The Ambiguity of Simplicity

C. Aghamohammadi
J.R. Mahoney
J.P. Crutchfield

SFI WORKING PAPER: 2016-02-005

Santa Fe Institute Working Paper 16-02-XXX
arXiv:1602.XXXXX

The Ambiguity of Simplicity

Cina Aghamohammadi,* John R. Mahoney,† and James P. Crutchfield‡
Complexity Sciences Center and Physics Department,
University of California at Davis, One Shields Avenue, Davis, CA 95616
(Dated: February 27, 2016)

A system's apparent simplicity depends on whether it is represented classically or quantally. This is not so surprising, as classical and quantum physics are descriptive frameworks built on different assumptions that capture, emphasize, and express different properties and mechanisms. What is surprising is that, as we demonstrate, simplicity is ambiguous: the *relative* simplicity between two systems can *change sign* when moving between classical and quantum descriptions. Thus, notions of absolute physical simplicity—minimal structure or memory—at best form a partial, not a total, order. This suggests that appeals to principles of physical simplicity, via Ockham's Razor or to the "elegance" of competing theories, may be fundamentally subjective, perhaps even beyond the purview of physics itself. It also raises challenging questions in model selection between classical and quantum descriptions. Fortunately, experiments are now beginning to probe measures of simplicity, creating the potential to directly test for ambiguity.

Keywords: stochastic process, hidden Markov model, ϵ-machine, causal states, mutual information

PACS numbers: 03.67.-a 05.30.-d 89.75.Kd 89.70.+c

We are to admit no more causes of natural things than such as are both true and sufficient to explain their appearances.

Isaac Newton, 1687
Philosophiæ Naturalis Principia Mathematica,
Book III, p. 398 [1]

Introduction Beyond his theory of gravitation, development of the calculus, and pioneering work in optics, Newton engendered a critical abstract transition that has resonated down through the centuries, guiding and even accelerating science's growth: Physics began to perceive the world as one subject to concise mathematical Laws. Above, Newton suggests that these Laws are not only a correct perception ("true and sufficient") but they are also *simple* ("admit no more causes"). By his dictates we should abandon the Ptolemaic epicycle machinery as a description of planetary motion for Newton's more elegant $F = ma$ and $F_g \propto m_1 m_2 / r^2$.

The desire for simplicity in a theory naturally leads us to consider *simplicity as a means for comparing* alternative theories. Here, we compare the parsimony of classical physics and quantum mechanics descriptions of stochastic processes. Classical versus quantum comparisons seem, of late, to be of much interest both for reasons of principle and of experiment. *Quantum supremacy* holds that quantum systems behave in ways beyond those that can be efficiently simulated by classical computers [2]. In a single cold 2D Fermi gas a spatial transition from core quantum mechanical states to classical emerges [3, 4]. And, the experimental ionization dynamics of highly excited electron states of a single Rydberg atom are well described by classical chaotic repeller dynamics [5]. The impression that one gleans is that it is an interesting time for the foundations of quantum mechanics.

The following adds a new perspective to these debates on the balance of classical and quantum theory, as concerns the simplicity of their descriptions.

To start, we consider a Nature full of stationary stochastic processes. A theory, then, is a mathematical object capable of yielding a process' behaviors and their probabilities. We can straightforwardly say that one process is more random than another via comparing their temperatures or their thermodynamic entropies. But how to compare them in terms of their structural simplicities? We make use of a well developed measure of simplicity in stochastic processes—the statistical complexity [6]. Measuring a process' internal memory, it allows for a concrete and interpretable answer to the question, which process is structurally simpler? Having applied this comparison to all processes [7], we can then lay out the whole space in a neat array, graded in a linear order from the simplest to the most complicated.

An interesting twist comes about if we add quantum mechanics to our modeling toolbox. Using descriptions that act on a quantum substrate offers new and surprising options. For example, it was shown that a quantum mechanical description can lead to a simpler representation than classical [8, 9]. Recently, this quantum advantage was verified experimentally [10]. We note in particular that the closed-form methods introduced in Ref. [11] to measure quantum simplicity obviate many distracting concerns about generality, approximation, and estimation. This rigor greatly focuses any ensuing debate on how to measure simplicity. Leveraging this analysis leads to what is most surprising: what appears to be a generic quantum simplification is no where near so straightforward. We show that the *relative simplicity* of classical and quantum descriptions can change. Specifically, there are stochastic processes, A and B, for which the classical theory says A is simpler than B, but quantum mechanics says B is simpler than A. What started out as a neat classical array is upended by a new quantum simplicity order.

To appreciate this, we first discuss in more detail what we mean by simplicity. Then, to couch the discussion in terms as physical (and familiar) as possible, we analyze the one-dimensional Ising spin chain, showing how it inherently contains such an ambiguity of simplicity. Going further, we demonstrate that the ambiguity of simplicity is robust: there exist parameter regions in which the ambiguity is stable against alternative quantum representations, that arguably would lead to different simplicity metrics. We show this first for the 2D Ising spin lattice and then establish it generally: the quantum advantage requires ambiguity. Finally, we draw out potential impacts for classical-quantum model selection and then propose experimental tests.

