

**Bridging the gap between classical and quantum many-body information dynamics**Andrea Pizzi <sup>1,2,\*</sup>, Daniel Malz <sup>3,4,†</sup>, Andreas Nunnenkamp <sup>5</sup> and Johannes Knolle <sup>4,6,7</sup><sup>1</sup>*Department of Physics, Harvard University, Cambridge, Massachusetts 02138, USA*<sup>2</sup>*Cavendish Laboratory, University of Cambridge, Cambridge CB3 0HE, United Kingdom*<sup>3</sup>*Max-Planck-Institute of Quantum Optics, Hans-Kopfermann-Straße 1, 85748 Garching, Germany*<sup>4</sup>*Munich Center for Quantum Science and Technology (MCQST), 80799 Munich, Germany*<sup>5</sup>*Faculty of Physics, University of Vienna, Boltzmanngasse 5, 1090 Vienna, Austria*<sup>6</sup>*Department of Physics, Technische Universität München, James-Franck-Straße 1, 85748 Garching, Germany*<sup>7</sup>*Blackett Laboratory, Imperial College London, London SW7 2AZ, United Kingdom* (Received 14 April 2022; revised 11 November 2022; accepted 21 November 2022; published 6 December 2022)

The fundamental question of how information spreads in closed quantum many-body systems is often addressed through the lens of the bipartite entanglement entropy, a quantity that describes correlations in a comprehensive (nonlocal) way. Among the most striking features of the entanglement entropy are its unbounded linear growth in the thermodynamic limit, its asymptotic extensivity in finite-size systems, and the possibility of measurement-induced phase transitions, all of which have no obvious classical counterpart. Here, we show how these key qualitative features emerge naturally also in classical information spreading, as long as one treats the classical many-body problem on par with the quantum one, that is, by explicitly accounting for the exponentially large classical probability distribution. Our analysis is supported by extensive numerics on prototypical cellular automata and Hamiltonian systems, for which we focus on the classical mutual information and also introduce a “classical entanglement entropy.” Our study sheds light on the nature of information spreading in classical and quantum systems, and opens avenues for quantum-inspired classical approaches across physics, information theory, and statistics.

DOI: [10.1103/PhysRevB.106.214303](https://doi.org/10.1103/PhysRevB.106.214303)**I. INTRODUCTION**

Many-body quantum systems display a huge variety of physical phenomena and may carry a vast amount of information in the exponentially many components of their wavefunction. Characterizing this information has become a major goal of modern quantum science, of prime relevance for quantum computing [1] and quantum simulation [2,3]. In recent years the study of closed many-body quantum systems has flourished in particular in the nonequilibrium regime, with key questions revolving around the dynamics of equilibration and information spreading [4,5]. These have become increasingly relevant in light of the experimental advances with nonequilibrium many-body systems kept in almost isolated conditions, e.g., in Rydberg atom arrays [6] or cold atoms in optical lattices [7].

One of the most prominent tools that has emerged from this field is the bipartite entanglement entropy (EE). As it can account for correlations in a comprehensive, nonlocal, multipoint way, the EE has been extensively adopted to monitor the dynamical entangling of the system’s parts in pure quantum systems [8–24]. Among the paradigms that have been unearthed are the linear and logarithmic unbounded growth of the EE in the thermodynamic limit for generic many-body

systems [21] and in many-body localized (MBL) ones [12,25], respectively, and its saturation to an extensive value for finite-size systems at long times. More recently, it has been shown that this extensivity allows for novel measurement-induced phase transitions (MIPTs), whereby the rate of local random measurements determines whether the asymptotic EE scales proportional to the volume of the considered system’s parts (“volume-law scaling”) or to the area of the boundary between them (“area-law scaling”). Since its discovery three years ago [26], the MIPT has received a remarkable amount of interest [27–35].

An obvious and conceptually fundamental question is: To what extent are these distinctive and celebrated many-body features of quantum information spreading purely quantum? One of the reasons why they have become well appreciated in the quantum world is that state-of-the-art numerical techniques can approximate [36] or even exactly describe [37,38] the exponentially large many-body wavefunction. In contrast, the probability distribution of classical nonequilibrium many-body systems, while also being exponentially large, is normally not explicitly accounted for. Rather, classical information spreading is mostly investigated in terms of spatiotemporal correlations, transport properties, and the spreading of perturbations [39–49], all of which are based on the study of few-body observables that can be effectively estimated *à la* Monte Carlo as averages over a handful of trajectories. According to the paradigm of thermalization, the expected value of these few-point observables reaches, at long

\*ap2076@cam.ac.uk

†daniel.malz@mpq.mpg.de

TABLE I. Bridging the classical-quantum gap in many-body information spreading. We summarize the key differences and, mostly, close analogies between classical and quantum information spreading. The setting we refer to is one in which initially local classical or quantum fluctuations become nonlocal under some time-reversible and local dynamics. In this table note that, for pure states, one can replace “quantum MI” by “quantum EE” because  $I_{A:B} = 2S_e$ . The substitution “classical MI” by “cEE” is also possible. The last column points to suitable sections, figures, or references in which each comparison can be best appreciated.

	Classical	Quantum	Ref.
Space	phase space	Hilbert space	
$N$ -Particle space size	$\sim e^{\mathcal{O}(N)}$	$\sim e^{\mathcal{O}(N)}$	
State of the system	probability distribution $p$	wavefunction $ \psi\rangle$	
The nonseparability of a bipartition is measured by $I_{A:B} = 0$ if	MI $I_{A:B}$	MI $I_{A:B}$	Sec. II
Few-point observables become <sup>a</sup>	$p_{A,B} = p_A p_B$	$ \psi\rangle =  \psi_A\rangle \psi_B\rangle$	Sec. II
The MI can be <sup>a</sup>	thermal	thermal	Sec. III
But the MI of a thermal ensemble is just <sup>a</sup>	volume law	volume law	Secs. IV, VI
The probability and wavefunction become <sup>a</sup>	area law	area law	Ref. [50]
Origin of effective randomization <sup>a</sup>	effectively randomized	effectively randomized	Sec. III
The effective randomization underlines <sup>a</sup>	chaos and incompressibility	chaotic spectrum and dephasing	Sec. III
For $N \rightarrow \infty$ the MI shows <sup>a</sup>	volume-law MI	volume-law MI	Ref. [51]
Site $j$ is measured according to	unbounded linear growth	unbounded linear growth	Figs. 2, 6
Right after a measurement, site $j$ is	the marginal probability $p_j$	the reduced density matrix $\rho_j$	Sec. V
Information spreading versus measurements can lead to	“factored out,” $p = p_j p_{\setminus j}$	“factored out,” $\rho = \rho_j \rho_{\setminus j}$	Sec. V
Computing the MI classically requires	MIPT	MIPT	Sec. V
Extracting the MI experimentally requires	exponentially large resources	exponentially large resources	Sec. VII
	exponentially many runs	exponentially many runs	Sec. VII

<sup>a</sup>Valid under the general circumstances specified in the main text.

times and under general circumstances, the value predicted by a thermal ensemble  $\propto e^{-\beta H}$  at suitable temperature  $\beta^{-1}$  [52]. But for instance, the bipartite mutual information (MI), that is often regarded as the natural classical analog of the bipartite EE, is many body in nature, requires the knowledge of the system’s full probability distribution, and at long times does not obviously reach the value predicted by a thermal ensemble.

Here, we argue that this mismatch in the description of quantum wavefunctions and classical probability distributions has led to a gap between our understanding of classical and quantum information spreading. We bridge this gap by (i) accounting for the dynamics of the full many-body probability distribution, thus treating the classical many-body problem on par with the quantum one, and (ii) focusing on many-body information measures akin to quantum EE, such as the classical MI and the “classical EE” (cEE), a complementary measure for classical nonseparability that we introduce. We find that the qualitative features of quantum and classical information spreading are remarkably similar, and show the first instances of asymptotic extensivity, unbounded linear growth, and MIPT of MI and cEE in a classical setting. The analogies between classical and quantum information spreading are summarized in Table I.

The rest of the paper is structured as follows. In Sec. II we introduce the main notation and definitions. In Sec. III we review the idea of thermalization following a quench to remark on the importance of retaining information on the full and exponentially large state of the system, both in the classical and quantum cases. In Sec. IV we show how the main features of quantum information spreading naturally emerge in classical cellular automata: the bipartite MI and cEE grow linearly in time until saturating to an extensive value. Emphasis is put on the role that time reversibility plays in these

effects. Upon interleaving the automaton dynamics with local measurements, we find and analyze a classical MIPT in Sec. V. The discussion moves then to continuous Hamiltonian dynamics in Sec. VI, where with analytical arguments and numerics we show the asymptotic MI to be generally extensive, reducing to area law only for initial conditions at an effective infinite temperature. We highlight a key difference between the MI and the cEE, namely, that the cEE remains asymptotically extensive even at infinite temperature, thus further narrowing the classical-quantum gap. In Sec. VII we argue that, beyond sharing the same phenomenology, classical and quantum many-body information spreading also require remarkably similar experimental and computational protocols to be observed, involving in both cases either exponentially large resources or exponentially many runs. We conclude in Sec. VIII with a discussion of the results and an outlook.

## II. NOTATION AND DEFINITIONS

Consider a bipartite system  $(A, B)$  consisting of either classical or quantum degrees of freedom. We are interested in quantifying the amount of information that the parts  $A$  and  $B$  carry on one another. The system is described by a probability distribution  $p$  and by a density matrix  $\rho$  in the classical and quantum cases, respectively. While we are here ultimately interested in the study of classical systems, to most easily appreciate the classical-quantum connections we will try to develop the formalism in clear analogy with the quantum one. Whether we are talking about a classical or a quantum system should be clear from context and notation.

The part  $A$  itself is described by the marginal probability distribution  $p_A = \text{Tr}_B p = \sum_B p_{A,B}$  or by the reduced density matrix  $\rho_A = \text{Tr}_B \rho$ , with  $\text{Tr}_B$  the partial trace over  $B$ . The classical entropy reads  $S = -\sum_{A,B} p_{A,B} \log p_{A,B}$ , whereas the

quantum (von Neumann) entropy yields  $S = -\text{Tr}[\rho \log \rho]$ . Similarly, the marginal entropy associated with  $A$  reads  $S_A = -\sum_A p_A \log p_A$  or  $S_A = -\text{Tr}_A[\rho_A \log \rho_A]$ , and analogously for  $S_B$ .

A central object in information theory, statistics, and statistical physics is the MI, quantifying by how much our ignorance on  $B$  is reduced when observing  $A$  [53]. Among its possible representations, one that holds for both the classical and quantum cases is

$$I_{A:B} = S_A + S_B - S. \quad (1)$$

The MI is non-negative, and vanishes if and only if the state is separable, that is, when  $p_{A,B} = p_A p_B$  or  $\rho = \rho_A \otimes \rho_B$ . Thus, the MI unambiguously diagnoses any statistical interdependence between  $A$  and  $B$ , that is, their nonseparability.

