<table>
<thead>
<tr>
<th><strong>Title</strong></th>
<th>Practical unitary simulator for non-markovian complex processes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author(s)</strong></td>
<td>Binder, Felix Christoph; Thompson, Jayne; Gu, Mile</td>
</tr>
<tr>
<td><strong>Citation</strong></td>
<td>Binder, F. C., Thompson, J., &amp; Gu, M. (2018). Practical Unitary Simulator for Non-Markovian Complex Processes. Physical Review Letters, 120(24), 240502-.</td>
</tr>
<tr>
<td><strong>Date</strong></td>
<td>2018</td>
</tr>
<tr>
<td><strong>URL</strong></td>
<td><a href="http://hdl.handle.net/10220/45507">http://hdl.handle.net/10220/45507</a></td>
</tr>
<tr>
<td><strong>Rights</strong></td>
<td>© 2018 American Physical Society. This paper was published in Physical Review Letters and is made available as an electronic reprint (preprint) with permission of American Physical Society. The published version is available at: [<a href="http://dx.doi.org/10.1103/PhysRevLett.120.240502">http://dx.doi.org/10.1103/PhysRevLett.120.240502</a>]. One print or electronic copy may be made for personal use only. Systematic or multiple reproduction, distribution to multiple locations via electronic or other means, duplication of any material in this paper for a fee or for commercial purposes, or modification of the content of the paper is prohibited and is subject to penalties under law.</td>
</tr>
</tbody>
</table>
Practical Unitary Simulator for Non-Markovian Complex Processes

Felix C. Binder,1,2,* Jayne Thompson,3 and Mile Gu1,3,2,†

1School of Physical and Mathematical Sciences, Nanyang Technological University, 637371 Singapore, Singapore
2Complexity Institute, Nanyang Technological University, 639673 Singapore, Singapore
3Centre for Quantum Technologies, National University of Singapore, 3 Science Drive 2, 117543 Singapore, Singapore

(Received 20 August 2017; revised manuscript received 9 February 2018; published 11 June 2018)

Stochastic processes are as ubiquitous throughout the quantitative sciences as they are notorious for being difficult to simulate and predict. In this Letter, we propose a unitary quantum simulator for discrete-time stochastic processes which requires less internal memory than any classical analogue throughout the simulation. The simulator’s internal memory requirements equal those of the best previous quantum models. However, in contrast to previous models, it only requires a (small) finite-dimensional Hilbert space. Moreover, since the simulator operates unitarily throughout, it avoids any unnecessary information loss. We provide a stepwise construction for simulators for a large class of stochastic processes hence directly opening the possibility for experimental implementations with current platforms for quantum computation. The results are illustrated for an example process.

DOI: 10.1103/PhysRevLett.120.240502

Stochastic processes permeate almost all quantitative sciences, describing the behavior of various dynamical systems when witnessed at discrete points in time. Predictive modeling and simulation of such processes is thus of great scientific relevance, allowing inference of their future behavior based on past observations. For complex processes, the amount of information such simulations track can become immense. There is thus significant interest in building simple models which can generate correct future statistics, while requiring only as much information about the past as necessary. Conceptually, such models align with the principle of Occam’s razor, contributing to improved understanding of a process by better isolating information relevant for future prediction. Practically, all information a simulator tracks about the past must be stored, thus bounding the corresponding memory requirement for stochastic simulation. Indeed, complexity theorists have studied this problem intensely in the field of computational mechanics [1–4]. This has resulted in a systematic method for constructing ε machines for general stochastic processes—devices that replicate the future statistics of a given process while storing provably less information about past observations compared to all other alternatives. The associated amount of past information they require—as quantified by information entropy—is known as the “statistical complexity,” a well established quantifier of structure that has been applied to diverse contexts, including neural spike trains [5], geomagnetism [6], stock indices [7,8], crystallography [9–11], and spin chains [12–16].

