussed how entropy ejected to the environment approaches 0 if the probability of <u>all but one</u> of a set of computational states being merged together approaches 0.
  - Corresponds to the case in this circuit example where the capacitor was initially at logic zero (subject to uncertainty due to thermal fluctuations).



$$W_0 = W_2 - W_1 = W_1(\mathrm{e}^{E_1/kT} - 1)$$

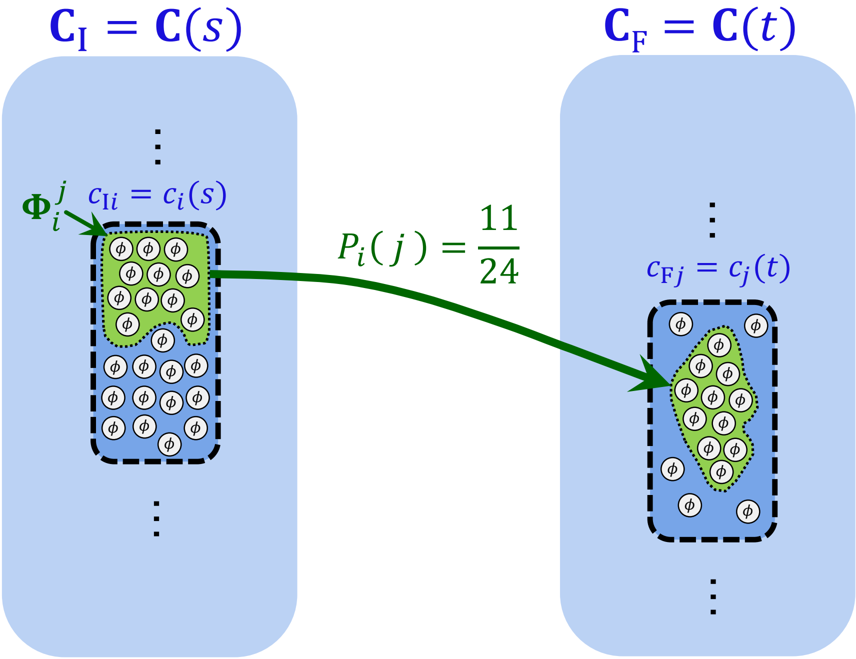
$$W_2 = W_1 \cdot \blacksquare W = W_1 \cdot \mathrm{e}^{E_1/kT}$$

- Here, the initial *computational* entropy is extremely tiny…
  - Prob. of being in logic 1 state initially is only $\mathrm{e}^{-E_1/kT}$, *e.g.*, 3.3×10^{-11} if $E_1 = 0.1$ aJ).
  - Entropy of initial computational state comes out to only $8.2 \times 10^{-10}k$.
    - Thus, only that tiny amount of entropy gets ejected to non-computational form in this case.

Note: Ejection/intake of entropy does not imply net entropy increase! (more on this later)