Generally, the marginal entropies  $S_A$  and  $S_B$  are not representative of the degree of interdependence of  $A$  and  $B$ , but just of the degree of uncertainty on  $A$  and  $B$  themselves:  $S_A$  and  $S_B$  can be positive while  $I_{A:B}$  vanishes. Still, special cases exist in which  $S_A$  and  $S_B$  are proportional to the MI. Most notably, for pure quantum states ( $\rho = |\psi\rangle\langle\psi|$ ) one has  $S = 0$ , and thus  $S_A = S_B = \frac{1}{2}I_{A:B}$ . Indeed, in this case  $S_A$  is known as EE. Since  $S_A$  is defined for generic mixed states, but can only be called EE when used on pure states, it is worth with a bit of redundancy to introduce a new symbol  $S_e$  for the EE, whose explicit expression reads

$$S_e = -\text{Tr}_A(\rho_A \log \rho_A), \quad \rho_A = \text{Tr}_B(|\psi\rangle\langle\psi|). \quad (2)$$

Because in a pure quantum system the MI coincides (up to a factor of 2) with the EE, it is natural to take the classical MI as a classical counterpart of the quantum EE. On top of this well-established classical analog of the EE, we propose an alternative and complementary one, the cEE. To define it, we first introduce the ‘‘classical reduced density matrix’’  $\rho_A$  as the matrix with entries

$$(\rho_A)_{A',A''} = \sum_B \sqrt{p_{A',B} p_{A'',B}}. \quad (3)$$

The diagonal elements of  $\rho_A$  yield the marginal probability distribution,  $(\rho_A)_{A,A} = p_A$ , and so  $\text{Tr}\rho_A = 1$ . Since  $\rho_A$  is positive semidefinite [54] and  $\text{Tr}\rho_A = 1$ , its eigenvalues  $\{\lambda_n\}$  represent a valid (positive and normalized) probability distribution, which we use to define the cEE as

$$S_e = -\text{Tr}[\rho_A \log \rho_A] = -\sum_n \lambda_n \log \lambda_n. \quad (4)$$

Note that the cEE  $S_e$  is nothing but the quantum EE associated to the ‘‘classical wavefunction’’ [55,56]  $|\psi\rangle$  with components

$$\langle A, B|\psi\rangle := \sqrt{p_{A,B}}, \quad (5)$$

which is correctly normalized,  $\langle\psi|\psi\rangle = \sum_{A,B} p_{A,B} = 1$ , and which lives in a fictitious Hilbert space generated by considering the configurations of  $(A, B)$  as a basis  $\{|A, B\rangle\}$ . By construction, many properties of the cEE thus directly follow from those of its quantum counterpart. Most importantly,  $S_e \geq 0$ , and  $S_e = 0$  if and only if  $I_{A:B} = 0$ , which makes the cEE a good witness of statistical dependence between  $A$  and  $B$ , in contrast to the marginal entropy  $S_A$  (the latter is computed on the diagonal of  $\rho_A$  rather than on its eigenvalues). Note that

quantum mechanics was helpful to come up with the definition of cEE above, but not necessary. Indeed, a scientist that completely ignores the existence of quantum mechanics could still compute the cEE, and interpret it as a measure of the nonseparability of  $A$  and  $B$ . Moreover, we note that, while for pure quantum states the MI and the EE coincide (up to a factor of 2), the classical MI and cEE are generally different. Indeed, while they both quantify statistical dependence between  $A$  and  $B$ , the cEE features scaling behaviors that the MI lacks, and that are crucial to connect to quantum information spreading.

As noted, the cEE in Eq. (4) coincides with the EE of a special set of wavefunctions, those with positive real components (for the chosen local basis). Indeed, the dynamics of the cEE in a classical system can be understood as the dynamics of the quantum EE in special quantum systems that do not change the phases of the wavefunction’s components, a setting studied, e.g., in Refs. [57–60], provided that the initial wavefunction is itself real and positive. Despite this, whether the general features of the EE in many-body quantum dynamics should also apply to the small corner of real positive wavefunctions is not obvious. Indeed, Fisher and Grover have shown that, for the case of random wavefunctions, which in the context of many-body dynamics considered here becomes relevant at long times, whether a wavefunction is complex or real can fundamentally affect its entanglement properties [61]. In particular, the Rényi entropies of order  $n > 1$  feature volume- and area-law scaling for complex and real wavefunctions, respectively, whereas they exhibit volume-law scaling in both cases for  $n \leq 1$ . While the former case shows that an extension quantum  $\rightarrow$  classical cannot be taken for granted, the latter opens the way to volume-law scaling of the cEE (that is, the Rényi entropy of order  $n \rightarrow 1$ ), and thus to its particularly rich physics.

### III. A LESSON FROM QUANTUM MECHANICS: THE IMPORTANCE OF TRACKING THE GLOBAL STATE

Touching upon a number of well-known ideas around the concepts of thermalization and convergence of ensembles, in this section we establish close parallels between classical and quantum quench dynamics. This highlights the need to account for the global state of the system when looking at the MI, a key conceptual step necessary to correctly frame the classical problem on par with the quantum one.

Consider a many-body system undergoing nondissipative dynamics generated by the Hamiltonian  $H$ . The global state of the system is described by a wavefunction  $|\psi\rangle$  and by a probability distribution  $p$  in the quantum and classical cases, respectively [62]. The wavefunction is represented with respect to a certain basis  $\{|n\rangle\}$  as  $|\psi\rangle = \sum_n \langle n|\psi\rangle |n\rangle$ , whereas the probability distribution is represented with respect to a certain set of coordinates  $x$  as  $p(x)$ . We are interested in the general scenario of local, many-body, and nonintegrable Hamiltonians  $H$ .

We assume the standard situation in which fluctuations are initially local, meaning that the various parts of the system are statistically independent. Specifically, we consider a product state  $|\psi(0)\rangle = |\psi_A(0)\rangle \otimes |\psi_B(0)\rangle$  and a disjoint probability distribution  $p(0) = p_A(0)p_B(0)$  for a quantum and a classical system, respectively, such that  $I_{A:B}(0) = S_e(0) = 0$ . The initial

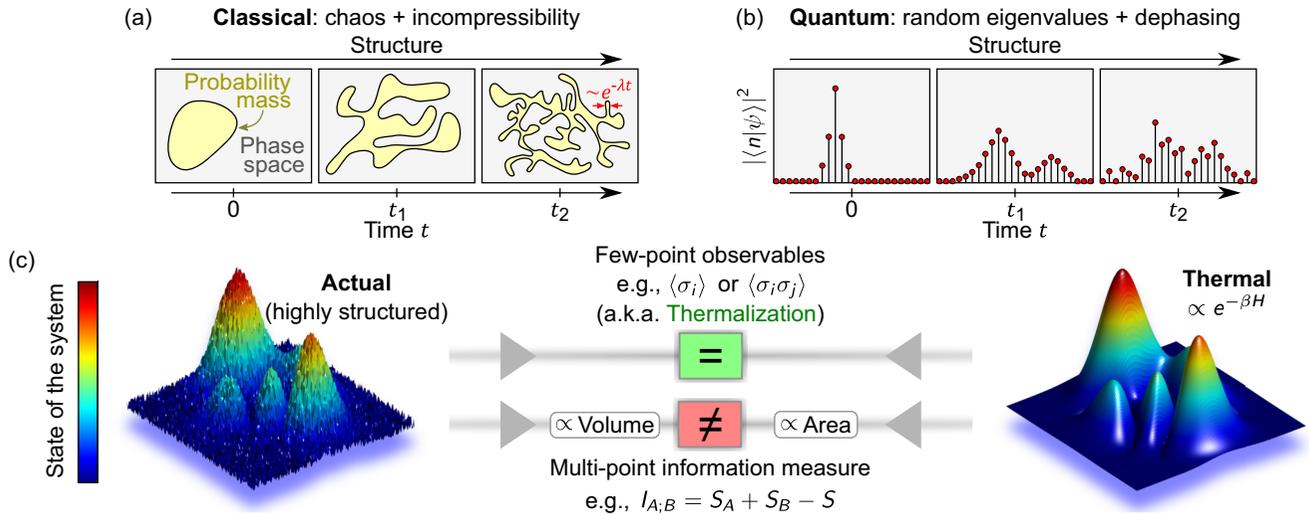


FIG. 1. Emergence of highly structured states in classical and quantum many-body dynamics. (a) Schematic representation of the evolution of the probability mass in the classical phase space under Hamiltonian dynamics. Due to chaos and incompressibility, the probability mass becomes more and more structured in phase space, with fine features of size  $\sim e^{-\lambda t}$ , with  $\lambda$  the Lyapunov exponent, effectively acquiring a random character. (b) Schematic representation of the evolution of the wavefunction, represented with respect to a computational basis  $\{|n\rangle\}$  sorted with respect to  $\langle n|H|n\rangle$ . The wavefunction becomes more and more structured due to the dephasing effects from an effectively random spectrum. (c) The state of the system, whether in the form of a classical probability distribution or quantum wavefunction, is at long times very highly structured in generic many-body dynamics, as pictorially represented on the left. By contrast, a thermal distribution  $\propto e^{-\beta H}$  washes out much of this fine structure (right). If the system has thermalized, few-body observables such as order parameters or correlation functions are correctly predicted by the thermal distribution  $\propto e^{-\beta H}$  (top), but multipoint information measures such as the MI are not (bottom). Specifically, thermal states have area-law MI whereas the exact state at late times has volume law. It is thus necessary to keep track of the exact state of the system to correctly account for information spreading, which is clear in the quantum setting, and which we here consider in the classical one.

quantum wavefunction (classical probability distribution) has a simple aspect in the Hilbert (phase) space. Of course, the shape depends on the specific choice of the basis (coordinates), but can be pictured as something simple if the latter is local, which we will assume [see left panels in Figs. 1(a) and 1(b)].

Under the nondissipative dynamics, the spreading of correlations breaks the local character of the wavefunction (probability distribution), which in the local basis becomes therefore more and more “structured.” In the quantum case, this can be understood from the wavefunction amplitudes  $|\langle n|\psi(t)\rangle|^2 = |\sum_E \langle n|E\rangle e^{-i\frac{Et}{\hbar}} \langle E|\psi(0)\rangle|^2$  acquiring an effectively random character due to the chaotic spectrum  $\{E\}$  of many-body nonintegrable Hamiltonians [63–67]. Indeed, the phases  $e^{-i\frac{Et}{\hbar}}$  at long times  $t$  become effectively random, resulting in probabilities  $|\langle n|\psi\rangle|^2$  that fluctuate in time and with no clear dependence on  $n$  [see the rightmost panel of Fig. 1(b)]. In the classical case, the dynamical randomization of the probability is instead due to incompressibility and chaos, which taken together imply the probability distribution to develop finer and finer features [52], or more and more structure, as time increases [see the leftmost panel of Fig. 1(a)].

It turns out that these fine features make only little difference for most observables of interest, according to the paradigm of thermalization. The latter stipulates that, in the time evolution after a quench, local few-point observables equilibrate at long times to their thermal value, which depends on the initial condition only through its energy. This general idea has long been established in classical physics

in terms of chaos and ergodicity [52] and more recently for closed quantum many-body systems in terms of the eigenstate thermalization hypothesis (ETH) [63–66]. The concept of thermalization is naturally associated with that of convergence of ensembles [52], according to which we can think of the distribution of the system (classical probability  $p$  or quantum density matrix  $\rho = |\psi\rangle\langle\psi|$ ) as approaching the stationary thermal one  $\propto e^{-\beta H}$ .

This point deserves much care, though. As a matter of fact, the idea of convergence of the ensembles only holds at the level of few-point observables, reduced density matrices, and marginal probability distributions. But as described above, the global state itself is far from stationary, let alone thermal, and this can have deep consequences on many-point observables. Most importantly, while the long-time value of the quantum MI fulfills volume-law scaling [8,13], which is rooted precisely in the effective randomness of the many-body wavefunction [51], the MI of a thermal state fulfills area-law scaling [50,68]. Indeed, a thermal distribution  $\rho_{th}$  washes out the fine structure of the state, thus missing a large (extensive) amount of MI, as pictorially illustrated in Fig. 1(c).