Classical and Quantum Simplicity We consider stationary, ergodic processes: each a bi-infinite sequence of random variables $X_{-\infty:\infty} = \ldots X_{-2}X_{-1}X_0X_1X_2\ldots$ where each random variable X_t (upper case) takes some value x_t (lower case) in a discrete alphabet set $\mathcal{A}$ and where all probabilities $\Pr(X_t, \ldots, X_{t+L})$ are invariant under time translation.

How is their degree of randomness quantified? Information theory [12] measures the uncertainty in a single observation via the *Shannon entropy*: $H[X_0] = -\sum_{x\in\mathcal{A}} \Pr(x)\log_2 \Pr(x)$ and the irreducible uncertainty per observation via the *entropy rate* [13]: $h_\mu = \lim_{L\to\infty} H[X_{0:L}]/L$. If we interpret the left half $X_{-\infty:0}$ as the "past" and the right half $X_{0:\infty}$ as the "future", we see that the entropy rate is the average uncertainty in the next observable given the entire past: $h_\mu = H[X_0|X_{-\infty:0}]$. Thus, as we take into account the correlations in the past, the unconditioned single-observation uncertainty $H[X_0]$ reduces to h_μ.

How reducible is the uncertainty of the future $X_{0:L}$? Naively, this should scale as $L(H[X_0] - h_\mu)$, but due to correlations within the future, it must be less. The answer comes in the mutual information between the past and the future, a quantity known as the *excess entropy* [14, and references therein]: $\mathbf{E} = \mathrm{I}\left[X_{-\infty:0} : X_{0:\infty}\right]$. In h_μ and $\mathbf{E}$, we have measures of randomness and of how much is predictable in a process, respectively.

Computational mechanics [15] supplements these with a direct measure of structure—the amount of process memory. Its main construct, the *ϵ-machine*, is a process's minimal, unifilar predictor. As such, we view a process' ϵ-machine as the "theory" of a process: a mechanism that exactly simulates a process' behaviors.

The ϵ-machine consists of *causal states* $\sigma \in \boldsymbol{\mathcal{S}}$ defined by an equivalence relation $\sim$ that groups histories, say $x_{-\infty:t}$ and $x_{-\infty:t'}$, that lead to the same future predictions $\Pr(X_{t:\infty}|\cdot)$: $x_{-\infty:t} \sim x_{-\infty:t'} \iff \Pr(X_{t:\infty}|x_{-\infty:t}) = \Pr(X_{t':\infty}|x_{-\infty:t'})$. In other words, if a simpler set of states is sufficient, then that set will be the preferred representation. And so, we see that the ϵ-machine is, in a well defined sense, a process' simplest theory.

Translating this notion of simplicity into a measurable quantity, we ask: *What is the minimum memory necessary to* implement *the maximal reduction of future uncertainty (by* $\mathbf{E}$ *bits)?* The answer is explicit when phrased in terms of the ϵ-machine: the historical information stored in the present. Quantitatively, this is the Shannon entropy of the causal state stationary distribution—the *statistical complexity*:

$$C_\mu = \mathrm{H}\left[\boldsymbol{\mathcal{S}}\right] = -\sum_{\sigma\in\boldsymbol{\mathcal{S}}} \pi_\sigma \log_2 \pi_\sigma \ , \quad (1)$$

where π_σ is the probability of causal state σ.

It is well known that the excess entropy is a lower-bound on the information size of the ϵ-machine: $\mathbf{E} \leq C_\mu$. In fact, this relation is only rarely an equality [16]. So, while $\mathbf{E}$ quantifies the amount to which a process is subject to explanation by an ϵ-machine theory, this simplest theory is typically larger, informationally speaking, than the predictability benefit it confers. That said, the ϵ-machine is the best (simplest) theory. Thus, we use C_μ to define our notion of classical simplicity. It provides an interpretable ordering of processes—process A is simpler than process B when $C_\mu^A < C_\mu^B$.

We may also consider the recently proposed quantum-machine representation of processes [8, 9, 11]. The quantum-machine consists of a set $\{|\eta_k(L)\rangle\}$ of pure *signal states* that are in one-to-one correspondence with the classical causal states $\sigma_k \in \boldsymbol{\mathcal{S}}$. Each signal state $|\eta_k(L)\rangle$ encodes the set of length-L words that may follow σ_k, as well as each corresponding conditional probability used for prediction from σ_k. Fixing L, we construct quantum states of the form:

$$|\eta_j(L)\rangle \equiv \sum_{w^L\in|\mathcal{A}|^L} \sum_{\sigma_k\in\boldsymbol{\mathcal{S}}} \sqrt{\Pr(w^L, \sigma_k|\sigma_j)} \ |w^L\rangle \, |\sigma_k\rangle \ , \quad (2)$$

where w^L denotes a length-L word and $\Pr(w^L, \sigma_k|\sigma_j) = \Pr(X_{0:L} = w^L, \mathcal{S}_L = \sigma_k|\mathcal{S}_0 = \sigma_j)$. Due to ϵ-machine unifilarity, a word w^L following a causal state σ_j leads to only one subsequent causal state. Thus, $\Pr(w^L, \sigma_k|\sigma_j) = \Pr(w^L|\sigma_j)$. The resulting Hilbert space is the product $\mathcal{H}_w \otimes \mathcal{H}_\sigma$. Factor space $\mathcal{H}_\sigma$ is of size $|\boldsymbol{\mathcal{S}}|$, the number of classical causal states, with basis elements $|\sigma_k\rangle$. Factor space $\mathcal{H}_w$ is of size $|\mathcal{A}|^L$, the number of length-L words, with basis elements $|w^L\rangle = |x_0\rangle \cdots |x_{L-1}\rangle$.

