Thermodynamic analog of integrate-and-fire neuronal networks by maximum entropy modelling

T. S. A. N. Simões[1], C. I. N. Sampaio Filho[2], H. J. Herrmann[2,3], J. S. Andrade Jr.[2], L. de Arcangelis[1]

1 Department of Mathematics and Physics, University of Campania "Luigi Vanvitelli", Viale Lincoln, 5, 81100 Caserta, Italy
2 Departamento de Física, Universidade Federal do Ceará, 60451-970, Fortaleza, Ceará, Brazil
3 PMMH, ESPCI, 7 quai St. Bernard, 75005 Paris, France

Abstract

Relying on maximum entropy arguments, certain aspects of time-averaged experimental neuronal data have been recently described using Ising-like models, allowing the study of neuronal networks under an analogous thermodynamical framework. Here, we apply for the first time the Maximum Entropy method to an Integrate-and-fire (IF) model that can be tuned at criticality, offering a controlled setting for a systematic study of criticality and finite-size effects in spontaneous neuronal activity, as opposed to experiments. We show that generalized Ising models that accurately predict the average local activities and correlation functions between neurons of the IF model networks in the critical state exhibit a spin glass phase for low temperatures, having mostly negative intrinsic fields and a bimodal distribution of interaction constants that tends to become unimodal for larger networks. Results appear to be affected by sample-to-sample connectivity variations and subsampling. Furthermore, we also found that networks with higher percentage of inhibitory neurons lead to Ising-like systems with reduced thermal fluctuations. Finally, considering only neuronal pairs associated with the largest correlation functions allows the study of larger system sizes.

Author summary

Brain activity, either stimulated or spontaneous, *in vivo* or *in vitro*, exhibits complex spatiotemporal behavior. Trying to make sense of it, several research groups have analyzed time-averaged experimental neuronal data using maximum entropy arguments, mapping the neuronal dynamics into a generalized Ising-like model and allowing to study neuronal data using tools typical of critical phenomena. However, the intricacy of real biological networks in experimental settings pose challenges in the precision and reliability of the neuronal measurements. Here, we apply for the first time the Maximum Entropy Method to an Integrate-and-fire model with synaptic plasticity, providing a foundation for a more systematic and comprehensive study of spontaneous brain activity. We show that generalized Ising models are able to reproduce the numerical time-averaged data of local activities and correlation functions of integrate-and-fire neurons and predict qualitatively higher-order quantities such as the three-point correlation functions across triplets of neurons. We show that subsampling affects the efficiency of the mapping and that the analogous thermodynamics functions

of the Ising-like models depend on sample-to-sample network variations and on the presence of inhibition in the neural network.

Introduction

Biological neural networks are highly complex systems [1], due to the large number of interacting degrees of freedom and the connectivity properties of its constituent elements. Neurons interact by generating action potentials ("firing" or "spiking"), with a biophysical mechanism first explained quantitatively by Hodgkin and Huxley in 1952 [2]. Their dynamics can be characterized by discretized time series of these firing patterns in binary notation [3, 4]. This description suggests an analogy with the binary spin states in Ising models. Indeed, Ising-like descriptions of brain activity can be traced back to fifty years ago, to the work by Little [5] and Hopfield [6], followed by the work of Amit et al. [7], where the authors used a spin glass model to describe neural networks. More recently [8–15], generalized Ising models have been used to describe the dynamics of recordings from stimulated neuronal activity, using the so-called Maximum Entropy Modelling (MEM) method [16]. The method consists in finding the least biased (or maximum entropy) probability distribution that is consistent with a given set of statistical measurements from the system under consideration but otherwise imposes no further constraints [17]. When this method is applied to describe the individual firing rates and correlation functions between neurons, the resulting statistical description is equivalent to the one of a specific Ising model with frustrated spins [8, 9]. Therefore, the MEM approach can be interpreted as an effective mapping of the dynamics from an out-of-equilibrium, multi-component system into a "Hamiltonian" one [11], specified by relatively few parameters when compared to the full set of possible states of the system in question, allowing to apply the framework of thermodynamics. Within the scope of neuroscience, MEM has been mostly used to analyse experimental data of neuronal recordings, such as visual inputs from cells of a salamander [8, 11, 12, 18–20] and rat retina [21], numerical simulations from a phenomenological model of retinal ganglion cells [22], responses from hippocampus place cells in rodent brains [13, 15], synchronized and desynchronized neuronal activity from the primary visual cortex of anaesthetised cats and awake monkeys [23], *in vivo* and *in vitro* neuronal activity from cortical tissue of rodents [24] and the nervous system of the nematod *C. Elegans* [14], an organism whose pattern of connectivity between all its 302 neurons is well known [25]. Overall, these studies have given insights regarding collective [8] and functional characteristics [19] of biological neural networks, as well as helping unraveling the information content of neuronal responses [11, 12, 26] and its analogous thermodynamic properties [12]. Particularly, it was shown that the thermal fluctuations present in these Ising systems tend to diverge with system size [9, 12, 14, 23, 24], a sign of critical behaviour. However, experimental measurements in real networks can pose challenges that may impact the relevance and range of applicability of the conclusions drawn using the MEM method. For instance, contemporary neuronal recording techniques are restricted by the duration and sampling rate of the recordings [20, 27] and a precise association between which neuron generated which spike is an open problem, known as "spike sorting" [28], possibly affecting the accuracy of the measured correlation functions [29]. Furthermore, precise estimations of the fraction of excitatory and inhibitory neuronal populations in real systems are also often difficult to obtain [30]. To mitigate these challenges, numerical models can offer a more controlled environment for studying neuronal activity, allowing one to systematically change many parameters of choice, tuning the activity state, make systematic size studies which do not rely on subsampling, and generate many equivalent independent samples, thus producing statistically relevant data.

In this context, we apply the MEM method to describe spontaneous neuronal activity generated by an Integrate-and-fire (IF) model implemented on a scale-free network with short- and long-term plasticity [31]. The IF model relies on a tuning parameter that regulates the dynamical state of the system [31, 32]. At appropriate values of the tuning parameter, the model generates bursts of firing activity with characteristic spatio-temporal statistics, known as neuronal avalanches, first observed in 2003, in acute slices of rat cortex [33], displaying sizes S and durations D that are power-law distributed, as $P(S) \propto S^{-1.5}$ and $P(D) \propto D^{-2}$, suggesting that biological neural networks operate near a critical point, as proposed in previous numerical and analytical studies [34–36], and is consistent with the divergence of thermal energy fluctuations observed in relatively small Ising-like models constructed from experimental neuronal data using the MEM method [9, 12, 14, 23]. Results suggest that the brain might act between a quiescence-like state and a state of hyperactivity, possibly offering several biological advantages such as optimal information transmission and storage [37]. Signs of criticality have been observed in a wide variety of animals, including humans [38–41], monkeys [41, 42], cats [23, 43], salamanders [44], turtles [45], worms [14, 46] and fish [47], a hallmark of universality [34].

Our aim here is to construct, using the MEM method mentioned above, pairwise, fully-connected Ising models of frustrated spins with local fields that replicate the average local activities and correlation functions between the neurons of the IF model in the critical state. Then we investigate spontaneous neuronal activity under a thermodynamic framework using our biologically relevant numerical IF model. In particular, we will investigate how the parameters of the Ising model and its thermodynamics properties change with the size and connectivity of the neuronal network. Additionally, we examine how different fractions of inhibitory neurons may influence the associated Ising models. We also consider partially-connected spin networks with only a subset of the total pairwise interactions, associated with the strongest neuronal pairwise correlations, to be able to analyse networks of sizes larger than the ones typically considered so far in the literature [27].

The following section is organized in 6 subsections. The first one analyses the firing dynamics generated by the IF model. The subsequent ones present the results from the mapping of the IF model into an Ising-like model using the MEM method and a study of its associated properties. In the penultimate section, we present a general discussion of the results and conclusion. In the final section, we describe the details of the IF model implementation and the MEM method.

Results

We consider scale-free neural networks of different size $N \in [20, 500]$ with short- and long-term plasticity, as well as a refractory mechanism, where neurons remain inactive for a single timestep immediately after firing (see Methods). Together with plastic adaptation, the refractory mechanism has been shown to play a crucial role in the appearance of critical behaviour and in shaping the topology of the network [48]. All measurements are performed on networks in the critical state. This is done by appropriately tuning the neurotransmitter recovery parameter δu_{rec} (see "Integrate-and-fire model" in Methods) and verifying that the distributions of neuronal avalanche sizes $P(S) \propto S^{-\tau_S}$ and durations $P(D) \propto D^{-\tau_D}$ decay as power laws with exponents $\tau_S = 1.5$ and $\tau_D = 2$ (see S1 Fig and S2 Fig), characteristic of the mean field self-organized branching process [49], and consistent with both numerical and experimental observations [31, 33, 40, 50]. Initially, we consider only fully-excitatory networks. Later, we will also consider networks with a fraction $p_{\mathrm{in}} > 0$ of inhibitory neurons. The time evolution of a system with $N = 120$ neurons in the critical state is

shown in the raster plot of Fig 1.

First, we analyse the firing statistics of the IF model networks. Then, we apply the MEM method to map the dynamics of the IF model networks of size $N \in [20, 120]$ into a pairwise Ising model, defined by N "local fields" h_i and $N \cdot (N-1)/2$ "interaction constants" J_{ij}, and study its thermodynamical properties. More specifically, this mapping is achieved by using the so-called Boltzmann Machine (BM) algorithm. This learning algorithm searches for the set of parameters h_i and J_{ij} that best fit the pairwise Ising model to the data of the IF model by comparison with the average local activities $\langle \sigma_i \rangle$ and correlation functions C_{ij} generated by each model (see "Maximum Entropy Modelling" in Methods).

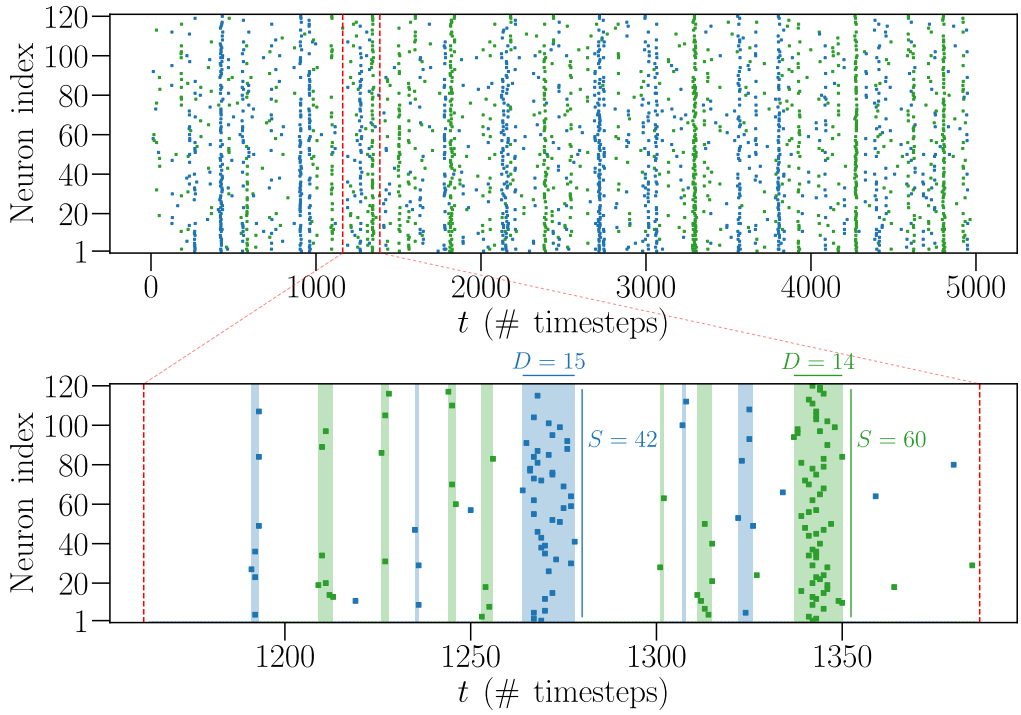


Fig 1. Time series for a system at criticality with $N = 120$ neurons. The top raster plot presents the time evolution of the firing states of $N = 120$ neurons over a certain number of timesteps, from a fully-excitatory IF network in the critical state. A coloured dot indicates that the respective neuron fired during that timestep, while an absence of a dot means it was inactive. The alternating green and blue colours are just a visual aid to distinguish between different avalanches. The bottom raster plot is a zoom-in of the top one, where the coloured shaded areas indicate avalanches of size $S > 1$ and duration D.

Firing statistics

To measure the IF network firing statistics, we consider a time bin $\Delta t_b = 5$ timesteps. We assign a binary value $\sigma_i^k \in \{-1, 1\}$ to each neuron at each time bin according to the following rule: $\sigma_i^k = 1$ if neuron i fired at least once during the k-th time bin, $\sigma_i^k = -1$ if the neuron i was inactive during the whole time bin k. We then calculate the average local activity $\langle \sigma_i \rangle$ of the N neurons over N_b time bins, defined as

$$\langle \sigma_i \rangle = \frac{1}{N_b} \sum_{k=1}^{N_b} \sigma_i^k \ , \tag{1}$$

as well as the average two-point activity $\langle \sigma_i \sigma_j \rangle$ between all $N \cdot (N-1)/2$ pairs of neurons i and j

$$\langle \sigma_i \sigma_j \rangle = \frac{1}{N_b} \sum_{k=1}^{N_b} \sigma_i^k \sigma_j^k \ . \tag{2}$$

From these quantities we can also define the two-point correlation functions C_{ij},

$$C_{ij} = \langle (\sigma_i - \langle \sigma_i \rangle) \cdot (\sigma_j - \langle \sigma_j \rangle) \rangle = \langle \sigma_i \sigma_j \rangle - \langle \sigma_i \rangle \langle \sigma_j \rangle \ , \tag{3}$$

where $\langle \sigma_i \rangle$ is closely related to the firing rate $r_i = (\langle \sigma_i \rangle + 1)/2\Delta t_b$ of neuron i [11], whereas C_{ij} quantifies the tendency for neurons i and j to fire together in the time interval Δt_b. Notice that, although likely related, the C_{ij} are completely distinct from the synaptic strengths w_{ij}. For instance, the C_{ij} are defined for each pair i and j and are symmetric by definition ($C_{ij} = C_{ji}$), whereas w_{ij} are not. For all simulations we use $N_b = 10^7$.

In Fig 2 we show the distributions of $\langle \sigma_i \rangle$ and C_{ij} obtained from N_c different network configurations (see S1 Fig for the respective avalanche statistics), for systems with different number of neurons N. We vary N_c according to N so that the total number of neurons considered $NN_c \sim 10000$. The average local activity $\langle \sigma_i \rangle$ is negative for all neurons (Fig 2a-c), implying that firing is a relatively rare event when considering individual neurons. In Fig 2d-f, we show that the correlation functions are small, but mostly non-zero for all neuron pairs, with their distributions peaked near zero, becoming sharper as the system size N increases. It is also noteworthy that, in experimental studies of the vertebrate retina [8], even weak pairwise correlations have been shown to likely become statistically significant in large networks, influencing the dynamics at the scale of the whole network. Another useful quantity that characterizes the firing dynamics is the probability $P(K)$ to observe $K \in [0, N]$ neurons firing

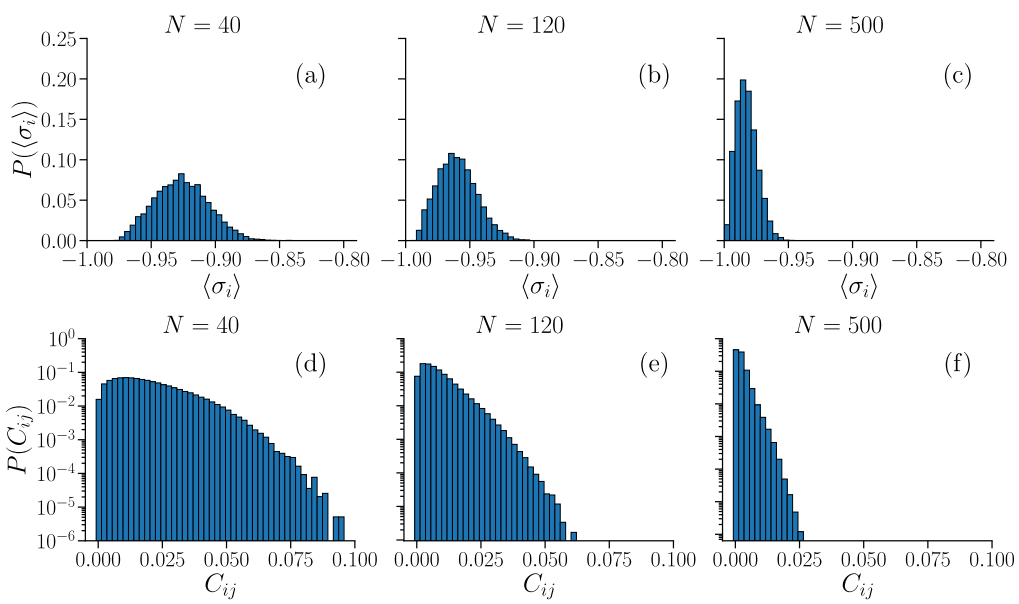


Fig 2. Distributions of the average local activity $\langle \sigma_i \rangle$ and of the correlation functions C_{ij} in IF networks. Average local activity $\langle \sigma_i \rangle$ (a-c) and correlation functions C_{ij} (d-f), calculated as averages over $N_b = 10^7$ time bins. Distributions are obtained for sizes $N = \{40, 120, 500\}$, for $N_c \sim 10000/N$ different fully-excitatory networks.

simultaneously within a time bin Δt_b,

$$P(K) = \frac{1}{N_b} \sum_{k=1}^{N_b} \delta_{K,K^k} \ , \tag{4}$$

where δ_{K,K^k} is the delta Kronecker function and $K^k = \sum_i^N \left(\sigma_i^k + 1\right)/2$ counts the number of neurons that fired during the k-th time bin. In Fig 3 we plot the probability (4) for several networks of different size N. For all N, $P(K)$ is well fitted by an exponential distribution in an intermediate regime. The exponential factor of the distributions consistently decreases as N increases, indicating that it is more likely to observe concurrent firing between neurons in larger networks.

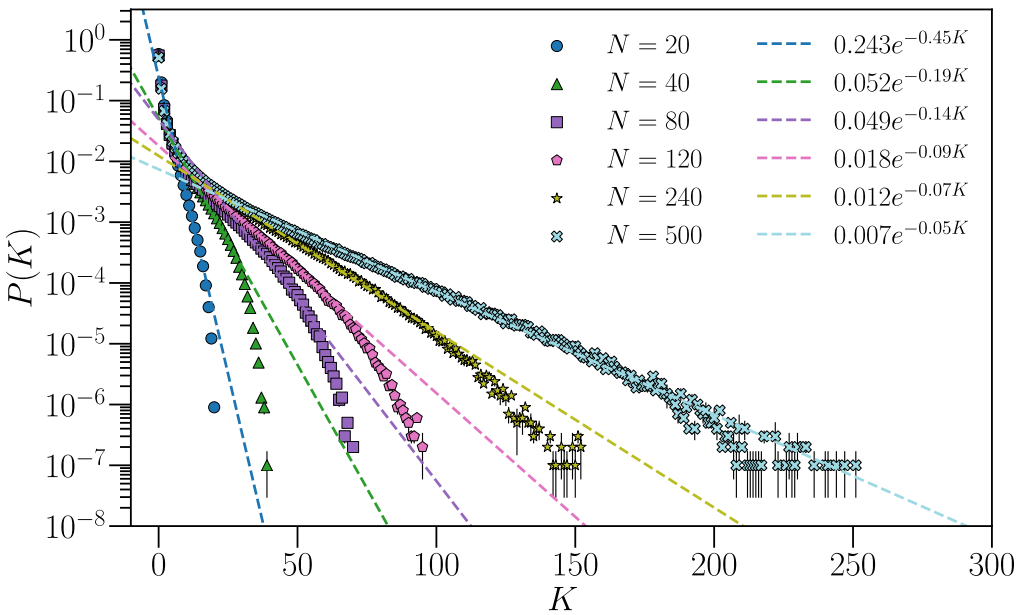


Fig 3. Probability $P(K)$ of observing K neurons firing simultaneously during a time bin $\Delta t_b = 5$ timesteps. Results are averages over N_c different fully-excitatory IF networks configurations of size N. N_c varies with N as in Fig 2. For each configuration, $P(K)$ is estimated by averaging over $N_b = 10^7$ bins. Error bars are given by the standard error, and most of them are smaller than the symbol size.

Results of the Boltzmann Machine learning

Next, we consider pairwise Ising models of different system sizes $N = \{20, 40, 80, 120\}$, fitted to reproduce the temporal averages generated by the IF model, using the Boltzmann Machine (BM) algorithm (see "Maximum Entropy Modelling" in Methods). Each spin i can either be in the up-state ($\sigma_i = 1$) or down-state ($\sigma_i = -1$), defining a particular spin configuration $\boldsymbol{\sigma} = \{\sigma_1, \sigma_2, ..., \sigma_N\}$. The Ising Hamiltonian $H(\boldsymbol{\sigma})$ is given by

$$H(\boldsymbol{\sigma}) = -\sum_i^N h_i \sigma_i - \sum_i^N \sum_{j<i}^N J_{ij} \sigma_i \sigma_j \ , \tag{5}$$

with h_i and J_{ij} the fitting parameters, acting as local fields on spin i and interaction constants between spins i and j, respectively. When appropriate, we use the superscript

(IF) and (BM) to distinguish between measurements using the IF model and those obtained with the BM using Monte Carlo. More specifically, $\langle ... \rangle^{(\mathrm{IF})}$ indicates an average over time bins in the IF model while $\langle ... \rangle^{(\mathrm{BM})}$ is an average over spin configurations of the Ising model.

In Fig 4 we present the values of $\{h_i\}$ and the distributions of the interaction constants $\{J_{ij}\}$, fitted to reproduce the temporal averages $\{\langle \sigma_i \rangle^{(\mathrm{IF})}\}$ and $\{\langle \sigma_i \sigma_j \rangle^{(\mathrm{IF})}\}$ of fully-excitatory IF networks of size N, tuned to the critical state. More specifically, Figs 4a-d show ranked plots of the fields h_i, where h_1 (h_N) corresponds to the spin associated with the neuron that fired the most (least) in the IF model. The fields h_i are mostly negative for all N. This is not surprising considering that a neuron in the IF model fires rarely, as discussed in previous sections, and the negative fields are a consequence of this sparse activity. We note, however, that, as the system size N increases, an increasingly larger fraction of spins have positive h_i, corresponding to the spins associated with the most active neurons. In Figs 4e-h we plot the distributions of the learned couplings J_{ij}. The distributions are bimodal, with the largest peak near zero and the second one at positive J_{ij}. However, the height of the second peak appears to decrease with the system size N, and the limiting distribution when $N \to \infty$ becomes consistent with a normal distribution centered around zero. It is interesting to recall that random, normally distributed couplings J_{ij} characterize the so-called Sherrington-Kirkpatrick (SK) model [51, 52], an Ising-like model which has been shown to exhibit both ferromagnetic and spin glass phases [51]. Unlike the SK model, however, the learned Ising models exhibit high heterogeneity in the distribution of the fields h_i, and since the learning technique fits simultaneously both the fields and interaction constants, the inference of the $\{J_{ij}\}$ should not be considered numerically decoupled from that of the $\{h_i\}$.