This example highlights the importance of keeping track of the exponentially large state of the system for studying many-body information spreading. While this is customary for quantum theories, it is not for classical ones, that instead either assume its convergence to a canonical ensemble or focus on few-body observables within Monte Carlo sampling. Individuating this difference as the main origin of the gap between our understanding of classical and quantum

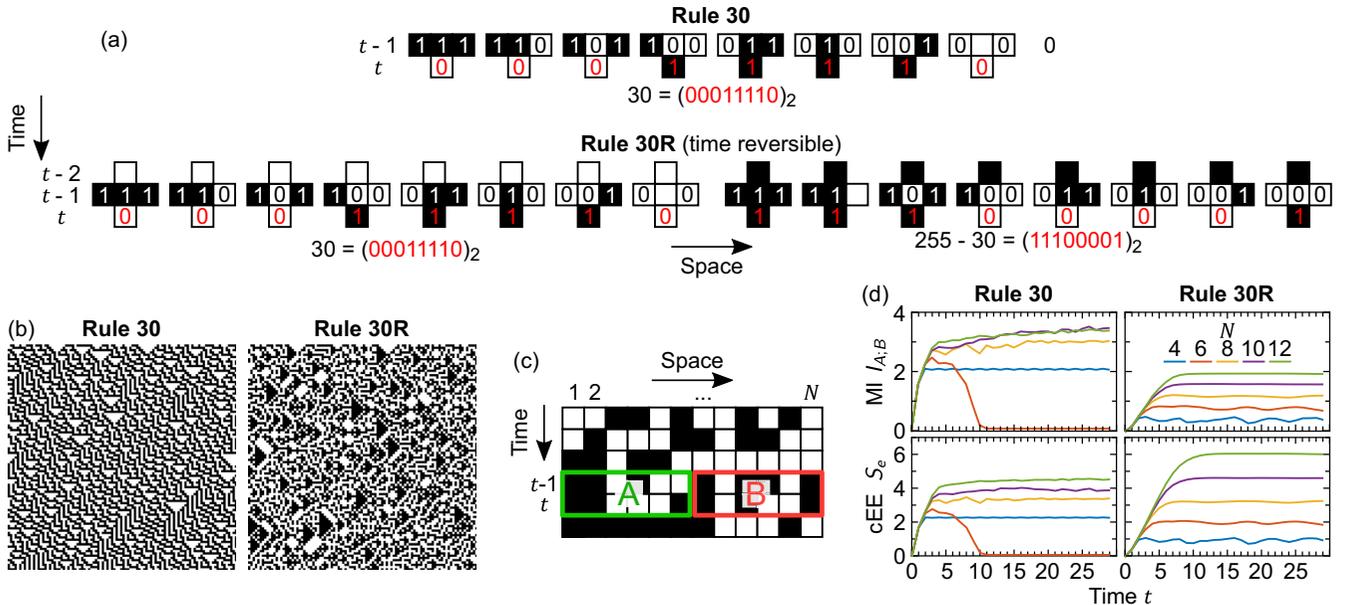


FIG. 2. Many-body information spreading in classical cellular automata. (a) Rule 30 cellular automaton (top) and its (second-order) reversible version Rule 30R (bottom). (b) Instances of the spacetime profiles of Rule 30 (left) and Rule 30R (right). (c) Graphical representation of the subsystems  $A$  and  $B$  that are used to compute MI and cEE. The state of the system at time  $t$  is unambiguously defined by the state of the bits at time  $t$  and at time  $t - 1$ . (d) Dynamics of MI and cEE for various system sizes  $N$ . In Rule 30 (left), the dynamics of MI and cEE does not exhibit typical features of classical and quantum nondissipative dynamics. By contrast, Rule 30R captures features such as the initial linear growth of entanglement  $\sim t$  and its eventual saturation to an extensive  $\sim N$  value (reached after a time  $t \sim N$ ). Indeed, this stresses that key ingredients behind this phenomenology are locality, chaos, and time reversibility, the latter of which is featured by Rule 30R but not by Rule 30. Here, we used  $q_0 = 0.7$ .

information spreading is the main conceptual finding of our work.

Before substantiating these ideas with numerics on Hamiltonian dynamics, which involves an extra phase-space discretization procedure and that we postpone to Sec. VI, let us take a step back and start by considering the simpler case of classical cellular automata.

#### IV. INFORMATION SPREADING IN CLASSICAL CELLULAR AUTOMATA

In this section, we apply our probabilistic framework to investigate information spreading in classical cellular automata. Imagine drawing an initial condition from a disjoint (in  $A$  and  $B$ ) initial probability distribution and to evolve it “blindly,” meaning without looking at it. If at time  $t$  we inspect the state of half of the system  $A$ , how much do we learn about  $B$ ? To address this question, we will consider how the many-body probability distribution evolves under the automaton dynamics and compute the MI and cEE. In addition to offering a discrete setting convenient for implementation, cellular automata help us to highlight the role played by time reversibility in the dynamics of these quantities.

Consider a system made of  $N$  bits,  $s = (s_1, s_2, \dots, s_N) \in \{0, 1\}^N$ . The bit string evolves in time under the action of some update rule,  $s(t+1) = \text{Rule}[s(t)]$ , with discrete time  $t = 0, 1, 2, \dots$ . For concreteness, we henceforth focus on Wolfram’s Rule 30 [69]. This rule shares two key features with the standard quantum settings for information spreading, namely, chaos and locality, but lacks a third: time reversibility.

This ingredient is decisive, because it ensures information to be preserved in time. In fact, of the 256 rules that can be obtained from local updates involving only the nearest-neighboring sites, none is both chaotic and time reversible [69]. One way of recovering time reversibility from a given rule, while preserving chaos, is to modify it as  $s(t+1) = \text{Rule}[s(t)] \otimes s(t-1)$ , where  $\otimes$  denotes a bitwise logical XOR operation [69,70]. Such a modified rule, that we will call Rule R, constitutes a “second-order” automaton, in which the state of the system at time  $t+1$  does not just depend on the state of the system at time  $t$ , but also on that at time  $t-1$ . Put differently, the reversibility of the automaton means that there exists an injective map connecting  $s(1)$  and  $s(0)$  to  $s(t)$  and  $s(t-1)$ , and vice versa. Rule 30, Rule 30R, and two instances of the spatiotemporal profiles that they generate are shown in Figs. 2(a) and 2(b).

In the case of first-order automata like Rule 30, the full information on the system is carried, at all times  $t$ , by a probability distribution  $p$  with support on the  $2^N$  possible bit-string configurations. In the case of a second-order automaton like Rule 30R, however, a similar distribution does not suffice. Indeed, what matters in this case is not just the configuration of the bit string at time  $t$ , but also that at time  $t-1$ . Consequently, to retain the full information on the system, one has to define the possible states  $\sigma$  as the  $2N$  bit strings composed of the bits  $s(t)$  and  $s(t-1)$  [see Fig. 2(c)]. The probability distribution  $p$  will correspondingly have  $2^{2N}$  components [71].

Initially, we take the bits  $\{s_j(t=0)\}$  and  $\{s_j(t=-1)\}$  to be independent and identically distributed (i.i.d.) random variables, equal to 1 with probability  $q_0$  and to zero otherwise.

That is, each of the  $4^N$  possible initial conditions  $\sigma_0$  has a probability  $p_{\sigma_0}(0)$  that reads

$$p_{\sigma_0}(0) = \prod_{j=1}^{2N} [\sigma_{0,j} q_0 + (1 - \sigma_{0,j})(1 - q_0)]. \quad (6)$$

The evolution of the probability vector from one time to the next is described by a map,

$$p(t+1) = F[p(t)], \quad (7)$$

which can be determined once and for all by checking how each two-time microstate  $\sigma$  is evolved by one application of the automaton. We note that, in the case of reversible automata, like Rule 30R, the map  $F$  acts as a permutation of the elements of  $p$ , because of injectivity. Indeed, in this case the probability does not change when ‘‘sitting’’ on a discrete trajectory  $\sigma_t$  starting in  $\sigma_0$ . In a formula,  $p_{\sigma_t}(t) = p_{\sigma_0}(0)$ , which we can view as the discrete version of the Liouville theorem for Hamiltonian systems [52].

By iterating Eq. (7), we evolve the many-body probability distribution  $p(t)$  and investigate the corresponding MI and cEE dynamics for various system sizes  $N$  in Fig. 2(d). In the nonreversible case of Rule 30, we find that the MI and the cEE generally grow until reaching a stationary value (except in the case  $N = 6$ , which appears somewhat special and that we attribute to particularly severe finite-size effects). The key features characterizing quantum information spreading, namely, linear growth and extensivity at short and long times, respectively, are absent. In striking contrast, these instead appear in the time-reversible automaton. In this case, we find that both the MI and the cEE grow linearly in time,  $S_e, I_{A:B} \propto t$  until a time  $\propto N$ , after which saturation to an extensive value  $\propto N$  takes over. Since the cEE is the EE obtained by treating the classical probability as a wavefunction [see Eq. (5)], we understand that the extensivity of the asymptotic cEE is due to an effective randomization of  $p$ , analogously to how in the quantum case it emerges from the effective randomness of the wavefunction [51,61]. Classical cellular automata therefore allow to appreciate time reversibility as a key ingredient underpinning the distinctive features of both classical and quantum information spreading.

## V. MEASUREMENT-INDUCED TRANSITION

A natural question thus emerges from the previous section, as it did in the quantum case: How does the tendency of chaotic dynamics to build extensive MI and cEE compete with the disentangling effect of local measurements? We turn to this question now, and show a classical MIPT, in which the rate of local measurements determines whether the asymptotic MI and cEE fulfill volume- or area-law scaling, in close analogy to its quantum counterpart [26].

### A. Protocol and classical measurements

Imagine as in Sec. IV to sample an initial condition  $\sigma_0$  and evolve it with the automaton. As long as we do not look at it, the state of the automaton at time  $t$  is described by a probability distribution  $p(t)$ , that starts from  $p(0)$  as in Eq. (6), and evolves as  $p(t+1) = F[p(t)]$  as in Eq. (7). Information

spreading generally builds MI between a given site  $j$  and the rest of the system, meaning  $p \neq p_j p_{\setminus j}$  and  $I_{j;\setminus j} > 0$ , with  $p_j = \text{Tr}_{\setminus j} p$  and  $p_{\setminus j} = \text{Tr}_j p$  the marginals of site  $j$  and of the rest of the system,  $\setminus j$ , respectively. Now, imagine that the ‘‘blind’’ evolution of the automaton is interleaved with random local measurements: each spacetime point  $(j, t)$  can be observed with probability  $p_m$  [Fig. 3(a)]. Measuring the state of site  $j$  at time  $t$  reduces our ignorance on the system and conditions the probability distribution [53]

$$p_{\sigma}(t^+) = \begin{cases} \frac{p_{\sigma}(t^-)}{p_j(m_{j,t}, t^-)} & \text{if } \sigma_j = m_{j,t} \\ 0 & \text{if } \sigma_j \neq m_{j,t}. \end{cases} \quad (8)$$

In Eqs. (8),  $t^-$  and  $t^+$  refer to the instants before and after the measurement, respectively;  $m_{j,t}$  is the measurement outcome, and  $p_j(m_{j,t}, t^-)$  its *a priori* probability. Simply put, to obtain  $p(t^+)$ , we remove all states incompatible with the measurement outcome and normalize the distribution again. In this process, the marginal of site  $j$  becomes  $p_{\sigma_j}(t^+) = \delta_{\sigma_j, m_{j,t}}$ , implying that the postmeasurement conditional many-body probability factorizes,  $p(t^+) = p_j(t^+) p_{\setminus j}(t^+)$ , and that the MI between site  $j$  and the rest of the system vanishes,  $I_{j;\setminus j} = 0$ .

After many measurements, the suppression of probability components in Eq. (8) would eventually lead at long times to the collapse of the state of the system to a single microstate  $\sigma^*$ ,  $p_{\sigma} \rightarrow \delta_{\sigma, \sigma^*}$ , for which no information can spread at all. More interesting is the situation in which the state of the system can never be known with infinite precision. This scenario is for instance natural when the classical bits are a toy model representation of the microscopic world, for which the unavoidable noise and limited measurement resolution make it impossible to fully determine the state of the system. Irrespective of the physical origin of the fluctuations, in our cellular automata model we will simply attribute them to what we call a ‘‘faulty measurement apparatus.’’ Specifically, we suppose that, after a measurement has been performed, the measured bit randomly flips with an error probability  $1 - q$ , with  $0 < q < 1$ . We use  $t^{++}$  to refer to the time right after the possible mistake has been introduced. We have that

$$p_{\sigma_j}(t^{++}) = q \delta_{\sigma_j, m_{j,t}} + (1 - q) \delta_{\sigma_j, 1 - m_{j,t}}. \quad (9)$$

The many-body probability distribution is modified accordingly:  $p(t^{++}) = p_j(t^{++}) p_{\setminus j}(t^{++})$ .

The local measurement protocol is closely analogous to the quantum case. There, the  $j$ th qubit is sampled according to its reduced density matrix  $\rho^{(j)} = \text{Tr}_{\setminus j} \rho$ : the qubit takes value 1 with probability  $\rho_{1,1}^{(j)}$ , and value 0 with probability  $\rho_{0,0}^{(j)}$ . The postmeasurement fluctuations at site  $j$  are intrinsic and due to Heisenberg uncertainty. As in the classical case, the effect of a quantum measurement is that of ‘‘factoring out’’ the state of  $j$  from that of the rest of the system, meaning  $|\psi(t^+)\rangle = |\psi_j(t^+)\rangle |\psi_{\setminus j}(t^+)\rangle$  or equivalently  $\rho(t^+) = \rho_j(t^+) \rho_{\setminus j}(t^+)$ , all of which are conditional on the measurement outcome.