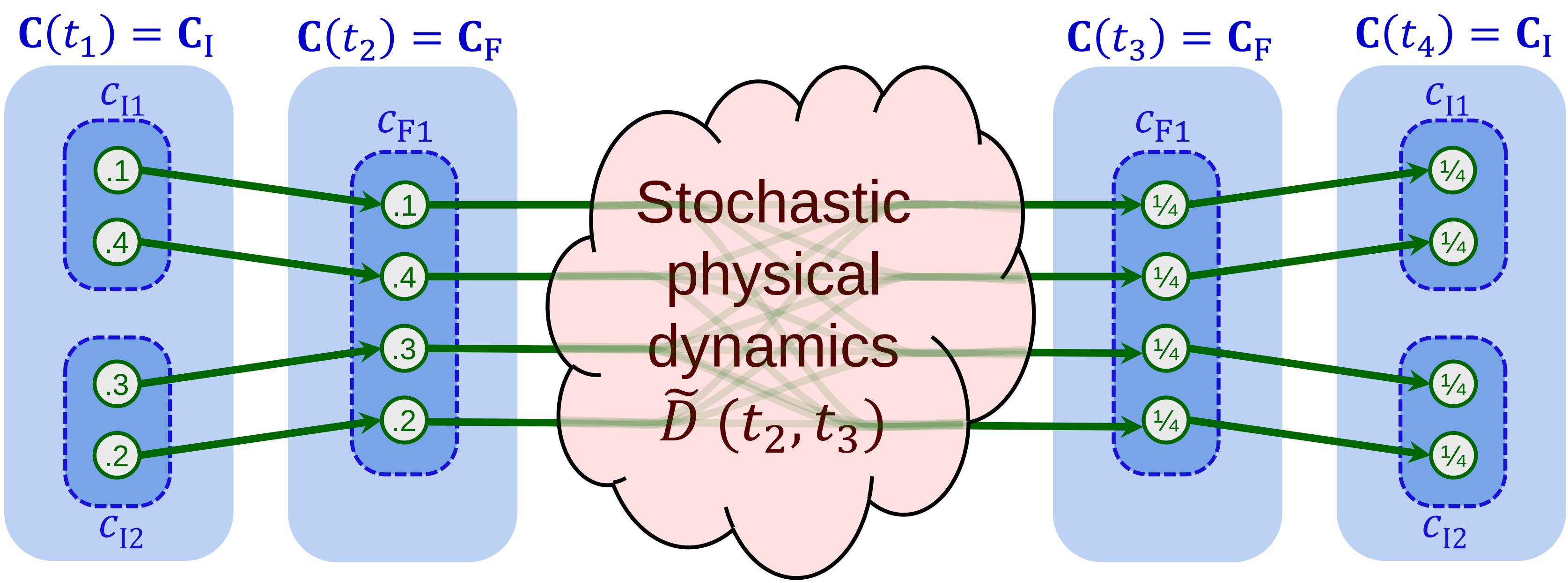
Stochastic computational operations

- These can occur in bijective dynamics if different subsets of the set of physical microstates making up the initial computational state transition to different new computational states…
  - In this example, initial computational state $c_{\mathrm{I}i}$ at time s has probability $P_i(j) = 11/24$ of transitioning to final computational state $c_{\mathrm{F}j}$
    - because 11 out of the 24 (equally-likely) microstates transition into there



Reversing Entropy Ejection (1/2)