Quantum technologies have great potential to reshape these studies. Recent advances indicate that for almost all stochastic processes, one can construct a quantum model that demands less past information than the simplest nonquantum counterpart, while generating statistically identical future predictions [17]. This has since led to surging activity at the interface between quantum physics and computational mechanics. Notable developments include the discovery that quantum and classical statistical complexity can exhibit very different scaling and qualitative behavior [13,14,16], a proposal of more efficient quantum models whose memory saving scale with long-range correlations present in a stochastic process [18,19], and a recent application to sampling of rare events [20]. Each of these reflects growing evidence that what we consider complex can fundamentally change in the quantum regime.

To fully establish such observations, however, one must also develop an understanding of the mechanisms in which quantum models for stochastic processes generate future predictions. For instance, in addition to a quantum model’s statistical complexity, it is interesting to also consider the number and dimension of physical systems dedicated to its working memory as well as the type of required interactions between these systems. Meanwhile, to achieve optimal results, the currently simplest known quantum models require measurement in a Hilbert space whose dimension scales with the cryptic order of a stochastic process—a property closely related to the presence of long-range correlations [18]. This space would become impractically large with increasing cryptic order.

This Letter aims to address these issues. We propose a quantum simulator for stochastic processes that includes an efficient quantum circuit for the generation of future predictions. It satisfies three particular desiderata: (1) The amount $C_q$ of past information required by the simulator equals that of the most efficient quantum models known so far. (2) The circuit is unitary, such that the internal entropy of
the simulator remains \( C_q \) throughout. (3) Future predictions are generated using only bipartite interactions between two systems of a bounded dimension (assuming bounded \( C_q \) and finite output alphabet).

Taken together, the presented results close a loophole whereby reduced entropic memory may not necessarily lead to practical savings of memory dimension in stochastic simulation. By showing that the simulator’s memory cost never exceeds \( C_q \) they complement the work by Riechers et al. who developed a method for computing \( C_q \) [19]. Lastly, they directly open the door for practical simulation of more complex stochastic processes in quantum laboratories.

**Computational mechanics.**—A stochastic process is a bi-infinite sequence of random variables \( X_t \), labeled by discrete time steps \( t \) [21] (Fig. 1(a)). It is governed by a probability distribution \( P(X, \bar{X}) \), where \( \bar{X} = \ldots, X_{-2}, X_{-1}, X_0 \) corresponds to the past of a process and \( \bar{X} = X_1, X_2, \ldots \) its future. Each \( X_j \) takes values \( x \) from an alphabet \( \mathcal{A} \). Strings of output symbols \( x_m: n := x_{m+1} x_{m+2} \ldots x_n \) form words. Each instance of a process has a specific past \( \bar{x} = x_{-\infty: 0} \), with each specific future \( \bar{x} = x_{0: \infty} \) occurring according to the conditional probability \( P(\bar{x} = \bar{x} | \bar{x}) \). Processes are assumed to be stationary, such that they are invariant with respect to time translations. We hence omit the corresponding index \( j \) [22].

A simulator replicates the conditional behavior of the corresponding process [Fig. 1(b)]. It is represented by a hidden Markov model (HMM) [23]: For each past \( \bar{x} \), the simulator’s memory can be configured in some internal state \( s_\bar{x} \), such that its future outputs are statistically indistinguishable from the stochastic process that is being modeled. Once configured in \( s_\bar{x} \), the simulator sequentially emits symbols governed by the expected conditional probability distribution \( P(\bar{x} | \bar{x}) \). To do so, it needs to update the state of its memory at each time step. For instance, on emitting a symbol \( r \), a simulator in state \( s_\bar{x} \) would transition to state \( s_\bar{x}' \) corresponding to the updated past \( \bar{x}' = \bar{x} r \).

Computational mechanics singles out a specific type of HMM called an \( e \)-machine by application of Occam’s razor; i.e., the internal states should contain as much information about the past as is required for accurate simulation of the future but nothing more [1–4]. This is formalized by a mapping \( e: \mathcal{A}^\infty \rightarrow \Sigma \) which partitions the set of pasts into a set of internal states \( \Sigma = \{s_1, s_2, \ldots\} \) according to the following equivalence relation:

\[
\bar{x} \sim_{e} \bar{x}', \text{ iff } P(\bar{x} | \bar{x}) = P(\bar{x}' | \bar{x}').
\] (1)

Accordingly, all pasts which imply identical future statistics correspond to the same internal state (called “causal state” in this context).