The resulting evolution of the probability distribution can be determined in a computer simulation according to the procedure schematized in Fig. 3(b). To measure the probability distribution  $p$  from an experiment, a procedure as in Fig. 3(c) is instead required, which unfolds as follows. An ensemble of initial states  $\sigma_0$  is sampled from the distribution  $p(0)$ , and evolved under the automaton rules. For the first of

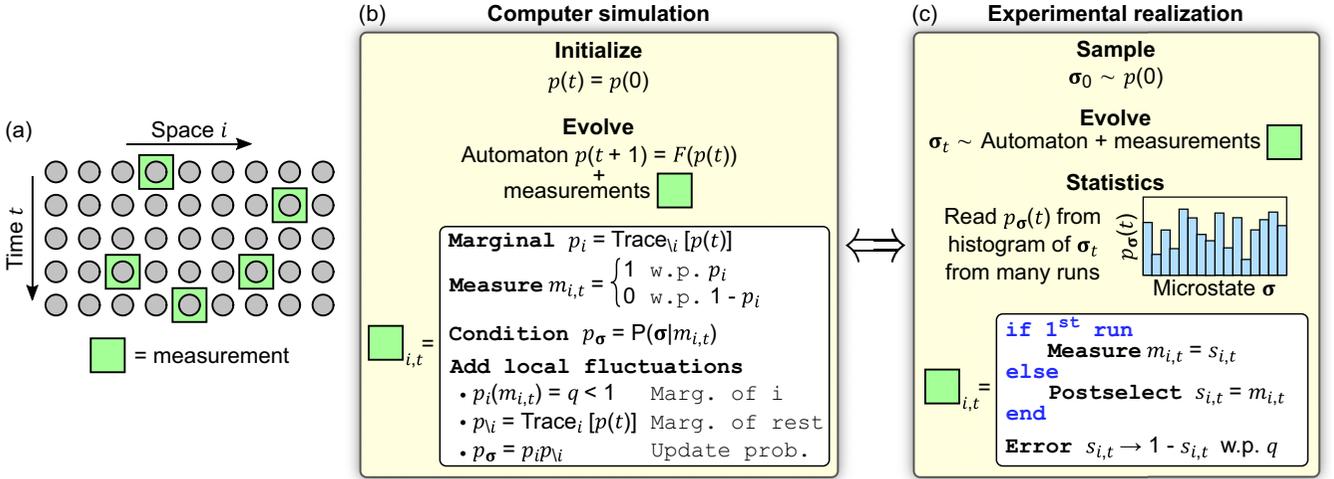


FIG. 3. Classical automaton with measurements: a schematic protocol illustration. (a) Each spacetime point can be measured (green boxes) with probability  $p_m$ . (b) The corresponding evolution of the many-body probability dynamics  $p(t)$  can be obtained in a computer simulation. The probability is initialized as  $p(t) = p(0)$ , and evolved under the automaton dynamics according to the map  $p(t+1) = F[p(t)]$  in Eq. (7). Measuring site  $i$  at time  $t$  corresponds to (i) computing the marginal distribution of the site  $p_i$ , (ii) drawing a measurement outcome according to  $p_i$ , (iii) conditioning the probability accordingly, and (iv) accounting for postmeasurement local fluctuations that flip the measured spin with probability  $q$ . The last point is in practice achieved by transforming the postmeasurement marginal from  $p_i(m_{i,t}) = 1 - p_i(1 - m_{i,t}) = 1$  to  $p_i(m_{i,t}) = q < 1$ , computing the marginal of the rest of the system  $p_{\setminus i}$ , and reconstructing the total distribution as  $p = p_i p_{\setminus i}$ . (c) Alternatively, the probability distribution could be obtained in a (numerical or actual) experiment by running the automaton on exponentially many initial configurations  $\sigma_0$  sampled from  $p(0)$ . The probability distribution  $p(t)$  at later times is obtained as a histogram of the trajectories over the exponentially many possible microstates  $\sigma$ . In the first run, measuring site  $i$  at time  $t$  consists of observing the state  $s_{i,t}$  of the measured bit and storing it in the measurement variable  $m_{i,t}$ . For all subsequent runs, the measurements are rather postselections: if  $s_{i,t}$  does not match the  $m_{i,t}$  measured in the first run, the trajectory is discarded. In any run, the postmeasurement local fluctuations are then accounted for by flipping the state of the measured bit with probability  $1 - q$ . Remarkably, as argued in Sec. VII, both the procedures in (b) and (c) are essentially identical to those needed in the quantum setting.

these runs, measuring a site  $i$  at time  $t$  means taking note of the state of  $s_i(t)$  of the bit, and saving it as  $m_{i,t}$ . The measurement outcome  $m_{i,t}$  is implicitly distributed according to the marginal  $p_i$ . The faulty measurement apparatus then introduces fluctuations as in Eq. (9), by leaving the bit  $s_i(t)$  unchanged or inverted with probabilities  $q$  and  $1 - q$ , respectively. For all the subsequent runs, measurements always happen in the same spacetime locations, and are not measurements but postselections. If  $s_{i,t}$  does not match the  $m_{i,t}$  measured in the first run, the trajectory is discarded, whereas otherwise the trajectory is maintained. Remarkably, this procedure is in close analogy to the one that should be followed in a quantum experiment, as we will argue in Sec. VII. Repeating this procedure exponentially many times, the probability  $p_\sigma(t)$  is estimated as the fraction of trajectories found in  $\sigma$  at time  $t$  and, in the limit of infinitely many runs, matches the probability obtained from the procedure in Fig. 3(b) discussed earlier.

## B. Numerical results

The above protocol is implemented in Fig. 4. In Fig. 4(a) we plot the dynamics of MI and cEE for various measurement rates  $p_m$ . After an initial transient, whose duration is larger for smaller  $p_m$  and that is partially cut by the chosen axes limits, the MI and the cEE reach a stationary value, up to finite-size temporal fluctuations. Asymptotic values for the MI and cEE are obtained as time averages and plotted in

Fig. 4(b) versus the measurement rate  $p_m$ , and for various system sizes  $N$ . Strikingly, for small  $p_m$ , the asymptotic values of MI and cEE scale with the system size  $N$ , whereas they do not for large  $p_m$ . To make the last statement quantitative, we use a linear fit to extract the scaling coefficients  $\alpha_{MI}$  and  $\alpha_{EE}$ , with  $I_{A:B}(\infty) \approx \alpha_{MI} N$  and  $S_\ell(\infty) \approx \alpha_{EE} N$ , and plot them versus  $p_m$  in Fig. 4(c). Indeed, these exhibit a transition from finite value to zero at  $p_m \approx 0.6$ , thus showing the first classical MIPT between area- and volume-law cEE and MI. These quantities closely follow the qualitative behavior of the quantum EE (see, e.g., Fig. 13 in Ref. [26]). In our simulations we take  $q = 0.75$  and  $q_0 = 0.8$  but the key qualitative features are not contingent on this choice. Generally, we do not expect the asymptotic MI and cEE (and, thus, the MIPT) to depend on the initial distribution.

In this section, we considered the most natural classical measurement procedure, that conditions the many-body probability distribution depending on the measurement outcome. We stress that similar results can be obtained for other local protocols that have the ability to decouple local fluctuations, that is, to suppress the mutual information between a local degree of freedom and the rest of the system, which can be achieved in many ways in both the classical and quantum settings. For instance, one could consider simple resettings, in which the measured bit is set to 1 with probability  $q$  and to 0 with probability  $1 - q$ . Away from the probabilistic framework adopted here, we also note that the idea of MIPT can be generalized to even simpler measurement protocols,

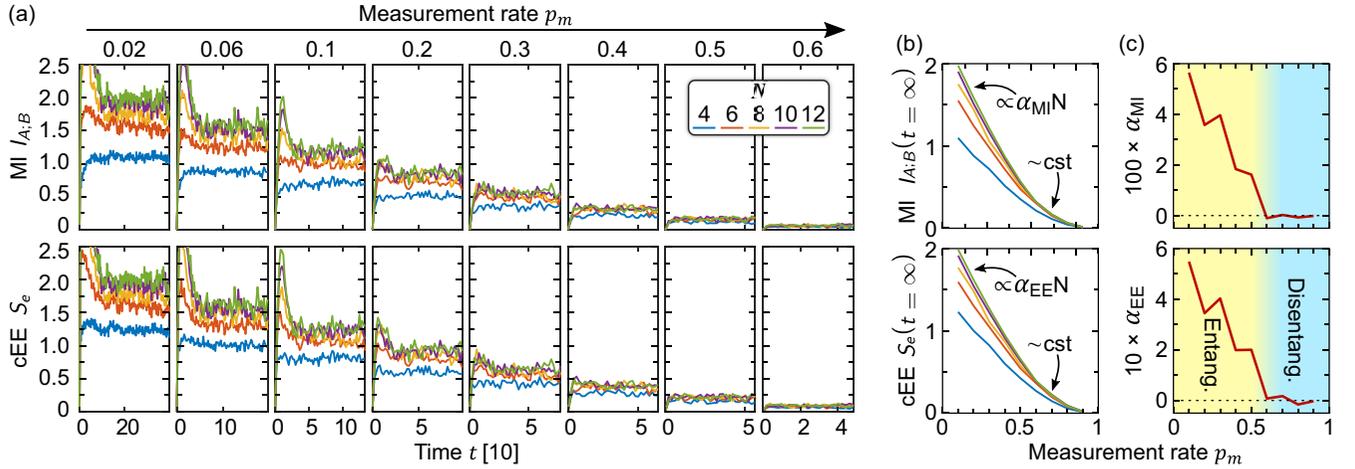


FIG. 4. Classical measurement-induced phase transition. (a) MI and cEE dynamics for various measurement rates  $p_m$  and system sizes  $N$ . After a very quick transient (partially cut), the MI and cEE saturate to a stationary value, up to temporal fluctuations due to finite-size effects. Crucially, the stationary value increases with  $N$  if  $p_m$  is small enough, and is insensitive of  $N$  otherwise. (b) The long-time asymptotic values  $I_{A,B}(t \rightarrow \infty)$  and  $S_e(t \rightarrow \infty)$  are obtained by averaging the results in (a) over the second half of the shown time interval, and plotted versus  $p_m$  for various system sizes  $N$ . The MI and cEE decay with  $p_m$ , reaching zero at  $p_m = 1$ . (c) The scaling coefficients  $\alpha$ , such that  $I_{A,B}(t \rightarrow \infty) \approx \alpha_{MI}N$  and  $S_e(\infty) \approx \alpha_{cEE}N$ , are obtained through linear fits for and highlight an entangling-disentangling MIPT (the fitting excludes  $N = 4$  and  $6$  to reduce finite-size effects). For  $p_m \lesssim 0.5$  the scaling coefficients are finite, indicating extensivity of the asymptotic MI and cEE and thus unbounded growth in the thermodynamic limit. By contrast, for  $p_m \gtrsim 0.5$  the long-time values of MI and cEE become insensitive to system size  $N$ . Here, we used  $q_0 = 0.8$ ,  $q = 0.75$ , and  $R = 100$ .

e.g., defined as locally projecting two trajectories on one another [72].

## VI. INFORMATION SPREADING IN MANY-BODY HAMILTONIAN DYNAMICS

Thanks to their discrete nature, cellular automata have proven particularly convenient for numerical implementation. But let us now go back to the case of Hamiltonian dynamics, which we have already considered in Sec. III as a motivation for keeping track of the full many-body probability. Going beyond the result for area-law MI in a thermal ensemble [50], in this section we find a simple expression for the asymptotic MI in Hamiltonian dynamics. We find it to be generally volume law, reducing to area law only at an infinite effective temperature, unlike the cEE. Introducing a careful phase-space discretization procedure we then numerically investigate information spreading in an interacting classical spin chain. We recover the salient features of information spreading in quantum quenches, and highlight a key difference between MI and cEE.