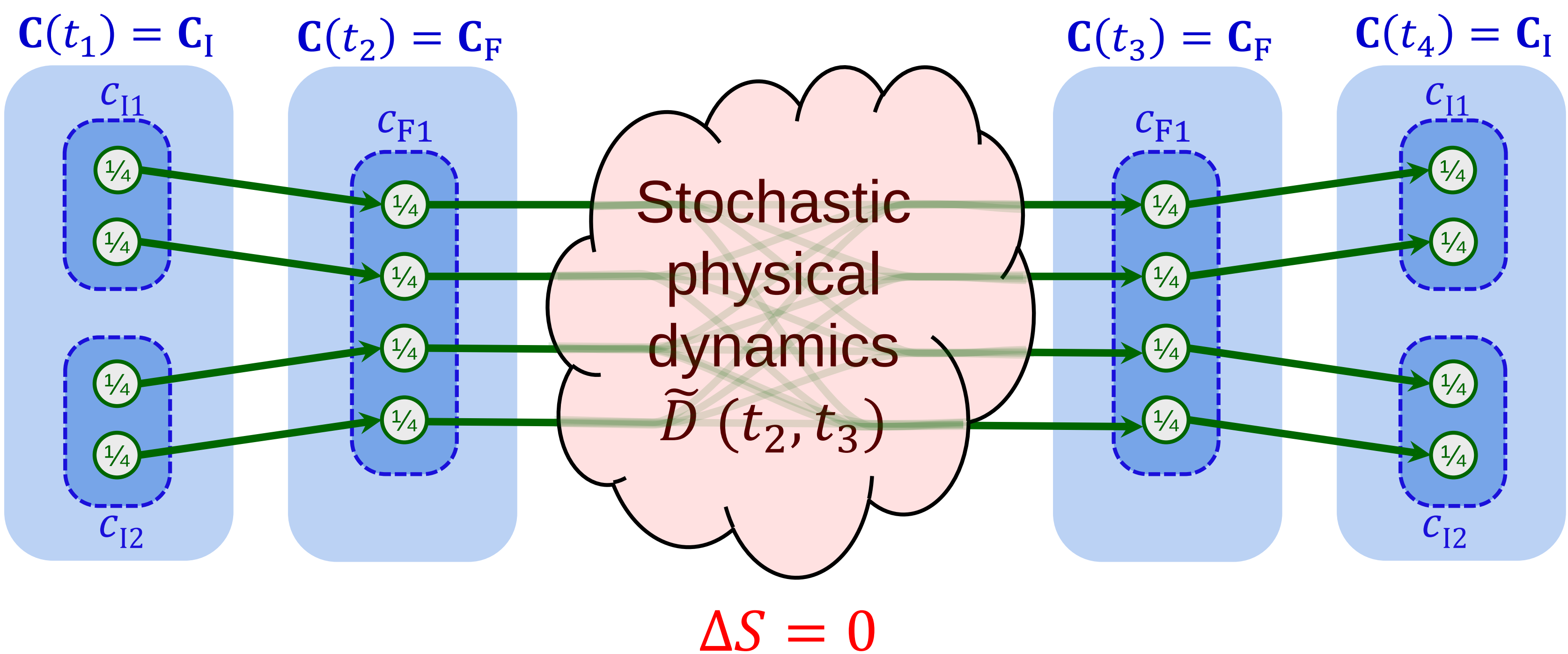
- The ejection of entropy can be reversible!
  - Here, we do a many-to-one operation followed by a stochastic operation to restore the computational entropy back to 1 bit
    - However, in this example, non-computational entropy increases by ~0.15 bits (some information about the initial physical state has been lost)



$$\Delta S = 0.106\ k = 0.153\ \text{bit}$$

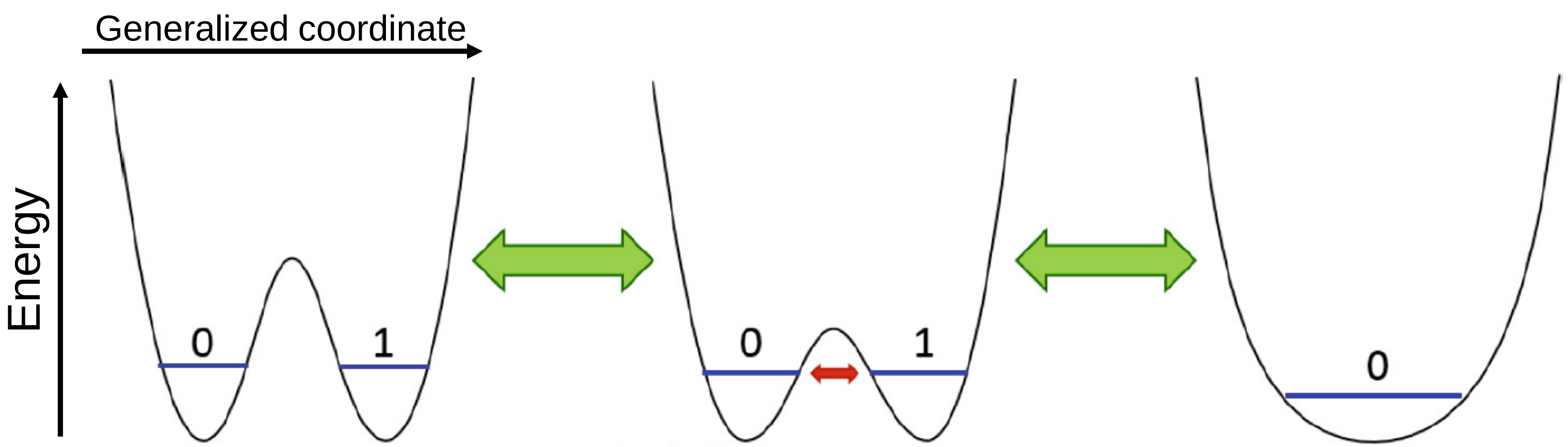
Reversing Entropy Ejection (2/2)

- The ejection of entropy can be reversible!
  - Here, we do a many-to-one operation followed by a stochastic operation to return the computational entropy to 1 bit
    - In ***<u>this</u>*** example, the initial non-computational entropy was *<u>already</u>* maximal, so there is no entropy increase!