The quantum measure of memory analogous to C_μ is the von Neumann entropy of the stationary state:

$$C_q = -\operatorname{Tr} \rho \log \rho \ , \quad (3)$$

where $\rho = \sum_i \pi_i \, |\eta_i\rangle \, \langle\eta_i|$. This quantum accounting of memory is generically less than the classical: $C_q \leq C_\mu$. Moreover, as with classical representations, the excess entropy provides a lower bound: $\mathbf{E} \leq C_q$, due to the Holevo bound [8, 17]. In fact, though rare in the space of processes, the classical and quantum informational sizes are equal exactly when both models are "maximally simple" or "ideal", that is, of size $\mathbf{E}$ bits: $C_q = \mathbf{E}$ and $C_\mu = \mathbf{E}$.

Ising Chain Simplicity To ground these ideas, let us consider the Ising spin chain, familiar from statistical physics [18], that historically played a critical role in understanding phase transitions [19], spin glasses [20], and lattice gasses [21]. Its impact has reached well beyond physics, too, to ecology [22], financial economics [23], and neuroscience [24]. Here, we first consider the one-dimensional nearest-neighbor Ising spin chain in the thermodynamic limit. The Hamiltonian is given by:

$$H = -\sum_{<i,j>} \left(J s_i s_j + b s_i\right) , \qquad (4)$$

where s_i, the spin at site i, takes on values $\{-1, +1\}$, J is the nearest-neighbor spin coupling constant, and b is the strength of the externally applied magnetic field.

One can measure each spin in the bi-infinite chain from left to right yielding the random variables $\ldots X_{-1}, X_0, X_1 \ldots$. In equilibrium this defines a stationary stochastic process that has been analyzed using computational mechanics [25]. Importantly, spins obey a conditional independence: $\Pr(X_{0:\infty}|x_{-\infty:0}) = \Pr(X_{0:\infty}|x_0)$. That is, the "future" spins (right half) depend not on the entire past (left half) but only on the most recent spin x_0. Therefore, spin configurations resulting from the Hamiltonian in Eq. (4) can be modeled by a simple two-state Markov chain consisting of up ($\uparrow$) and down ($\downarrow$) states with self-transition probabilities [25]: $p \equiv \Pr(\uparrow \mid \uparrow) = N_+/D$ and $q \equiv \Pr(\downarrow \mid \downarrow) = N_-/D$, where $N_\pm = \exp\beta(J \pm b)$ and:

$$\begin{aligned} D = &\exp(\beta J)\cosh(\beta b) \\ &+ \sqrt{\exp(-2\beta J) + \exp(2\beta J)\sinh(\beta b)^2} , \end{aligned}$$

with $\beta = 1/(k_B T)$.

Calculating the ϵ-machine via the causal-state equivalence relation is straightforward. There are exactly two causal states; except when $p = 1 - q$ where we find only one causal state. The conclusion is that the two-state Markov chain process is minimally represented by the ϵ-machine in Fig. 1. Using Eq. (1), the statistical complexity is directly calculated as a function of p and q. Figure 2 shows that C_μ is a monotonically increasing function of temperature T: $1 - C_\mu \propto T^{-2}$ at high T. In particular, for the three processes chosen at temperatures $T_\alpha < T_\gamma < T_\delta$: $C_\mu^\alpha < C_\mu^\gamma < C_\mu^\delta$.

Consider now the quantum representation of the spin configurations. Each causal state σ_1 and σ_2 is mapped

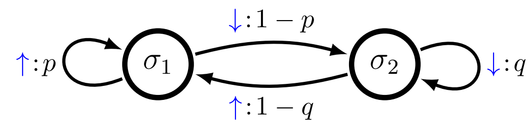

FIG. 1. The ϵ-machine for the nearest-neighbor Ising spin chain has two causal states σ_1 and σ_2. If the last observed spin x_0 is up ($s_0 = +1$) the current state is σ_1 and if it's down ($s_0 = -1$) is σ_2. If the current state is σ_1, with probability p the next spin observed is up and, if the current state is σ_2, with probability q the next spin observed is down.