### A. Asymptotic mutual information in Hamiltonian dynamics

First, we provide a heuristic expression for the asymptotic long-time MI in many-body Hamiltonian dynamics. First of all, Hamiltonian dynamics involves continuous variables, and thus a continuous probability density  $p$  over the phase space. The first point to clarify is therefore how to make sense of the definitions in Sec. II, that are for discrete distributions instead. A discretization of the phase space can be performed dividing it into little volumes  $\Delta$ . In the continuous limit  $\Delta \rightarrow 0$ ,

it is straightforward to show that  $S \approx \mathcal{S} - \log \Delta$ , with  $\mathcal{S} = \int dx p(x) \log(p(x))$  the differential entropy, for which we use a calligraphic notation and involve integration over the whole many-body phase space [53]. Contrary to what one might naively expect, the differential entropy  $\mathcal{S}$  is therefore *not* the continuous limit of the discrete entropy  $S$ , which diverges as  $\log \Delta$  instead. Fortunately, this divergence does not affect the MI, which instead does have a well-defined continuous limit [53]. Indeed, in the expression of the MI in terms of the entropies, Eq. (1), the diverging terms of the entropies cancel, meaning that one can equivalently write  $I_{A,B} = S_A + S_B - S$  or  $I_{A,B} = S_A + S_B - \mathcal{S}$ , in the limit  $\Delta \rightarrow 0$ .

This remark being made, we can now find an expression for the long-time asymptotic value of the MI in Hamiltonian dynamics. Due to Liouville's theorem, the integral of any function of  $p$  over the whole phase space is constant,  $\frac{d}{dt} \int dx f(p(x)) = 0$ . In particular, the differential entropy is conserved,  $\mathcal{S}(t) = \mathcal{S}(0)$ . While, strictly speaking, the total probability distribution does not thermalize due to chaos and incompressibility (see Sec. III); we expect the marginals to do so [73]. As a result, we can assume at sufficiently long times  $t$  that  $S_A(t) + S_B(t) \approx S^\beta + s_{al}$ , where  $S^\beta$  denotes the entropy of a thermal ensemble, whereas  $s_{al}$  is an area-law correction. The inverse temperature  $\beta$  is found from the usual consistency condition  $\langle H \rangle_0 = \langle H \rangle_\beta$ , that is,  $\int dx p(t=0, x) H(x) = \int dx z^{-1} e^{-\beta H(x)} H(x)$ , with  $z = \int dx e^{-\beta H(x)}$  the partition function. From Eq. (1) we thus obtain that the MI at sufficiently long times reaches the asymptotic value

$$I_{A,B}(\infty) \approx S^\beta - \mathcal{S}(0) + s_{al}. \quad (10)$$

Equation (10) states that the asymptotic MI between two halves  $A$  and  $B$  of a many-body Hamiltonian system is equal

to the difference between the initial entropy and a thermal entropy at temperature  $\beta^{-1}$  set by the initial state  $p(0)$ , up to a boundary correction  $s_{al}$ . Note that the entropies in Eq. (10) can equivalently be meant as differential entropies or as discrete entropies in the continuous limit, since the diverging terms cancel anyway, as discussed above.

An important observation is that, by definition, the thermal distribution is the one that maximizes the entropy at a given temperature, and so  $S^\beta \geq S(0)$ , which is consistent with the positivity of the MI and in which the equality holds only for initial states with infinite temperature, for which  $\beta = 0$  and  $\langle H(x) \rangle_0 = \frac{\int dx H(x)}{\int dx}$ , or for thermal initial distributions. Here, we use the term “infinite-temperature state” in reference not to the infinite-temperature thermal distribution, but to any probability distribution whose expected energy coincides with the one that the infinite-temperature thermal distribution would predict. From Eq. (10) we can thus conclude that initial nonthermal distributions  $p(0)$  corresponding to a finite effective temperature  $0 < \beta < \infty$  generally lead to a volume-law asymptotic MI,  $I_{A:B}(\infty) \approx S^\beta - S(0) = \mathcal{O}(N)$ , whereas initial distributions  $p(0)$  that are either thermal or at an effective infinite temperature  $\beta = 0$  result in an area-law asymptotic MI,  $I_{A:B}(\infty) \approx s_{al} \ll N$ .

Further, to estimate the latter, we can compute the MI associated to a random probability distribution, in the same spirit of the calculation of the EE for a random wavefunction [51]. Let us assume that the components of the probability  $p_\sigma$  are i.i.d. random numbers with average  $\langle p_\sigma \rangle = [\sum_\sigma 1]^{-1}$  [74]. For large  $N$  and by virtue of the central limit theorem, the total entropy reads  $S \approx -\frac{\langle p_\sigma \log p_\sigma \rangle}{\langle p_\sigma \rangle}$ . The marginals become instead uniform,  $p_{\sigma_A} \rightarrow \langle p_{\sigma_A} \rangle$  and  $p_{\sigma_B} \rightarrow \langle p_{\sigma_B} \rangle$ , for which the marginal entropies are maximal,  $S_A + S_B \approx -\log \langle p_\sigma \rangle$ . Plugging these results into Eq. (1) quickly leads to the MI at infinite time and effective temperature:

$$I_{A:B}^\infty \approx \left\langle \frac{p_\sigma}{\langle p_\sigma \rangle} \log \frac{p_\sigma}{\langle p_\sigma \rangle} \right\rangle. \quad (11)$$

Equation (10), Eq. (11), and the prediction of volume- and area-law scaling of the asymptotic MI in many-body Hamiltonian systems at finite and infinite effective temperature initial conditions, respectively, are among the major findings of this work, which we now verify numerically.

## B. Case study: Heisenberg chain in a magnetic field

We numerically exemplify the above ideas for a paradigmatic model in the context of chaos, information spreading, and transport: the Heisenberg chain in a magnetic field. Consider  $N$  classical spins  $\mathbf{S}_j = (S_j^x, S_j^y, S_j^z)$  with unitary modulus  $|\mathbf{S}_j|^2 = 1$  and  $j = 1, 2, \dots, N$ , and with Hamiltonian

$$H = \sum_{j=1}^N (\mathbf{S}_j \cdot \mathbf{S}_{j+1} + h_j S_j^z), \quad (12)$$

containing an isotropic nearest-neighbor interaction and a disordered magnetic field along  $z$ , the coefficients  $\{h_j\}$  being independent random numbers drawn uniformly in  $[-W, W]$ . The system undergoes Hamiltonian dynamics,  $S_i^\alpha = \{S_i^\alpha, H\}$ , where  $\{\dots\}$  denotes Poisson brackets and

$\{S_i^\alpha, S_j^\beta\} = \delta_{i,j} \epsilon_{\alpha,\beta,\gamma} S_i^\gamma$ , with  $\delta_{i,j}$  the Kronecker delta,  $\epsilon_{\alpha,\beta,\gamma}$  the Levi-Civita antisymmetric symbol, and  $\alpha, \beta$ , and  $\gamma$  in  $\{x, y, z\}$ . Periodic boundary conditions are assumed, and results can be averaged over  $R$  independent disorder realizations, as relevant. Again, the bipartition  $(A, B)$  is chosen to correspond to the left and right halves of the chain ( $N$  is assumed even with no loss of generality).

The phase space of the system is that parametrized by  $2N$  angles, two per spin, and is continuous as in any Hamiltonian system. As such, it requires discretization, which we perform in the minimal nontrivial way, that is, by splitting each solid angle in  $M = 2$  hemispheres as illustrated Fig. 5(a). Two points that can be taken as representative for these hemispheres are the poles, which we may call “single-body reference points” and which we tag with a discrete variable  $\sigma_j$  taking  $M$  values, say,  $\sigma_j = \pm 1$  for the  $\pm \mathbf{x}$  hemispheres, respectively. The many-body phase space is correspondingly divided into  $M^N$  “pockets,” each of which is represented by a many-body reference point (in which the spins point either towards  $+\mathbf{x}$  or towards  $-\mathbf{x}$ ), and tagged by a bit string  $\sigma = (\sigma_1, \sigma_2, \dots, \sigma_N)$  [see Fig. 5(b)]. Explicitly, the  $\sigma$ th pocket contains all the spin configurations such that  $\text{sign}(S_1^x, S_2^x, \dots, S_N^x) = \sigma$ .

Our goal is to find the dynamics of the discrete probability distribution  $p_\sigma(t)$ . The idea is the following. First, we consider the  $M^N$  many-body reference points as possible initial conditions. Specifically, we consider each spin to initially point either along  $+\mathbf{x}$  or along  $-\mathbf{x}$ , with probability  $q_0$  and  $1 - q_0$ , respectively, so that an initial configuration tagged  $\sigma_0$  has probability

$$p_{\sigma_0} = \prod_{j=1}^N [\sigma_{0,j} q_0 + (1 - \sigma_{0,j})(1 - q_0)]. \quad (13)$$

Again, the initial factorizability of the probability distribution makes our stochastic initial condition the one-to-one correspondent of quantum product states, and implies  $I_{A:B}(0) = S_e(0) = 0$ . Second, we evolve the  $M^N$  trajectories integrating the ordinary differential equations (ODEs) of Hamiltonian dynamics with an ODE solver, which necessarily makes them depart from the reference points they started from, and explore the continuous phase space. Third, each trajectory is projected onto the discrete space to define  $\sigma_t = \text{sign}[S_1^x(t), S_2^x(t), \dots, S_N^x(t)]$  at time  $t$ . Fourth, the probability  $p_\sigma(t)$  is obtained as

$$p_\sigma(t) = \sum_{\sigma_0: \sigma_t = \sigma} p_{\sigma_0}(0), \quad (14)$$

that is, as the sum of the initial probabilities  $p_{\sigma_0}(0)$  of those states  $\sigma_0$  that, at a time  $t$ , are found in the  $\sigma$ th pocket of the phase space. Fifth and final, the thus-obtained vector of discretized probabilities  $p$  is used to compute quantities of interest, including the MI and cEE.

For the trivial points  $q_0 = 0, 1$  the spins are all initially perfectly polarized along either  $-\mathbf{x}$  or  $+\mathbf{x}$ , respectively, and remain so at all times, with no dynamics happening, no spreading of correlations,  $p(t) = p(0)$ , and  $I_{A:B} = S_e = 0$ . On the other hand, the value  $q_0 = 0.5$  corresponds to an infinite-temperature state,  $\langle H \rangle_0 = 0 = \langle H \rangle_{\beta=0}$ .