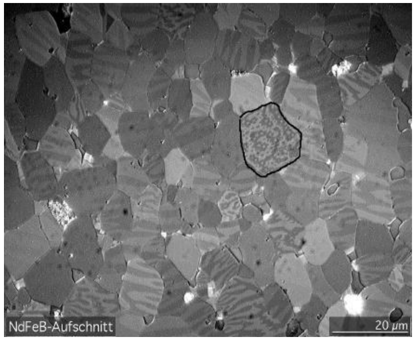
Bistable potential well implementation

- A simple class of physical implementations of the entropy ejection and intake process on the previous slide
  - Use two degenerate states separated by a potential barrier
    - *E.g.*, in quantum dots, superconducting circuits, many other systems
  - Lowering the barrier partially will allow states to equilibrate
  - Lowering the barrier fully will completely merge the states
  - Going left and then right re-randomizes the digital state
    - If done adiabatically, there is no increase in entropy in this process

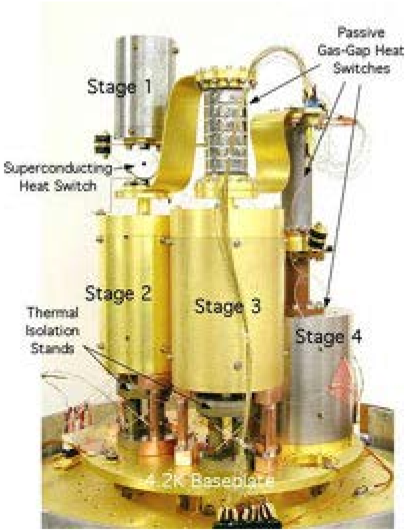


$$\mathbf{C}(s) = \mathbf{C}_{\mathrm{I}} = \{c_{\mathrm{I}0}, c_{\mathrm{I}1}\} \qquad \mathbf{C}\left(\frac{s+t}{2}\right) = \left\{c_0\left(\tfrac{s+t}{2}\right), c_1\left(\tfrac{s+t}{2}\right)\right\} \qquad \mathbf{C}(t) = \mathbf{C}_F = \{c_{\mathrm{F}0}\}$$

