Physics Letters A ••• (••••) •••–•••

Contents lists available at ScienceDirect

Physics Letters A

www.elsevier.com/locate/pla

The ambiguity of simplicity in quantum and classical simulation

Cina Aghamohammadi, John R. Mahoney, James P. Crutchfield

Complexity Sciences Center and Physics Department, University of California at Davis, One Shields Avenue, Davis, CA 95616, United States

ARTICLE INFO

Article history:
Received 27 September 2016
Received in revised form 19 December 2016
Accepted 19 December 2016
Available online xxxx
Communicated by C.R. Doering

Keywords:
Quantum information
Information theory
Stochastic process
Hidden Markov model
ϵ-Machine

ABSTRACT

A system's perceived 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. Here, we examine the minimum required memory for simulation. We see that the notions of absolute physical simplicity 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. Recent rapid progress in quantum computation and quantum simulation suggest that the ambiguity of simplicity will strongly impact statistical inference and, in particular, model selection.

> 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]]

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 but they are also *simple*. Consequently, one should abandon the Ptolemaic epicycles for Newton's 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 two descriptions of stochastic processes — one classical and one quantum. Classical versus quantum comparisons have, of late, captured our attention 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]. A single cold 2D Fermi gas supports coexistence of both quantum mechanical states at its core and classical states on its periphery [3,4]. The overriding impression is that now is an interesting time for the foundations of quantum mechanics. The following adds a new phenomenon to these debates on the balance of classical and quantum theories, 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' probabilities. We can straightforwardly say that one process is more random than another via comparing their temperatures or 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 — a measure of internal memory [5] or the minimum required memory to simulate a process. It provides a concrete and interpretable answer to the question, which process is structurally simpler? By applying this comparison, we may order all processes from the simplest to the most complicated [6].

With recent progress in quantum computation [7–9], an interesting twist comes about if we add quantum mechanics to our modeling toolbox. Descriptions that act on a quantum substrate offer new and surprising options. For example, it was shown that a quantum mechanical description can lead to a simpler representation [10–14] and even in some cases infinitely simpler [15,16]. Recently, this quantum advantage was verified experimentally [17]. Proceeding with these methods, we discover what is most surprising: the *relative simplicity* of classical and quantum descriptions can change. Specifically, there are stochastic processes, A and B,

E-mail addresses: caghamohammadi@ucdavis.edu (C. Aghamohammadi), jrmahoney@ucdavis.edu (J.R. Mahoney), chaos@ucdavis.edu (J.P. Crutchfield).

http://dx.doi.org/10.1016/j.physleta.2016.12.036

for which 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. This means quantizing a simple classical model may not be as simple as quantizing a more complicated classical model. As a consequence model selection is complicated by the addition of a quantum model class.

2. 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 takes some value x_t in a discrete alphabet set $\mathcal{A}$ and where all probabilities $\Pr(X_t, \ldots, X_{t+L})$ are time-invariant.

How is their degree of randomness quantified? Information theory [18] measures the uncertainty in a single observation X_0 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* [19]: $h_\mu = \lim_{L\to\infty} H[X_{0:L}]/L$. If we interpret the left half $X_{-\infty:0} = \ldots X_{-2}X_{-1}$ as the "past" and the right half $X_{0:\infty} = X_0X_1X_2\ldots$ 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 past correlations, the naive uncertainty $H[X_0]$ reduces to h_μ.

How reducible is our uncertainty in the future $X_{0:\infty}$ knowing the past $X_{-\infty:0}$? The answer is given by the mutual information between the past and the future — the *excess entropy* [20]: $\mathbf{E} = \mathrm{I}[X_{-\infty:0} : X_{0:\infty}]$. With h_μ and $\mathbf{E}$, we measure randomness and predictability, respectively.

Let's say we want to simulate a given process. To do this we write a computer code that follows an algorithm and allocate the memory the algorithm needs. For a given process *computational mechanics* [21] identifies the optimal algorithm — the process' *ϵ-machine*. This is a *unifilar hidden Markov model* [22] that uses only the minimum required memory for simulation. We view a process' ϵ-machine as the "theory" of a process in that it specifies a mechanism that exactly simulates a process' behaviors. In this way, computational mechanics supplements $\mathbf{E}$ and h_μ with a measure of structure — the minimum required amount memory to simulate the given process.