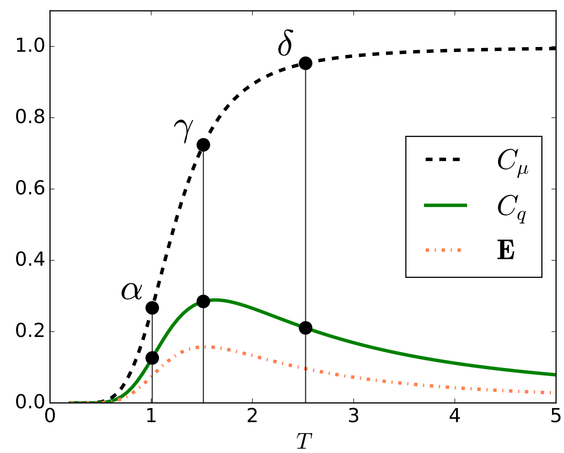

FIG. 2. Classical and quantum measures of Ising chain simplicity: Statistical complexity C_μ, quantum state complexity C_q, and excess entropy $\mathbf{E}$ versus temperature T in units of J/k_B at $b = 0.3$ and $J = 1$. ($C_\mu(T)$ and $\mathbf{E}(T)$ after Ref. [26] and $C_q(T)$ after Ref. [27].) Three particular spin processes are highlighted α, γ, and δ at temperatures T_α, T_γ, and T_δ.

to a pure quantum state that resides in a spin one-half space [27]:

$$\begin{aligned} |\sigma_1\rangle &= \sqrt{p}\,|\uparrow\rangle + \sqrt{1-p}\,|\downarrow\rangle \\ |\sigma_2\rangle &= \sqrt{1-q}\,|\uparrow\rangle + \sqrt{q}\,|\downarrow\rangle \ . \end{aligned} \qquad (5)$$

Intuitively, the quantum overlap accounts for the fact that the conditional predictions $\Pr(X_{0:\infty}|\sigma_1)$ and $\Pr(X_{0:\infty}|\sigma_2)$ share some subset of future outcomes. The density matrix for the ensemble is then:

$$\rho = \pi_1 |\sigma_1\rangle\langle\sigma_1| + \pi_2 |\sigma_2\rangle\langle\sigma_2| \ . \qquad (6)$$

Computing the quantum analog $C_q = -\operatorname{Tr}\rho\log\rho$ as a function of temperature, Fig. 2 shows that this quantum size is generically well below the classical size C_μ. Thus, the quantum theory for the Ising chain is simpler than the classical: $C_q^\alpha < C_\mu^\alpha$, $C_q^\gamma < C_\mu^\gamma$, and $C_q^\delta < C_\mu^\delta$. Given the broad progress of late in quantum information and computation [28, 29], it is notable, but perhaps no longer so surprising, that there exists such a quantum represen-

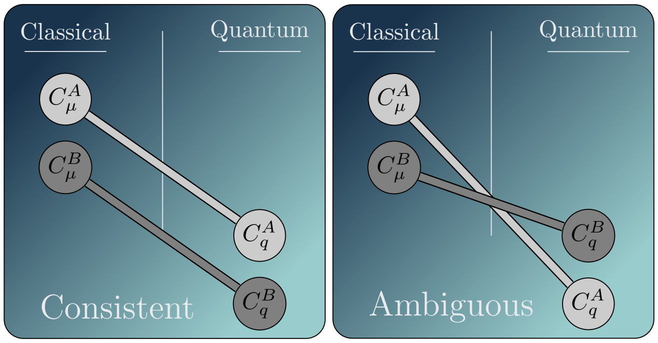

FIG. 3. (left) Classical and quantum rankings provide a consistent interpretation of which process is simpler. (right) Rankings reverse. And so, the question of simplicity is ambiguous.

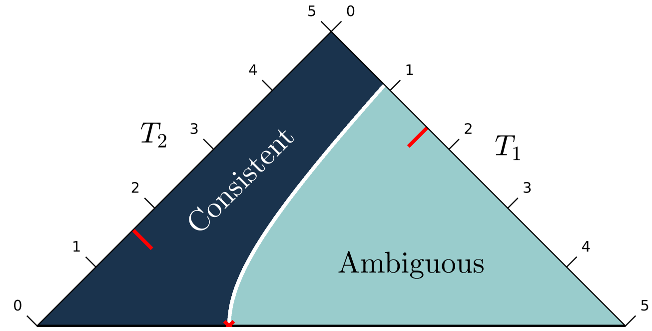

FIG. 4. Ambiguity diagram for Ising spin chain: Each point corresponds to a pair of Ising spin chains at temperatures T_1 and T_2 with $J = 1$ and $b = 0.3$. Consistency is found near the ($T = 0$) axes, while ambiguity dominates the remainder of parameter space. Curved boundary between these two regions ends at a temperature corresponding to $\max(C_q)$: $T_{C_q} \simeq 1.63$ (marked as a red dash).

tational advantage.

Ambiguity of Simplicity Absolute sizes aside, what can we say about the associated process *rankings*? How does the notion of "simpler" survive the transition from classical to quantum description?

Observe (Fig. 2) that, unlike the classical measure C_μ, the quantum simplicity C_q is not monotonic in the family of processes reached via increasing temperature: $C_q^\alpha < C_q^\delta < C_q^\gamma$. Moreover, the maximum C_q occurs at temperature $T_{C_q} \simeq 1.63$ while the excess entropy is maximized at temperature $T_{\mathbf{E}} \simeq 1.53$. Though a straightforward observation at this point, this basic feature provides the kernel for drawing out several counterintuitive consequences.

First, what is the consequence of nonmonotonicity itself? Take the processes α and γ in Fig. 2. Classically and quantally, α is simpler than γ. In contrast, for processes γ and δ we find that γ is simpler than δ classically, while δ is simpler than γ quantally.