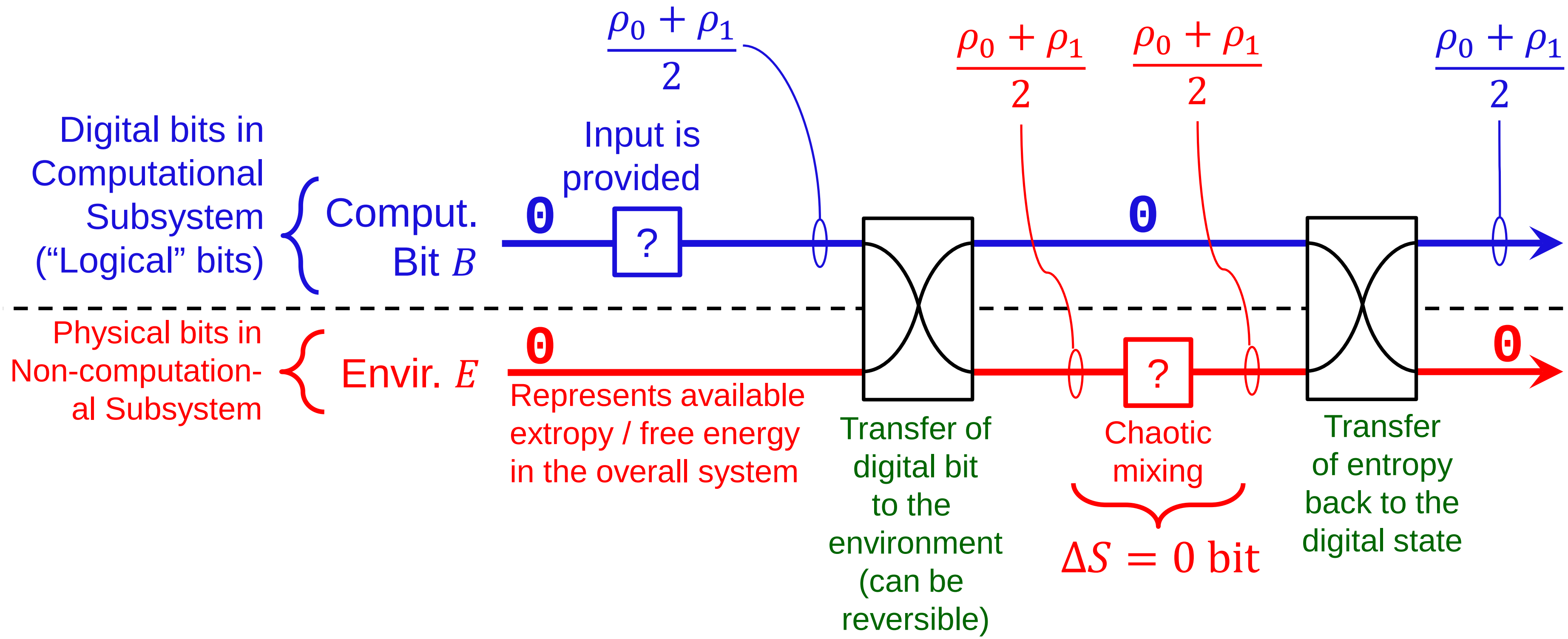
Adiabatic demagnetization / paramagnetic refrigeration

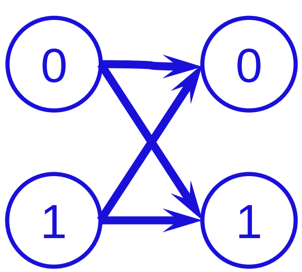


- A physical phenomenon that has been well studied for a very long time, and that nicely illustrates the reversibility of entropy ejection
  - Utilized in practice in cryogenic applications!
- The randomly-oriented magnetic domains in a sample of paramagnetic material can be considered to contain entropy in a "frozen," "digital," "computational" form…
  - When you apply a magnetic field and align the domains, this ejects the entropy from the domains and heats their surrounding environment…
  - But if you *remove* the magnetic field adiabatically, and allow the domains to re-randomize themselves, this *takes in* entropy from the thermal environment, locking it into this "digital" form, and *cools* the thermal environment!