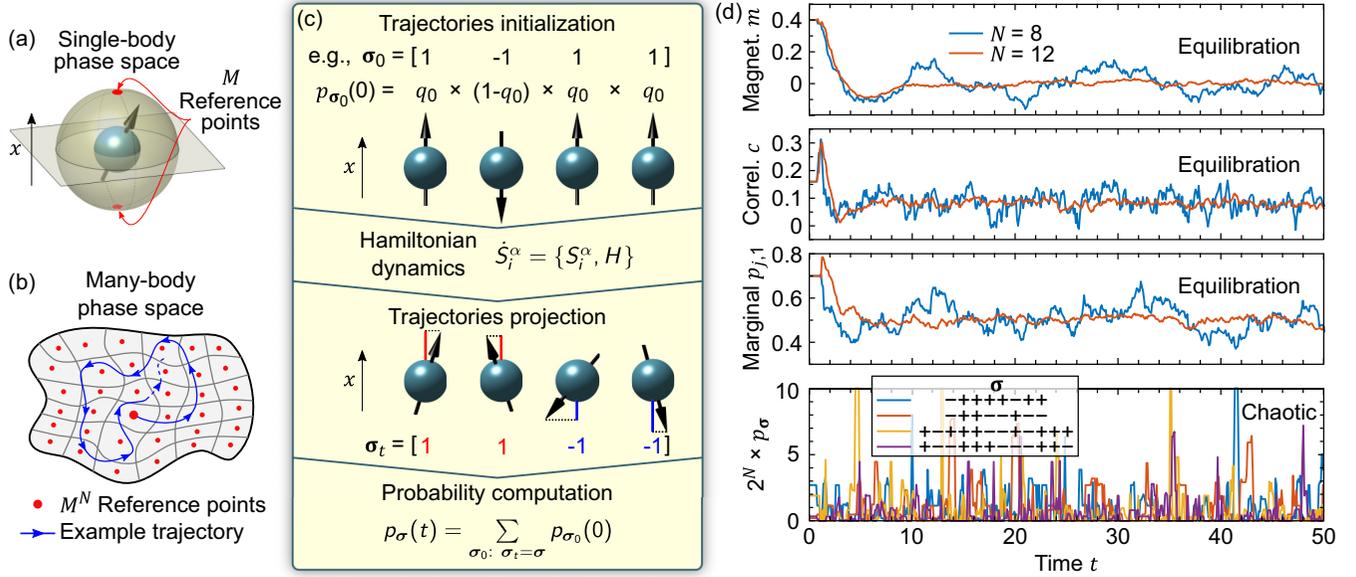


FIG. 5. Phase-space discretization for Hamiltonian dynamics. (a) Discretization of the single-body phase space of a classical spin. The sphere is divided into two hemispheres, each identified by a reference point (at the pole, in red). (b) Discretization of the many-body phase space, that is divided into  $M^N$  pockets, each of which is identified by a many-body reference point (red dots). A schematic trajectory starting from a reference point (larger red dot) is shown in blue. (c) Protocol for obtaining the dynamics of the discrete probability, in the illustrative case of  $N = 4$  spins. As initial conditions, we consider the  $2^N$  possible spin states in which each spin is either aligned or antialigned along  $x$  [corresponding to the reference points in (b)]. Initial states are labeled with a bit string  $\sigma$  and their associated probability  $p_{\sigma_0}$  factorizes. The trajectories are evolved under Hamiltonian dynamics, and their coordinates in the reduced phase space, obtained as  $\sigma_t = \text{sign}(S_1^x, \dots, S_N^x)$ , are used to compute the probability distribution  $p_\sigma(t)$ . (d) The ensemble averages of few-point observables, such as magnetization  $m$ , correlation  $c$ , and one-site marginal distribution  $p_{\sigma_1}(t)$ , equilibrate in time to the thermal value, up to temporal fluctuations that decrease with the system size  $N$ . By contrast, the many-body probability  $p_\sigma(t)$  itself does not equilibrate, independent of  $N$ . Here, we considered the spin model in Eq. (12), with parameters  $J = 1$ ,  $W = 2$ ,  $R = 1$ , and  $q_0 = 0.7$ .

First and foremost, we aim at verifying that the discretization protocol that we proposed leads to the desired key features of many-body dynamics, and in particular to the contrast between the equilibration of few-point observables and the nonequilibration of the probability itself, *à la* Liouville. In Fig. 5(c) we show, for increasing system sizes  $N$ , the dynamics of the magnetization  $m(t) = \frac{1}{N} \sum_{\sigma} p_{\sigma}(t) \sum_{j=1}^N \sigma_j$ , of a correlator  $c(t) = \frac{1}{N} \sum_{\sigma} p_{\sigma}(t) \sum_{j=1}^N \sigma_j \sigma_{j+1}$ , of a marginal probability  $p_{\sigma_1}(t)$ , and of the whole many-body probability  $p_{\sigma}(t)$  for some randomly chosen discretized states  $\sigma$ . Few-point observables and marginal probabilities quickly equilibrate, with asymptotic temporal fluctuations as a finite-size result and decreasing with system size  $N$  (even at a fixed and minimal discretization resolution  $M = 2$ ), whereas the many-body probability itself maintains (up to a factor of  $2^N$ , due to normalization) chaotic fluctuations in time, without clear dependence on the state index  $\sigma$ .

These key features being verified, we are now in the position to finally investigate information spreading in many-body Hamiltonian dynamics through the lens of MI and cEE, whose time traces we plot in Fig. 6 for various system sizes  $N$  and initial single-spin probability  $q_0$ . Initially vanishing, both the MI and cEE grow linearly in time, signaling the ballistic spreading of correlations, before saturating to an asymptotic value.

Let us first focus on the MI. At infinite temperature ( $q_0 = 0.5$ ), we find that the asymptotic value of the MI becomes

insensitive of  $N$  for  $N \gtrsim 10$  [see Fig. 6(a)]. The infinite-temperature asymptotic value of the MI is understood from Eq. (11). Assuming that any of the  $2^N$  trajectories can be found in any of the  $2^N$  pockets of the phase space with equal likelihood, we have that  $p_{\sigma} = 2^{-N} x_{\sigma}$ , with  $x_{\sigma}$  a random variable with binomial distribution of parameters  $2^N$  and  $2^{-N}$ . This converges, for  $N \rightarrow \infty$ , to a Poisson distribution of parameter  $\lambda = 1$ , i.e.,  $x_{\sigma} = k$  with probability  $\frac{1}{ek!}$ . From Eq. (11), we thus get that the asymptotic MI at infinite effective temperature is

$$I_{A:B}^{\infty} \approx \sum_{k=1}^{\infty} \frac{\log(k+1)}{ek!} \approx 0.5734, \quad (15)$$

that perfectly matches the numerics in Fig. 6(a). By contrast, at finite temperature ( $q_0 = 0.7$ ) the asymptotic MI is proportional to  $N$ , i.e., extensive [see Fig. 6(b)]. Our numerics thus confirms the analytical prediction in Eq. (10): the asymptotic MI fulfills area- and volume-law scaling at infinite and finite temperature, respectively.

As for the cEE, we instead find that both at finite and infinite temperature the asymptotic value fulfills volume-law scaling [see Figs. 6(c) and 6(d)]. Indeed, extensivity of the cEE is understood not dissimilarly from the quantum case: a wavefunction with randomized components (in the computational basis) leads to an extensive cEE [51,61]. In the classical case, the randomness of the classical wavefunction, in Eq. (5),

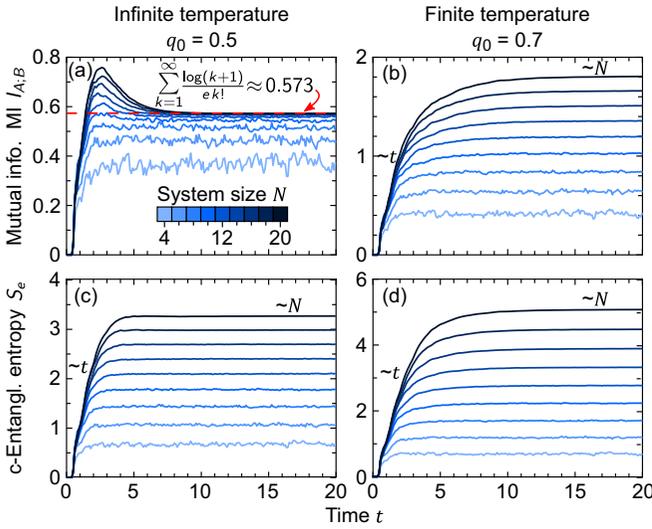


FIG. 6. Information dynamics in Hamiltonian many-body system. The MI (top) and the cEE (bottom) are plotted versus time for various system sizes  $N$ . Two representative single-spin initial probabilities are considered,  $q_0 = 0.5$  (left) and  $q_0 = 0.7$  (right), corresponding to infinite and finite effective temperatures, respectively. In all the considered cases, the quantities of interest initially grow linearly in time, signaling the ballistic spreading of correlations. At sufficiently long times, saturation to an asymptotic value sets in. (a) At infinite temperature, the asymptotic MI becomes insensitive to  $N$  for  $N \gtrsim 10$ , reaching the value predicted in Eq. (15) and fulfilling area-law scaling, while (b) at finite temperature it fulfills a volume-law scaling instead. [(c), (d)] On the other hand, the asymptotic cEE is extensive both at finite and infinite temperature. Here, we considered the Heisenberg chain in Eq. (12) for  $J = 1$ ,  $W = 3$ , and  $R = 50$ .

ultimately emerges as a result of chaos and incompressibility, which render the probability distribution highly structured, unlike a thermal one.

The different scaling of the asymptotic value of the MI and the cEE with system size at an infinite effective temperature is a major difference that we find between the two information measures. In fact, in the analog quantum case, pure states at an infinite effective temperature do result in volume-law scaling asymptotic EE [51], which highlights the necessity for the cEE to be introduced as the close classical counterpart of the EE, able to capture features that the classical MI misses.

We perform a more in-depth scaling analysis in Fig. 7. In Fig. 7(a) we plot  $I_{A:B}(t = \infty)$  and  $S_e(t = \infty)$  versus the single-spin initial probability  $q_0$ . As expected, in the trivial limits  $q_0 = 0, 1$ , corresponding to initially fully polarized spins and no dynamics, the MI and cEE remain zero at all times. For any other values of  $q_0$ , nontrivial dynamics occurs, information spreads, and the asymptotic MI and cEE take a finite value. However, while volume-law scaling is observed for the cEE at any  $0 < q_0 < 1$ , the MI is extensive only for  $q_0 \neq 0, 0.5, 1$ , whereas it fulfills area-law scaling at infinite temperature for  $q_0 = 0.5$ . Figure 7(b) helps to better appreciate the specific analytic form of the scaling, which is linear in system size  $N$ . With a linear fit we can extract the scaling coefficient  $\alpha$ , which we plot in Fig. 7(c) versus the single-spin initial probability  $q_0$ , providing a comprehensive confirmation of our assessments concerning the scaling of MI and cEE.

## VII. MANY-BODY INFORMATION IN PRACTICE: BRIDGING THE PROCEDURAL GAP

As we have shown, many of the key features of many-body information spreading are remarkably similar in the quantum

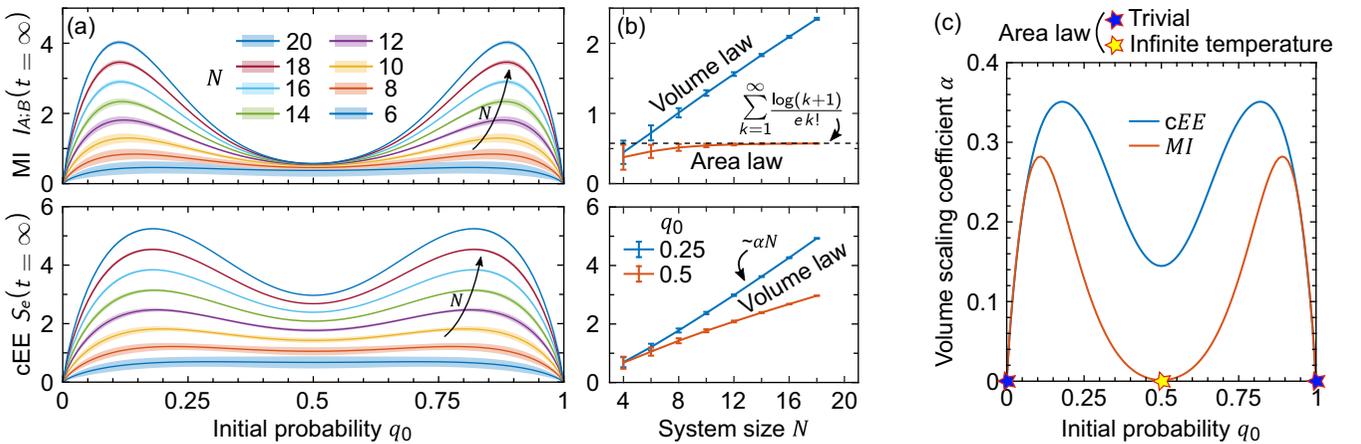


FIG. 7. Long-time scaling analysis in many-body Hamiltonian dynamics. (a) The asymptotic MI  $I_{A:B}(t = \infty)$  and cEE  $S_e(t = \infty)$  are plotted versus the single-spin initial probability  $q_0$  for various system sizes  $N$ . Solid lines and shaded bands represent mean and uncertainty (over the late time range  $20 < t < 30$  and disorder realizations). At  $q_0 = 0, 1$  the dynamics freezes and no information spreads, leaving  $I_{A:B} = S_e = 0$ . At infinite temperature,  $q_0 = 0.5$ , the asymptotic MI and cEE present local minima, but with a key difference: the former becomes independent of  $N$  for  $N \gtrsim 10$ , whereas the latter grows with  $N$ . For other values of  $q_0$ , both the asymptotic MI and cEE scale with  $N$ . (b) The scaling is better characterized for representative values of  $q_0$ , and found to be linear at large  $N$ ,  $\sim \alpha N$ . The infinite-temperature value of the MI predicted in Eq. (15) is reported for comparison. (c) The ultimate appreciation of the scaling properties is achieved by plotting the scaling coefficient  $\alpha$  versus the initial probability  $q_0$ , as obtained from a linear fit (of the three largest values of  $N$  only, to reduce small-size effects). Finite and vanishing values of  $\alpha$  denote volume- and area-law scaling, respectively. The latter occurs at the trivial points  $q_0 = 0, 1$  and, only for the MI, at infinite temperature  $q_0 = 0.5$ . Here, we used  $J = 1$ ,  $W = 3$ , and  $R = 50$ .

and classical settings. The two, however, are fundamentally different because the many-body wavefunction  $|\psi\rangle$  can be encoded in a physical system, whereas the classical many-body probability distribution cannot [75]. However, we argue here that this difference is of limited practical relevance when measuring or computing the EE, cEE, or MI.