To test the quality of the learning process, in Fig 5 we compare the $\{\langle \sigma_i \rangle^{(\mathrm{BM})}\}$ and $\{C_{ij}^{(\mathrm{BM})}\}$ of the Ising-like model with the neural network data of the IF model $\{\langle \sigma_i \rangle^{(\mathrm{IF})}\}$ and $\{C_{ij}^{(\mathrm{IF})}\}$. As described above (by virtue of equations (36) and (37), in the "Methods" section), if the learning is successful, these quantities should be identical since these are

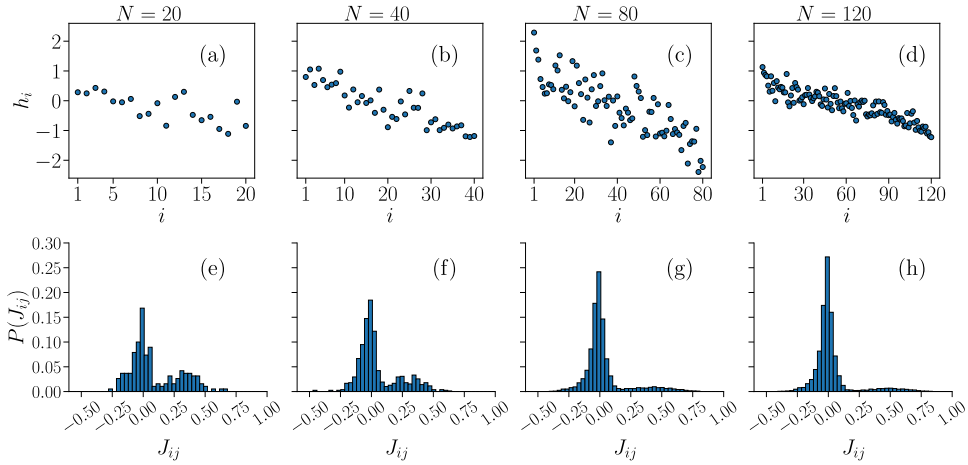


Fig 4. Sets of learned fields h_i and interaction constants J_{ij} of the pairwise Ising model that reproduce the time averages of the IF model. Plots of the fields h_i (a-d), sorted by the average local activity $\langle \sigma_i \rangle^{(\mathrm{IF})}$ of the associated neuron i, in order of decreasing $\langle \sigma_i \rangle^{(\mathrm{IF})}$, and distributions of the interaction constants J_{ij} (e-h), learned by the BM to reproduce the average local activity $\langle \sigma_i \rangle$ and two-point correlation functions C_{ij} of a fully-excitatory IF neural network at criticality, with $N = \{20, 40, 80, 120\}$ neurons.

precisely the constraints imposed in the learning procedure. This is confirmed for both $\langle \sigma_i \rangle$ and C_{ij}, for all N, with all points following closely the bisector $y = x$.

An interesting quantity to assess the predictive power of the model are the three-point correlation functions T_{ijk} between all $N \cdot (N-1) \cdot (N-2)/6$ triplets of neurons i, j and k,

$$\begin{aligned} T_{ijk} &= \langle (\sigma_i - \langle \sigma_i \rangle) \cdot (\sigma_j - \langle \sigma_j \rangle) \cdot (\sigma_k - \langle \sigma_k \rangle) \rangle = \\ &= \langle \sigma_i \sigma_j \sigma_k \rangle - \langle \sigma_i \rangle \langle \sigma_j \sigma_k \rangle - \langle \sigma_j \rangle \langle \sigma_i \sigma_k \rangle - \langle \sigma_k \rangle \langle \sigma_i \sigma_j \rangle + 2 \langle \sigma_i \rangle \langle \sigma_j \rangle \langle \sigma_k \rangle \ , \end{aligned} \tag{6}$$

where the averages $\langle ... \rangle$ are defined analogously as in equations (1) or (3). In Fig 6 we compare the triplets between the IF and the Ising models, for systems with $N = \{20, 40, 80, 120\}$ considered previously. Even though the T_{ijk} are not constrained by the learning procedure, the generalized Ising model still captures their systematic qualitative behaviour. Unlike $\langle \sigma_i \rangle$ and C_{ij}, however, the triplets appear to be consistently overestimated by the Ising model, when compared to their IF model counterpart. We note that this behaviour was also seen in experimental recordings from stimulated neurons in a salamander retina [9], albeit to a less significant degree.

Another useful measure to assess how well the Ising-like model describes the IF model data is the probability for simultaneous firing $P(K)$ within a small time window, as defined in equation (4). In the Ising model, the sum over time bins in equation (4) is replaced by a sum over spin configurations generated by Monte Carlo simulation, where K now counts the number of up-spins in a single configuration in the Monte Carlo evolution. We compare this quantity between the IF and Ising models in Fig 7 for systems of size $N = \{20, 40, 80, 120\}$. The Ising model seems to predict the $P(K)$ correctly only when K is very small, of the order $K/N \lesssim 0.10$, with noticeable discrepancies seen also for this range of K, particularly for the probability of no activity ($K = 0$), in systems of size $N > 20$. For larger K, there is an intermediate range where the Ising model underestimates $P(K)$, followed by a region in which $P(K)$ is greatly

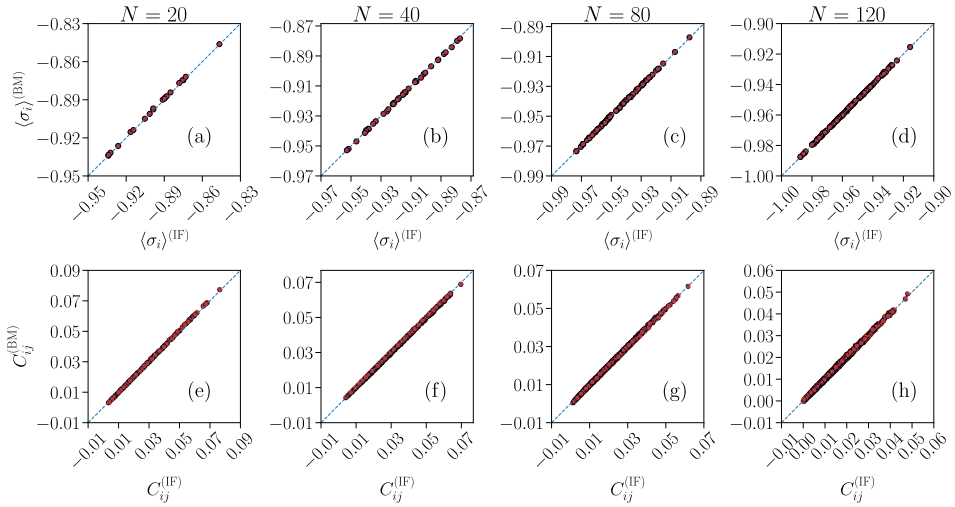


Fig 5. Quality test for the BM learning process. Comparison between the average local activity $\langle \sigma_i \rangle$ (a-d) and correlation functions C_{ij} (e-h) of the Ising-like model (y-axes) and IF model network (x-axes), for system sizes $N = \{20, 40, 80, 120\}$. The blue dashed lines are the bisector $y = x$. Results are averages over $N_b = 10^7$ time bins for the IF model (IF), and averages over $M_c = 3 \cdot 10^6$ spin configurations for the Ising model (BM). Error bars are given by the standard error, and are smaller or equal to the symbol size.

overestimated. This effect seems to increase with the system size N. This particular behaviour was also observed when training a pairwise Ising model to reproduce the time averages of stimulated neuronal activity from a salamander retina [11].

Thermodynamics of the pairwise Ising models

Having mapped the neural network to a pairwise Ising model, we can start to analyse its properties. Given a system with N spins, for each spin configuration we can measure the magnetization $M(\boldsymbol{\sigma}) = \sum_i^N \sigma_i$ and the energy $E(\boldsymbol{\sigma}) = -H(\boldsymbol{\sigma}) = \sum_i^N h_i \sigma_i + \sum_i^N \sum_{j<i}^N J_{ij} \sigma_i \sigma_j$. From the fluctuations of these quantities, according to the fluctuation-dissipation theorem, we can calculate the susceptibility χ and the specific heat C_v of the Ising system as a function of

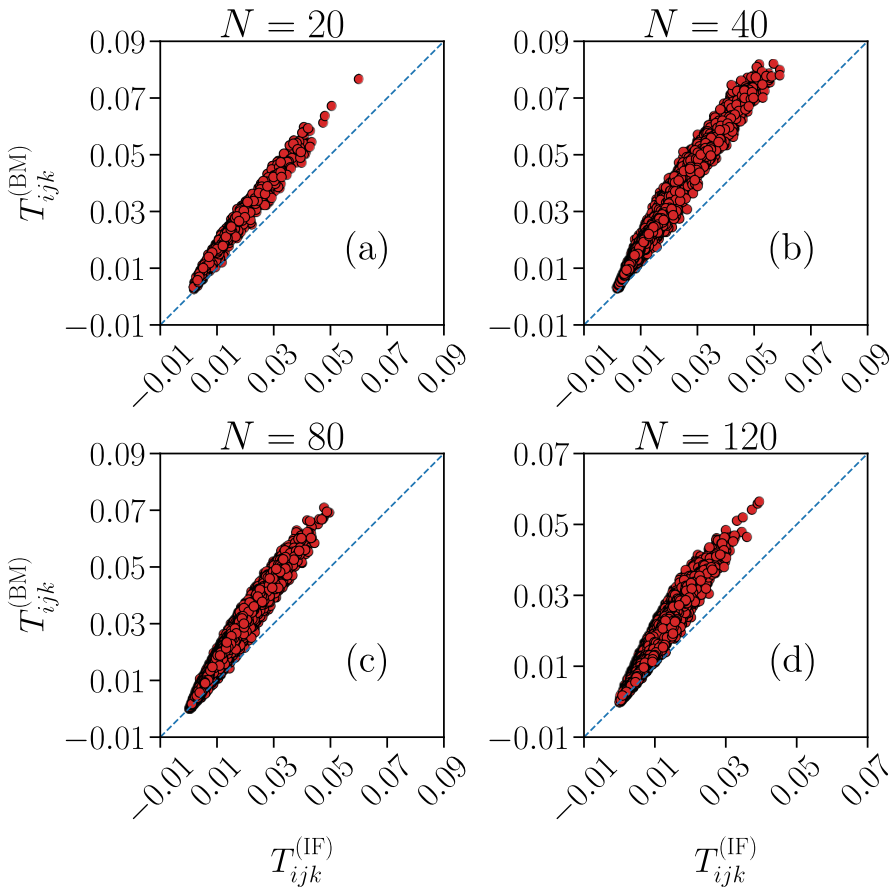


Fig 6. Predictive capability of the Ising model for the three-point correlation functions T_{ijk}. Comparison between the three-point correlation functions T_{ijk} of the Ising-like model (y-axes) and the IF model network (x-axes), for system sizes $N = \{20, 40, 80, 120\}$. The blue dashed lines are the bisector $y = x$. Results are averages over $N_b = 10^7$ time bins for the IF model (IF), and averages over $M_c = 3 \cdot 10^6$ spin configurations for the Ising model (BM). Error bars are given by the standard error, and are smaller or equal to the symbol size.

temperature T,

$$\chi = \frac{1}{T} \cdot \left(\langle M^2 \rangle - \langle M \rangle^2 \right) \ , \tag{7}$$

$$C_v = \frac{1}{T^2} \cdot \left(\langle E^2 \rangle - \langle E \rangle^2 \right) \ , \tag{8}$$

where $\langle ... \rangle$ indicates an average over spin configurations generated by Monte Carlo simulations. By changing T, one can probe the thermodynamic properties of these spin systems. We start from a random spin configuration and thermalize it as described in the Methods section. Notice, however, that the MEM method only claims that the Ising model is representative of the IF model for $T = 1 \equiv T_0$, since that is the temperature used during the BM learning procedure. Since there is no evident analogy between the control parameter of the learned Ising models, T, and that of the IF model, δu_{rec},

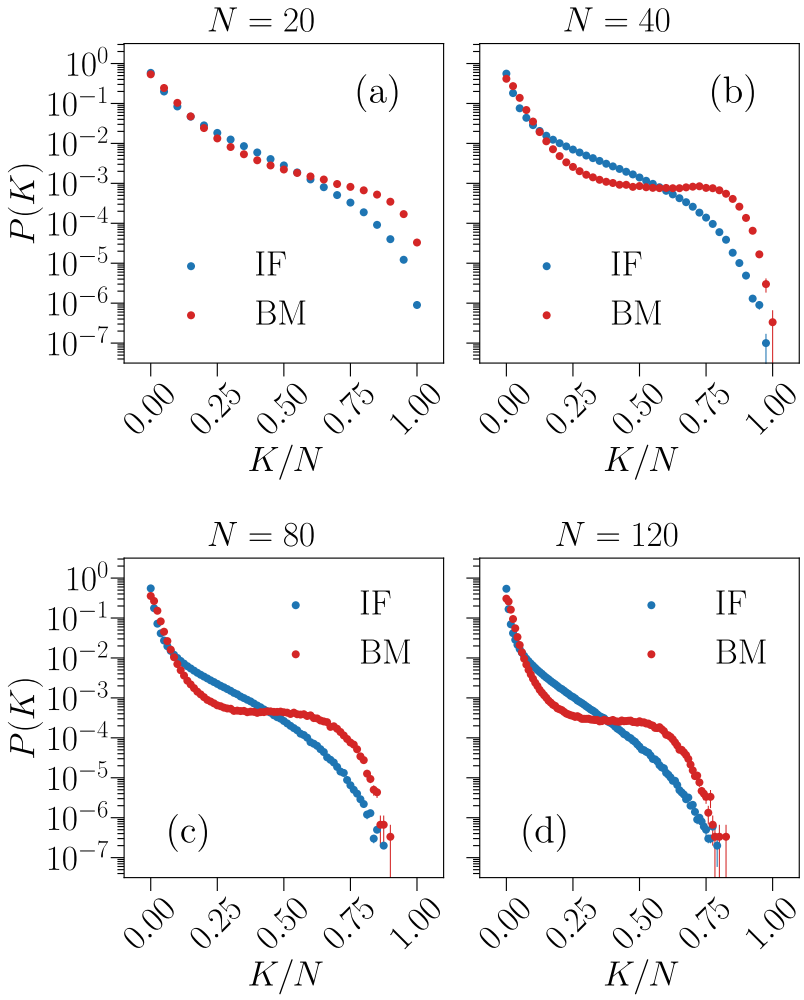


Fig 7. Predictive capability of the Ising model for the probability of simultaneous firing $P(K)$. Comparison between the simultaneous firing/up-state probability $P(K)$ of the Ising-like model (red) and the IF model network (blue), for system sizes $N = \{20, 40, 80, 120\}$. Results are averages over $N_b = 10^7$ time bins for the IF model (IF), and averages over $M_c = 3 \cdot 10^6$ spin configurations for the Ising model (BM). Error bars are given by the standard error, and most are smaller or equal to the symbol size.

temperatures $T \neq T_0$ have no obvious physical meaning in these learned Ising models, besides being a useful parameter to check if there is anything special about this T_0 [12].

In Fig 8 we plot the average magnetization per spin $m = \langle M \rangle / N$, the susceptibility χ and the specific heat C_v as a function of the temperature T for systems of size $N = \{20, 40, 80, 120\}$. For each N, we consider five Ising systems with distinct parameters learned from different IF networks of identical size at criticality, to assess the sample-to-sample variations for IF networks with different specific connectivities between the neurons.

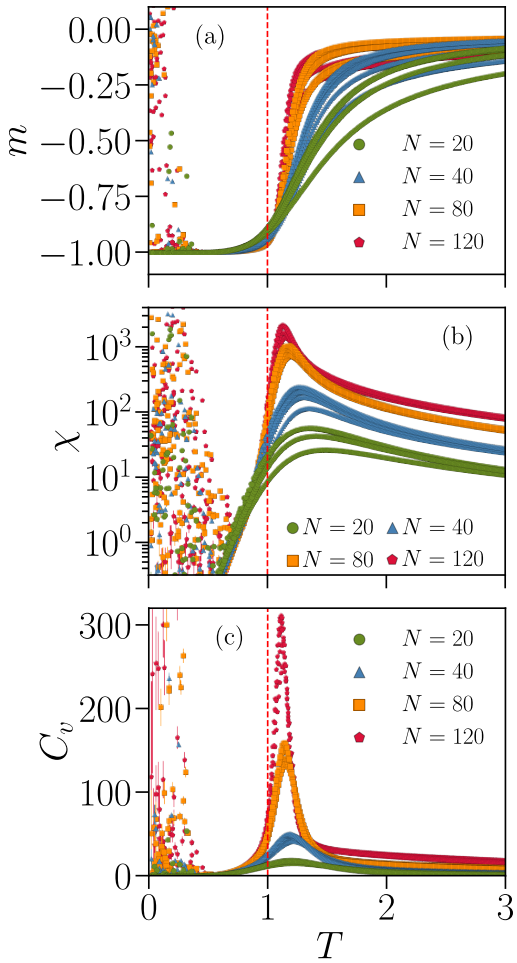


Fig 8. Thermodynamic functions of Ising-like models associated with fully-excitatory IF networks of different system sizes N. Average magnetization per spin m (a), susceptibility χ (b) and specific heat C_v (c) as a function of the temperature $T \in [0.1, 3.0]$, for systems with $N = \{20, 40, 80, 120\}$ spins. Different curves with the same colour and symbol correspond to Ising systems with parameters fitted to different IF network configurations with the same size N but different specific connectivities. The vertical dashed lines indicate $T = T_0 = 1$, the "default" temperature used in the BM learning procedure to fit the respective Ising parameters to each IF network. The cloud of random values observed for $T < 1$ suggests the presence of a spin-glass phase, where the thermal energy is insufficient to drive the system away from the initial random spin configuration. Results are averages over $M_c = 3 \cdot 10^6$ spin configurations. Error bars are given by the standard error and are overall smaller or equal to the symbol size.

The maxima of the susceptibility and specific heat appear near the "default" temperature $T_0 = 1$ associated with the IF model neuronal data, and its height increases with N, indicating that the neural networks tend to maximize thermodynamic fluctuations. This phenomenon, where fluctuations diverge with increasing system size, is a hallmark of a system operating near a critical point [53].

Looking now at different network configurations of the same size N, we can clearly detect large variations across configurations. Particularly, for different IF model networks with the same N, the maximum value of the susceptibility χ varies by a factor ~ 2 for sizes $N = \{20, 40\}$ and ~ 1.5 for $N = \{80, 120\}$.

Another interesting result is the cloud of random values observed for $T < 1$ for all thermodynamic quantities. This may be an indication that, for low temperatures, the Ising model transitions into a spin-glass phase. Since this phase usually exhibits complex energy landscapes [52], thermal fluctuations in this temperature regime might not be sufficient to drive the system away from the randomly chosen initial configuration and into the ground state. Indeed, starting with all $\sigma_i = -1$ as initial configuration for the Monte Carlo simulations removes this effect entirely (see S3 Fig).

Subnetworks

In experiments, neuronal recordings are usually performed over a subset of neurons out of the total population [11, 15]. As such, it is of practical interest to study how network subsampling might affect the analysis.

A very recent study [15] shows that the properties of firing statistics change with the spatial distribution of the subsampled neuronal patch. Taking this into account, we will consider two types of subnetworks, also studied in the aforementioned work: a subset of neurons that are spatially close together, Fig 9a, or picked at random positions, Fig 9b. For the former case, we select the neurons that are closest to the center of the cubic lattice.

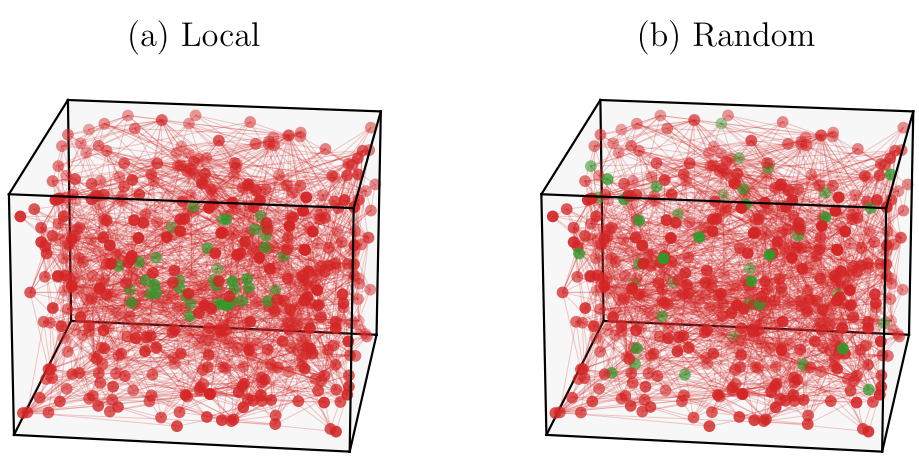


Fig 9. Schematic representations of IF model subnetworks with different spatial distribution. Subnetworks (green) with neurons closely packed together (a) or picked at random positions (b). Subnetworks have $n = 40$ neurons in a system of size $N = 500$.

In Fig 10 we plot the results of the simultaneous firing probability $P(K)$ and triplets T_{ijk} measured in subnetworks of $n = 40$ neurons in systems with $N = 500$ total neurons, for both spatial distributions described previously. We also compare these quantities with the predictions of the corresponding Ising models. The learning efficiency for the

parameters of these models is similar to the cases using full networks, as assessed by inspecting the constrained temporal averages, $\langle \sigma_i \rangle$ and C_{ij} (see S4 Fig).

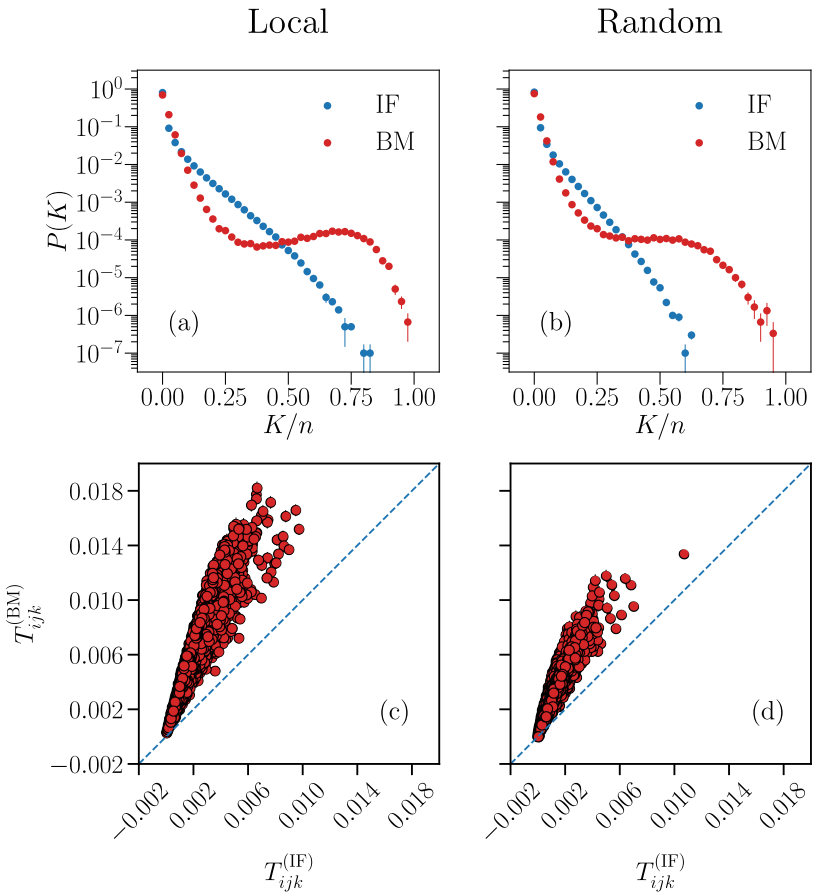


Fig 10. Predictive capability of the Ising model for IF model subnetworks. Comparison between Monte Carlo sampling of the Ising model (BM) and the neural network data (IF) of the probability $P(K)$ (a-b) and triplets T_{ijk} (c-d), for subnetworks with $n = 40$ neurons, with a local (a and c) and random (b and d) spatial distribution, in a system with $N = 500$ neurons. The blue dashed lines in the bottom plots are the bisector $y = x$. Results are averages over $N_b = 10^7$ time bins for the IF model (IF), and averages over $M_c = 3 \cdot 10^6$ spin configurations for the Ising model (BM). Error bars are given by the standard error, and are overall smaller or equal to the symbol size.

As in the case of entire networks, the Ising models fail to predict $P(K)$ for large K, with even larger discrepancies (Fig 10). For the IF network, the $P(K)$ distribution decays exponentially for both subnetworks, as for the entire network case. Analogously, the probability $P(K)$ for the learned Ising models shows an initial exponential decay for low K, for both spatial distributions. Conversely, at $K/n \approx 0.35$, the behaviour of $P(K)$ from the Ising model is slightly different between the two spatial distributions. For the local case, the probability for observing simultaneously up-spins reaches a second local maximum at $K/n \approx 0.75$. Notice, however, that this is a small effect, magnified by the logarithmic scale. On the other hand, for the random case, a plateau in $P(K)$ of the Ising model is observed in the intermediate regime of K, followed by a subsequent decay. These differences between the IF and Ising model, compared to the full network case (Fig 7), could be due to the fact that we are neglecting influences from neurons that are not included in the subnetworks, and therefore not encoded in the

Ising model parameters.