Thermodynamically Reversible Erasure of an Uncorrelated Bit

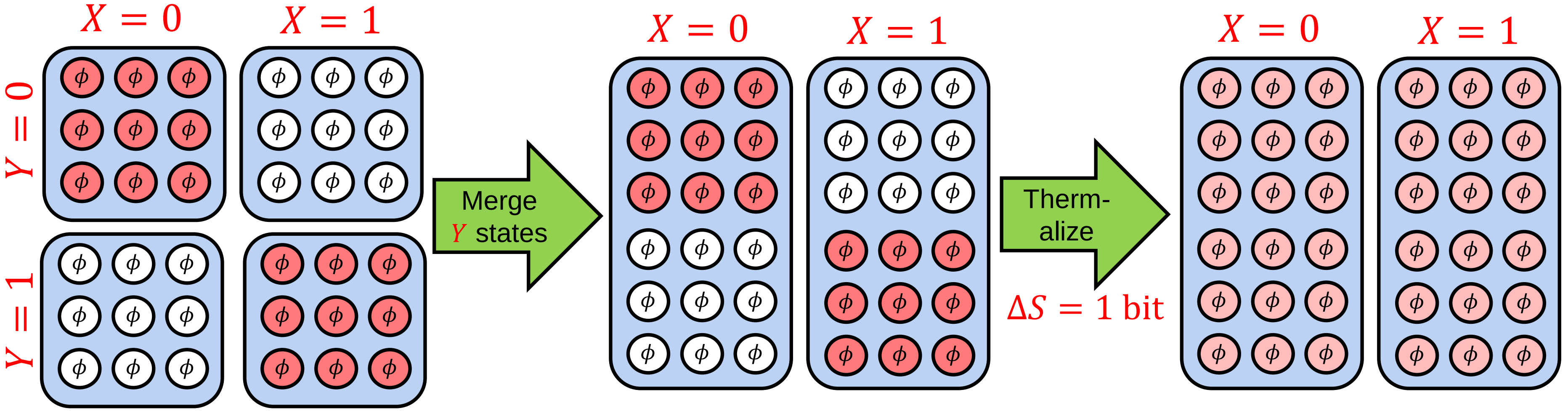




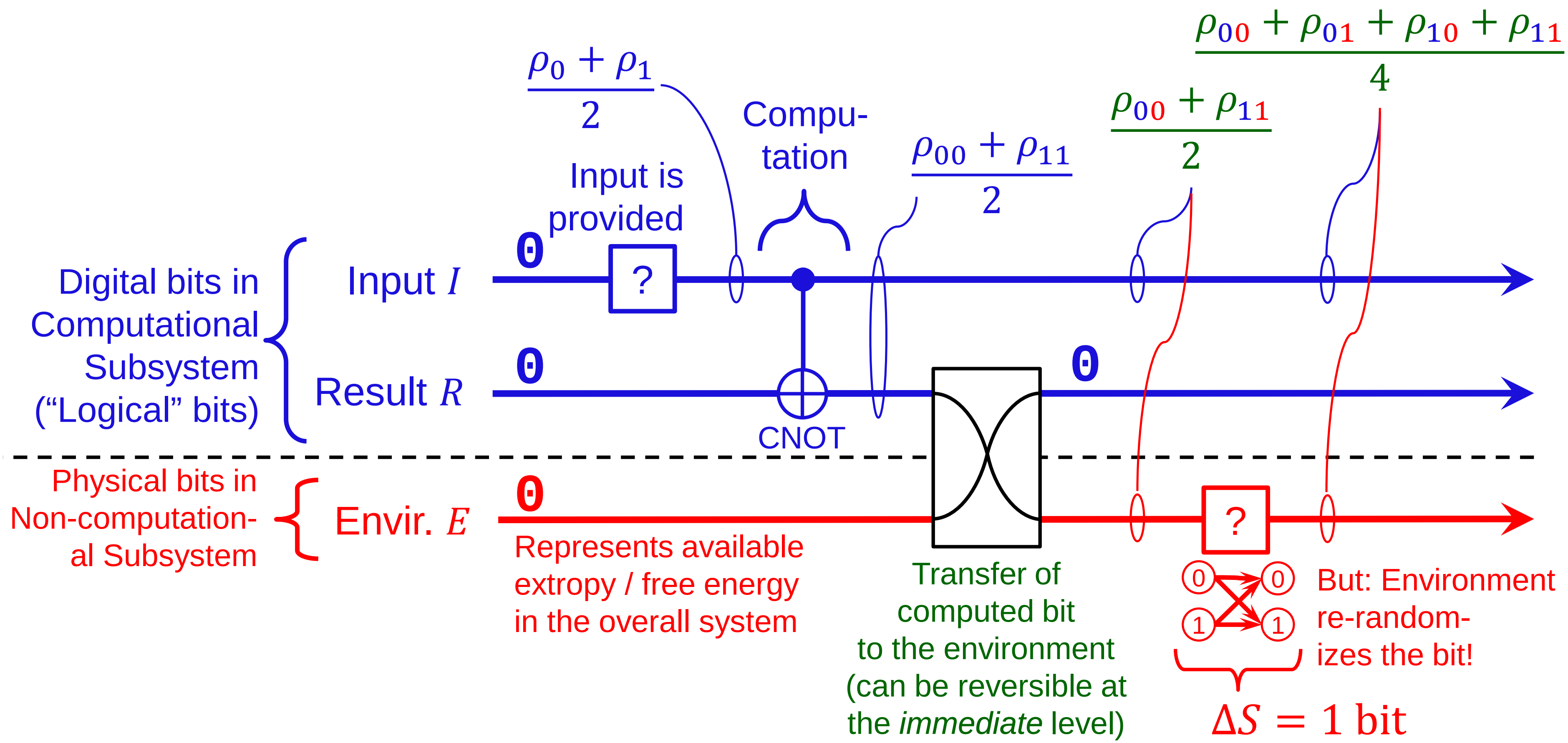
- Overall map including mixing is non-injective
- No autocorrelation between initial & final state
- Not “logically reversible” in traditional sense