For illustration, consider the alternating process—a string of alternating symbols 1 and 0. Its future may be replicated by storing only the last symbol. This process hence has two causal states, in correspondence with the last observed value.

Equation (1) entails a property called “unifilarity.” It implies that the current causal state, together with the next output symbol, determines the subsequent causal state with certainty, i.e., by some function \( \lambda(i, x) \), which takes the value of the next causal state’s label. Hence, the conditional transition probabilities \( P(x, j | i) \) — corresponding to a transition from state \( s_i \) to state \( s_j \) upon emission of symbol \( x \) — can be expressed in terms of \( P(x | i) \), the probability of emitting symbol \( x \) from causal state \( i \),

\[
P(x, j | i) = P(x | i) \delta_{j, \lambda(i, x)},
\] (2)

where \( \delta_{i,k} \) is the Kronecker \( \delta \).

In practice, an \( e \)-machine’s defining quantities \( \mathcal{A}, \Sigma, P(x | i) \), and \( \lambda(i, x) \) may be obtained from a given data string by a suitable reconstruction algorithm [1,24,25].

It follows from stationarity that \( P(x | i) \) and \( \lambda(i, x) \) do not change over time. Moreover, the stationary probability \( \pi_j \) of being in causal state \( s_j \) is time independent.
This allows the definition of the statistical complexity
\[ C_\mu := - \sum_i \pi_i \log \pi_i. \]

It represents the average amount of information an \( \epsilon \) machine needs to store about the past [1,2,26]. Alternatively, the topological complexity
\[ C^0_\mu := \log |\Sigma| \]
quantifies the memory of an \( \epsilon \) machine in terms of the number of causal states \( |\Sigma| \) [1].

The \( q \) simulators and their internal states.—We now describe \( q \) simulators—fully-quantum simulators for stationary stochastic processes that extend \( \epsilon \)-machines to the quantum domain. Illustrated in Fig. 1(c), they are a specific type of simulator comprising a single quantum system as internal memory and functioning by repeated application of the same unitary operation \( U \).

A \( q \) simulator for a given stochastic process features a set of internal states \( \{ |\sigma_j\rangle \} \), each associated with the process’ corresponding causal state. These states may be used to unitarily generate a future pattern of any length, one symbol at a time [Fig. 1(c)].

Starting in \( |\sigma_j\rangle \), the total state of the circuit after the first application of \( U \) is given by
\[ |1_i\rangle := U|\sigma_j\rangle|0\rangle = \sum_x \sqrt{P(x,i)} |\sigma_j(x,i)\rangle |x\rangle, \]
where the sum goes over all symbols \( x \). Using Eq. (2), we may equally write \( |1_i\rangle = \sum_i \sum_j \sqrt{P(x,i)} |\sigma_j\rangle |x\rangle \).

More generally, after \( L \) applications, the total state corresponding to the first \( L + 1 \) registers in Fig. 1(c) is
\[ |L_i\rangle := U^L|\sigma_j\rangle|0\rangle^{\otimes L} = \sum_{x_0:L} \sqrt{P(x_0:L,i)} |\sigma_j(x_0:L)\rangle |x_0:L\rangle. \]

Here, \( \lambda(i,x_{0:L}) \) extends the previous notation in a natural way from single symbols to words, and \( U \) only acts on two registers at a time [see Fig. 1(c)]. Identity operations on the remaining registers are omitted.

When the symbol registers are measured in the computational basis (i.e., given by \( \{ |x\rangle \} \), each word \( x_{0:L} \), corresponding to a single state \( |\sigma_j(x_{0:L})\rangle \) in the sum, is measured with the desired probability \( P(x_0:L,i) \). This extends to mixtures of different \( |L_i\rangle \).