The behavior of the triplets is similar to the case of entire networks (Fig 6), wherein the triplets are systematically overestimated by the Ising model. We note also that overall the triplets have much smaller values, when compared to the ones found for the full network with $N = 40$. Similar results were found for both $P(K)$ and the T_{ijk} for subnetworks of larger size $n = 80$ in the same IF network with $N = 500$ (not shown). The fact that both cases of spatial distributions yield Ising models with similar accuracy in reproducing the original IF model data is unexpected, as this is seemingly in contrast to recent experimental studies on mouse brains [15], where the generalized Ising model predicted better the results from subgroups of ~ 100 neurons that were spatially clustered together. It should be noted, however, that the experimental study pertains to stimulated neuronal activity due to visual stimuli, whereas our model simulates spontaneous neuronal activity [31], in the absence of any external stimulation.

Networks with inhibitory neurons

Inhibitory neurons hamper neuronal activity, consequently affecting the dynamics of the network [54, 55]. Increasing the percentage p_{in} of inhibitory neurons in a neural network moves the system into a subcritical regime [55]. A way to keep the system close to the critical state when $p_{\text{in}} > 0$ is by increasing the value of the tuning parameter δu_{rec}. Here we consider systems of size $N = 80$ and $p_{\text{in}} > 0$, in the critical state (see S2 Fig), and study how the presence of inhibition might affect the properties of the associated Ising-like models. The agreement of the average local activities and correlation functions between the Ising model and the IF networks with inhibitory neurons is the same as for fully-excitatory ones (see S5 Fig).

In Fig 11 we show the distributions of fields h_i and interaction constants J_{ij} learned for IF networks with $p_{\text{in}} = \{0\%, 10\%, 20\%\}$ inhibitory neurons.

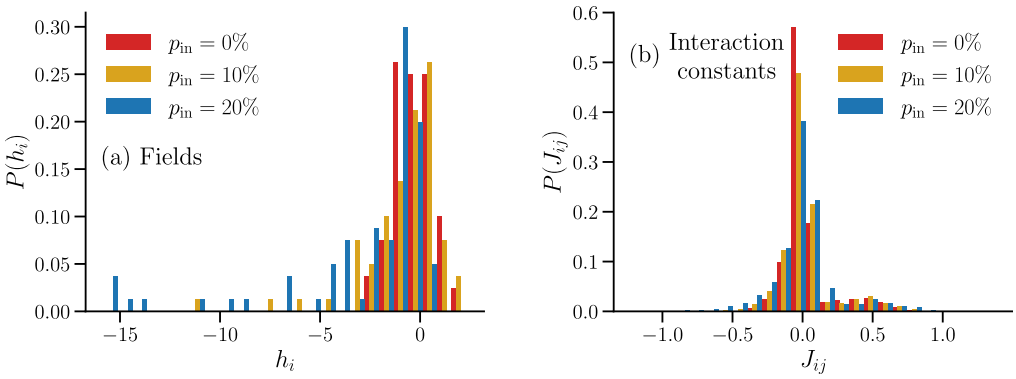


Fig 11. Distributions of the learned fields h_i and interaction constants J_{ij} of a pairwise Ising model for neural networks with different fractions of inhibitory neurons p_{in}. Fields h_i (a) and interaction constants J_{ij} (b) associated with IF networks with $p_{\text{in}} = \{0\%, 10\%, 20\%\}$ and $N = 80$, obtained after $N_{\text{BM}} = 60000$ iterations of the BM.

As seen in Fig 11a, the distribution of fields h_i acquires a tail towards negative h_i for $p_{\text{in}} > 0$, becoming more pronounced as p_{in} increases. Interestingly, these h_i are associated with excitatory neurons (see S6 Fig). A possible explanation is that, in the IF network, the neurons with the smallest average local activity $\langle \sigma_i \rangle$ are not necessarily inhibitory neurons, but rather neurons with incoming connections from inhibitory ones, which in turn are more likely to be excitatory since $p_{\text{in}} < 50\%$. This indicates that the $\{h_i\}$ do not encode the information about whether a neuron is excitatory or inhibitory.

In Fig 11b we see that the peak of the distributions of the interaction constants J_{ij}, decreases with $p_{\rm in}$, with a more pronounced tail towards negative J_{ij} observed for $p_{\rm in} = 20\%$. Interestingly, the presence of the second peak of $P(J_{ij})$ at positive J_{ij} seems to be robust with respect to changes in $p_{\rm in}$.

In Fig 12 we plot as a function of the temperature T the thermodynamic functions m, χ and C_v of the Ising systems with the parameters of Fig 11. In Fig 12a, the plateau for m at high T in the system associated to $p_{\rm in} = 20\%$ is lower than the one associated to the fully-excitatory network, indicating a higher tendency for spins to be in the down state ($\sigma_i = -1$) for the former case in the paramagnetic phase, which could stem from the inhibitory neurons. This is consistent with the observation of stronger negative fields h_i when $p_{\rm in} = 20\%$, as seen in Fig 11a. In Figs 12b and 12c, we see that the values of the maxima of χ and C_v decrease with $p_{\rm in}$, while their position with respect to the temperature remains unchanged. This suggests that introducing inhibition in the IF networks reduces thermal fluctuations in the associated Ising model.

Partially-connected pairwise Ising models

One of the main disadvantages of BM learning is its intense CPU time demand [27, 56], restricting the network sizes one could potentially analyse. To try to circumvent this problem, we will consider Ising models with pruned links [19], whereby we remove the couplings associated with the weakest correlated pairs of neurons. Specifically, we set $J_{ij} = 0$ if the corresponding $C_{ij}^{\rm (IF)}$ is below a certain threshold, defined as a fraction of the largest measured correlation function, i.e. if $C_{ij}^{\rm (IF)} < \eta \max\left(C_{ij}^{\rm (IF)}\right)$, where $\eta \in [0, 1]$ sets the threshold. Thus, the BM only needs to learn a subset of the total number of couplings J_{ij}, possibly accelerating the convergence process and consequently allowing the study of networks larger than the ones considered so far. Furthermore, an obvious speed-up is also achieved by making use of the fact that a subset of the $\{J_{ij}\}$ are zero, and therefore can be disregarded in the double sum of equation (5) during the Monte Carlo simulations. We will consider three thresholds $\eta \in \{0.10, 0.15, 0.20\}$, using a fully-excitatory system of size $N = 180$ in the critical state. Computing only a fraction of the interaction terms allowed to obtain the BM results for the $N = 180$ system in a less or comparable CPU time with respect to the one required by the $N = 120$ systems where we considered the full set $\{J_{ij}\}$, with a greater CPU speed-up achieved the larger the threshold η is.

In Fig 13 we present the comparison between the correlation functions of the IF model and the partially-connected Ising model for the three different thresholds. We see that, even though more than 40% couplings J_{ij} have been removed in the Ising model for the lowest threshold considered, $\eta = 0.10$, and more than 70% for the largest one, $\eta = 0.20$, the overall data is still reconstructed by the partially-connected Ising model, with a decrease in quality of the fit for the lowest C_{ij} (grey dots in the plots) since they were not considered in the learning process of the BM. On the other hand, all $\langle \sigma_i \rangle$ are well predicted by the Ising model (see S7 Fig) since we fit all the $N = 180$ fields $\{h_i\}$. In Fig 14 we plot the learned fields and the distributions of the learned non-zero interaction constants of the partially-connected Ising model for the three different thresholds η. While the set of fields $\{h_i\}$ remains qualitatively similar for the three thresholds η, with a majority of negative fields as in the case of the fully-connected Ising models, the distributions of the non-zero J_{ij} change significantly with η, with the peak at positive J_{ij} increasing and the one at $J_{ij} \approx 0$ decreasing as we remove progressively more J_{ij}. Since we remove only the J_{ij} associated with the smallest correlation functions C_{ij}, this seemingly indicates that the second peak at $J_{ij} > 0$ also seen in the distributions for the fully-connected Ising models (Fig 4) is associated with the subset of the largest C_{ij}.

As usual, we can also analyze how the Ising model might predict quantities that are

not being constrained by the BM algorithm. In Fig 15 we present the comparison between the tree-point correlation functions T_{ijk} measured in the partially-connected Ising model and the IF model. As in the fully-connected case, the Ising model consistently overestimates the values of T_{ijk}. The quality of the fit for the smallest threshold considered, $\eta = 0.10$, is comparable to the fully-connected Ising models (Fig 6). However, as η increases, the quality of the fit declines, particularly for the smallest $T_{ijk}^{(\mathrm{IF})}$. Next, we present in Fig 16 the results for the probability $P(K)$ of simultaneous firing. The predictive capability of the partially-connected Ising model for small K is similar to that of fully-connected ones (Fig 7), while significant differences can be seen for large K. As the threshold η is increased, the Ising model seems to predict larger

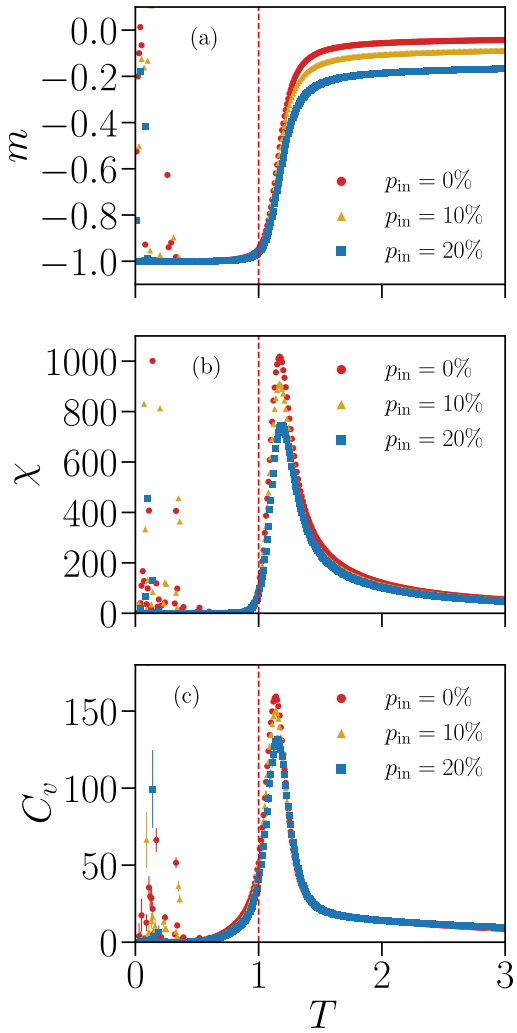


Fig 12. Thermodynamic functions of Ising-like models associated with IF model networks with different fractions of inhibitory neurons $p_{\mathbf{in}}$. Average magnetization per spin m (a), susceptibility χ (b) and specific heat C_v (c) as a function of the temperature $T \in [0.1, 3.0]$, simulated using the learned parameters shown in Fig 11 for different fractions of inhibitory neurons $p_{\mathrm{in}} = \{0\%, 10\%, 20\%\}$ and $N = 80$. As in Fig 8, the cloud of random values for $T < 1$ suggests the presence of a spin-glass phase. Results are averages over $M_c = 3 \cdot 10^6$ spin configurations. Error bars are given by the standard error and are overall smaller or equal to the symbol size.

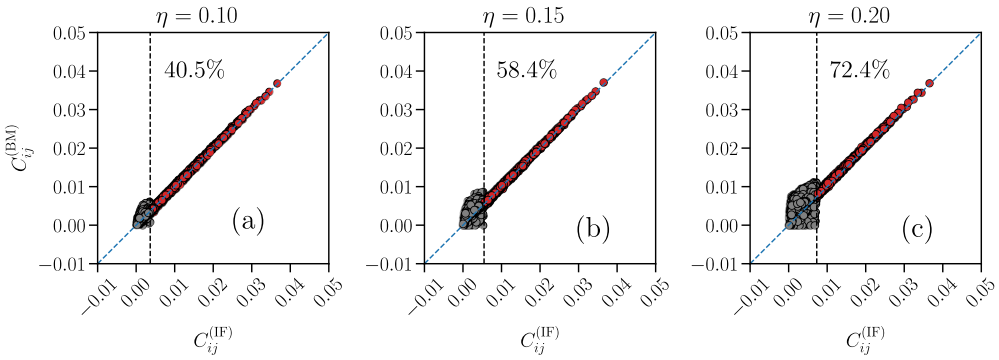


Fig 13. Quality test for the BM learning process with a partially-connected Ising model. Comparison between the correlation functions C_{ij} of the partially-connected Ising-like model (y-axes) and IF model network (x-axes), for a fully-excitatory system at criticality of size $N = 180$ and three different thresholds $\eta = \{0.10, 0.15, 0.20\}$ for the removal of the J_{ij}. If $C_{ij}^{(\mathrm{IF})} < \eta \max(C_{ij}^{(\mathrm{IF})})$ (grey dots), the corresponding interaction constant is set to $J_{ij} = 0$. The vertical dashed lines indicate the value $\eta \max(C_{ij}^{(\mathrm{IF})})$, with the corresponding approximate percentage of removed J_{ij} reported at the right of this line. The blue dashed lines are the bisector $y = x$. Results are averages over $N_b = 10^7$ time bins for the IF model (IF), and averages over $M_c = 3 \cdot 10^7$ spin configurations for the Ising model (BM). Error bars are given by the standard error, and are smaller or equal to the symbol size.

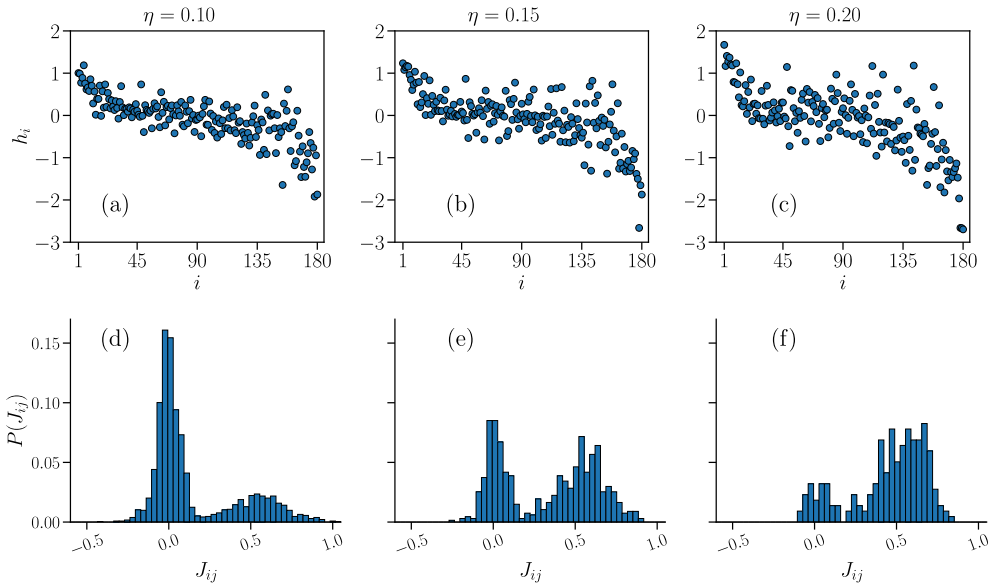


Fig 14. Sets of learned fields h_i and non-zero interaction constants J_{ij} of the partially-connected Ising model. Plots of the fields h_i (a-c), sorted by the average local activity $\langle \sigma_i \rangle^{(\mathrm{IF})}$ of the associated neuron i, in order of decreasing $\langle \sigma_i \rangle^{(\mathrm{IF})}$, and distributions of the non-zero interaction constants J_{ij} (d-f), that reproduce the respective data of the Ising model presented in Fig 13 for the three different thresholds $\eta = \{0.10, 0.15, 0.20\}$, for a fully-excitatory system at criticality of size $N = 180$.

simultaneous activity at the scale of the whole network, where $K \approx N$, indicated by a local maximum in $P(K)$. This seems to indicate that the small J_{ij} that we are

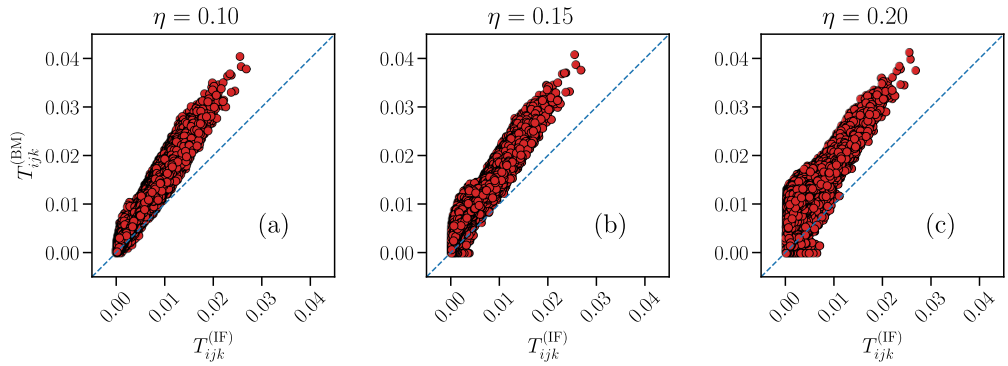


Fig 15. Predictive capability of the partially-connected Ising model for the three-point correlation functions T_{ijk}. Comparison between the three-point correlation functions T_{ijk} of the partially-connected Ising-like model (y-axes) and the IF network (x-axes), for the three thresholds $\eta = \{0.10, 0.15, 0.20\}$, for a fully-excitatory system at criticality of size $N = 180$. Results are averages over $N_b = 10^7$ time bins for the IF model (IF), and averages over $M_c = 3 \cdot 10^7$ spin configurations for the Ising model (BM). Error bars are given by the standard error, and most are smaller or equal to the symbol size.

disregarding encode the information concerning the large K regime of $P(K)$.

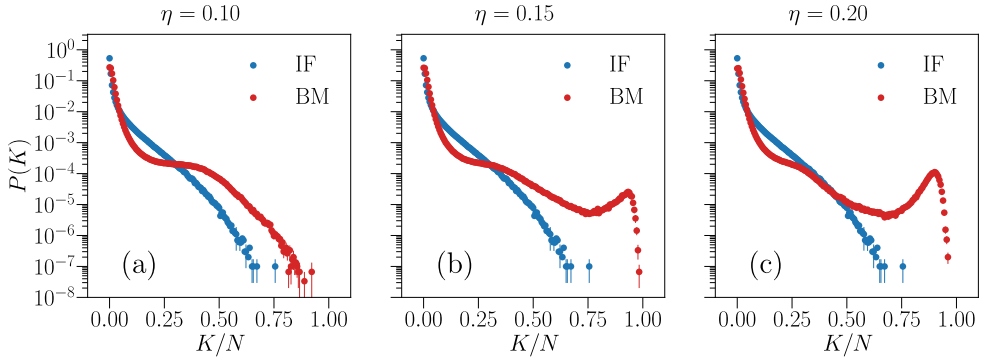


Fig 16. Predictive capability of the partially-connected Ising model for the probability of simultaneous firing $P(K)$. Comparison between the simultaneous firing/up-state probability $P(K)$ of the Ising-like model (red) and the IF model network (blue), for the three thresholds $\eta = \{0.10, 0.15, 0.20\}$, for a fully-excitatory system at criticality of size $N = 180$. Results are averages over $N_b = 10^7$ time bins for the IF model (IF), and averages over $M_c = 3 \cdot 10^7$ spin configurations for the Ising model (BM). Error bars are given by the standard error, and most are smaller or equal to the symbol size.

Finally, in Fig 17 we plot the thermodynamic functions m, χ and C_v as a function of the temperature T for the partially-connected Ising systems, considering the three cases $\eta = \{0.10, 0.15, 0.20\}$. Results are quite robust with respect to η, taking into account that there is a difference of $\approx 31.9\%$ of removed J_{ij} between the smallest and largest thresholds η considered. The tendency for the susceptibility χ and specific heat C_v to diverge with N is also still clearly visible when comparing to the results for a smaller system size with $N = 120$ (black symbols), which in turn considers the full set $\{J_{ij}\}$.

Discussion

By means of the Maximum Entropy method (MEM), we mapped the local and pairwise information of IF complex neural networks into Ising models with frustrated spins. Independently of the system size N, the local fields h_i are mostly negative (Fig 4a-d) and the distribution of interaction constants J_{ij} (Fig 4e-h) is bimodal for small systems, but tends to a normal distribution centered around zero as N increases. These Ising systems display a spin glass phase at low temperatures, independently of N (Fig 8), and the susceptibility and specific heat tend to diverge with the system size, with a maximum near $T = T_0 = 1$, the effective temperature used in the BM learning to fit the

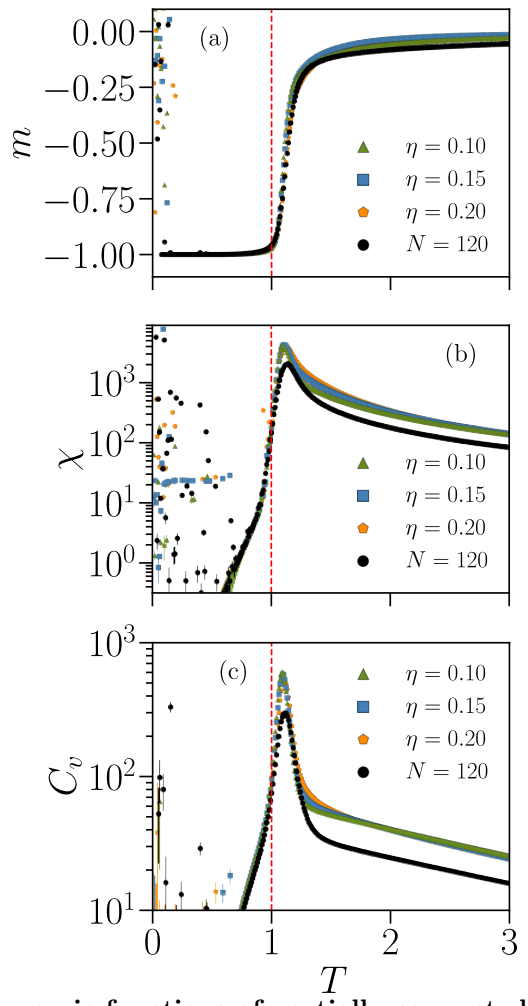


Fig 17. Thermodynamic functions of partially-connected Ising models associated with an IF model network with $N = 180$. Average magnetization per spin m (a), susceptibility χ (b) and specific heat C_v (c) as a function of the temperature $T \in [0.1, 3.0]$, simulated using the learned parameters shown in Fig 14 for the three different thresholds $\eta = \{0.10, 0.15, 0.20\}$, for a fully-excitatory IF network at criticality of size $N = 180$. The black symbols show the results for a learned Ising model associated with an IF network of size $N = 120$, considering the full set $\{J_{ij}\}$. As in Figs 8 and 12, the cloud of random values for $T < 1$ suggests the presence of a spin-glass phase. Results are averages over $M_c = 3 \cdot 10^6$ spin configurations. Error bars are given by the standard error and are overall smaller or equal to the symbol size.

IF network neuronal data. This is an indication that the Ising model analogs of the IF networks operate near a critical point. Furthermore, at least for $N \leq 120$, the height of the maximum in the susceptibility is sensible to changes in the details of the connectivity of the network (Fig 8b). We remark that the Ising models do not predict well unconstrained quantities such as the three-point correlation functions (Fig 6) and the probability of simultaneous firing even for whole networks (Fig 7). This discrepancy is enhanced when considering only a subset of the neurons (Fig 10), likely due to inputs from neurons outside the subnetwork which are not considered in the MEM mapping. This indicates that additional caution should be taken when analysing real neural networks using the MEM approach, which often consider only a subset of the total neuronal population. The presence of inhibitory neurons in the IF networks leads to reduced thermal fluctuations in the associated Ising models, evidenced by a decrease of the maxima near the critical point (Fig 12).

Ultimately, the potential of the MEM approach using a numerical model of neural activity is limited by its CPU time demand when considering large system sizes $N \gtrapprox 120$. To circumvent this issue, we consider a partially-connected Ising model with only a subset of the total pairs of interaction constants $\{J_{ij}\}$ obtained considering only the largest correlation functions measured in the IF network, allowing to study a system of size $N = 180$ (Figs 13-17). Similar MEM procedures allow the study of much larger sizes, such as the so-called restricted Boltzmann Machine [57] or a random projection model [58], but the resulting maximum entropy distributions from these approaches are no longer analogous to that of an Ising model, and the thermodynamical interpretation is no longer applicable [27]. A very recent study [59] shows that it might be possible to consider large networks considering only connections which contain the maximum mutual information among all pairs of neurons. Since the IF model can also tune systems off the critical state, this approach could potentially be used to study thermodynamic properties of networks in the sub- and supercritical state, brain states associated with pathological conditions [60].