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 $\Pr(X_{t:\infty}|\cdot)$: $x_{-\infty:t} \sim x_{-\infty:t'} \iff \Pr(X_{t:\infty}|x_{-\infty:t}) = \Pr(X_{t':\infty}|x_{-\infty:t'})$. From this, one concludes that a process' ϵ-machine is, in a well defined sense, its simplest predictive theory.

Translating this notion of simplicity into a measurable quantity, we ask: *What is the minimum memory necessary to* implement *optimal prediction?* The answer is the historical information stored in the ϵ-machine. Quantitatively, this is the Shannon entropy of the causal-state stationary distribution $\{\pi_\sigma\}$, the *statistical complexity*:

$$C_\mu = \mathrm{H}[\boldsymbol{\mathcal{S}}] = -\sum_{\sigma\in\boldsymbol{\mathcal{S}}} \pi_\sigma \log_2 \pi_\sigma \ , \tag{1}$$

It is well known that the excess entropy is a lower-bound on this structural measure: $\mathbf{E} \le C_\mu$. In fact, this relation is only rarely an equality [23]. And so, while $\mathbf{E}$ quantifies the amount to which a process is subject to explanation by its ϵ-machine "theory", this simplest theory is typically larger, informationally speaking (C_μ), 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 [10–12]. The quantum-machine con-

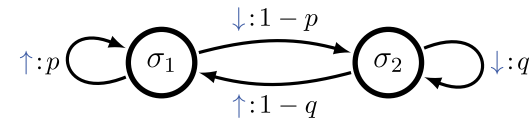

Fig. 1. The ϵ-machine for the nearest-neighbor Ising spin chain has two causal states σ_1 and σ_2. If the last observed spin x_0 is up ($s_0 = +1$) the current state is σ_1 and if it's down ($s_0 = -1$) is σ_2. If the current state is σ_1, with probability p the next spin observed is up and, if the current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

sists 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. Fixing L, we construct quantum states:

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

where w^L denotes a length-L word and $\Pr(w^L, \sigma_k|\sigma_j) = \Pr(X_{0:L} = w^L, \mathcal{S}_L = \sigma_k|\mathcal{S}_0 = \sigma_j)$. 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 $|\mathcal{A}|^L$, with basis elements $|w^L\rangle = |x_0\rangle\cdots|x_{L-1}\rangle$.

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

$$C_q = -\mathrm{Tr}(\rho\log\rho) \ , \tag{3}$$

where $\rho = \sum_i \pi_i|\eta_i\rangle\langle\eta_i|$. This quantum analog of memory is generically less than the classical: $C_q \le C_\mu$. Also, due to the Holevo bound [10,24], $\mathbf{E} \le C_q$. Though rare in process space, the classical and quantum informational sizes are equal exactly when both models are "maximally simple": $\mathbf{E} = C_q = C_\mu$.

3. Ising chain simplicity

The Ising spin-chain Hamiltonian is given by:

$$H = -\sum_{<i,j>} (Js_is_j + bs_i) \ , \tag{4}$$

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

In equilibrium the bi-infinite chain of spin random variables defines a stationary stochastic process which 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. The conclusion (see Supp. Materials) 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. Fig. 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 these spin configurations. Each causal state is mapped to a pure quantum state that resides in a spin one-half space [13]:

$$|\sigma_1\rangle = \sqrt{p}|\uparrow\rangle + \sqrt{1-p}|\downarrow\rangle$$
$$|\sigma_2\rangle = \sqrt{1-q}|\uparrow\rangle + \sqrt{q}|\downarrow\rangle \ . \tag{5}$$

(We use a more compact spin up/down notation, rather than the quantum machine notation of Eq. (2).) Intuitively, the quantum overlap accounts for the fact that the conditional predictions

Please cite this article in press as: C. Aghamohammadi et al., The ambiguity of simplicity in quantum and classical simulation, Phys. Lett. A (2017), http://dx.doi.org/10.1016/j.physleta.2016.12.036