In this way, even the familiar 1D Ising spin chain illustrates what is a general phenomenon—the ambiguity of simplicity. How general? Consider two generic processes A and B, for which no change in ranking occurs under the quantum lens. This indicates a *consistency* between the two representational viewpoints, at least with respect to processes A and B: $C_\mu^A > C_\mu^B \Leftrightarrow C_q^A > C_q^B$. Figure 3(left) illustrates this circumstance. Suppose, though, that viewed through our classical lens B appears simpler than A but, as for the spin chain at high temperature, our quantum lens reverses the ranking of A and B. We refer to this phenomenon as *ambiguity*. See Fig. 3(right). One concludes that the basic question—"Which process is simpler?"—no longer has a well defined answer.

How generic are consistency and ambiguity in the Ising spin chain parameter space? In Fig. 4 we construct an ambiguity diagram that compares all pairs of processes at temperatures T_1 and T_2 in the range $[0, 5]$. There, we fix the magnetic field $b = 0.3$ and coupling constant $J = 1$. We find that the only consistent pairs are those within a shrinking envelop around the axes ($T_1 = 0$ and $T_2 = 0$). The bulk of parameter space, then, contains ambiguously ranked pairs. The singular feature of the diagram is the leftmost point along the boundary between the two regimes. This occurs at the temperature $T_{C_q} \simeq 1.63$ where we find the maximum value of C_q. Monotonicity of C_μ ensures that a transition from the consistent region to the ambiguous one is controlled by the reordering of C_q values and not by C_μ values.

Robustness of ambiguity One can object that this ambiguity is merely an artifact of the particular quantum model-size measure C_q or of the assumptions in constructing the quantum states from a process' ϵ-machine. This is a valid concern, especially since minimality of the above quantum-machine representation (or any other quantum representation, for that matter) has not been established. Critically, as we now prove, the essence of ambiguity does not depend on this contingency.

Denote by $\widetilde{C_q}$ the memory measure of an optimal quantum model $\widetilde{Q}$ built according to some hypothetical, alternative quantum representational scheme. Since $\widetilde{C_q}$, like C_q, is also bounded between $\mathbf{E}$ and C_μ ([8, 17]), we can define sufficient criteria for consistency and ambiguity between $\widetilde{C_q}$ and C_μ. (For the following and without loss of generality, we also assume that the hypothetical model $\widetilde{Q}$ is at least as efficient as our original quantum-machine: $\widetilde{C_q} \leq C_q$.)

Assume that for processes A and B, B is classically simpler: $C_\mu^B < C_\mu^A$. Then, since $\mathbf{E} \leq \widetilde{C_q} \leq C_q$, the stronger criterion $\mathbf{E}^A > C_q^B$ ensures that any $\widetilde{Q}$ must yield consistency in classical and quantum ordering and is therefore, what we call, *certainly consistent*. See Fig. 5(left). Similarly, if $\mathbf{E}^B > C_q^A$, we know that any $\widetilde{Q}$ must yield an ambiguous ordering and is *certainly ambiguous*. See

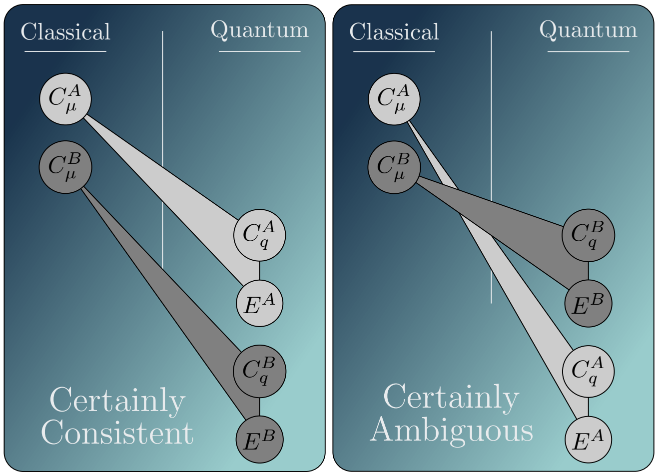

FIG. 5. Constraining hypothetical, as-yet-unknown frameworks for building quantum models $\widetilde{Q}$: Appealing to size measures C_q and $\mathbf{E}$ and without knowing any further details about $\widetilde{Q}$, we can still identify processes for which classical and quantum simplicity orderings must *certainly* be consistent or ambiguous. Cases exist that fall into neither of these stricter categories.

Fig. 5(right).

Figure 6 illustrates these stricter relations within the same Ising parameter region used in Fig. 4. The central region does not satisfy either strict constraint. As expected, the certainly consistent (ambiguous) area is a proper subset of the consistent (ambiguous) area.

One concludes that no matter what future improvements may be found in quantum representations, these "certain" subregions are robust and will have known consistency or ambiguity. This is a strong statement about how one can or cannot systematically rank the simplicity of systems classically and quantally. Again, the basic Ising spin chain is sufficiently rich to illustrate these these new phenomena.

Discussion How common is ambiguity? First, what can we say about ambiguity in the analogous (nearest-neighbor, ferromagnetic) two-dimensional Ising system? At the extreme $T = 0$ and for any nonzero value of external field, the ground state will be in uniform alignment with the field. This means that any random variable constructed from spin variables must have vanishing entropy. Lacking a complete computational mechanics of structure in two-dimensional patterns [30], it is still clear that any analog of statistical complexity (and thereby C_q) will vanish at $T = 0$ for such uniform configurations.