Methods

Integrate-and-fire model

The model considers N neurons placed randomly inside a cubic space of side $L = \sqrt[3]{N/\rho}$, where $\rho = 0.016$ [55] is the density of the neurons. Connections are directed and weighted, with dynamic synaptic strengths $w_{ij} \in [0, 1]$ between pre- and post-synaptic neurons i and j. Each neuron has at least one incoming connection, and the distribution of out-going degrees $k_{\text{out}} \in [2, 20]$ is a power law, i.e. $P(k_{\text{out}}) \propto k_{\text{out}}^{-2}$, following experimental measurements in functional networks [61]. The probability that two neurons are connected decays exponentially with their euclidean distance r, $P(r) \propto e^{-r/r_0}$, where $r_0 = 5$ is a characteristic length [62]. Neurons can be excitatory or inhibitory, with fraction p_{in}. We implement synaptic plasticity, in which the strengths w_{ij} can change dynamically with time. We consider both short-term and long-term plasticity (STP and LTP for short, respectively). STP refers to the modification of synaptic strengths on a short timescale, of the order of milliseconds. This is mainly caused by the depletion of synaptic resources over time required for synaptic transmission [63]. We denote by $u_i(t) \in [0, 1]$ the normalized amount of neurotransmitters available to neuron i at time t, representing the so-called readily releasable pool of neurotransmitter vesicles [64, 65]. On the other hand, LTP is a Hebbian-like process which strengthens or weakens connections strengths $g_{ij} \in (0, 1]$ depending on their usage over time, acting on much longer timescales, ranging from minutes to hours or even years [66]. We therefore separate the synaptic strengths

$w_{ij}(t) = u_i(t)g_{ij}$ into the two components $u_i(t)$ and g_{ij}, to emphasize the role of these two plasticity mechanisms. The time dependence of $w_{ij}(t)$ comes only from the STP term $u_i(t)$, since the LTP strength terms g_{ij} can be regarded as constant compared to the timescale of the dynamics.

Each neuron i is characterized by a membrane potential v_i. A neuron i will fire at some time t when its potential v_i surpasses a threshold $v_c = 1$. Activity will then propagate to all post-synaptic neurons j with incoming connections from i according to the following equations [31]:

$$v_j(t+1) = v_j(t) \pm v_i(t)u_i(t)g_{ij} \ , \tag{9}$$

$$u_i(t+1) = u_i(t) \cdot (1 - \delta u) \ , \tag{10}$$

$$v_i(t+1) = 0 \ , \tag{11}$$

where $+$ and $-$ stands for excitatory and inhibitory pre-synaptic neuron, respectively, and $\delta u = 0.05$ represents the fractional amount of neurotransmitter released at each neuronal firing [67]. The timestep unit corresponds roughly to the joint interval of synaptic and axonal delay, i.e. the time interval between the generation of the action potential at the pre-synaptic neuron and the membrane potential change at the post-synaptic one, and is of the order of 10 milliseconds [68, 69]. Equation (9) models the signal transmission from pre- to post-synaptic neurons i and j, respectively. Equation (10) models the STP dynamics, while equation (11) resets the membrane potential of the pre-synaptic neuron into its resting state. We set a minimum value $v_{\min} = -1$ for the membrane potential of each neuron, to prevent the possibility of a v_i being systematically decreased to overly negative values when $p_{\text{in}} > 0$. After firing, a neuron enters into a refractory period of a single timestep $t_r = 1$ during which it is incapable of receiving or eliciting any activity. Equations (9)-(11) can lead to the generation of neuronal avalanches [31], periods of cascading activity separated by intervals of quiescence. Their size S is defined by the number of neurons which fire during the avalanche, while their duration D is the number of numerical timesteps it lasted. To keep the activity ongoing, we implement a small external stimulation. Namely, at every timestep, even during avalanches, a voltage input $\delta v = 0.1 v_c$ is added to a randomly chosen neuron. From equation (10) we see that the amount of synaptic resources gradually depletes as a neuron fires, eventually rendering it incapable of transmitting further signals. In real systems, this is counteracted by a recovery mechanism, where the readily releasable pool is slowly recharged over the span of seconds [70]. Since activity propagation is of the order of milliseconds, we assume a separation of timescales [71], and recharge simultaneously the u_i of all neurons by a certain amount δu_{rec} only at the end of every avalanche,

$$u_i(t) \to u_i(t) + \delta u_{\text{rec}} \ . \tag{12}$$

By changing the value of δu_{rec} one can adjust the dynamical state of the system [31, 55]. At a specific value of δu_{rec}, the network can be tuned to a critical state, producing avalanches whose size distribution behaves as a power law with a cut-off that scales with the system size N. Identical systems with δu_{rec} below or above this tuned value would exhibit behaviour typical of sub- and supercritical systems, respectively. In a subcritical system, activity is extinguished too quickly, leading to an exponential decay in the size distribution, while for a supercritical one, an excess of large avalanches is observed, characterized instead by a local maximum of the size distribution near the cut-off [31].

Before performing any measurements, we let the dynamics evolve for a certain number of timesteps in order to shape the distribution of strengths g_{ij} by LTP. We start by setting them uniformly distributed in the interval $g_{ij} \in [0.04, 0.06]$. According to the

rules of Hebbian plasticity [72–74], if some neuron i frequently stimulates another neuron j, then the synapse from i to j will be strengthened. In this case, we increase the strength of the synapses g_{ij} proportionally to the voltage variation induced in the post-synaptic neuron j due to i as,

$$g_{ij}(t+1) = g_{ij}(t) + \delta g_j \ , \tag{13}$$

where $\delta g_j = \beta |v_j(t+1) - v_j(t)|$ and $\beta = 0.04$ sets the rate of this adaptation. On the other hand, synapses that are rarely active tend to weaken over time [74]. Therefore, at the end of each avalanche, we decrease all terms g_{ij} by the average increase in strength per synapse,

$$g_{ij}(t+1) = g_{ij}(t) - \frac{1}{N_s} \sum \delta g_j \ , \tag{14}$$

where N_s is the number of synapses. In real networks, synapses that weaken consistently are eventually pruned. To avoid modifying the scale-free structure of the network, we modify the strengths g_{ij} either for a fixed number $N_{\text{aval}} = 10^4$ of avalanches or until a strength g_{ij} first reaches a minimum value $g_{\min} = 10^{-5}$, where we then set that strength to $g_{ij} = g_{\min}$.

Maximum Entropy Modelling

The binarization of the IF model firing dynamics introduces the notion of the probability $P_{\text{IF}}(\boldsymbol{\sigma})$ to observe, during a time bin of duration Δt_b, any of the 2^N possible patterns of firing states $\boldsymbol{\sigma} = \{\sigma_1, \sigma_2, ..., \sigma_N\}$ in a network of size N, with each $\sigma_i \in \{-1, 1\}$. We are interested in defining this $P_{\text{IF}}(\boldsymbol{\sigma})$ in a way that is consistent with the expectation values measured in the IF model for a given network. Restating the problem more generally, we have a random variable $\boldsymbol{\sigma}$ defined by a probability distribution $P(\boldsymbol{\sigma})$ which is unknown *a priori*. This distribution $P(\boldsymbol{\sigma})$ must, of course, obey the normalization condition,

$$\sum_{\boldsymbol{\sigma}} P(\boldsymbol{\sigma}) = 1 \ , \tag{15}$$

where $\sum_{\boldsymbol{\sigma}}$ indicates a sum over all possible outcomes of $\boldsymbol{\sigma}$. We want to find an expression for $P(\boldsymbol{\sigma})$ given the average value of some function $f(\boldsymbol{\sigma})$ of the random variable $\boldsymbol{\sigma}$,

$$\langle f \rangle = \sum_{\boldsymbol{\sigma}} P(\boldsymbol{\sigma}) f(\boldsymbol{\sigma}) \ . \tag{16}$$

The maximum entropy principle [17] states that the best choice for $P(\boldsymbol{\sigma})$ out of all possible distributions which are consistent with previous knowledge (in this case, the average $\langle f \rangle$) is the one that has the maximum entropy, i.e. the least biased one. In other words, we should use the probability distribution that maximizes the entropy [75],

$$\mathcal{S} = -a \sum_{\boldsymbol{\sigma}} P(\boldsymbol{\sigma}) \ln[P(\boldsymbol{\sigma})] \ , \tag{17}$$

while subject to constraints (15) and (16), where a is a positive constant that defines the units of $\mathcal{S}$, which we set to $a = 1$ without loss of generality. To solve this problem, we can make use of the method of Lagrangian multipliers [17]. For Lagrangian multipliers λ_0 and λ_1 associated with the constraints (15) and (16), respectively, the Lagrangian reads,

$$\mathcal{L}[P(\boldsymbol{\sigma})] = -\sum_{\boldsymbol{\sigma}} P(\boldsymbol{\sigma}) \ln[P(\boldsymbol{\sigma})] + \lambda_0 \cdot \left(\sum_{\boldsymbol{\sigma}} P(\boldsymbol{\sigma}) - 1 \right) + \lambda_1 \cdot \left(\sum_{\boldsymbol{\sigma}} P(\boldsymbol{\sigma}) f(\boldsymbol{\sigma}) - \langle f \rangle \right) \ . \tag{18}$$

To proceed, it is useful to explicitly define each probability $p_i \equiv P(\boldsymbol{\sigma} = \boldsymbol{\sigma}^i)$, where $\boldsymbol{\sigma}^i$ is a particular outcome of the random variable $\boldsymbol{\sigma}$. Since the Lagrangian (18) depends on a function, $P(\boldsymbol{\sigma})$, $\mathcal{L}$ is a functional, and to search for a local extremum with respect to $P(\boldsymbol{\sigma})$ we require the following functional (or variational) derivative [76],

$$\begin{aligned}\frac{\delta \mathcal{L}}{\delta P(\boldsymbol{\sigma})}(\boldsymbol{\sigma}^i) &\equiv \lim_{\varepsilon \to\ 0} \frac{1}{\varepsilon} \cdot \left(\mathcal{L}[P(\boldsymbol{\sigma}) + \varepsilon \delta_{\boldsymbol{\sigma},\boldsymbol{\sigma}^i}] - \mathcal{L}[P(\boldsymbol{\sigma})]\right) = \\ &= \frac{d\mathcal{L}}{dP(\boldsymbol{\sigma})} \delta_{\boldsymbol{\sigma},\boldsymbol{\sigma}^i} \ , \end{aligned} \tag{19}$$

where ε is a real number and $\delta_{\boldsymbol{\sigma},\boldsymbol{\sigma}^i}$ is the delta Kronecker function. The equality follows from multiplying and dividing the right side by $\delta_{\boldsymbol{\sigma},\boldsymbol{\sigma}^i}$ and then taking the limit $\varepsilon \to 0$. Equation (19) expresses the change in the functional $\mathcal{L}$ due to changes in the function $P(\boldsymbol{\sigma})$ at the particular point $\boldsymbol{\sigma} = \boldsymbol{\sigma}^i$. The extrema condition equation then implies

$$\begin{aligned} &\frac{d\mathcal{L}}{dP(\boldsymbol{\sigma})} \delta_{\boldsymbol{\sigma},\boldsymbol{\sigma}^i} = 0 \ , \\ &\implies -(\ln[p_i] + 1) + \lambda_0 + \lambda_1 f(\boldsymbol{\sigma}^i) = 0 \ , \\ &\implies p_i = p_0 e^{\lambda_1 f(\boldsymbol{\sigma}^i)} \end{aligned} \tag{20}$$

with $p_0 \equiv e^{1+\lambda_0}$. Since the last result is identical for all other probabilities $p_{j \neq i}$, we have for the probability distribution $P(\boldsymbol{\sigma})$,

$$P(\boldsymbol{\sigma}) = p_0 e^{\lambda_1 f(\boldsymbol{\sigma})} \ . \tag{21}$$

The factor p_0 is determined by the normalization condition (15),

$$p_0 = \frac{1}{\sum_{\boldsymbol{\sigma}} e^{\lambda_1 f(\boldsymbol{\sigma})}} \equiv \frac{1}{Z} \ , \tag{22}$$

so

$$P(\boldsymbol{\sigma}) = \frac{e^{\lambda_1 f(\boldsymbol{\sigma})}}{Z} \ . \tag{23}$$

The last result can be easily generalized for the case that we have instead C constraints associated with the average value of C functions of the random variable $\boldsymbol{\sigma}$, $\{f_1(\boldsymbol{\sigma}), f_2(\boldsymbol{\sigma}), ..., f_C(\boldsymbol{\sigma})\}$. In this case, besides λ_0, associated with the normalization condition, one has C Lagrangian multipliers $\{\lambda_1, \lambda_2, ..., \lambda_C\}$, and the last term in the Lagrangian (18) needs to be modified by including a sum over the C constraints. This then gives

$$P(\boldsymbol{\sigma}) = \frac{1}{Z} e^{\sum_{c=1}^{C} \lambda_c f_c(\boldsymbol{\sigma})} \ , \tag{24}$$

$$Z = \sum_{\boldsymbol{\sigma}} e^{\sum_{c=1}^{C} \lambda_c f_c(\boldsymbol{\sigma})} \ . \tag{25}$$

Going back now to the particular case of the IF model, for a given network with N neurons, as a first approximation, we can aim at constraining the single-neuron $\langle \sigma_i \rangle$ and pairwise $\langle \sigma_i \sigma_j \rangle$ information, and, by extension, the correlation functions C_{ij} (see "Firing statistics" in the Results section). This corresponds to $N + N \cdot (N-1)/2$ constraints, giving N Lagrangian multipliers associated with each $f_i(\boldsymbol{\sigma}) = \sigma_i$, that we will denote as h_i, and $N \cdot (N-1)/2$ multipliers for each different pair $f_{ij}(\boldsymbol{\sigma}) = \sigma_i \sigma_j$, which we will

denote in turn as J_{ij}. For this case, the maximum entropy probability (24) reads,

$$P(\boldsymbol{\sigma}) = \frac{1}{Z} e^{-H(\boldsymbol{\sigma})} \ , \tag{26}$$

$$Z = \sum_{\boldsymbol{\sigma}} e^{-H(\boldsymbol{\sigma})} \ , \tag{27}$$

$$H(\boldsymbol{\sigma}) = -\sum_{i}^{N} h_i \sigma_i - \sum_{i}^{N} \sum_{j<i}^{N} J_{ij} \sigma_i \sigma_j \ . \tag{28}$$

Equations (26)-(28) are the expressions for the least-biased probability distribution that is consistent with the measured values of $\langle \sigma_i \rangle$ and C_{ij} in the IF model. This is the Boltzmann distribution, and is mathematically equivalent to the distribution of spin configurations of a generalized Ising model [9],

$$P_{\text{Ising}}(\boldsymbol{\sigma}) = \frac{1}{Z_{\text{Ising}}} e^{-\frac{H(\boldsymbol{\sigma})}{T}} \ , \tag{29}$$

$$Z_{\text{Ising}} = \sum_{\boldsymbol{\sigma}} e^{-\frac{H(\boldsymbol{\sigma})}{T}} \ , \tag{30}$$

with temperature $T = 1$ in natural units (Boltzmann constant $k_B = 1$), and where $H(\boldsymbol{\sigma})$ is the Hamiltonian or energy function and Z_{Ising} is the partition function, whose sum runs over all 2^N possible configurations of spin states $\boldsymbol{\sigma}$. This prompts us to interpret this maximum entropy description as a mapping from the temporal averages of the neuronal network to that of a fully-connected spin lattice. Under this conceptual view, h_i is analogous to a local external field acting on spin i, whereas J_{ij} is an interaction constant between spins i and j. The task is now to find the values of h_i and J_{ij} that reproduce the measured expectation values from the IF model. This is a particular example of the inverse Ising problem [77–79], also known as Boltzmann Machine (BM) learning [80], which consists in inferring the Hamiltonian of a certain complex multi-component system from its observed statistics. In principle, each parameter h_i and J_{ij} can be determined from the derivative of the logarithm of the partition function (27),

$$\langle \sigma_i \rangle = \frac{\partial}{\partial h_i} \ln[Z] \ , \tag{31}$$

$$\langle \sigma_i \sigma_j \rangle = \frac{\partial}{\partial J_{ij}} \ln[Z] \ . \tag{32}$$

However, the number of terms in Z grows exponentially with N, as 2^N, and an analytical approach becomes intractable when $N \gtrsim 20$. To proceed, notice that we are trying to describe an empirical distribution $P_{\text{IF}}(\boldsymbol{\sigma})$, as observed from the measured temporal averages of the IF model, using the analytical description $P(\boldsymbol{\sigma})$, as given by equation (26). We want to choose the $\{h_i\}$ and $\{J_{ij}\}$ that minimize the loss in information when using $P(\boldsymbol{\sigma})$ as a proxy for $P_{\text{IF}}(\boldsymbol{\sigma})$. Specifically, this means we want the set of $\{h_i\}$ and $\{J_{ij}\}$ that minimize the so-called Kullback-Leibler divergence [81] between these distributions,

$$D_{\text{KL}}\left(P_{\text{IF}}(\boldsymbol{\sigma}) || P(\boldsymbol{\sigma})\right) = \sum_{\boldsymbol{\sigma}} P_{\text{IF}}(\boldsymbol{\sigma}) \ln \left[\frac{P_{\text{IF}}(\boldsymbol{\sigma})}{P(\boldsymbol{\sigma})} \right] \ . \tag{33}$$

From the minimum condition equations one can show that,

$$\frac{\partial D_{\text{KL}}}{\partial h_i} = 0 \implies \langle \sigma_i \rangle^{(\text{IF})} = \langle \sigma_i \rangle^{(\text{BM})} \ , \tag{34}$$

$$\frac{\partial D_{\text{KL}}}{\partial J_{ij}} = 0 \implies \langle \sigma_i \sigma_j \rangle^{(\text{IF})} = \langle \sigma_i \sigma_j \rangle^{(\text{BM})} \ , \tag{35}$$

where $\langle\sigma_i\rangle^{(\mathrm{IF})} \equiv \sum_{\boldsymbol{\sigma}} \sigma_i P_{\mathrm{IF}}(\boldsymbol{\sigma})$ are the empirical averages measured in the IF model and $\langle\sigma_i\rangle^{(\mathrm{BM})} \equiv \sum_{\boldsymbol{\sigma}} \sigma_i P(\boldsymbol{\sigma})$ are the ones estimated from the Boltzmann distribution (26), and analogously for the $\langle\sigma_i\sigma_j\rangle$. Therefore, a suitable method to search for the fields h_i and interaction constants J_{ij} of the Hamiltonian (28) is the following iterative scheme [9],

$$h_i(x+1) = h_i(x) - \eta_h(x) \cdot \left(\langle\sigma_i\rangle^{(\mathrm{BM})} - \langle\sigma_i\rangle^{(\mathrm{IF})} \right) \ , \tag{36}$$

$$J_{ij}(x+1) = J_{ij}(x) - \eta_J(x) \cdot \left(\langle\sigma_i\sigma_j\rangle^{(\mathrm{BM})} - \langle\sigma_i\sigma_j\rangle^{(\mathrm{IF})} \right) \ , \tag{37}$$

where $\eta_h(x) = 2\eta_J(x) \propto x^{-\alpha}$ are decreasing learning rates, with $\alpha = 0.4$ for $N \leq 40$, $\alpha = 0.6$ if $40 < N \leq 120$, and $\alpha = 1.0$ otherwise. We set a smaller learning rate for the $\{J_{ij}\}$ since their number ($\sim N^2$) is much larger when compared to the $\{h_i\}$ (N), so we update their values at a slower rate to avoid divergences during the learning procedure. At each iteration x, the data sets $\{\langle\sigma_i\rangle^{(\mathrm{IF})}\}$ and $\{\langle\sigma_i\sigma_j\rangle^{(\mathrm{IF})}\}$ generated by the IF model are compared to those estimated by sampling the distribution (26), $\{\langle\sigma_i\rangle^{(\mathrm{BM})}\}$ and $\{\langle\sigma_i\sigma_j\rangle^{(\mathrm{BM})}\}$, using the sets of fields $\{h_i(x)\}$ and coupling constants $\{J_{ij}(x)\}$. We start with $h_i(x=1) = \langle\sigma_i\rangle^{(\mathrm{IF})}$ and $J_{ij}(x=1) = 0$ and then iterate equations (36) and (37) typically until $x = N_{\mathrm{BM}} \sim 60000$. At each iteration, $\{\langle\sigma_i\rangle^{(\mathrm{BM})}\}$ and $\{\langle\sigma_i\sigma_j\rangle^{(\mathrm{BM})}\}$ are estimated from Monte Carlo simulations, using the Metropolis algorithm [82, 83], by averaging over $M_c = 3 \cdot 10^5$ spin configurations. We disregard the first $150N$ configurations for systems with $N \leq 120$, or up to $10^5 N^2$ configurations for the $N = 180$ case, in order to reduce correlations with the initial state, and use only every $2N$-th configuration for averaging, to reduce autocorrelations. At the end of the learning routine, we study the generalized Ising models with the set of fitted parameters $\{h_i\}$ and $\{J_{ij}\}$ by sampling the distribution (26), averaging over an increased amount of spin configurations $M_c = 3 \cdot 10^6$ for systems with $N \leq 120$, or up to $M_c = 3 \cdot 10^7$ configurations for the system with $N = 180$, to reduce error bars.

Acknowledgements

L.d.A. acknowledges support from the Italian MUR project PRIN2017WZFTZP and from NEXTGENERATIONEU (NGEU) funded by the Ministry of University and Research (MUR), National Recovery and Resilience Plan (NRRP), and project MNESYS (PE0000006)— A multiscale integrated approach to the study of the nervous system in health and disease (DN. 1553 11.10.2022).

References

1. Koch C, Laurent G. Complexity and the Nervous System. Science. 1999;284(5411):96–98. doi:10.1126/science.284.5411.96.

2. Hodgkin AL, Huxley AF. A quantitative description of membrane current and its application to conduction and excitation in nerve. The Journal of Physiology. 1952;117(4):500. doi:10.1113/jphysiol.1952.sp004764.

3. Rieke F, Warland D, de Ruyter van Steveninck R, Bialek W. Spikes: exploring the neural code. Cambridge, MA, USA: MIT Press; 1999.

4. Dayan P, Abbott LF. Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems. The MIT Press; 2005.

5. Little WA. The existence of persistent states in the brain. Mathematical Biosciences. 1974;19(1):101–120. doi:10.1016/0025-5564(74)90031-5.

6. Hopfield JJ. Neural networks and physical systems with emergent collective computational abilities. Proceedings of the National Academy of Sciences. 1982;79(8):2554–2558. doi:10.1073/pnas.79.8.2554.

7. Amit DJ, Gutfreund H, Sompolinsky H. Spin-glass models of neural networks. Physical Review A. 1985;32(2):1007–1018. doi:10.1103/PhysRevA.32.1007.

8. Schneidman E, Berry MJ, Segev R, Bialek W. Weak pairwise correlations imply strongly correlated network states in a neural population. Nature. 2006;440(7087):1007–1012. doi:10.1038/nature04701.

9. Tkacik G, Schneidman E, Berry II MJ, Bialek W. Spin glass models for a network of real neurons. arXiv:09125409 [q-bio]. 2009;.

10. Tkačik G, Marre O, Mora T, Amodei D, II MJB, Bialek W. The simplest maximum entropy model for collective behavior in a neural network. Journal of Statistical Mechanics: Theory and Experiment. 2013;2013(03):P03011. doi:10.1088/1742-5468/2013/03/P03011.

11. Tkačik G, Marre O, Amodei D, Schneidman E, Bialek W, Ii MJB. Searching for Collective Behavior in a Large Network of Sensory Neurons. PLOS Computational Biology. 2014;10(1):e1003408. doi:10.1371/journal.pcbi.1003408.

12. Tkačik G, Mora T, Marre O, Amodei D, Palmer SE, Berry MJ, et al. Thermodynamics and signatures of criticality in a network of neurons. Proceedings of the National Academy of Sciences. 2015;112(37):11508–11513. doi:10.1073/pnas.1514188112.

13. Meshulam L, Gauthier JL, Brody CD, Tank DW, Bialek W. Collective Behavior of Place and Non-place Neurons in the Hippocampal Network. Neuron. 2017;96(5):1178–1191.e4. doi:10.1016/j.neuron.2017.10.027.