It remains to be shown that the desired operator \( U \) actually exists. Considering states \( |L_i\rangle, |L_j\rangle, |M_i\rangle, \) and \( |M_j\rangle \), as defined by Eq. (6), this is the case iff \( \langle L_i|L_j\rangle = \langle M_i|M_j\rangle \) for all \( L, M, i, j, \) or—more simply—iff
\[ \langle \sigma_i|\sigma_j\rangle = \langle 1_i|1_j\rangle \quad \forall \ i, j. \]
procedure [For some processes, the internal states $|\sigma_j\rangle$ may be linearly dependent. This does not impede the generality of our construction: We may set $U$ to be the identity operation on that subspace which is not spanned by $\{|\sigma_j\rangle\}$ [34].]

$|\sigma_1\rangle = |e_1\rangle; \quad |\sigma_2\rangle = c_{12}|e_1\rangle + \sqrt{1-c_{12}^2}|e_2\rangle$, etc. \hspace{.5cm} (10)

(III) Solve Eq. (5) for those columns of $U$ which correspond to input states $|e_j\rangle|0\rangle$. This determines the matrix elements $U_{ij,j'0} \equiv \langle e_i|\langle j|U|e_{j'}\rangle|0\rangle$. (IV) Obtain the remaining columns of $U$ by a Gram-Schmidt procedure starting from the already determined columns.

This prescription is constructive and works for all discrete-valued stationary processes with a finite number of causal states (see Supplemental Material D of Ref. [28]).

Example: upset-gambler process. —We demonstrate our construction explicitly for the upset-gambler process. The process is distinguished as being particularly simple but yet useful for the uppset-gambler process. The nonzero transition probabilities are, in shorthand, $P(0, s_A|s_A) = p$, $P(1, s_A|s_A) = 1 - p$, $P(0, s_A|s_B) = q$, and $P(1, s_A|s_B) = 1 - q$. It is assumed that $0 < p < 1$, $0 < q < 1$, and $p < q$. Further details are included in Supplemental Material B of Ref. [28].

Replacing arabic with roman numerals for the states of the upset-gambler process, we have

$|I_A\rangle = \sqrt{1-p}|\sigma_A\rangle|1\rangle + \sqrt{p}|\sigma_B\rangle|0\rangle$ and \hspace{.5cm} (11)

$|I_B\rangle = \sqrt{q}|\sigma_A\rangle|0\rangle + \sqrt{1-q}|\sigma_B\rangle|1\rangle$. \hspace{.5cm} (12)

Equations (7) and (10) allow the internal states to be expressed in terms of any orthonormal qubit basis $\{|e_j\rangle\}$,

$|\sigma_A\rangle = |e_1\rangle; \quad |\sigma_B\rangle = c\sqrt{p}\bar{q}|e_1\rangle + c\xi|e_2\rangle$, \hspace{.5cm} (13)

with $\bar{p}\equiv 1-p$, $\bar{q}\equiv 1-q$, $c\equiv (1-\sqrt{pq})^{-1}$, $\xi\equiv \sqrt{q}\sqrt{p}$. This in turn allows for the specification of the first and third column of the desired unitary matrix $U$, written in the basis $\{|e_1\rangle|0\rangle, |e_1\rangle|1\rangle, |e_2\rangle|0\rangle, |e_2\rangle|1\rangle\}$,

$U = \begin{pmatrix}
  c\sqrt{p}\bar{q} & 1 - c\bar{p}\bar{q} & 
  \sqrt{\bar{p}} & -\sqrt{\bar{p}} \\
  c\xi\sqrt{\bar{p}} & -c\sqrt{p}\bar{q} & 
  0 & 0
\end{pmatrix}$. \hspace{.5cm} (14)

Here, $\#$ stands for undetermined matrix entries. The only condition for determining these entries is that the four columns of the matrix must form an orthonormal basis. For instance, a simple choice would be $(0,0,0,1)^T$ for the second column and $((\sqrt{p}\bar{q}, -\sqrt{p}\bar{q}, \sqrt{\bar{p}}, 0)^T$ for the fourth.