At very high T, though, spins become increasingly uncorrelated and the probability distribution over configurations approaches uniformity, *but is not exactly uniform.* That is, for any sufficiently high finite temperature, the system is has some, perhaps weak, correlation and so is not memoryless. Causal states in this regime remain

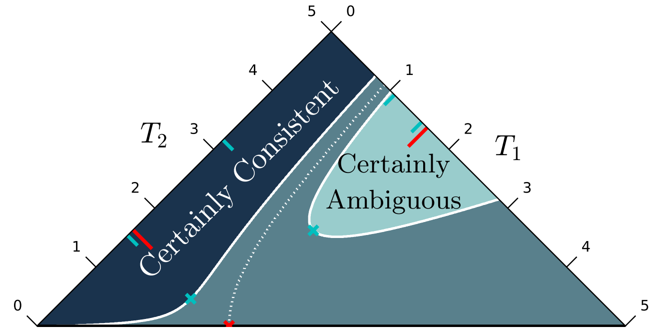

FIG. 6. Certain ambiguity diagram: Each point corresponds to a pair of Ising spin chains at temperatures T_1 and T_2 with $J = 1$ and $b = 0.3$. Dashed line marks previous certainty/ambiguity border of Fig. 4. Certain consistency (ambiguity) is a proper subset of consistent (ambiguous). Local extrema of new boundaries at temperatures corresponding to $(C_q = \max(\mathbf{E}), \max(\mathbf{E}))$ and $(\max(\mathbf{E}), C_q = \max(\mathbf{E}))$ (marked with short blue lines). Long red lines mark the same values as in Fig. 4.

probabilistically distinct. So, as with the 1D case, at very high temperature ($T \gg 1$, but $T \neq \infty$) $C_\mu(T)$ is not zero. What can we say about C_q in this limit? For high $T \gg 1$ spin randomness makes the quantum states $\{|\eta\rangle\}$ (Eq. 2) more and more indistinguishable. And so, their increasing overlaps $\langle \eta_i | \eta_j \rangle \to 1$, driving C_q to zero monotonically. The conclusion is that for the 2D Ising, at $T \approx 0$ and $T \gg 1$, we have the same qualitative picture for the simplicity measures as in Fig. 2. This brief argument says that ambiguity exists in the 2D Ising spin model as well.

Perhaps the ambiguity of simplicity is special to spin systems. The appendix shows that it is, in fact, a much more general phenomenon, by introducing a set of easily satisfied conditions such that two simplicity functions over a set of structured objects must yield ambiguous ordering. In particular, taking the space of all ϵ-machines as a set and C_μ and $\widetilde{C_q}$ as the two measures , we find that these conditions are satisfied. The general consequence is that either the two measures selected are trivially equal or ambiguity must exist. In other words, if the world is not ambiguous, quantum mechanics cannot simplify its explanation. One concludes that ambiguity is necessary for quantum simplification.

Closing Remarks We now see that comparing classical physics and quantum mechanics descriptions of the world calls into question our basic belief in the simplicity of physical theories. However, monitoring model simplicity (and therefore model ordering) is far from being the sole domain of physics. It is key in a variety of contemporary statistical inference tasks, specifically in model selection [31].

Imagine two competing models A and B of some finite data $\mathcal{D}$. In Bayesian inference, one widely employed methodology, choosing one over model another requires us to calculate the posterior probabilities that each generated $\mathcal{D}$. This requires specifying a prior probability distribution over A and B at the outset [32]. Such priors are commonly constructed to favor simpler models. Indeed, there is a long history of methods to avoid overfitting to data that directly incorporate simplicity measures into model selection, including Akaike's Information Criterion [33], Boltzmann Information Criterion [34], Minimum Description Length [35], and Minimum Message Length [36].

Classically, we may find that A is simpler than B. This fact then enters our inference through the model prior, favoring A. Given that the two likelihoods $\Pr(\mathcal{D}|A)$ and $\Pr(\mathcal{D}|B)$ are the same or similar enough, the inference identifies A as preferred. As we showed, the tables may turn dramatically when presented with quantum data; we might find there that B is much simpler. We must then reconcile the fact that had we constructed the model prior using our quantum lens, B would have yielded as the preferred model.

We introduced the ambiguity of simplicity focusing on classical and quantum descriptions of classical processes. Quantum supremacy [2] suggests we go further to probe how (and if) ambiguity manifests when modeling quantum processes. This can be probed in the 1D quantum Heisenberg spin chain [37], for example. Measuring each spin within the bi-infinite chain in the z-direction yields a stochastic process—one that can be described classically or quantally. The Heisenberg spin chain is realized experimentally in the quasi-1D magnetic order found in antiferromagnetic $KCuF_3$ crystals [38–40]. One can then adapt the methods of 1D chaotic crystallography [41] to extract the ϵ-machine and quantum-machine descriptions of the quantum crystalline structure from the neutron scattering measurements. These and perhaps other experiments will provide an entreé to analyzing the ambiguity of simplicity in quantum systems.