14. Chen X, Randi F, Leifer AM, Bialek W. Searching for collective behavior in a small brain. Physical Review E. 2019;99(5):052418. doi:10.1103/PhysRevE.99.052418.

15. Meshulam L, Gauthier JL, Brody CD, Tank DW, Bialek W. Successes and failures of simple statistical physics models for a network of real neurons; 2023. Available from: http://arxiv.org/abs/2112.14735.

16. Nguyen HC, Zecchina R, Berg J. Inverse statistical problems: from the inverse Ising problem to data science. Advances in Physics. 2017;66(3):197–261. doi:10.1080/00018732.2017.1341604.

17. Jaynes ET. Information Theory and Statistical Mechanics. Physical Review. 1957;106(4):62–79. doi:10.1103/PhysRev.106.620.

18. Cocco S, Leibler S, Monasson R. Neuronal couplings between retinal ganglion cells inferred by efficient inverse statistical physics methods. Proceedings of the National Academy of Sciences. 2009;106(33):14058–14062. doi:10.1073/pnas.0906705106.

19. Ganmor E, Segev R, Schneidman E. The Architecture of Functional Interaction Networks in the Retina. Journal of Neuroscience. 2011;31(8):3044–3054. doi:10.1523/JNEUROSCI.3682-10.2011.

20. Humplik J, Tkačik G. Probabilistic models for neural populations that naturally capture global coupling and criticality. PLOS Computational Biology. 2017;13(9):e1005763. doi:10.1371/journal.pcbi.1005763.

21. Delamare G, Ferrari U. Time-Dependent Maximum Entropy Model for Populations of Retinal Ganglion Cells†; 2022. Available from: https://www.biorxiv.org/content/10.1101/2022.07.13.498395v1.

22. Nonnenmacher M, Behrens C, Berens P, Bethge M, Macke JH. Signatures of criticality arise from random subsampling in simple population models. PLOS Computational Biology. 2017;13(10):e1005718. doi:10.1371/journal.pcbi.1005718.

23. Hahn G, Ponce-Alvarez A, Monier C, Benvenuti G, Kumar A, Chavane F, et al. Spontaneous cortical activity is transiently poised close to criticality. PLOS Computational Biology. 2017;13(5):e1005543. doi:10.1371/journal.pcbi.1005543.

24. Sampaio Filho CIN, de Arcangelis LdA, Herrmann HJ, Plenz DP, Kells P, Ribeiro TL, et al.. Ising-like model replicating time-averaged spiking behaviour of in vivo and in vitro neuronal networks; 2023.

25. White JG, Southgate E, Thomson JN, Brenner S. The structure of the nervous system of the nematode Caenorhabditis elegans. Philosophical Transactions of the Royal Society of London B, Biological Sciences. 1986;314(1165):1–340. doi:10.1098/rstb.1986.0056.

26. Granot-Atedgi E, Tkačik G, Segev R, Schneidman E. Stimulus-dependent Maximum Entropy Models of Neural Population Codes. PLOS Computational Biology. 2013;9(3):e1002922. doi:10.1371/journal.pcbi.1002922.

27. Gardella C, Marre O, Mora T. Modeling the Correlated Activity of Neural Populations: A Review. Neural Computation. 2019;31(2):233–269. doi:10.1162/neco_a_01154.

28. Buccino AP, Garcia S, Yger P. Spike sorting: new trends and challenges of the era of high-density probes. Progress in Biomedical Engineering. 2022;4(2):022005. doi:10.1088/2516-1091/ac6b96.

29. Ventura V, Gerkin RC. Accurately estimating neuronal correlation requires a new spike-sorting paradigm. Proceedings of the National Academy of Sciences. 2012;109(19):7230–7235. doi:10.1073/pnas.1115236109.

30. Pastore VP, Massobrio P, Godjoski A, Martinoia S. Identification of excitatory-inhibitory links and network topology in large-scale neuronal assemblies from multi-electrode recordings. PLOS Computational Biology. 2018;14(8):e1006381. doi:10.1371/journal.pcbi.1006381.

31. Michiels van Kessenich L, Luković M, de Arcangelis L, Herrmann HJ. Critical neural networks with short- and long-term plasticity. Physical Review E. 2018;97(3):032312. doi:10.1103/PhysRevE.97.032312.

32. Eurich CW, Herrmann JM, Ernst UA. Finite-size effects of avalanche dynamics. Physical Review E. 2002;66(6):066137. doi:10.1103/PhysRevE.66.066137.

33. Beggs JM, Plenz D. Neuronal Avalanches in Neocortical Circuits. The Journal of Neuroscience. 2003;23(35):11167–11177. doi:10.1523/JNEUROSCI.23-35-11167.2003.

34. de Arcangelis L, Perrone-Capano C, Herrmann HJ. Self-Organized Criticality Model for Brain Plasticity. Physical Review Letters. 2006;96(2):028107. doi:10.1103/PhysRevLett.96.028107.

35. Levina A, Herrmann JM, Geisel T. Dynamical synapses causing self-organized criticality in neural networks. Nature Physics. 2007;3(12):857–860. doi:10.1038/nphys758.

36. Chialvo DR, Bak P. Learning from mistakes. Neuroscience. 1999;90(4):1137–1148. doi:10.1016/S0306-4522(98)00472-2.

37. Shew WL, Yang H, Yu S, Roy R, Plenz D. Information Capacity and Transmission Are Maximized in Balanced Cortical Networks with Neuronal Avalanches. Journal of Neuroscience. 2011;31(1):55–63. doi:10.1523/JNEUROSCI.4637-10.2011.

38. Haimovici A, Tagliazucchi E, Balenzuela P, Chialvo DR. Brain Organization into Resting State Networks Emerges at Criticality on a Model of the Human Connectome. Physical Review Letters. 2013;110(17):178101. doi:10.1103/PhysRevLett.110.178101.

39. Priesemann V, Valderrama M, Wibral M, Quyen MLV. Neuronal Avalanches Differ from Wakefulness to Deep Sleep – Evidence from Intracranial Depth Recordings in Humans. PLOS Computational Biology. 2013;9(3):e1002985. doi:10.1371/journal.pcbi.1002985.

40. Shriki O, Alstott J, Carver F, Holroyd T, Henson RNA, Smith ML, et al. Neuronal Avalanches in the Resting MEG of the Human Brain. Journal of Neuroscience. 2013;33(16):7079–7090. doi:10.1523/JNEUROSCI.4286-12.2013.

41. Yu S, Yang H, Shriki O, Plenz D. Universal organization of resting brain activity at the thermodynamic critical point. Frontiers in Systems Neuroscience. 2013;7.

42. Petermann T, Thiagarajan TC, Lebedev MA, Nicolelis MAL, Chialvo DR, Plenz D. Spontaneous cortical activity in awake monkeys composed of neuronal avalanches. Proceedings of the National Academy of Sciences. 2009;106(37):15921–15926. doi:10.1073/pnas.0904089106.

43. Hahn G, Petermann T, Havenith MN, Yu S, Singer W, Plenz D, et al. Neuronal Avalanches in Spontaneous Activity In Vivo. Journal of Neurophysiology. 2010;104(6):3312–3322. doi:10.1152/jn.00953.2009.

44. Mora T, Deny S, Marre O. Dynamical Criticality in the Collective Activity of a Population of Retinal Neurons. Physical Review Letters. 2015;114(7):078105. doi:10.1103/PhysRevLett.114.078105.

45. Shew WL, Clawson WP, Pobst J, Karimipanah Y, Wright NC, Wessel R. Adaptation to sensory input tunes visual cortex to criticality. Nature Physics. 2015;11(8):659–663. doi:10.1038/nphys3370.

46. Aguilera M, Alquézar C, Izquierdo EJ. Signatures of criticality in a maximum entropy model of the C. elegans brain during free behaviour. MIT Press; 2017. p. 29–35. Available from: https://dx.doi.org/10.1162/isal_a_010.

47. Ponce-Alvarez A, Jouary A, Privat M, Deco G, Sumbre G. Whole-Brain Neuronal Activity Displays Crackling Noise Dynamics. Neuron. 2018;100(6):1446–1459.e6. doi:10.1016/j.neuron.2018.10.045.

48. Michiels van Kessenich L, de Arcangelis L, Herrmann HJ. Synaptic plasticity and neuronal refractory time cause scaling behaviour of neuronal avalanches. Scientific Reports. 2016;6(1):32071. doi:10.1038/srep32071.

49. Zapperi S, Lauritsen KB, Stanley HE. Self-Organized Branching Processes: Mean-Field Theory for Avalanches. Physical Review Letters. 1995;75(22):4071–4074. doi:10.1103/PhysRevLett.75.4071.

50. Mazzoni A, Broccard FD, Garcia-Perez E, Bonifazi P, Ruaro ME, Torre V. On the Dynamics of the Spontaneous Activity in Neuronal Networks. PLoS ONE. 2007;2(5):e439. doi:10.1371/journal.pone.0000439.

51. Sherrington D, Kirkpatrick S. Solvable Model of a Spin-Glass. Physical Review Letters. 1975;35(26):1792–1796. doi:10.1103/PhysRevLett.35.1792.

52. Mezard M, Parisi G, Virasoro M. Spin Glass Theory and Beyond: An Introduction to the Replica Method and Its Applications. vol. 9 of World Scientific Lecture Notes in Physics. WORLD SCIENTIFIC; 1986. Available from: http://www.worldscientific.com/worldscibooks/10.1142/0271.

53. Böttcher L, Herrmann HJ. Computational Statistical Physics. Cambridge: Cambridge University Press; 2021. Available from: https://www.cambridge.org/core/books/computational-statistical-physics/A094EE101BEC246EC313EB9638F85EA5.

54. Raimo D, Sarracino A, de Arcangelis L. Role of inhibitory neurons in temporal correlations of critical and supercritical spontaneous activity. Physica A: Statistical Mechanics and its Applications. 2021;565:125555. doi:10.1016/j.physa.2020.125555.

55. Nandi MK, Sarracino A, Herrmann HJ, de Arcangelis L. Scaling of avalanche shape and activity power spectrum in neuronal networks. Physical Review E. 2022;106(2):024304. doi:10.1103/PhysRevE.106.024304.

56. Yeh FC, Tang A, Hobbs JP, Hottowy P, Dabrowski W, Sher A, et al. Maximum Entropy Approaches to Living Neural Networks. Entropy. 2010;12(1):89–106. doi:10.3390/e12010089.

57. van der Plas TL, Tubiana J, Le Goc G, Migault G, Kunst M, Baier H, et al. Neural assemblies uncovered by generative modeling explain whole-brain activity statistics and reflect structural connectivity. eLife. 2023;12:e83139. doi:10.7554/eLife.83139.

58. Maoz O, Tkačik G, Esteki MS, Kiani R, Schneidman E. Learning probabilistic neural representations with randomly connected circuits. Proceedings of the National Academy of Sciences. 2020;117(40):25066–25073. doi:10.1073/pnas.1912804117.

59. Lynn CW, Yu Q, Pang R, Bialek W, Palmer SE. Exactly solvable statistical physics models for large neuronal populations; 2023. Available from: http://arxiv.org/abs/2310.10860.

60. Berger D, Varriale E, van Kessenich LM, Herrmann HJ, de Arcangelis L. Three cooperative mechanisms required for recovery after brain damage. Scientific Reports. 2019;9(1):15858. doi:10.1038/s41598-019-50946-y.

61. Eguíluz VM, Chialvo DR, Cecchi GA, Baliki M, Apkarian AV. Scale-Free Brain Functional Networks. Physical Review Letters. 2005;94(1):018102. doi:10.1103/PhysRevLett.94.018102.

62. Roerig B, Chen B. Relationships of Local Inhibitory and Excitatory Circuits to Orientation Preference Maps in Ferret Visual Cortex. Cerebral Cortex. 2002;12(2):187–198. doi:10.1093/cercor/12.2.187.

63. Zucker RS, Regehr WG. Short-Term Synaptic Plasticity. Annual Review of Physiology. 2002;64(1):355–405. doi:10.1146/annurev.physiol.64.092501.114547.

64. Rosenmund C, Stevens CF. Definition of the Readily Releasable Pool of Vesicles at Hippocampal Synapses. Neuron. 1996;16(6):1197–1207. doi:10.1016/S0896-6273(00)80146-4.

65. Kaeser PS, Regehr WG. The readily releasable pool of synaptic vesicles. Current Opinion in Neurobiology. 2017;43:63–70. doi:10.1016/j.conb.2016.12.012.

66. Abraham WC. How long will long-term potentiation last? Philosophical Transactions of the Royal Society of London Series B: Biological Sciences. 2003;358(1432):735–744. doi:10.1098/rstb.2002.1222.

67. Ikeda K, Bekkers JM. Counting the number of releasable synaptic vesicles in a presynaptic terminal. Proceedings of the National Academy of Sciences. 2009;106(8):2945–2950. doi:10.1073/pnas.0811017106.

68. Lisman JE, Raghavachari S, Tsien RW. The sequence of events that underlie quantal transmission at central glutamatergic synapses. Nature Reviews Neuroscience. 2007;8(8):597–609. doi:10.1038/nrn2191.

69. Lemaréchal JD, Jedynak M, Trebaul L, Boyer A, Tadel F, Bhattacharjee M, et al. A brain atlas of axonal and synaptic delays based on modelling of cortico-cortical evoked potentials. Brain. 2021;145(5):1653–1667. doi:10.1093/brain/awab362.

70. Markram H, Tsodyks M. Redistribution of synaptic efficacy between neocortical pyramidal neurons. Nature. 1996;382(6594):807–810. doi:10.1038/382807a0.

71. Zeraati R, Priesemann V, Levina A. Self-Organization Toward Criticality by Synaptic Plasticity. Frontiers in Physics. 2021;9.

72. Hebb DO. The Organization of Behavior. New York: John Wiley & Sons, Inc.; 1949.

73. Bliss TVP, Lømo T. Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. The Journal of Physiology. 1973;232(2):331–356. doi:10.1113/jphysiol.1973.sp010273.

74. Bi Gq, Poo Mm. Synaptic Modifications in Cultured Hippocampal Neurons: Dependence on Spike Timing, Synaptic Strength, and Postsynaptic Cell Type. Journal of Neuroscience. 1998;18(24):10464–10472. doi:10.1523/JNEUROSCI.18-24-10464.1998.

75. Shannon CE. A mathematical theory of communication. The Bell System Technical Journal. 1948;27(3):379–423. doi:10.1002/j.1538-7305.1948.tb01338.x.

76. Gelfand IM, Fomin SV. Calculus of Variations. Dover Publications, New York; 2000.

77. Keller JB, Zumino B. Determination of Intermolecular Potentials from Thermodynamic Data and the Law of Corresponding States. The Journal of Chemical Physics. 1959;30(5):1351–1353. doi:10.1063/1.1730184.

78. Aurell E, Ekeberg M. Inverse Ising Inference Using All the Data. Physical Review Letters. 2012;108(9):090201. doi:10.1103/PhysRevLett.108.090201.

79. Castellana M, Bialek W. Inverse Spin Glass and Related Maximum Entropy Problems. Physical Review Letters. 2014;113(11):117204. doi:10.1103/PhysRevLett.113.117204.

80. E Rumelhart D, L McClelland J. Learning and Relearning in Boltzmann Machines. In: Parallel Distributed Processing: Explorations in the Microstructure of Cognition: Foundations. MIT Press; 1987. p. 282–317.

81. Kullback S, Leibler RA. On Information and Sufficiency. The Annals of Mathematical Statistics. 1951;22(1):79–86. doi:10.1214/aoms/1177729694.

82. Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E. Equation of State Calculations by Fast Computing Machines. The Journal of Chemical Physics. 1953;21(6):1087–1092. doi:10.1063/1.1699114.

83. Binder K, editor. Monte Carlo Methods in Statistical Physics. vol. 7 of Topics in Current Physics. 2nd ed. Berlin, Heidelberg: Springer; 1986. Available from: https://link.springer.com/10.1007/978-3-642-82803-4.

Supporting information

S1 Fig. Avalanche size and duration distributions of critical fully-excitatory IF networks.

S1 Table. Values of $\delta u_{\mathbf{rec}}$ for the fully-excitatory systems considered in S1 Fig

S2 Fig. Avalanche size and duration distributions of critical IF networks with different fractions of inhibitory neurons.

S2 Table. Values of $\delta u_{\mathbf{rec}}$ for the systems with $N = 80$ considered in S2 Fig

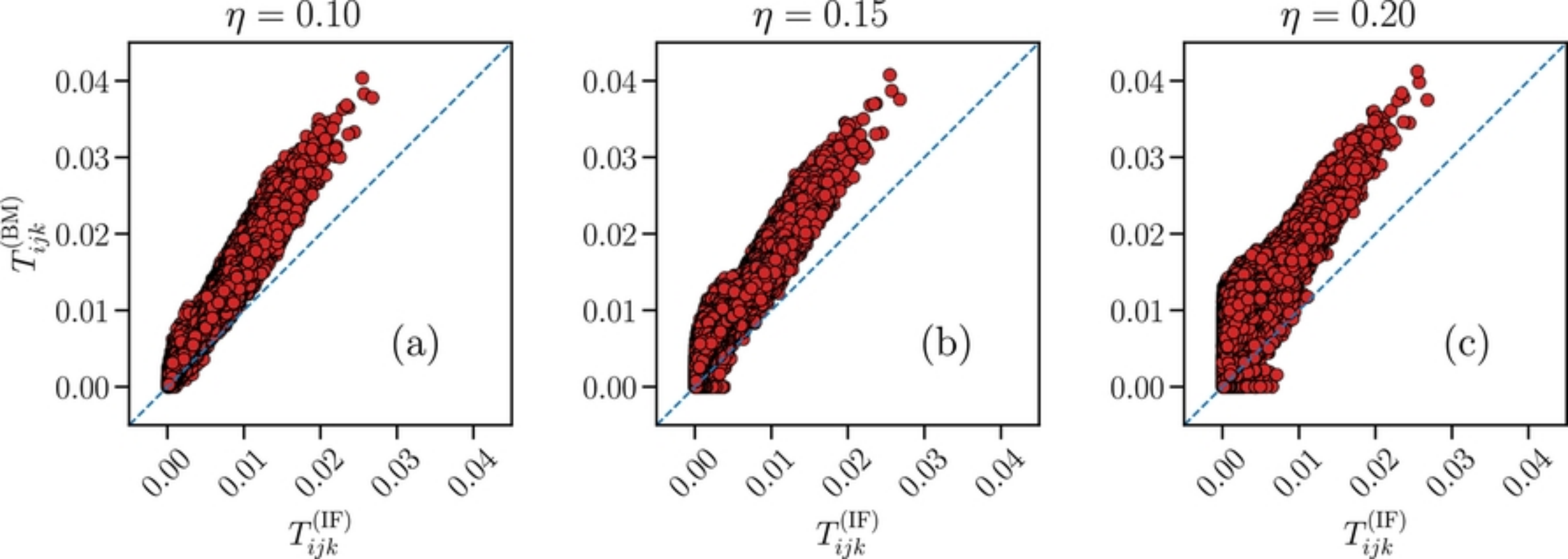
S3 Fig. Dependence with the initial spin configuration of the thermodynamic functions of Ising-like models associated with fully-excitatory IF networks.

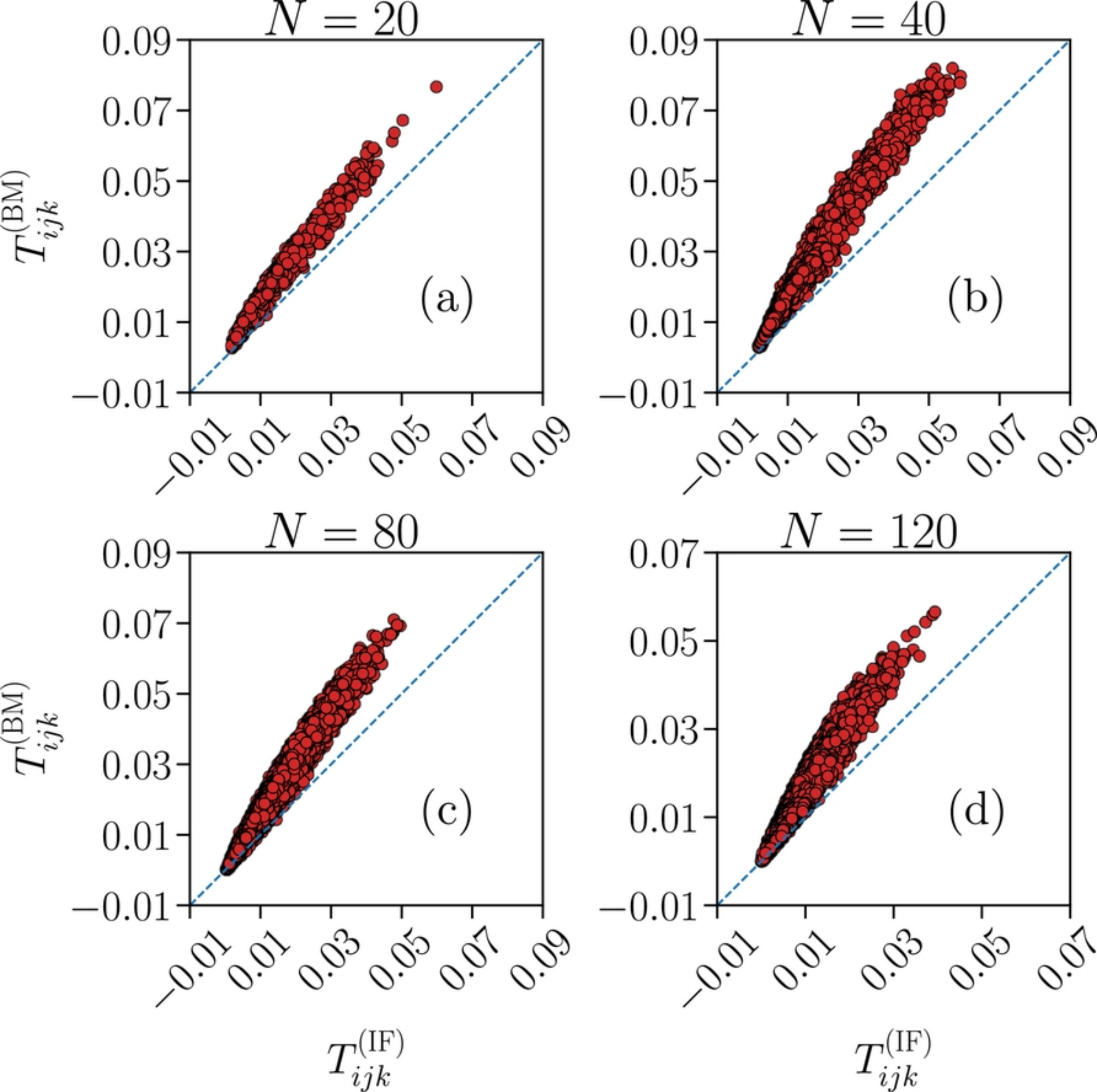
S4 Fig. Quality test for the BM learning process for IF subnetworks.

S5 Fig. Quality test for the BM learning process for IF subnetworks with different fractions of inhibitory neurons.