To obtain the quantum EE, two routes can be taken. The first is to evolve the wavefunction on a classical computer, and use it to compute the EE. This requires exponentially large computational resources (memory and time). The second is to evolve the wavefunction on a physical quantum system, e.g., a quantum computer, which does not require exponentially large resources. However, computing the EE requires full tomography of the system and thus exponentially many runs [1]. Furthermore, once reconstructed, the wavefunction should again be stored on a classical computer, where with exponentially large resources one can then compute the EE.

The classical scenario is completely analogous, to the point that we can almost duplicate the previous paragraph: In the classical case, two routes can be taken to compute the dynamics of the MI (or cEE). The first is to evolve the probability distribution on a classical computer through Eq. (7), and use it to compute the MI (or cEE). This route requires exponentially large computational resources (memory and time). The second is to run the actual classical dynamics itself on a physical implementation of a cellular automaton. Extracting the many-body probability distribution requires configuration statistics over exponentially many runs. Furthermore, once reconstructed, the probability distribution should again be stored on a computer, where with exponentially large resources one can then compute the MI (or cEE).

This remarkably close procedural classical-quantum analogy continues when introducing measurements as in Sec. V, which again requires one of the two routes above. In particular, to obtain the EE in an experiment, the following would be necessary. A first run of the experiment would return a list of measurement outcomes  $\{m_j(t)\}$  and a physical state encoding the many-body wavefunction. To access the latter and compute the EE, however, one would need to run the experiments exponentially many times, making sure that the final to-be-measured wavefunction is always the same. That is, in the following runs one should not only repeat the measurements at the same spacetime locations, but also make sure that their outcome matches that of the first run,  $\{m_j(t)\}$ , which requires postselection. Indeed, this procedure perfectly matches the classical one described in Sec. V and Fig. 3(c).

## VIII. DISCUSSION AND CONCLUSION

In this work, we have shown that many of the celebrated features of quantum many-body information spreading can also appear in a classical setting, provided that one explicitly accounts for the whole exponentially large probability distribution, without making any assumption on its equilibration. Within this framework, we have shown that key phenomenology of quantum EE dynamics, namely, linear growth until saturation to an extensive value, also occurs in both the clas-

sical MI and the “classical EE” (cEE) defined above. These quantities both unambiguously quantify nonseparability, as they both vanish if and only if the two halves of the system  $A$  and  $B$  are statistically independent. The MI is a natural and widely recognized classical analog of the EE, whereas the cEE is a new and more original quantity that we introduce to measure nonseparability in the form of an entropy. We have studied the dynamics of both these objects in cellular automata and Hamiltonian systems. For the former, we highlighted time reversibility to play a key role in information spreading, and found the first instance of a classical MIPT, in direct correspondence with its quantum counterpart. As for Hamiltonian dynamics, we instead provided a simple expression for the asymptotic MI, which yields volume- and area-law scaling for states at an effective finite and infinite temperature, respectively. We verified this prediction numerically in the case of the Heisenberg chain, and showed it is in contrast to the cEE, whose asymptotic value is extensive even at infinite effective temperature, like in the quantum case. Finally, we argued that the study of classical and quantum information spreading requires in practice analogous procedures, involving exponentially large resources in both cases.

Since the EE is defined for quantum systems, one may naively assume that effects that are described or measured through it are also purely quantum. Of course, the interplay of carefully engineered interference effects and many-body entanglement can lead to striking quantum effects, for instance enabling quantum computing [1]. By contrast, the entanglement generated in the chaotic dynamics relevant in the context of thermalization is generally “uncontrolled.” Rather than allowing to efficiently perform a computational task, such entanglement requires exponentially large resources or many runs for its characterization, effectively resulting in physics that is much more classical than one could have expected. Indeed, here we have shown that a whole range of dynamical behaviors of the EE also emerges in the context of classical correlation spreading.

The idea of sampling and evolving many classical initial conditions adopted here is reminiscent of phase-space methods like the truncated Wigner approximation (TWA) [76–80]. The mission of the latter is to efficiently simulate the dynamics of quantum many-body systems by evolving an ensemble of classical trajectories, for which the sampling of the initial conditions is constructed to most faithfully mimic quantum fluctuations. A classical-quantum link for EE has also been established for quantum systems with a well-defined classical limit, for which the EE dynamics can be understood in relation to the presence (or lack) of underlying classical chaos [81].

The philosophy of our work contrasts with that mission: while we got inspired by the machinery of quantum mechanics to frame the classical problem on par with the quantum one, our goal was not to construct a classical model that could quantitatively emulate quantum physics, nor to link the classical and quantum descriptions in the semiclassical limit. Instead, we aimed at showing that the main qualitative features of many-body information spreading are analogous in the classical and quantum settings. If the classical-quantum quantitative agreement of few-point observables in the TWA is by construction, in the sense that it requires devising

a classical protocol that emulates the quantum one, the classical-quantum qualitative agreement of multipoint information measures in this work is intrinsic, in the sense that it just happens, from the most natural setting of initially independent local fluctuations that become nonlocally correlated through dynamics. In light of this philosophy, our classical treatment did not aim at being a simple and efficient approximation of quantum mechanics, but conversely at showing that classical and quantum information spreading can be equally complex and rich.

Looking forward, our work opens a broad spectrum of research questions. Indeed, within the framework and with the tools that we have outlined here, the study of classical information spreading could be as fruitful as it has been in the quantum case. For instance, a very interesting question regards the role of strong disorder in classical information spreading. In the quantum case, disorder can lead to MBL, which is characterized by a peculiar slow growth  $\sim \log t$  of the EE, until reaching asymptotic extensivity [12]. In the classical case MBL is not possible but transport is nonetheless strongly suppressed [82], and a slow growth of the cEE is expected. The precise characterization of the functional form of such a growth is certainly worth studying. On a related line, further research should investigate what role the counterparts of integrability [83,84] and conserved quantities [49] play in classical information spreading. The cEE could then shed new light on classical prethermalization [85–88] and prethermal discrete-time crystals [89–91]. As well, it is now natural to expect that the research thread on quantum MIPT [26–35] would carry over to the classical realm, in which the study of different kinds of automata, initial conditions, and measurement protocols would be worthwhile.

The cEE introduced here is the EE associated to the classical wavefunction  $|\psi\rangle$  in Eq. (5). As a result, there exists a quantum monitored circuit that is equivalent to the quantum circuit we have studied (this is true for any discrete reversible Markov chain). It is therefore natural to ask about other Rényi entropies. Interestingly, those behave markedly differently for complex and real wavefunctions [61], which suggests clear limits to the analogy between quantum and classical infor-

mation spreading. Delineating this further would be a fruitful topic for future research.

The many-body probabilistic framework that we established could then be used to firmly address the question of what the strict classical counterparts of other information-related quantities from current quantum research are. For instance, it could be used to define a classical out-of-time order correlator (OTOC) going beyond the decorrelator in Refs. [43,47] (based on just two copies of the system, rather than on a probability distribution over exponentially many of them).

Finally, on a higher and open-ended level, our work opens the way to quantum-inspired advances in physics, information theory, and statistics. Indeed, while quantum EE can only be applied to pure quantum systems, the cEE that we introduced in Eq. (4) has potentially a very broad range of applicability. Beyond the information spreading considered here, it is suited for *any* probability distribution. Indeed, the cEE can capture some features that the MI misses (e.g., here, extensivity at infinite temperature), and could, thus, prove a powerful quantum-inspired tool to characterize new aspects of classical correlations.

#### ACKNOWLEDGMENTS

We thank C. Castelnovo, S. Choi, P. Claeys, A. G. Green, M. Hofstetter, I. Kaminer, A. Lamacraft, L. Pacchiardi, and N. Y. Yao for insightful discussions related to this work. We acknowledge support from the Imperial-TUM flagship partnership. A.P. acknowledges support from the Royal Society and hospitality at TUM. D.M. acknowledges funding from ERC Advanced Grant QUENOCOBA under the EU Horizon 2020 program (Grant Agreement No. 742102). This research is part of the Munich Quantum Valley, which is supported by the Bavarian state government with funds from the Hightech Agenda Bayern Plus.

A.P. conceived the study, performed the calculations, and drafted the manuscript. All the authors contributed with continued discussion of ideas and results, and with the finalization of the writing.

- 
- [1] M. A. Nielsen and I. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, UK, 2002).
  - [2] I. M. Georgescu, S. Ashhab, and F. Nori, *Rev. Mod. Phys.* **86**, 153 (2014).
  - [3] E. Altman, K. R. Brown, G. Carleo, L. D. Carr, E. Demler, C. Chin, B. DeMarco, S. E. Economou, M. A. Eriksson, K.-M. C. Fu, M. Greiner, K. R. A. Hazzard, R. G. Hulet, A. J. Kollar, B. L. Lev, M. D. Lukin, R. Ma, X. Mi, S. Misra, C. Monroe, K. Murch, Z. Nazario, K. K. Ni, A. C. Potter, P. Roushan, M. Saffman, M. Schleier-Smith, I. Siddiqi, R. Simmonds, M. Singh, I. B. Spielman, K. Temme, D. S. Weiss, J. Vuckovic, V. Vuletic, J. Ye, and M. Zwierlein, *PRX Quantum* **2**, 017003 (2021).
  - [4] A. Polkovnikov, K. Sengupta, A. Silva, and M. Vengalattore, *Rev. Mod. Phys.* **83**, 863 (2011).
  - [5] C. Gogolin and J. Eisert, *Rep. Prog. Phys.* **79**, 056001 (2016).
  - [6] A. Browaeys and T. Lahaye, *Nat. Phys.* **16**, 132 (2020).
  - [7] C. Gross and I. Bloch, *Science* **357**, 995 (2017).
  - [8] P. Calabrese and J. Cardy, *J. Stat. Mech.: Theory Exp.* (2005) P04010.
  - [9] G. De Chiara, S. Montangero, P. Calabrese, and R. Fazio, *J. Stat. Mech.: Theory Exp.* (2006) P03001.
  - [10] L. Amico, R. Fazio, A. Osterloh, and V. Vedral, *Rev. Mod. Phys.* **80**, 517 (2008).
  - [11] P. Calabrese and J. Cardy, *J. Phys. A: Math. Theor.* **42**, 504005 (2009).
  - [12] J. H. Bardarson, F. Pollmann, and J. E. Moore, *Phys. Rev. Lett.* **109**, 017202 (2012).
  - [13] H. Kim and D. A. Huse, *Phys. Rev. Lett.* **111**, 127205 (2013).
  - [14] W. W. Ho and D. A. Abanin, *Phys. Rev. B* **95**, 094302 (2017).