Role of Correlations

- In light of the foregoing points, this is essential for understanding the true reason for the entropy increase in Landauer's principle!
  - Suppose we have two one-bit computational state variables, X and Y, and suppose that initially X is *random*, but we know that $Y = X$.
    - *E.g.*, this would be the case if Y was computed earlier using $Y := X$.
  - Thus, the joint system XY contains 1 bit of entropy, but also 1 bit of *known information* that is shared between X and Y (*i.e.*, mutual information).
  - If Y is then erased, this *known* computational information is ejected to become (briefly) known *non-computational* information, which then rapidly becomes thermalized, (permanently) increasing entropy by 1 bit's worth ($k \ln 2$).
    - This is why erasing *computed* bits (in particular) is *not* thermodynamically reversible!

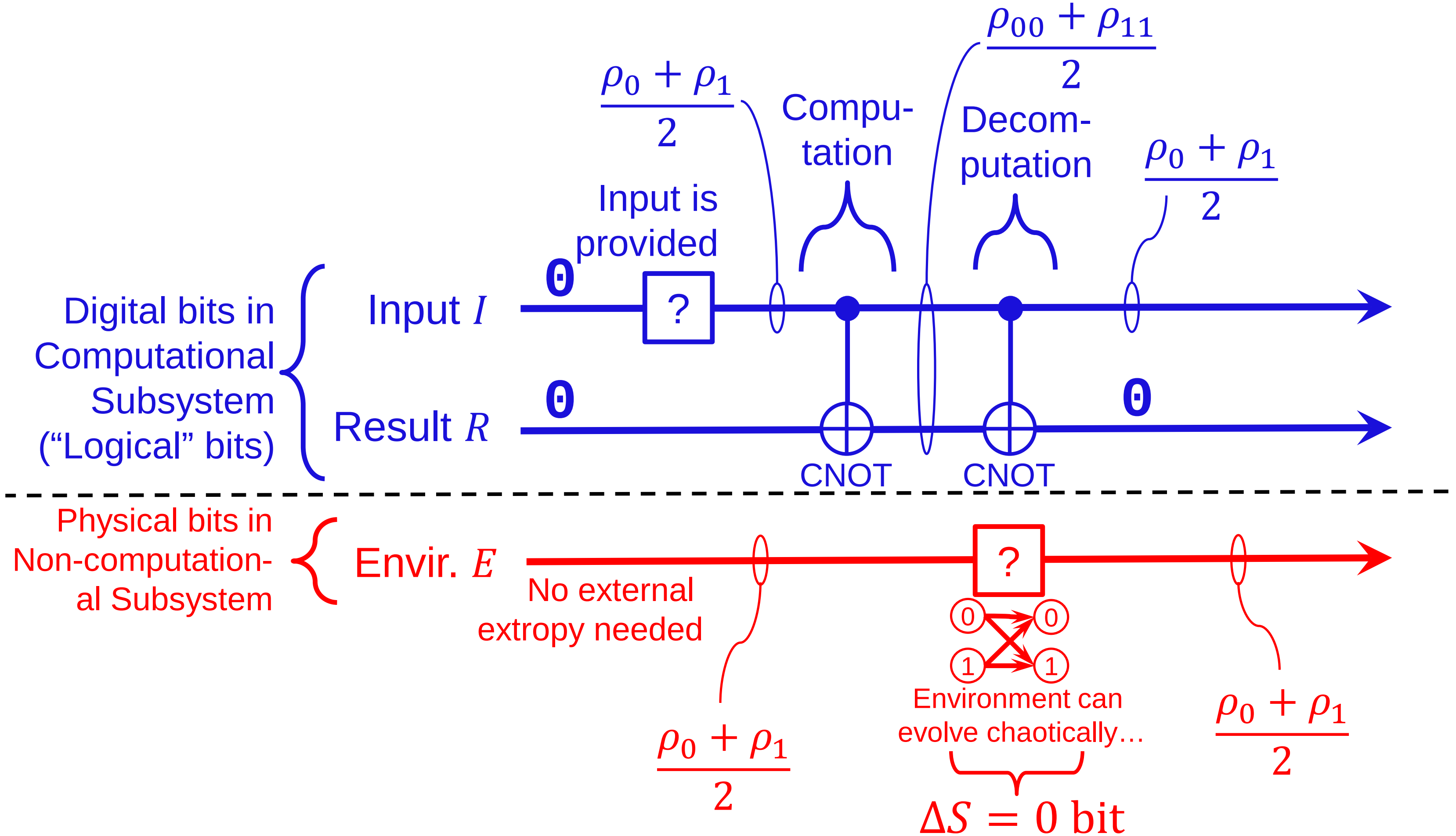


Logically Irreversible, Oblivious Erasure of a Correlated Bit



Moving a computed, correlated bit to an (*unpredictable!*) thermal environment *necessarily, inevitably* loses its correlations, and thus increases entropy!

Logically Reversible, Non-oblivious Decomputation of a Correlated Bit



Decomputing correlated bits, instead of ejecting them to the thermal environment, avoids losing correlations & increasing entropy!

This is why reversible computing (and **only** it!) can avoid Landauer’s limit…

Several recent empirical validations of Landauer's Principle

Redundant with the already very well-established, century-old basic facts of statistical physics, but hey, the doubters and skeptics are very stubborn people!

- Bérut *et al.*, 2012 (*Nature*)
  - Colloidal particle in a modulated double-well potential