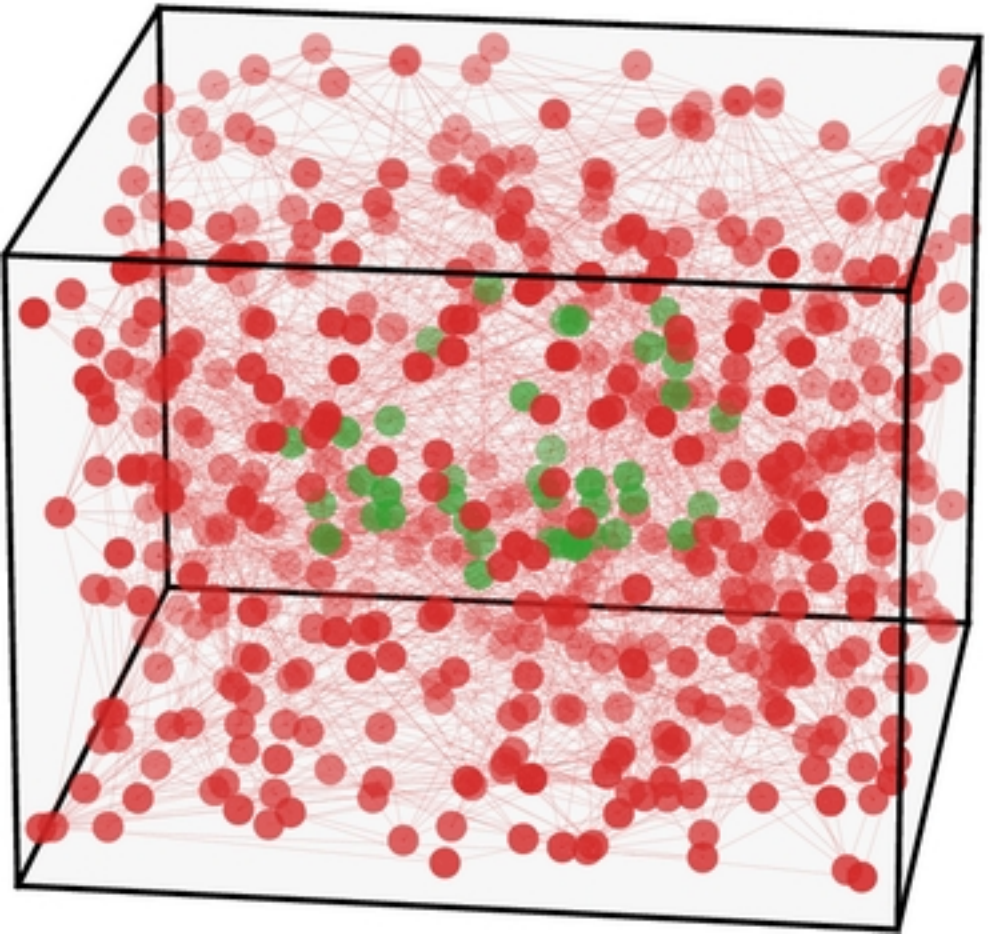
S6 Fig. Learned fields h_i of the pairwise Ising model associated with an IF network with inhibitory neurons.

S7 Fig. Quality test for the BM learning process with a partially-connected Ising model of the average local activities.

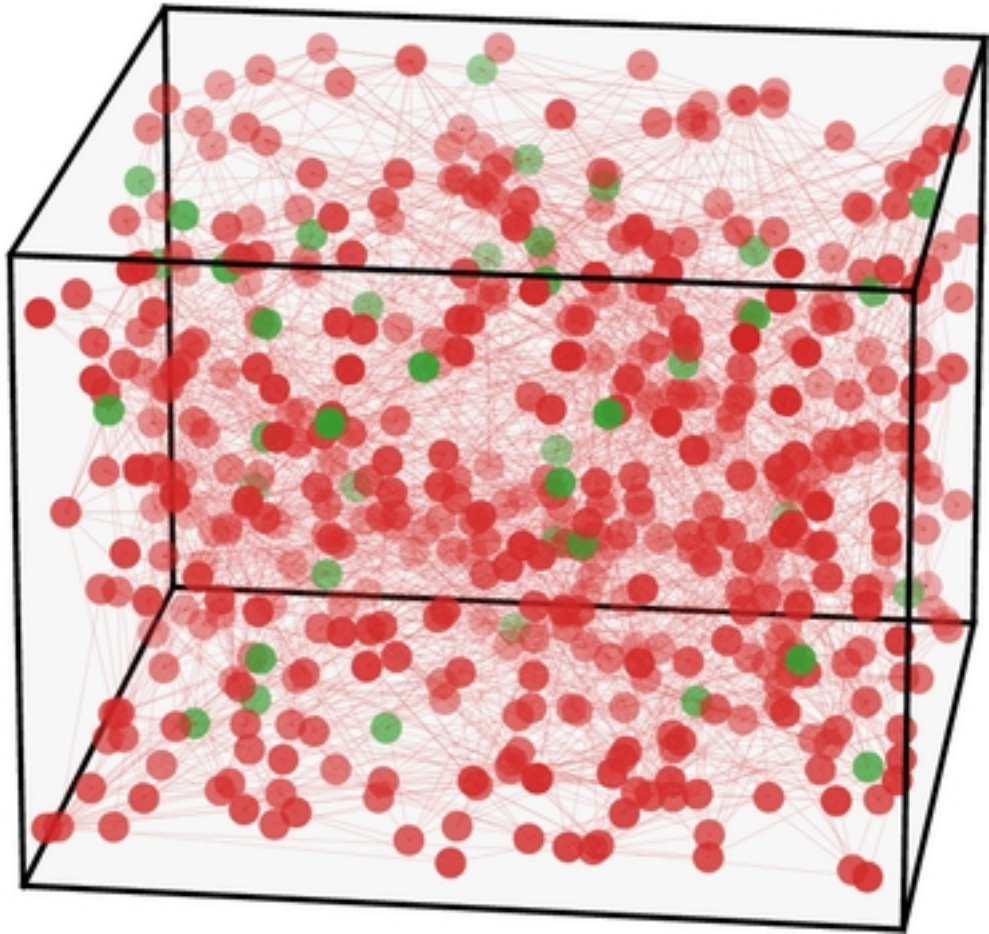

$\eta = 0.10$
$\eta = 0.15$
$\eta = 0.20$
(a)
(b)
(c)
$T_{ijk}^{(\mathrm{BM})}$
$T_{ijk}^{(\mathrm{IF})}$
0.00
0.01
0.02
0.03
0.04

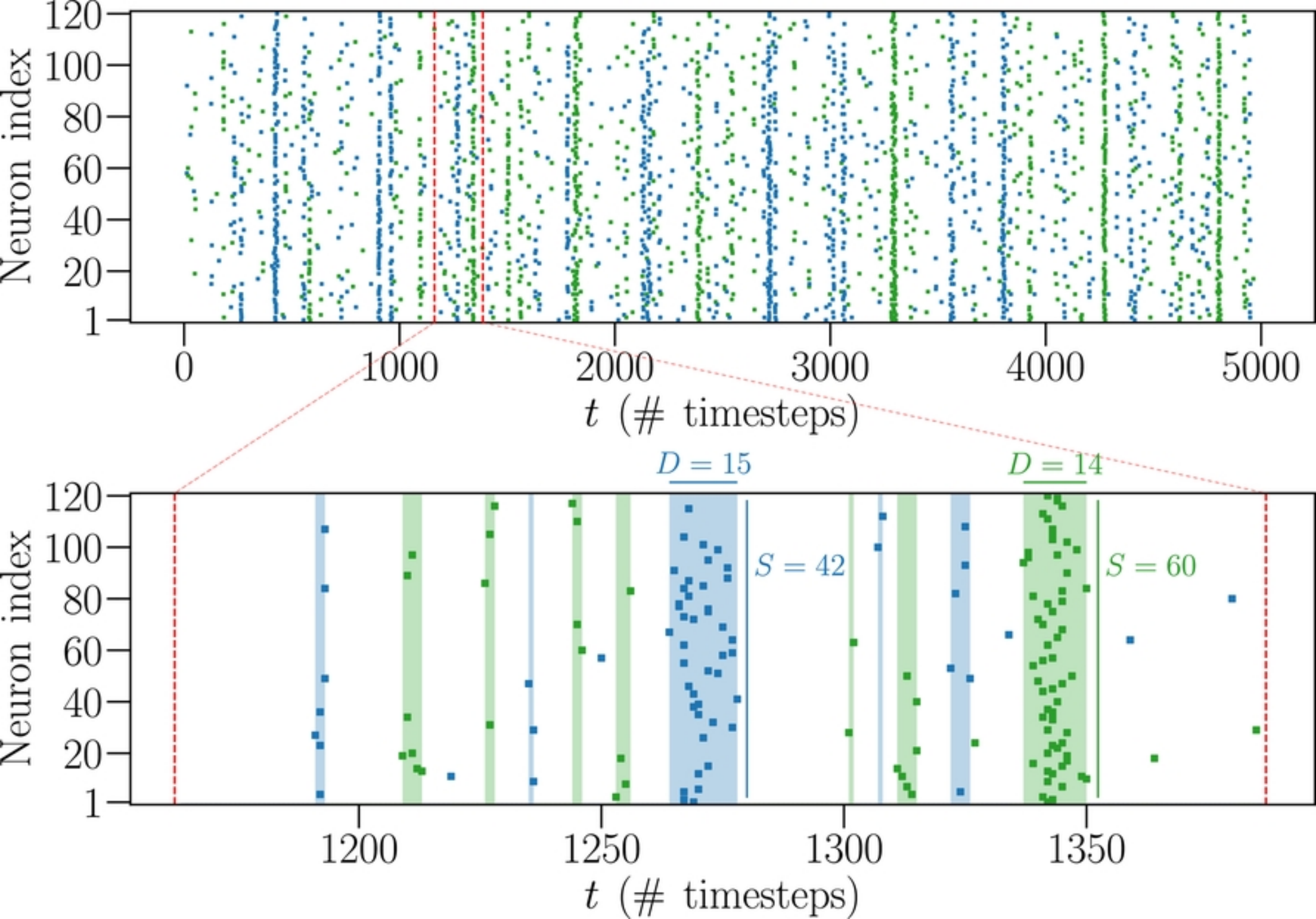

$N = 20$
$N = 40$
$N = 80$
$N = 120$
(a)
(b)
(c)
(d)
$T_{ijk}^{(\mathrm{BM})}$
$T_{ijk}^{(\mathrm{IF})}$
0.09
0.07
0.05
0.03
0.01
−0.01

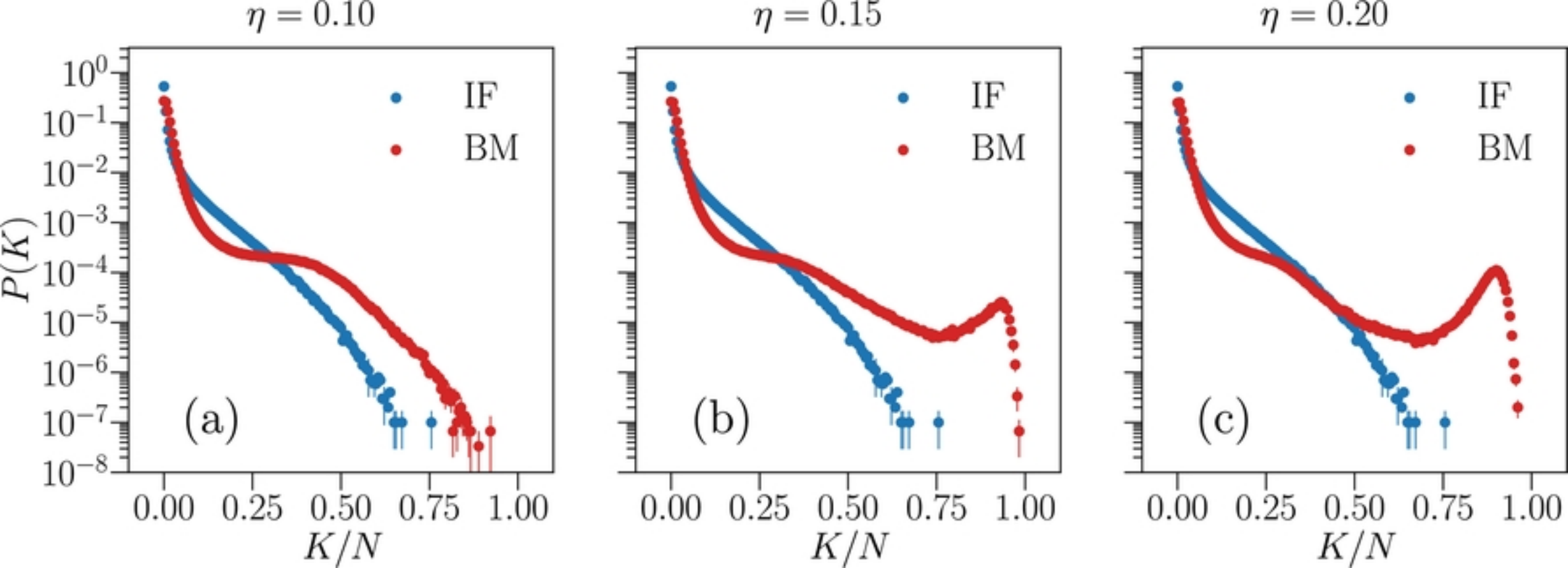
(a) Local

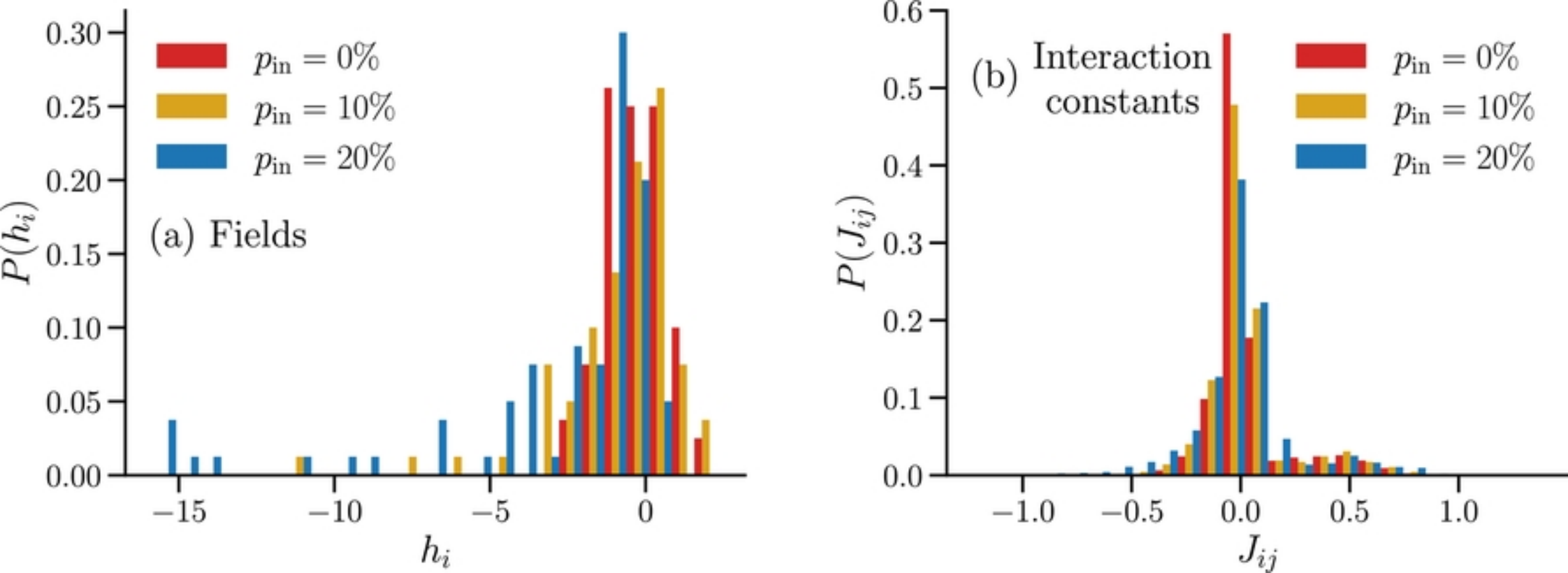


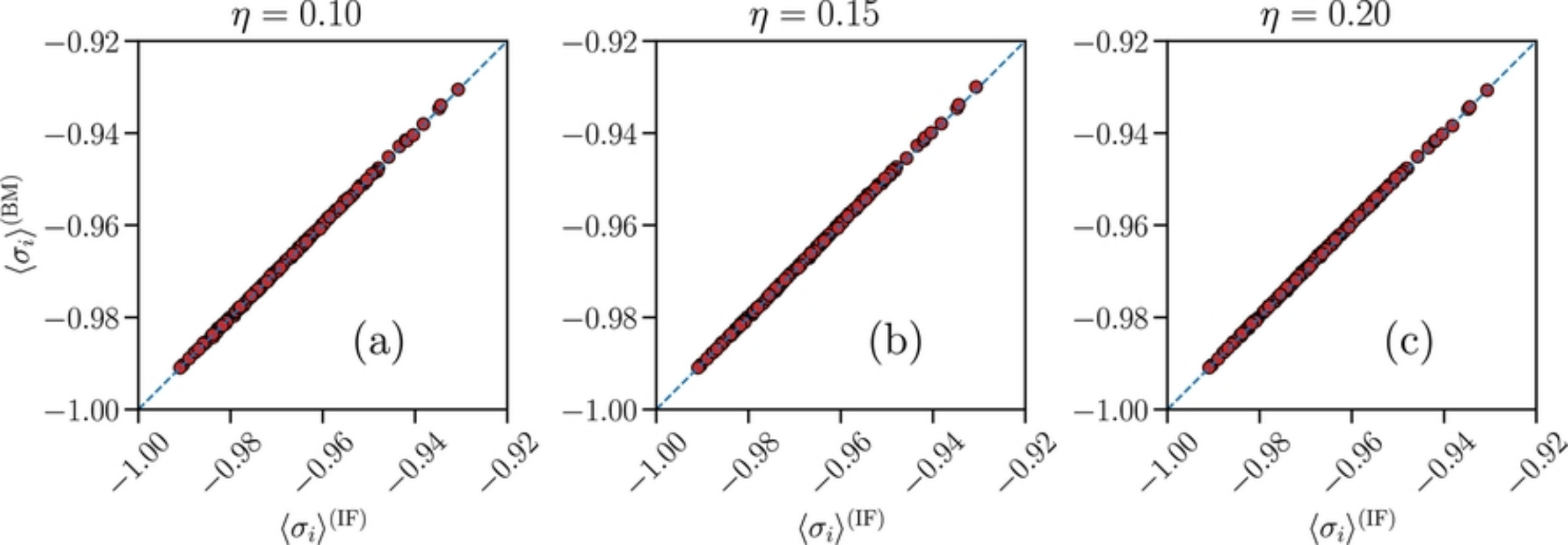
(b) Random

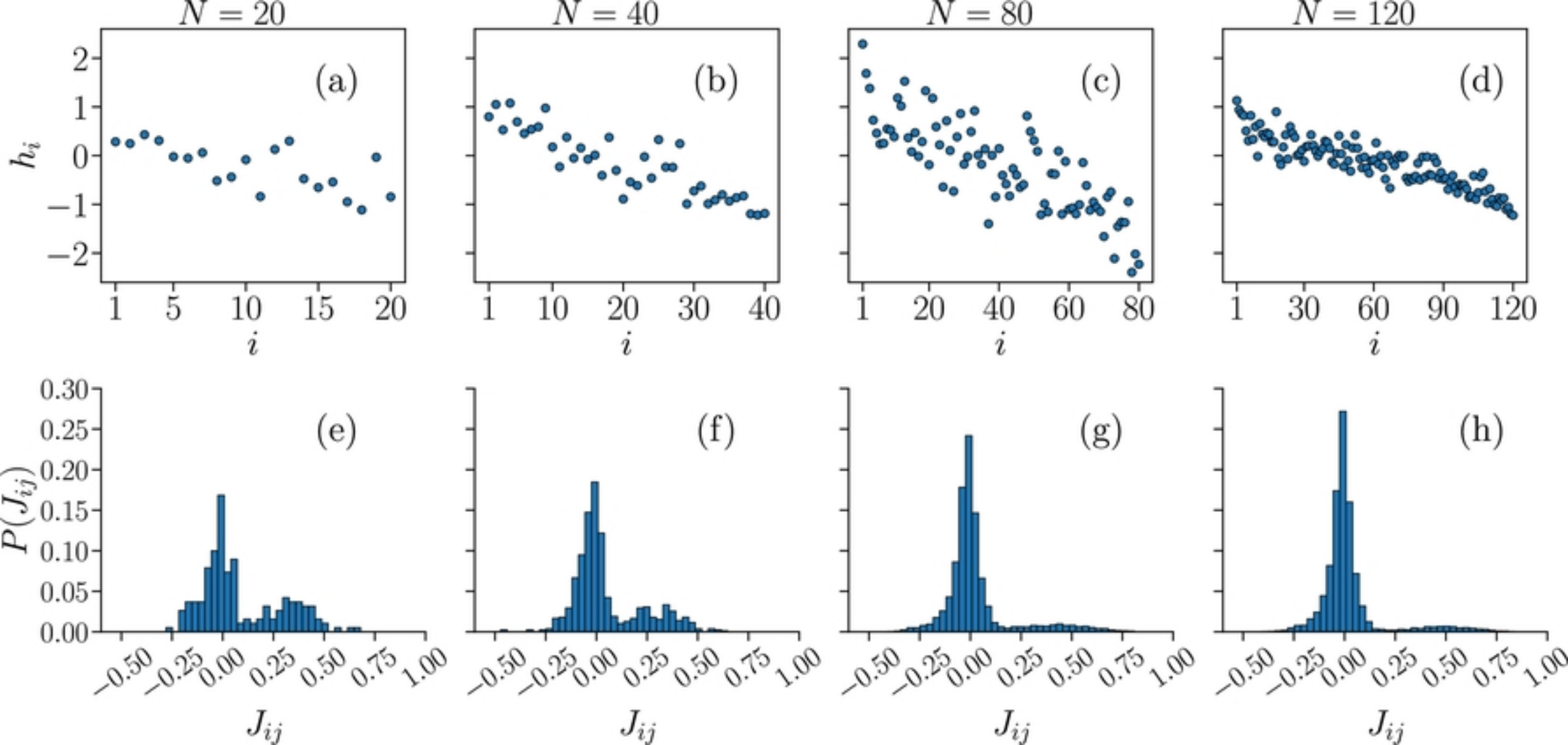


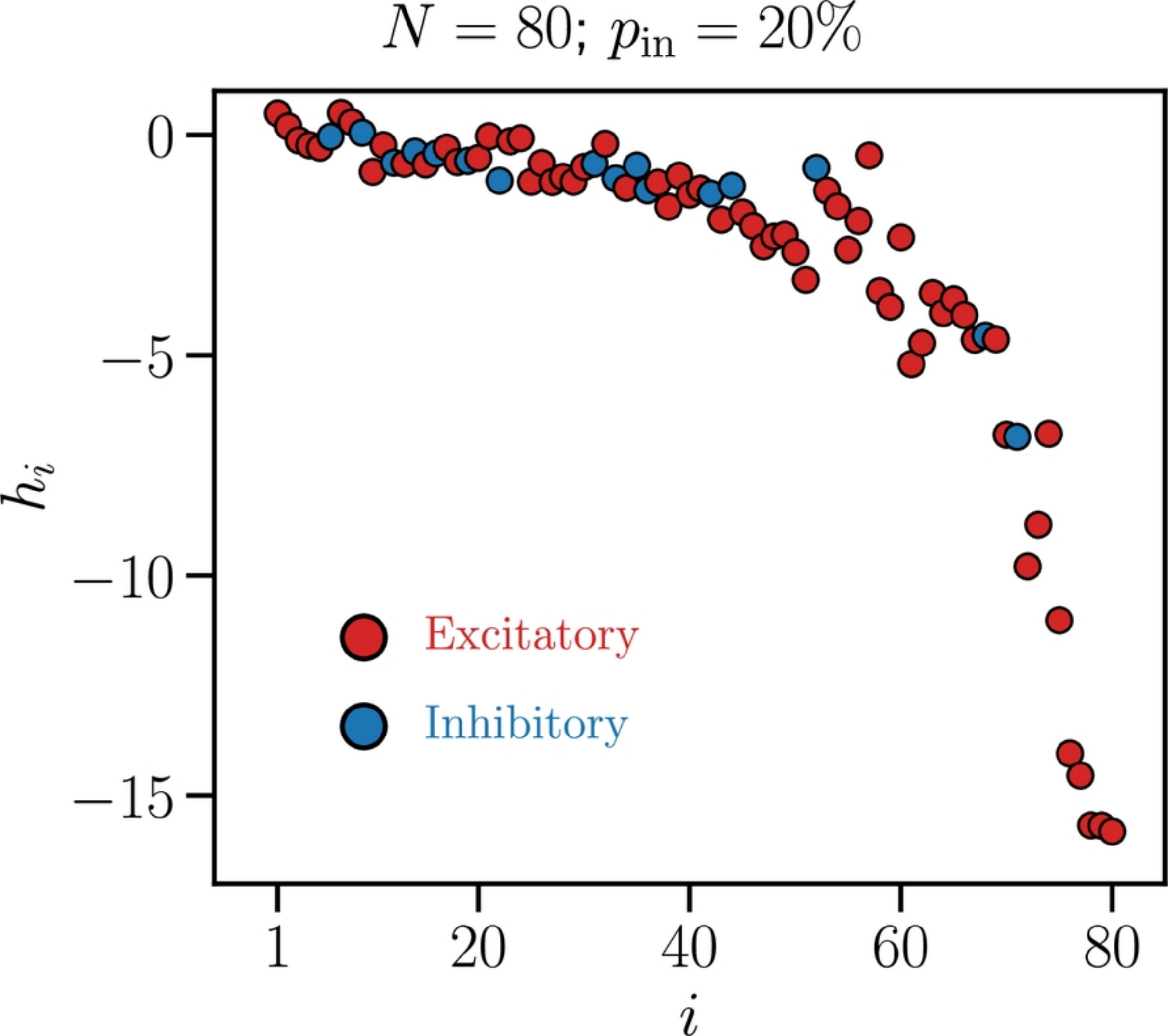

Neuron index
120
100
80
60
40
20
1
0
1000
2000
3000
4000
5000
t (# timesteps)
D = 15
D = 14
S = 42
S = 60
1200
1250
1300
1350