Using this two-qubit unitary operator, the upset-gambler process can be simulated with internal entropy $C_q$ throughout. The latter may be easily computed using Eqs. (8) and (13) and the stationary probabilities $\pi_A = 1/(1+p)$, $\pi_B = p/(1+p)$.

Relation to previous works. —Here, $q$ simulators coincide with previous models for the case of Markovian processes [17–19] where unitary implementation has already been demonstrated [16,35]. For non-Markovian processes, however, the currently simplest known quantum models face a trade-off between memory reduction due to quantum encoding and increasing size of the simulator’s working memory (proportional to the number $L$ of $|A|$-dimensional systems constituting said memory) [18,19]. This trade-off can be seen in those models’ quantum machine complexity.

$\tilde{C}_q(L) := S \left( \sum_i \pi_i \langle \eta_i(L) | \langle \eta_i(L) \rangle \right)$, \hspace{.5cm} where \hspace{.5cm} (15)

$|\eta_i(L)\rangle := \sum_{x_{0:L}} \sqrt{P(x_{0:L}|i)} \langle \lambda(i,x_{0:L}) | x_{0:L}\rangle$. \hspace{.5cm} (16)

Here, orthonormal states $|j\rangle$ [i.e., $\langle \lambda(i,x_{0:L}) |$] correspond to classical causal states $s_j$.

While this has not been proven in general, there is strong numerical indication that $\tilde{C}_q(L)$ decreases with increasing encoding length $L$ [19]. No further reduction occurs when $L$ exceeds the cryptic order, formally defined as the smallest $k$ for which the following conditional Shannon entropy vanishes:

$H(S_k|\tilde{X}) = 0$. \hspace{.5cm} (17)

Here, $S_k$ is the random variable corresponding to the causal state after $k$ time steps. Hence, $\tilde{C}_q(k) = \tilde{C}_q(M) \forall M > k$. In the case of infinite cryptic order, $\tilde{C}_q(L)$ approaches $\lim_{L\to\infty}\tilde{C}_q(L)$ exponentially [18,19].

Supplemental Material C of Ref. [28] contains the proof that $C_q = \lim_{L\to\infty}\tilde{C}_q(L)$. Importantly, $q$ simulators only require a single system of dimension $|A|$ as working memory to achieve the same memory savings. This is shown in Fig. 2 for the example of the upset-gambler process.

In addition to a trade-off between memory savings and dimension of the internal working memory, previous models for non-Markovian processes implicitly rely on a measure-and-prepare logic [17–20]. They encode causal states in distinct orthogonal states of the simulator’s internal memory [i.e., the states $|j\rangle$ in Eq. (16)] and require projective measurement and a fresh encoding for the next simulation step. For instance, if causal state $|j\rangle$ were
complexity measurement and preparation, the memory required to store the state \( \eta_j(L) \) would be prepared. Between measurement and preparation, the memory required to store the outcome consequently equals the classical statistical complexity \( C_p \). Here, \( q \) simulators add to these models by allowing for unitary simulation even of non-Markovian processes, hence confirming that \( C_q \approx \mu_0 \) for \( Q \approx 0.10 \) which it reaches asymptotically for \( L \to \infty \). The \( q \) simulator operates at \( C_q \) throughout (green dashed line). For an illustration of the process’ full parameter range, see Supplemental Material B of Ref. [28].

measured, the state \( \eta_j(L) \) would be prepared. Between measurement and preparation, the memory required to store the outcome consequently equals the classical statistical complexity \( C_p \). Here, \( q \) simulators add to these models by allowing for unitary simulation even of non-Markovian processes, hence confirming that \( C_q = \lim_{L \to \infty} \tilde{C}_q(L) \) correctly identifies the amount of quantum memory throughout the simulation and not just at intermediate points. Since \( q \) simulators operate by sequential application of the same unitary interaction at each time step, they require no additional processing of the memory system.