ACKNOWLEDGMENTS

We thank A. Aghamohammadi, M. Anvari, R. D'Souza, R. James, F. Pourbabaee, and D. Varn for useful conversations. JPC thanks the Santa Fe Institute for its hospitality during visits as an External Faculty member. This material is based upon work supported by, or in part by, the John Templeton Foundation and the U. S. Army Research Laboratory and the U. S. Army Research Office under contracts W911NF-13-1-0390 and W911NF-13-1-0340.

* caghamohammadi@ucdavis.edu

† jrmahoney@ucdavis.edu

‡ chaos@ucdavis.edu

[1] Sir Isaac Newton. *The Mathematical Principles of Natural Philosophy.* H. D. Symonds, 1729. A. Motte, trans. 1

[2] J. Preskill. Quantum computing and the entanglement frontier. *arXiv:1203.5813*, 2012. 1, 6

[3] I. Boettcher, L. Bayha, D. Kedar, P. A. Murthy, M. Neidig, M. G. Ries, A. N. Wenz, G. Zurn, S. Jochim, and T. Enss. Equation of state of ultracold Fermions in the 2D BEC-BCS crossover region. *Phys. Rev. Lett*, 116:045303, 2016. 1

[4] K. Fenech, P. Dyke, T. Peppler, M. G. Lingham, S. Hoinka, H. Hu, , and C. J. Vale. Thermodynamics of an attractive 2d Fermi gas. *Phys. Rev. Lett.*, 116:045302, 2016. 1

[5] K. Burke, K. A. Mitchell andB. Wyker, S. Ye, and F. B. Dunning. Demonstration of turnstiles as a chaotic ionization mechanism in Rydberg atoms. *Phys. Rev. Lett.*, 107:113002, 2011. 1

[6] J. P. Crutchfield and K. Young. Inferring statistical complexity. *Phys. Rev. Let.*, 63:105–108, 1989. 1

[7] D. P. Feldman, C. S. McTague, and J. P. Crutchfield. The organization of intrinsic computation: Complexity-entropy diagrams and the diversity of natural information processing. *CHAOS*, 18(4):043106, 2008. 1

[8] M. Gu, K. Wiesner, E. Rieper, and V. Vedral. Quantum mechanics can reduce the complexity of classical models. *Nature Comm.*, 3(762):1–5, 2012. 1, 2, 3, 4

[9] J. R. Mahoney, C. Aghamohammadi, and J. P. Crutchfield. Occam's quantum strop: Synchronizing and compressing classical cryptic processes via a quantum channel. *Scientific Reports*, 6:20495, 2016. 1, 2

[10] M. S. Palsson, M. Gu, J. Ho, H. M. Wiseman, and G. J. Pryde. Experimental quantum processing enhancement in modelling stochastic processes. *arXiv preprint arXiv:1602.05683*, 2016. 1

[11] P. M. Riechers, J. R. Mahoney, C. Aghamohammadi, and J. P. Crutchfield. A closed-form shave from Occam's quantum razor: Exact results for quantum compression. *arXiv:1510.08186*, 2015. 1, 2

[12] T. M. Cover and J. A. Thomas. *Elements of Information Theory.* Wiley-Interscience, New York, second edition, 2006. 2

[13] A. N. Kolmogorov. A new metric invariant of transient dynamical systems and automorphisms in Lebesgue spaces. *Dokl. Akad. Nauk. SSSR*, 119:861, 1958. (Russian) Math. Rev. vol. 21, no. 2035a. 2

[14] J. P. Crutchfield and D. P. Feldman. Regularities unseen, randomness observed: Levels of entropy convergence. *CHAOS*, 13(1):25–54, 2003. 2

[15] J. P. Crutchfield. Between order and chaos. *Nature Physics*, 8(January):17–24, 2012. 2

[16] R. G. James, J. R. Mahoney, C. J. Ellison, and J. P. Crutchfield. Many roads to synchrony: Natural time scales and their algorithms. *Phys. Rev. E*, 89:042135, Apr 2014. 2

[17] A. S. Holevo. Bounds for the quantity of information transmitted by a quantum communication channel. *Problems Inform. Trans.*, 9:177–183, 1973. 3, 4

[18] S.. G. Brush. History of the Lenz-Ising model. *Rev. Mod. Phys.*, 39(4):883, 1967. 3

[19] R. Peierls. On Ising's model of ferromagnetism. *Math. Proc. the Cambridge Phil. Soc.*, 32(03):477–481, 1936. 3