- [15] A. Nahum, J. Ruhman, S. Vijay, and J. Haah, *Phys. Rev. X* **7**, 031016 (2017).
- [16] A. Nahum, S. Vijay, and J. Haah, *Phys. Rev. X* **8**, 021014 (2018).
- [17] V. Khemani, A. Vishwanath, and D. A. Huse, *Phys. Rev. X* **8**, 031057 (2018).
- [18] C. W. von Keyserlingk, T. Rakovszky, F. Pollmann, and S. L. Sondhi, *Phys. Rev. X* **8**, 021013 (2018).
- [19] S. Gopalakrishnan, D. A. Huse, V. Khemani, and R. Vasseur, *Phys. Rev. B* **98**, 220303(R) (2018).
- [20] V. Alba and P. Calabrese, *SciPost Phys.* **4**, 017 (2018).
- [21] D. A. Abanin, E. Altman, I. Bloch, and M. Serbyn, *Rev. Mod. Phys.* **91**, 021001 (2019).
- [22] T. Zhou and A. Nahum, *Phys. Rev. B* **99**, 174205 (2019).
- [23] B. Bertini, P. Kos, and T. Prosen, *Phys. Rev. X* **9**, 021033 (2019).
- [24] M. Ippoliti, T. Rakovszky, and V. Khemani, *Phys. Rev. X* **12**, 011045 (2022).
- [25] M. Serbyn, Z. Papić, and D. A. Abanin, *Phys. Rev. Lett.* **110**, 260601 (2013).
- [26] B. Skinner, J. Ruhman, and A. Nahum, *Phys. Rev. X* **9**, 031009 (2019).
- [27] Y. Li, X. Chen, and M. P. A. Fisher, *Phys. Rev. B* **100**, 134306 (2019).
- [28] Y. Bao, S. Choi, and E. Altman, *Phys. Rev. B* **101**, 104301 (2020).
- [29] S. Choi, Y. Bao, X.-L. Qi, and E. Altman, *Phys. Rev. Lett.* **125**, 030505 (2020).
- [30] Q. Tang and W. Zhu, *Phys. Rev. Res.* **2**, 013022 (2020).
- [31] C.-M. Jian, Y.-Z. You, R. Vasseur, and A. W. W. Ludwig, *Phys. Rev. B* **101**, 104302 (2020).
- [32] A. Zabalo, M. J. Gullans, J. H. Wilson, S. Gopalakrishnan, D. A. Huse, and J. H. Pixley, *Phys. Rev. B* **101**, 060301(R) (2020).
- [33] A. Nahum, S. Roy, B. Skinner, and J. Ruhman, *PRX Quantum* **2**, 010352 (2021).
- [34] M. Ippoliti, M. J. Gullans, S. Gopalakrishnan, D. A. Huse, and V. Khemani, *Phys. Rev. X* **11**, 011030 (2021).
- [35] M. Block, Y. Bao, S. Choi, E. Altman, and N. Y. Yao, *Phys. Rev. Lett.* **128**, 010604 (2022).
- [36] U. Schollwöck, *Ann. Phys.* **326**, 96 (2011).
- [37] A. Weiß and H. Fehske, *Computational Many-Particle Physics* (Springer, Berlin, 2008), pp. 529–544.
- [38] P. Weinberg and M. Bukov, *SciPost Phys.* **2**, 003 (2017).
- [39] J. A. Vastano and H. L. Swinney, *Phys. Rev. Lett.* **60**, 1773 (1988).
- [40] S. Lepri, A. Politi, and A. Torcini, *J. Stat. Phys.* **82**, 1429 (1996).
- [41] G. Giacomelli, R. Hegger, A. Politi, and M. Vassalli, *Phys. Rev. Lett.* **85**, 3616 (2000).
- [42] A. Das, S. Chakrabarty, A. Dhar, A. Kundu, D. A. Huse, R. Moessner, S. S. Ray, and S. Bhattacharjee, *Phys. Rev. Lett.* **121**, 024101 (2018).
- [43] T. Bilitewski, S. Bhattacharjee, and R. Moessner, *Phys. Rev. Lett.* **121**, 250602 (2018).
- [44] V. Khemani, D. A. Huse, and A. Nahum, *Phys. Rev. B* **98**, 144304 (2018).
- [45] M. Kumar, A. Kundu, M. Kulkarni, D. A. Huse, and A. Dhar, *Phys. Rev. E* **102**, 022130 (2020).
- [46] T. Bilitewski, S. Bhattacharjee, and R. Moessner, *Phys. Rev. B* **103**, 174302 (2021).
- [47] S.-W. Liu, J. Willsher, T. Bilitewski, J.-J. Li, A. Smith, K. Christensen, R. Moessner, and J. Knolle, *Phys. Rev. B* **103**, 094109 (2021).
- [48] A. J. McRoberts, T. Bilitewski, M. Haque, and R. Moessner, *Phys. Rev. B* **105**, L100403 (2022).
- [49] A. Deger, S. Roy, and A. Lazarides, *Phys. Rev. Lett.* **129**, 160601 (2022).
- [50] M. M. Wolf, F. Verstraete, M. B. Hastings, and J. I. Cirac, *Phys. Rev. Lett.* **100**, 070502 (2008).
- [51] D. N. Page, *Phys. Rev. Lett.* **71**, 1291 (1993).
- [52] M. Kardar, *Statistical Physics of Particles* (Cambridge University Press, Cambridge, UK, 2007).
- [53] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, *Wiley Series in Telecommunications and Signal Processing* (Wiley-Interscience, New York, 2006).
- [54] The reduced density matrix can be expressed as  $\rho_A = (P^{(1/2)})^T P^{(1/2)}$ , where  $P^{(1/2)}$  is a “squared probability matrix” with entries  $P_{A,B}^{(1/2)} = \sqrt{p_{A,B}}$ . Given a vector  $z$ , we have that  $z^T \rho_A z = z^T (P^{(1/2)})^T P^{(1/2)} z$ . Calling  $y = P^{(1/2)} z$ , we therefore have that  $z^T \rho_A z = y^T y \geq 0$ , that is, that  $\rho_A$  is positive semidefinite.
- [55] C. Wetterich, *Nucl. Phys. B* **917**, 241 (2017).
- [56] C. Wetterich, [arXiv:2011.02867](https://arxiv.org/abs/2011.02867).
- [57] J. Iaconis, S. Vijay, and R. Nandkishore, *Phys. Rev. B* **100**, 214301 (2019).
- [58] K. Klobas, B. Bertini, and L. Piroli, *Phys. Rev. Lett.* **126**, 160602 (2021).
- [59] V. Alba, J. Dubail, and M. Medenjak, *Phys. Rev. Lett.* **122**, 250603 (2019).
- [60] K. Klobas and B. Bertini, *SciPost Phys.* **11**, 107 (2021).
- [61] T. Grover and M. P. A. Fisher, *Phys. Rev. A* **92**, 042308 (2015).
- [62] Strictly speaking, the classical system is described by a continuous probability density on phase space. This will be discussed later in Sec. VI, but should not disturb the reading now.
- [63] J. M. Deutsch, *Phys. Rev. A* **43**, 2046 (1991).
- [64] M. Srednicki, *Phys. Rev. E* **50**, 888 (1994).
- [65] M. Rigol, V. Dunjko, and M. Olshanii, *Nature (London)* **452**, 854 (2008).
- [66] J. M. Deutsch, *Rep. Prog. Phys.* **81**, 082001 (2018).
- [67] Y. Y. Atas, E. Bogomolny, O. Giraud, and G. Roux, *Phys. Rev. Lett.* **110**, 084101 (2013).
- [68] Note that only the MI can be used to make this point, because the EE is only defined for a pure state. Consider the limit case of an infinite-temperature thermal state. The density matrix of the system is the identity matrix, and the von Neumann entropies  $S_A$  and  $S_B$  of its subparts are thus extensive. This, however, should not be interpreted as a success of the thermal distribution to capture the extensivity of the EE. In fact,  $A$  and  $B$  in an infinite-temperature state are completely independent, and  $S_A$  does not carry any information on the EE.
- [69] S. Wolfram, *A New Kind of Science* (Wolfram Media Inc., Champaign, IL, 2002).
- [70] M. Aldana, S. Coppersmith, and L. P. Kadanoff, *Perspectives and Problems in Nonlinear Science* (Springer, Berlin, 2003), pp. 23–89.
- [71] Note that to simplify numerics one could alternatively consider, also for second-order cellular automata like Rule 30R, a shorter probability distribution defined on the  $2^N$  single-time  $N$ -bit strings only. This would correspond to a phase-space

- reduction, which is sensible provided that only  $2^N$  possible initial conditions are considered [e.g., by setting  $s_j(t = -1) = 0$  while allowing for random  $s_j(t = 0)$ ]. This idea will become clearer in Sec. VI for Hamiltonian dynamics, where such a reduction is necessary for phase-space discretization purposes. While this is possible, we have adopted a cellular automaton here precisely so we do not have to deal with any phase-space reduction.
- [72] J. Willsher, S.-W. Liu, R. Moessner, and J. Knolle, *Phys. Rev. B* **106**, 024305 (2022).
- [73] The continuous marginal  $p_A(A) = \int dB p(A, B)$  can be split as a sum of integrals over small volumes  $\{\mathcal{B}_n\}$  of phase space at constant  $A$ ,  $p_A(A) = \sum_n \int_{\mathcal{B}_n} dB p(A, B)$ . One can increase the number of volumes  $\mathcal{B}_n$  to make their volume small enough, so that the energy  $H$  within each of them can be considered constant (that is, with variation small compared to the involved energy scales). The probability distribution  $p(A, B)$  fluctuates over a phase-space length scale  $\sim e^{-\lambda t}$  that becomes arbitrarily small and eventually much smaller than that of  $\mathcal{B}_n$ . Having arbitrarily many fluctuations within  $\mathcal{B}_n$ , we expect the probability to self-average, meaning that  $\int_{\mathcal{B}_n} dB p(A, B) \approx \int_{\mathcal{B}_n} dB p_{th}(A, B)$ . Rebuilding the integral one thus gets  $p_A(A) = \int dB p_{th}(A, B)$ , that is, that the marginal thermalizes.
- [74] Strictly speaking, the components of the probability are not independent, because they are tied by the normalization  $\sum_\sigma p_\sigma$ . However, for  $N \rightarrow \infty$ , that is, for a diverging number of possible microstates, the normalization constraint becomes weaker and weaker, and the components of the probability distribution can be considered as effectively i.i.d.
- [75] As discussed in Sec. IV, the probability distribution is implicitly evolved by the system as long as the state of the system is not observed. This is in analogy with the quantum case, in which the evolution of the wavefunction proceeds undisturbed as long as one does not measure the system. The two cases, however, differ in that the quantum wavefunction is composed of a superposition of states, which is physical rather than reflecting our ignorance on the system.
- [76] A. Polkovnikov, S. Sachdev, and S. M. Girvin, *Phys. Rev. A* **66**, 053607 (2002).
- [77] A. Sinatra, C. Lobo, and Y. Castin, *J. Phys. B: At. Mol. Opt. Phys.* **35**, 3599 (2002).
- [78] A. Polkovnikov, *Phys. Rev. A* **68**, 053604 (2003).
- [79] A. Polkovnikov, *Ann. Phys.* **325**, 1790 (2010).
- [80] J. Schachenmayer, A. Pikovski, and A. M. Rey, *Phys. Rev. X* **5**, 011022 (2015).
- [81] A. Lerose and S. Pappalardi, *Phys. Rev. A* **102**, 032404 (2020).
- [82] V. Oganesyan, A. Pal, and D. A. Huse, *Phys. Rev. B* **80**, 115104 (2009).
- [83] B. Buča, K. Klobas, and T. Prosen, *J. Stat. Mech.: Theory Exp.* (2021) 074001.
- [84] J. W. P. Wilkinson, T. Prosen, and J. P. Garrahan, *Phys. Rev. E* **105**, 034124 (2022).
- [85] A. Rajak, R. Citro, and E. G. Dalla Torre, *J. Phys. A: Math. Theor.* **51**, 465001 (2018).
- [86] T. Mori, *Phys. Rev. B* **98**, 104303 (2018).
- [87] A. Rajak, I. Dana, and E. G. Dalla Torre, *Phys. Rev. B* **100**, 100302(R) (2019).
- [88] O. Howell, P. Weinberg, D. Sels, A. Polkovnikov, and M. Bukov, *Phys. Rev. Lett.* **122**, 010602 (2019).
- [89] A. Pizzi, A. Nunnenkamp, and J. Knolle, *Phys. Rev. Lett.* **127**, 140602 (2021).
- [90] A. Pizzi, A. Nunnenkamp, and J. Knolle, *Phys. Rev. B* **104**, 094308 (2021).
- [91] B. Ye, F. Machado, and N. Y. Yao, *Phys. Rev. Lett.* **127**, 140603 (2021).