  - Heat dissipation in bit erasure approached Landauer $kT \ln 2$ limit
- Orlov *et al.*, 2012 (*Japanese Journal of Applied Physics*)
  - Adiabatic charge transfer across a resistor
  - Verified that adiabatic transfers can dissipate $< kT \ln 2$
- Jun *et al.*, 2014 (*Physical Review Letters*)
  - Higher-precision version of Bérut experiment
  - Validated the limit, and that reversible transformations avoid it
- Yan *et al.*, 2018 (*Physical Review Letters*)
  - Quantum-mechanical experiment
  - Validated Landauer's Principle holds at single-atom level

Conclusion

- Landauer's principle really does follow directly as a simple and rigorous logical consequence of extremely fundamental insights of statistical physics that have been known for at least a century now…
  - The thermodynamic definitions of entropy and temperature
  - The reversibility of microphysics / second law of thermodynamics
  - Boltzmann & Planck's statistical understanding of physical entropy…
    - Von Neumann and Shannon only reformulated/reapplied it for specific domains!
      - Nothing fundamental about the Boltzmann/Planck definition was changed!
  - *Nothing else is needed* for the argument, except simple mathematics…
    - No other empirical inputs are even required to prove it!
    - We don't even need to make any equilibrium assumptions!
- However, to appreciate a number of subtleties of the proof is necessary if one wishes to help educate the skeptics… In particular:
  - The phenomenon of entropy intake in stochastic operations
  - The necessity of accounting for correlations
- In my opinion, the major barriers for our field are still *educational*:
  - Clarify these basic concepts → Engineering development will follow…