[20] S. F. Edwards and P. W. Anderson. Theory of spin glasses. *J. Physics F: Metal Physics*, 5(5):965, 1975. 3
[21] T.-D. Lee and C.-N. Yang. Statistical theory of equations of state and phase transitions. II. Lattice gas and Ising model. *Phys. Rev.*, 87(3):410, 1952. 3
[22] A. Hastings, J. Machta, and A. E Noble. *Emergent long-range synchronization of oscillating ecological populations without external forcing described by Ising universality*, volume 6. Nature Comm., 2015. 3
[23] D. Sornette. Physics and financial economics (1776–2014): Puzzles, Ising, and agent-based models. *Rep. Prog. Physics*, 77(6):062001, 2014. 3
[24] E. Schneidman, M. J. Berry, R. Segev, and W. Bialek. Weak pairwise correlations imply strongly correlated network states in a neural population. *Nature*, 440(7087):1007–1012, 2006. 3
[25] D. P. Feldman and J. P. Crutchfield. Discovering noncritical organization: Statistical mechanical, information theoretic, and computational views of patterns in simple one-dimensional spin systems. 1998. Santa Fe Institute Working Paper 98-04-026. 3
[26] D. P. Feldman. *Computational Mechanics of Classical Spin Systems*. PhD thesis, University of California, Davis, 1998. Published by University Microfilms Intl, Ann Arbor, Michigan. 3, 7
[27] W. Y. Suen, J. Thompson, A. J. Garner, V. Vedral, and M. Gu. *The classical-quantum divergence of complexity in the Ising spin chain*. arXiv preprint arXiv:1511.05738, 2015. 3
[28] M. A. Nielsen and I. L. Chuang. *Quantum computation and quantum information*. Cambridge University Press, 2010. 3
[29] J. Preskill. Lecture notes for Physics 229: Quantum information and computation. *California Institute of Technology*, page 16, 1998. 3
[30] Though see Ref. [26, 42]. 5
[31] K. P. Burnham and D. R. Anderson. *Model selection and multimodel inference: A practical information-theoretic approach*. Springer Science & Business Media, 2003. 5
[32] C. C Strelioff and J. P Crutchfield. Bayesian structural inference for hidden processes. *Phys. Rev. E*, 89(4):042119, 2014. 6
[33] H. Akaike. A new look at the statistical model identification. *IEEE Trans. Auto. Control*, 19(6):716–723, 1974. 6
[34] G. Schwarz. Estimating the dimension of a model. *Ann. Stat.*, 6(2):461–464, 1978. 6
[35] J. Rissanen. *Stochastic complexity in statistical inquiry*, volume 15. World scientific, 1998. 6
[36] C. S. Wallace and F. P. Freeman. Estimation and inference by compact coding. *J. R. Statist. Soc. B*, 49:240, 1987. 6
[37] R. J. Baxter. *Exactly solved models in statistical mechanics*. Academic Press, 1989. 6
[38] D. A. Tennant, R. A. Cowley, S. E. Nagler, and A. M. Tsvelik. Measurement of the spin-excitation continuum in one-dimensional $KCuF_3$ using neutron scattering. *Phys. Rev. B*, 52:13368–13380, 1995. 6
[39] D. A. Tennant, S. E. Nagler, D. Welz, G. Shirane, and K. Yamada. Effects of coupling between chains on the magnetic excitation spectrum of $KCuF_3$. *Phys. Rev. B*, 52:13381–13389, 1995.
[40] J. M. Maillet. Heisenberg spin chains: from quantum groups to neutron scattering experiments. *Seminaire Poincare*, X:139–177, 2007. 6
[41] D. P. Varn and J. P. Crutchfield. Chaotic crystallography: How the physics of information reveals structural order in materials. *Curr. Opin. Chem. Eng.*, 7:47–56, 2015. 6
[42] D. P. Feldman and J. P. Crutchfield. Structural information in two-dimensional patterns: Entropy convergence and excess entropy. *Phys. Rev. E*, 67(5):051103, 2003. 7

Appendix

First, we lay bare the mathematical argument and then we interpret it in terms of the physical setting of the main text.

Consider a set of objects S and two functions over the set $F_1 : S \to G$ and $F_2 : S \to G$.

If there exists $s_1, s_2 \in S$, such that $F_1(s_1) > F_1(s_2)$ and $F_2(s_1) < F_2(s_2)$, then we say these functions are *ambiguous* over S.

We define three conditions for the set and functions.

Condition A The two functions map onto the whole space G: $F_1(S) = G$ and $F_2(S) = G$.

Condition B For all $g \in G$ there exists $x \in S$ such that $F_1(x) = F_2(x) = g$.

Condition C Assume $\preceq$ is a dense, total order on space G.

Theorem 1. *Given two functions F_1 and F_2 that map set S to space G and satisfy Conditions A, B, and C: No ambiguity implies that for all $x \in S$, $F_1(x) = F_2(x)$.*

Proof. *We prove the contrapositive by contradiction. Assume there exists $x \in S$ such that $F_1(x) \neq F_2(x)$. Without loss of generality, let $F_1(x) \succeq F_2(x)$. Since $\preceq$ is a dense total order on G, there is $g \in G$ such that $F_1(x) \succeq g \succeq F_2(x)$. By Condition B, there exists $y \in S$ such that $F_1(y) = F_2(y) = g$. Trivially then, $F_1(x) \succeq F_1(y)$ and $F_2(x) \preceq F_2(y)$. This demonstrates ambiguity and completes the proof.*

We can interpret this in the setting of stationary processes with measures C_μ and $\widetilde{C_q}$ and discuss the space of all possible quantum sizes. More specifically, consider the case $F_1 = C_\mu$ and $F_2 = \widetilde{C_q}$. We know that for any value $y \in \mathbb{R}$, there exists an ϵ-machine with $C_\mu = \widetilde{C_q} = \mathbf{E} = y$. This satisfies the assumption. Then, our results say that if the world is not ambiguous, the two measures are equivalent. In other words, the quantum advantage $\widetilde{C_q}$ *requires* ambiguity.