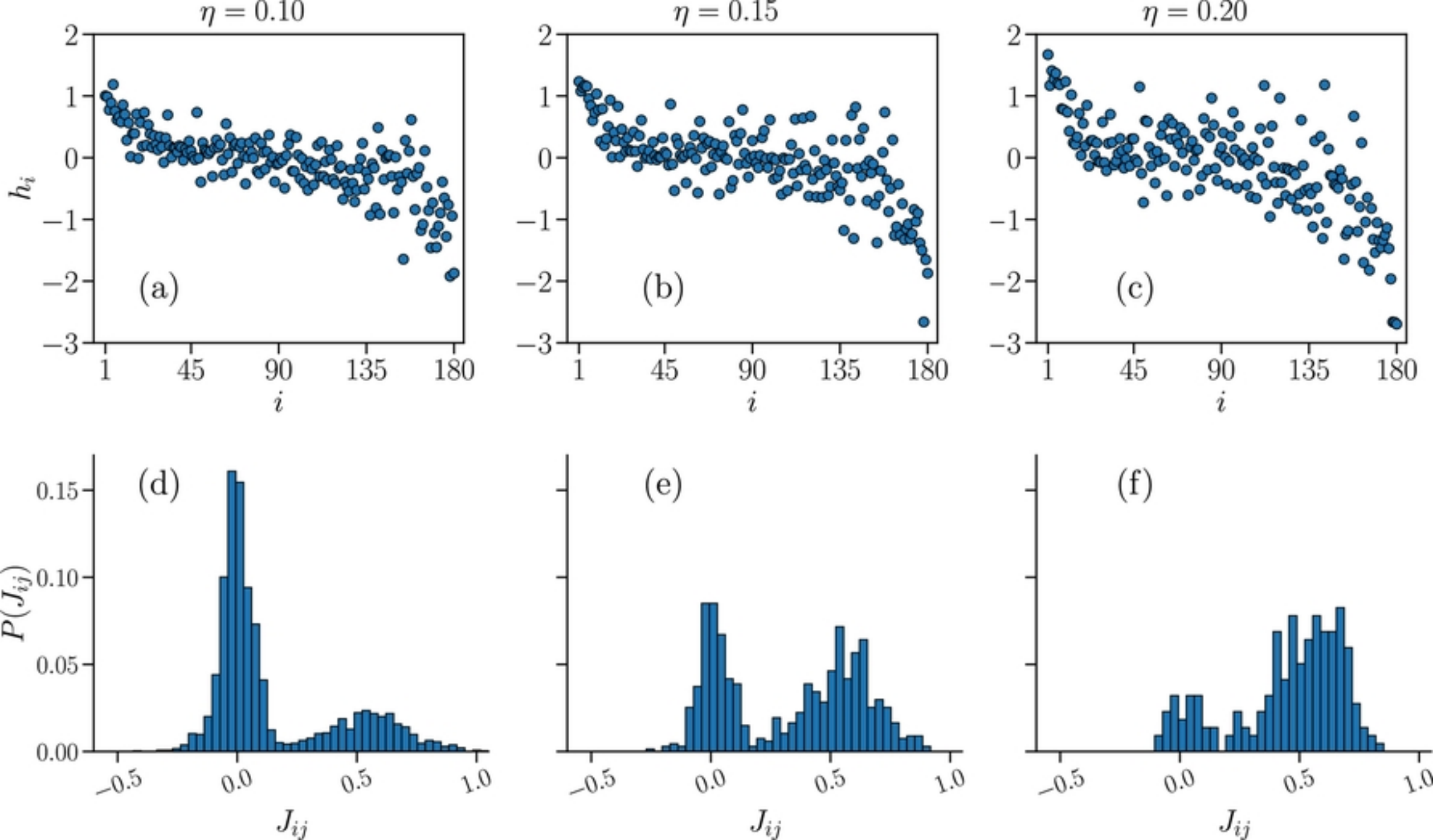

$\eta = 0.10$
$\eta = 0.15$
$\eta = 0.20$
IF
BM
(a)
(b)
(c)
$P(K)$
K/N
10^0
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{-7}
10^{-8}
0.00
0.25
0.50
0.75
1.00

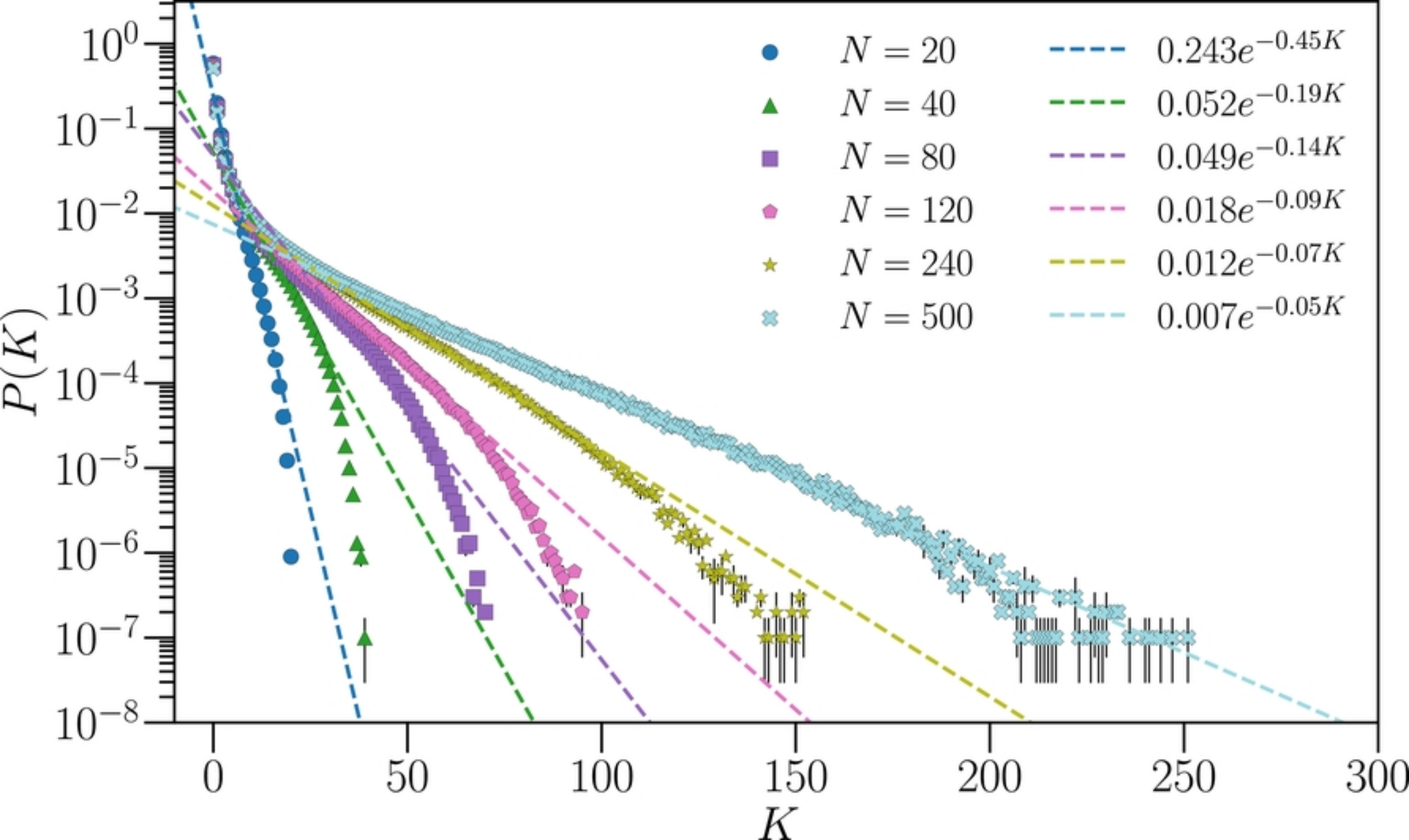

(a) Fields
$p_{\text{in}} = 0\%$
$p_{\text{in}} = 10\%$
$p_{\text{in}} = 20\%$
$P(h_i)$
h_i
0.30
0.25
0.20
0.15
0.10
0.05
0.00
−15
−10
−5
0
(b) Interaction constants
$p_{\text{in}} = 0\%$
$p_{\text{in}} = 10\%$
$p_{\text{in}} = 20\%$
$P(J_{ij})$
J_{ij}
0.6
0.5
0.4
0.3
0.2
0.1
0.0
−1.0
−0.5
0.0
0.5
1.0

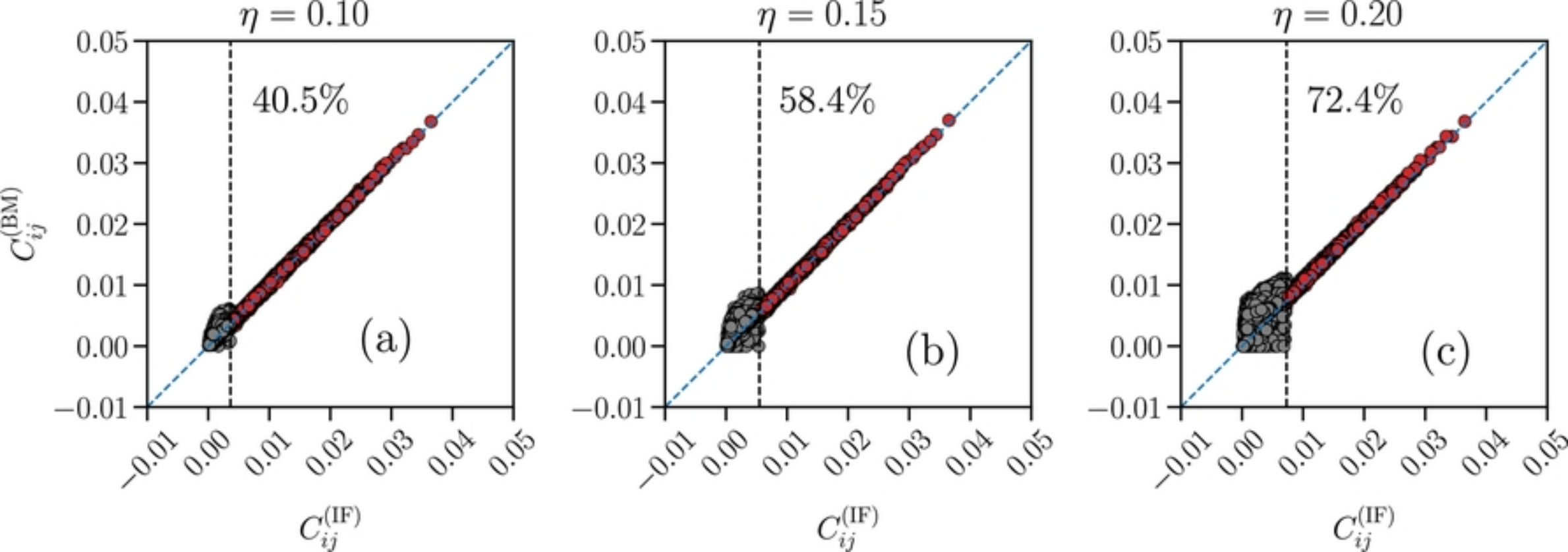

$\eta = 0.10$
$\eta = 0.15$
$\eta = 0.20$
(a)
(b)
(c)
$\langle \sigma_i \rangle^{(\mathrm{BM})}$
$\langle \sigma_i \rangle^{(\mathrm{IF})}$
−0.92
−0.94
−0.96
−0.98
−1.00

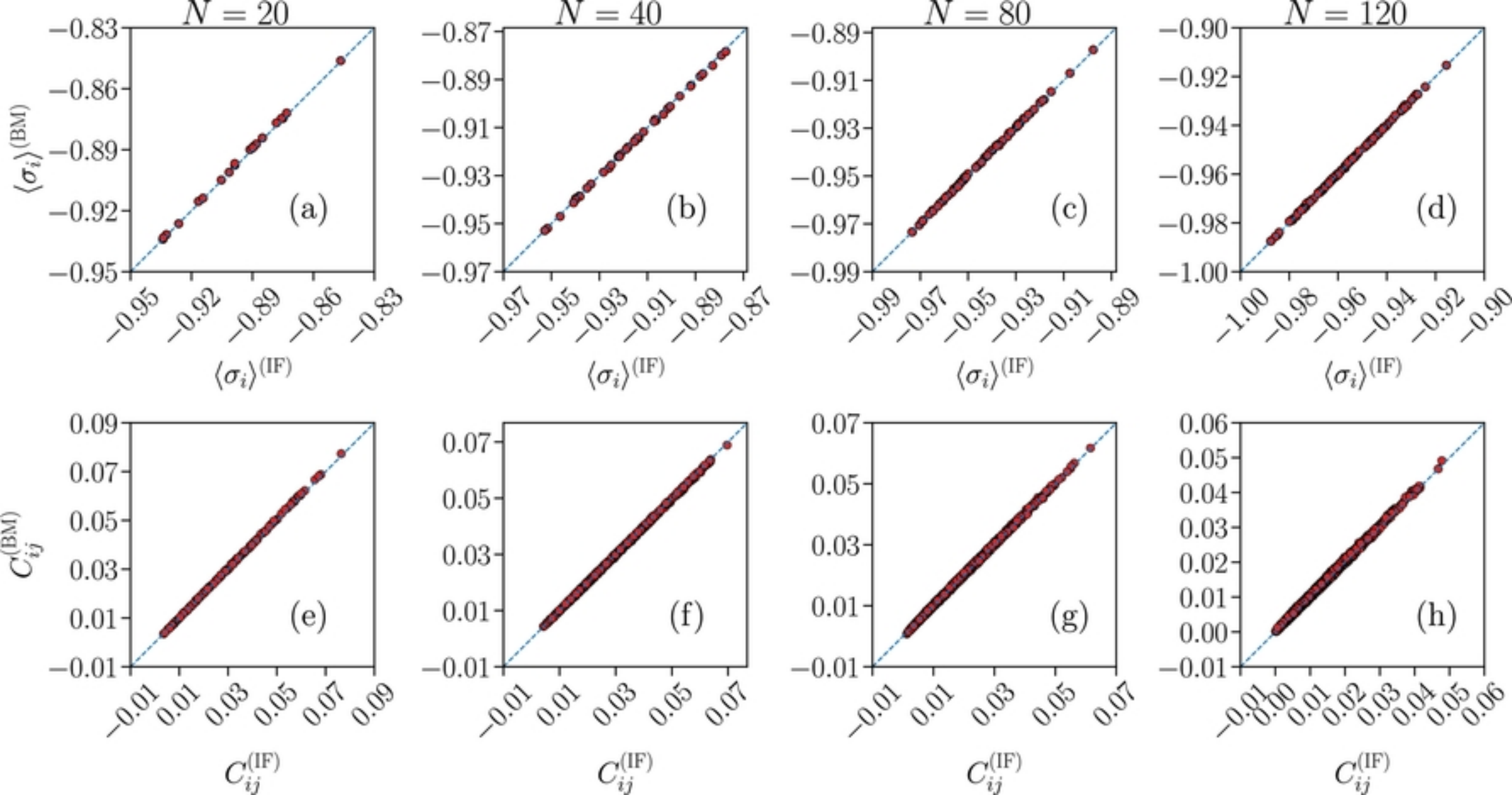

$N = 20$
$N = 40$
$N = 80$
$N = 120$
(a)
(b)
(c)
(d)
h_i
i
(e)
(f)
(g)
(h)
$P(J_{ij})$
J_{ij}

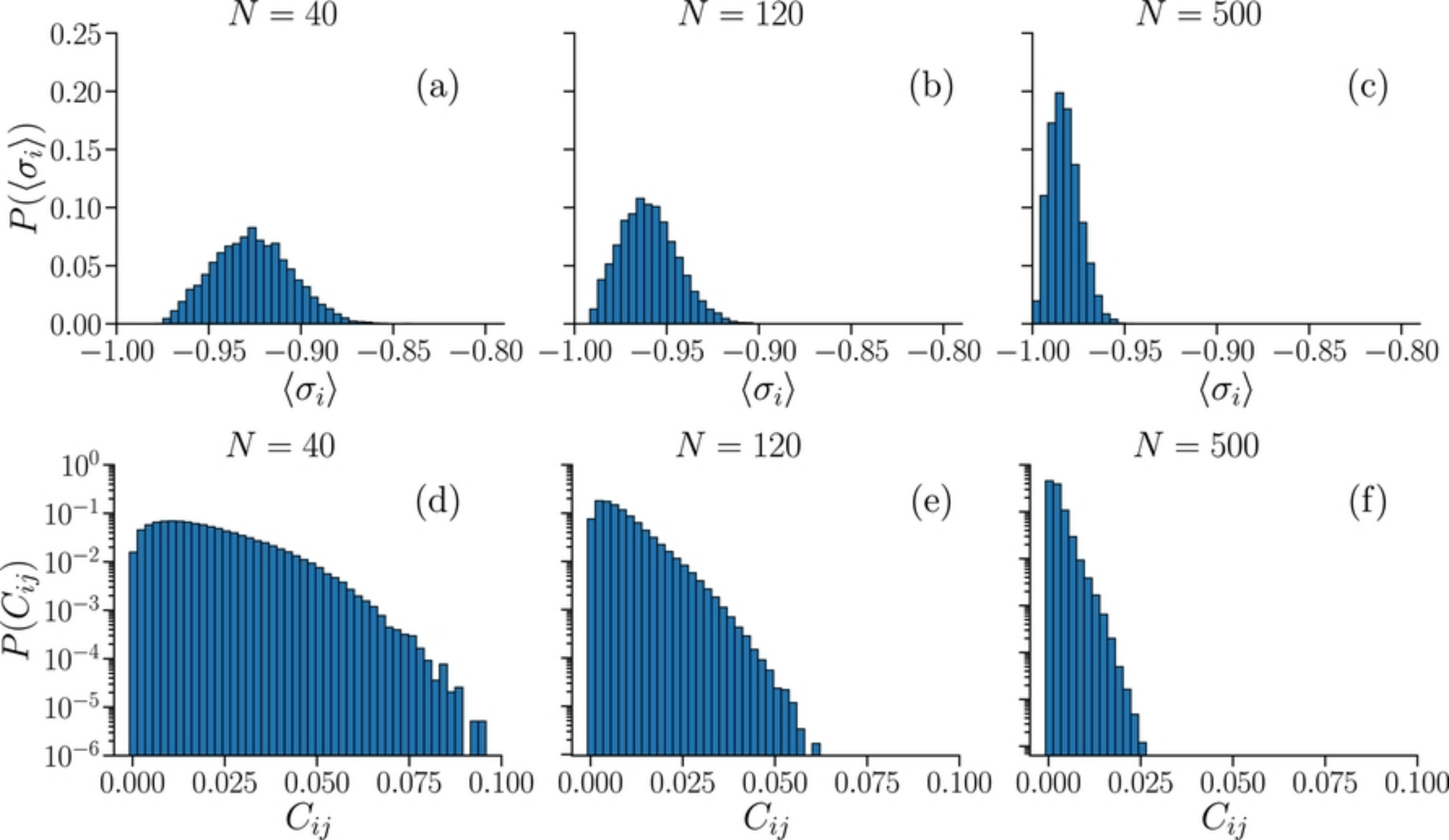

$N = 80$; $p_{\mathrm{in}} = 20\%$
0
−5
−10
−15
h_i
Excitatory
Inhibitory
1
20
40
60
80
i

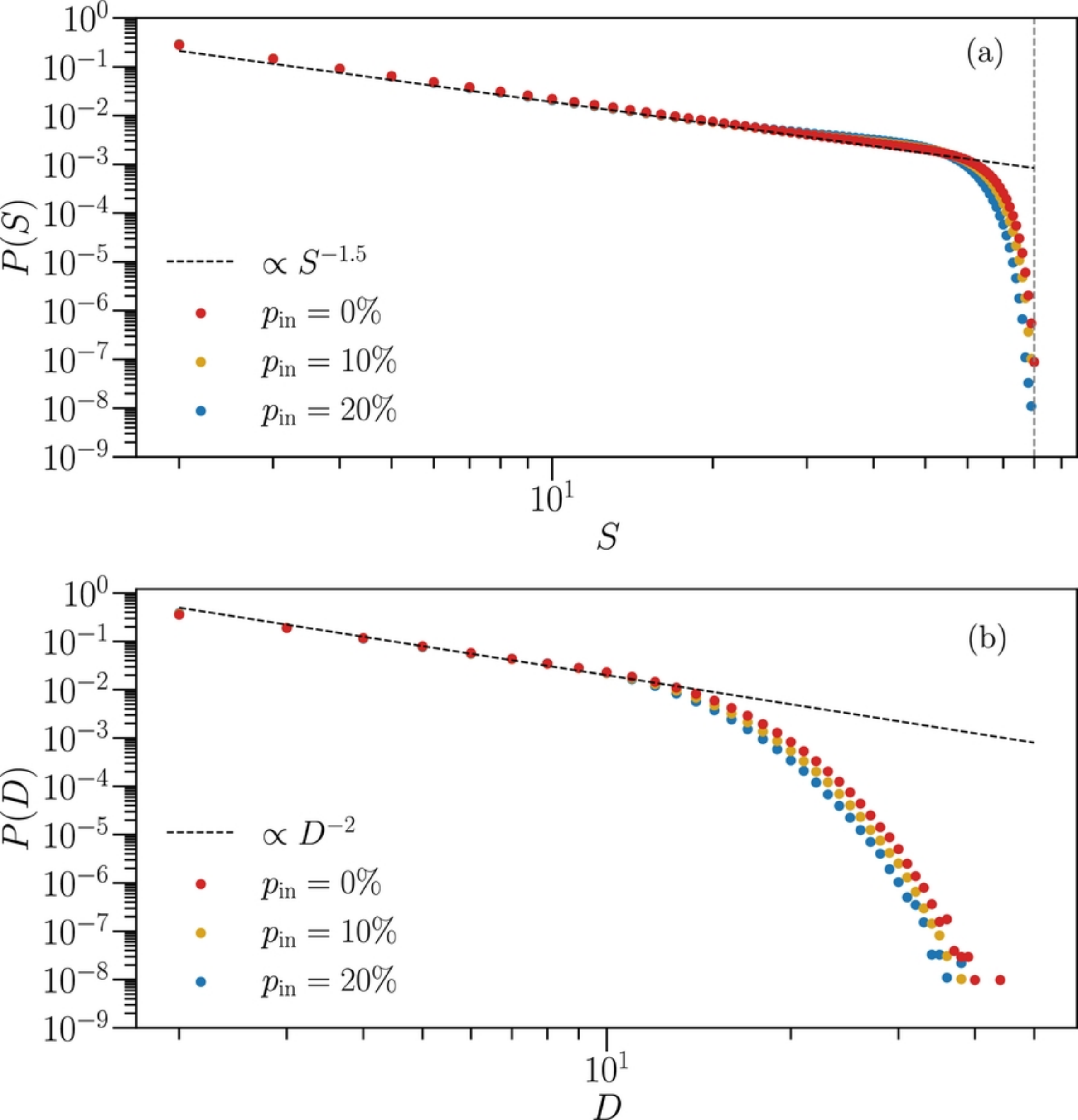

$\eta = 0.10$
$\eta = 0.15$
$\eta = 0.20$
h_i
i
(a)
(b)
(c)
(d)
(e)
(f)
$P(J_{ij})$
J_{ij}