Discussion.—We have introduced unitary quantum simulators of stochastic processes which store less information than any classical counterpart. The resulting \( q \) simulator encodes relevant information about a process’ past directly into states retained in quantum memory. At each time step \( t \), the \( q \) simulator unitarily interacts its memory and an incoming ancillary system, such that later measurement of that system in a standard basis yields a statistically correct output \( x_t \) and simultaneously collapses the memory into the correct quantum state for continued simulation at the next time step. Our framework is constructive: given a stochastic process, one can solve for both the quantum memory states and the exact form of the unitary gate allowing future simulation. The resulting simulator’s required memory \( C_q \) aligns with the best previously known quantum constructions [17–19]. Our new simulation scheme then improves upon these by providing a means to generate desired predictions, one step at a time, by repeated unitary interaction such that the memory of the simulator never exceeds \( C_q \) at any time during the operation. It hence avoids trade-offs between entropic memory reduction and dimension of the memory state space. Complementing Ref. [19], our results may be used to analytically compute \( C_q \). The same construction holds when the single-shot memory quantifier \( C_q \) is considered, further reducing the memory Hilbert space for processes which exhibit \( C_q^0 < C_q^0 \).

These results pave the way to new experimental realizations. Quantum advantage in stochastic simulation has been demonstrated for the special case of Markovian processes [16,35], but dimensional scaling has so far made simulations of non-Markovian counterparts significantly more challenging. Application of our results to an infinite Markov order process has resulted in a \( q \) simulator that requires only a single two-qubit unitary gate and a single qubit of memory, bringing it well within reach of present experimental technologies.

It would be interesting to extend these results to continuous-time processes [31,36] and to interactive systems which take information from their environment and use it to generate suitable output responses [37,38]. In both cases, quantum models have been demonstrated to be more powerful than all classical counterparts. Our techniques offer a promising approach towards building unitary quantum simulators for such generalized scenarios.

These further research avenues will benefit from recent results that relate stochastic processes to so-called matrix product states which are used in the description of quantum spin chains [39].

The authors thank A. Monras, K. Wiesner, C. Yang, V. Narasimhachar, C. Di Franco, A. Garner, T. Elliott, W. Y. Suen, N. Tischler, J. Mahoney, and J. Crutchfield for valuable comments and acknowledge support by the National Research Foundation of Singapore (Fellowship NRF-NRFF2016-02), the John Templeton Foundation (Grant No. 54914), Foundational Questions Institute grant, “Observer-dependent complexity: the quantum-classical divergence over ‘what is complex?’,” and the Singapore Ministry of Education Tier 1 RG190/17.

quantum@felix-binder.net

\[ C_q = \lim_{L \to \infty} \tilde{C}_q(L) \]

\[ C_q^0 < C_q^0 \]

\[ \tilde{C}_q(L) \]

\[ \mu_0 \]

\[ \eta_j(L) \]

\[ \tilde{C}_q(L) \]

\[ \lim_{L \to \infty} \tilde{C}_q(L) \]

\[ q \]

\[ C_q \approx \mu_0 \]

\[ C_p \]

\[ x_t \]

\[ \text{PhD thesis, University of Wisconsin-Madison, 2001.} \]

\[ \text{Neural Comput. 22, 121 (2010).} \]

\[ \text{Neural Comput.} \]

\[ \text{Phys. Rev. Lett. 63, 105 (1989).} \]

\[ \text{J. Stat. Phys. 104, 817 (2001).} \]

\[ \text{J. Phys. 8, 17 (2012).} \]

\[ \text{Neural Comput. 22, 121 (2010).} \]

Formally, stationarity implies that for any word $w_L$ consisting of $L$ symbols $P(X_0:L = w_L) = P(X_1;L = w_L)\forall L$. This does not contradict the fact that generally $P(X_0:L = w_L|X_{-\infty:0}) \neq P(X_1;L = w_L|X_{-\infty:1})$ for $X_{-\infty:0} \neq X_{-\infty:1}$.