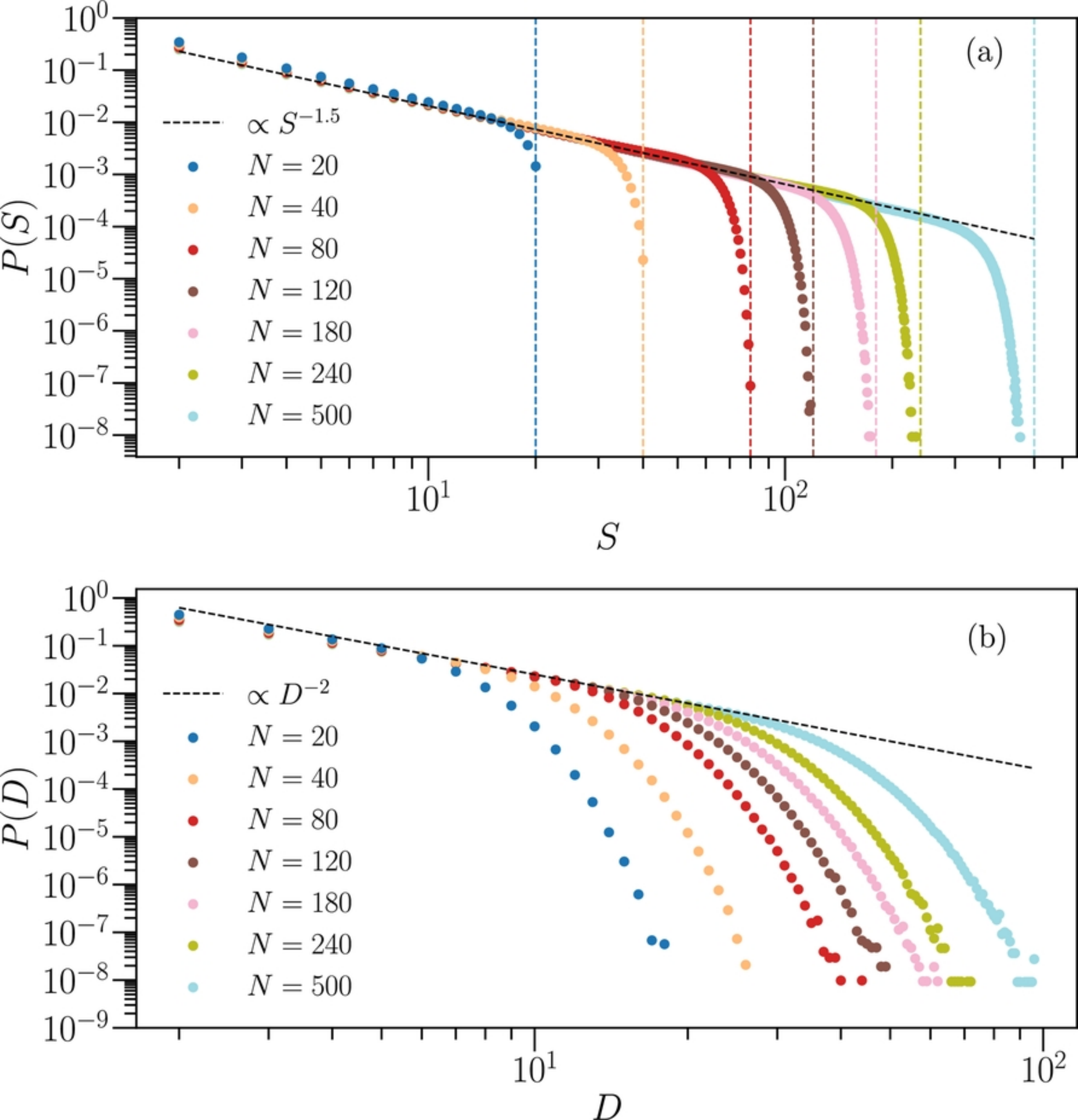

$N = 20$
$N = 40$
$N = 80$
$N = 120$
$N = 240$
$N = 500$
$0.243e^{-0.45K}$
$0.052e^{-0.19K}$
$0.049e^{-0.14K}$
$0.018e^{-0.09K}$
$0.012e^{-0.07K}$
$0.007e^{-0.05K}$
$P(K)$
K
10^0
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{-7}
10^{-8}
0
50
100
150
200
250
300

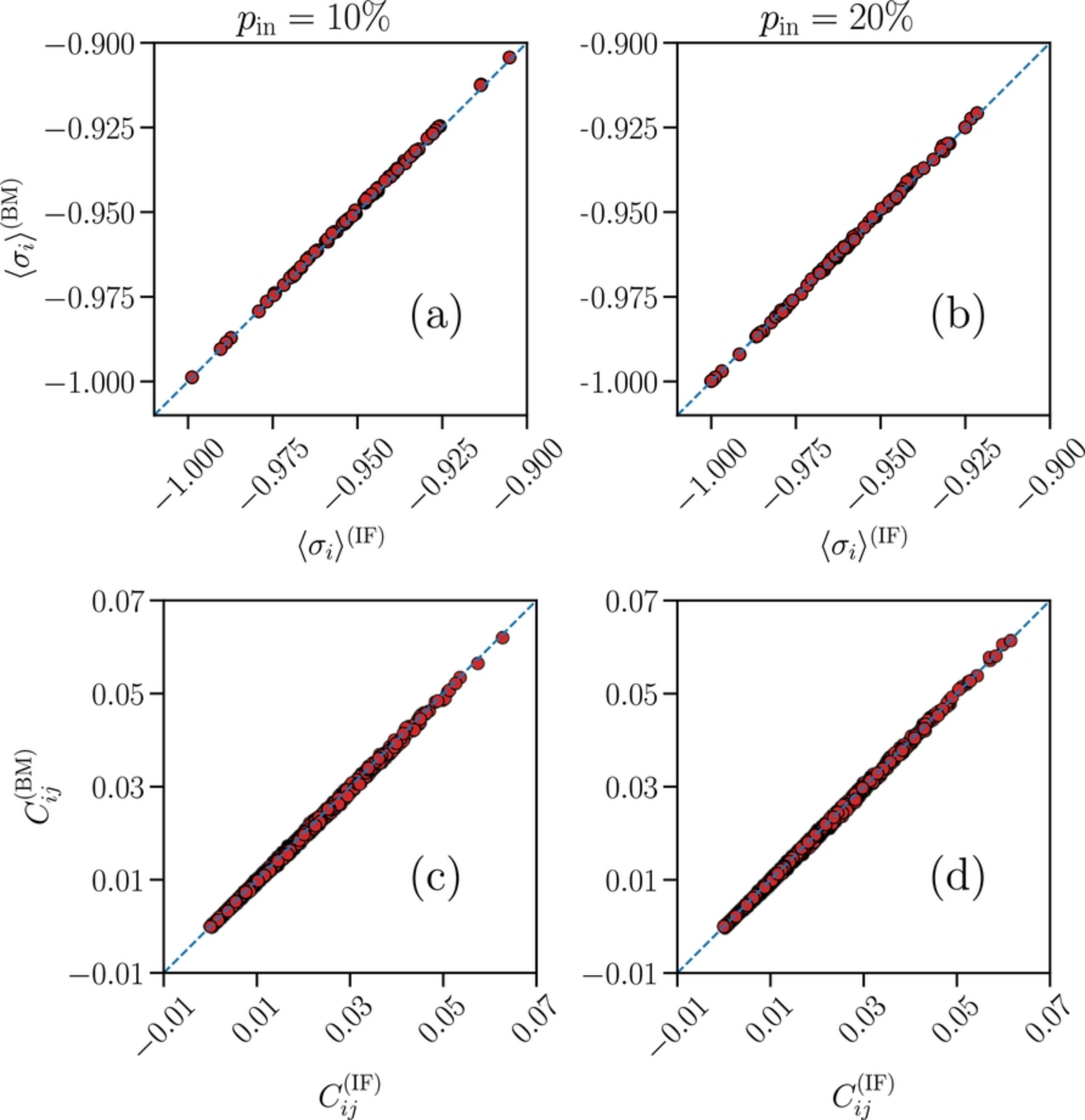

$\eta = 0.10$
$\eta = 0.15$
$\eta = 0.20$
40.5%
58.4%
72.4%
(a)
(b)
(c)
$C_{ij}^{(\mathrm{BM})}$
$C_{ij}^{(\mathrm{IF})}$
0.05
0.04
0.03
0.02
0.01
0.00
−0.01

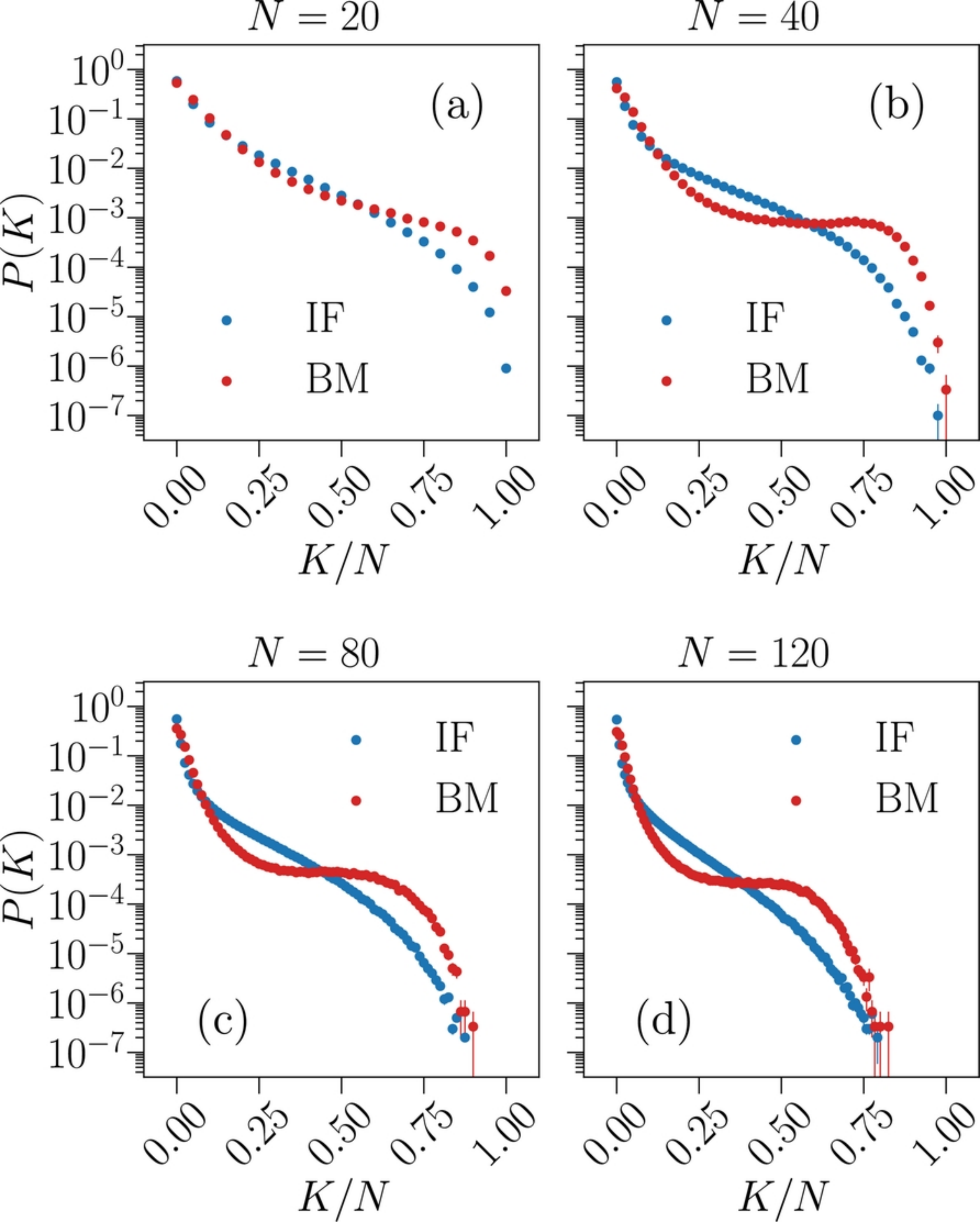

$N = 20$
$N = 40$
$N = 80$
$N = 120$
$\langle \sigma_i \rangle^{(\mathrm{BM})}$
$\langle \sigma_i \rangle^{(\mathrm{IF})}$
(a)
(b)
(c)
(d)
$C_{ij}^{(\mathrm{BM})}$
$C_{ij}^{(\mathrm{IF})}$
(e)
(f)
(g)
(h)

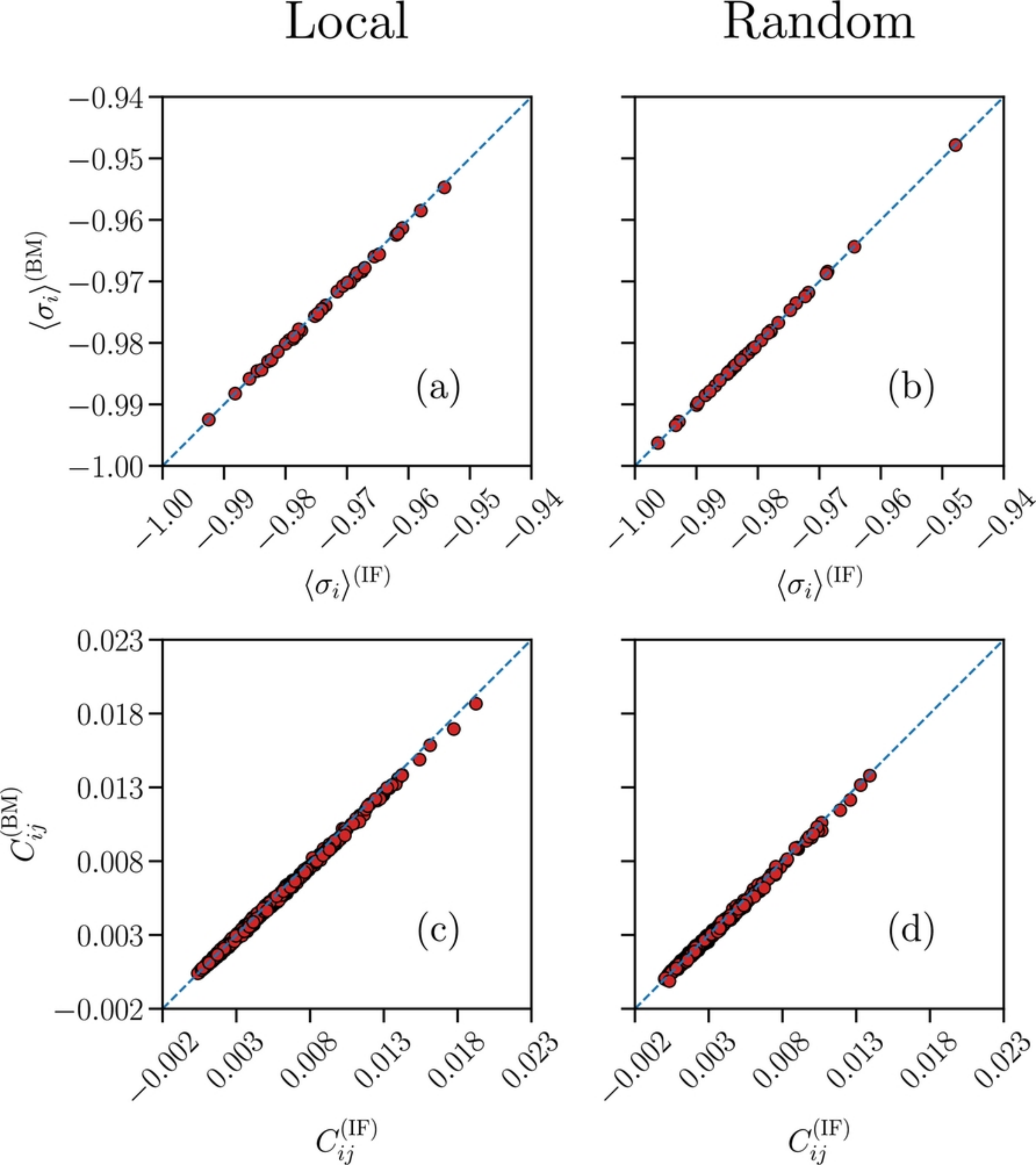

$N = 40$
$N = 120$
$N = 500$
(a)
(b)
(c)
(d)
(e)
(f)
$P(\langle\sigma_i\rangle)$
$\langle\sigma_i\rangle$
$P(C_{ij})$
C_{ij}
0.25
0.20
0.15
0.10
0.05
0.00
−1.00
−0.95
−0.90
−0.85
−0.80
10^0
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
0.000
0.025
0.050
0.075
0.100

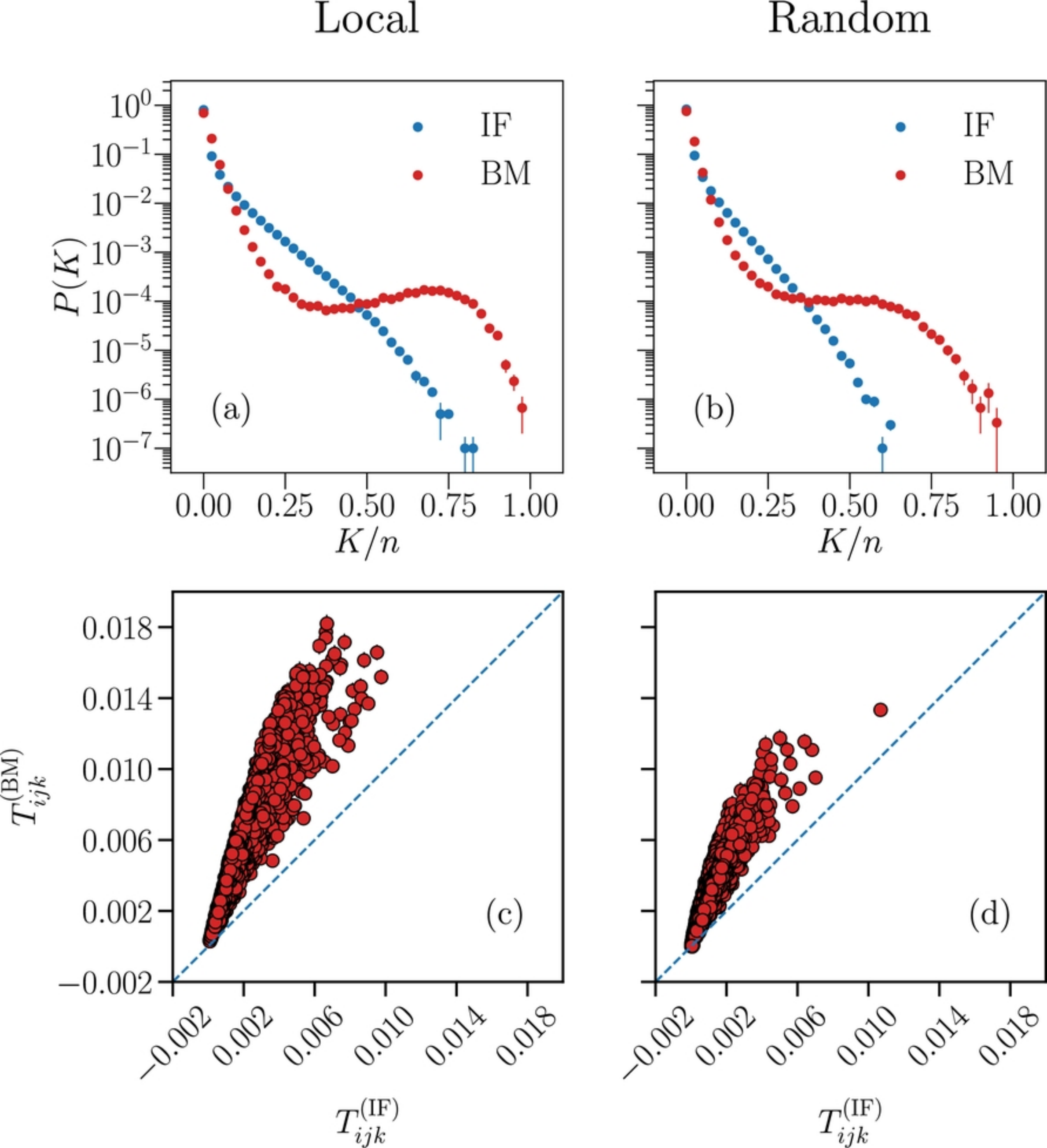

(a)
$P(S)$
S
$\propto S^{-1.5}$
$p_{\text{in}} = 0\%$
$p_{\text{in}} = 10\%$
$p_{\text{in}} = 20\%$
(b)
$P(D)$
D
$\propto D^{-2}$
$p_{\text{in}} = 0\%$
$p_{\text{in}} = 10\%$
$p_{\text{in}} = 20\%$

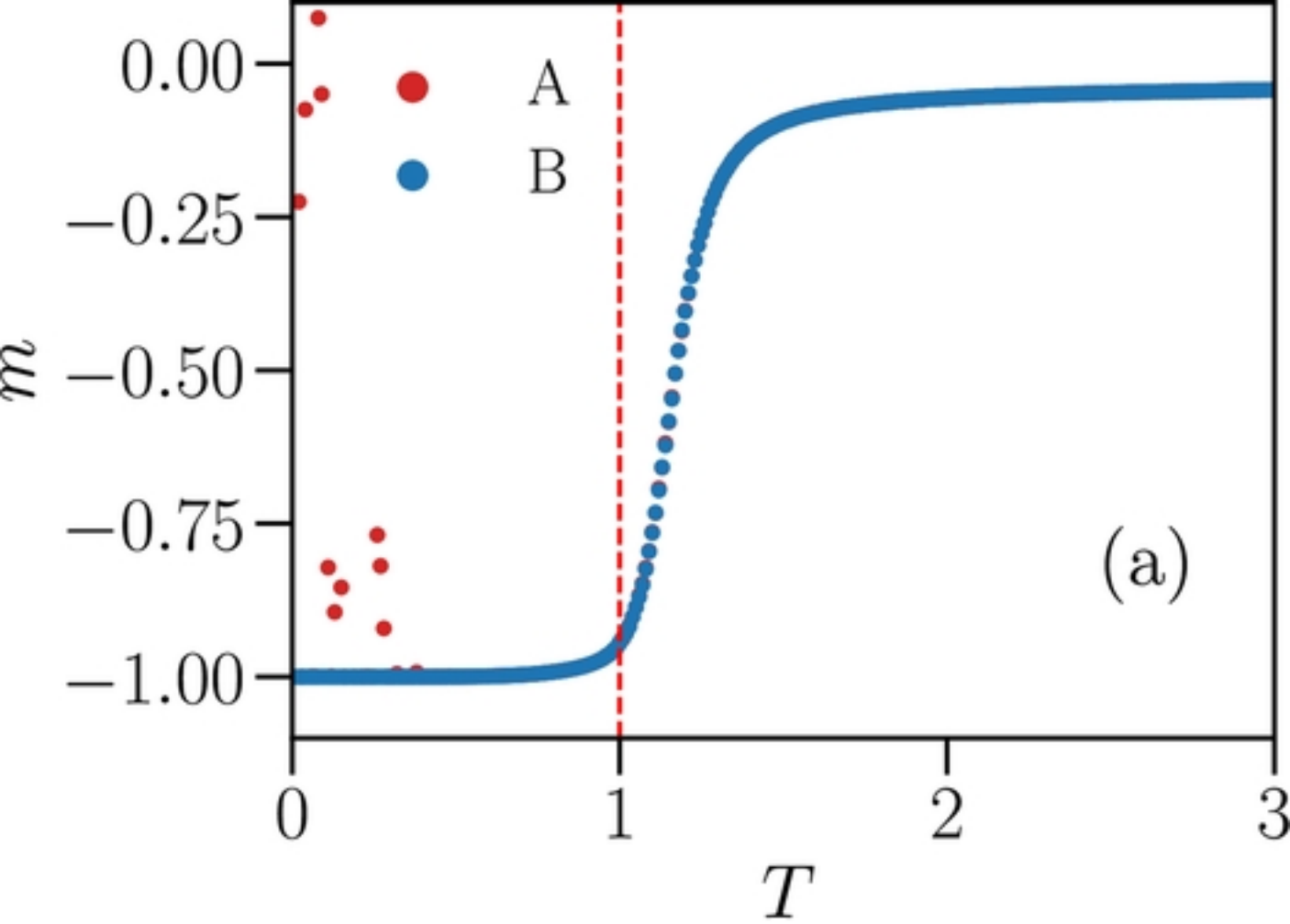

(a)
$\propto S^{-1.5}$
$N = 20$
$N = 40$
$N = 80$
$N = 120$
$N = 180$
$N = 240$
$N = 500$
$P(S)$
S
(b)
$\propto D^{-2}$
$N = 20$
$N = 40$
$N = 80$
$N = 120$
$N = 180$
$N = 240$
$N = 500$
$P(D)$
D

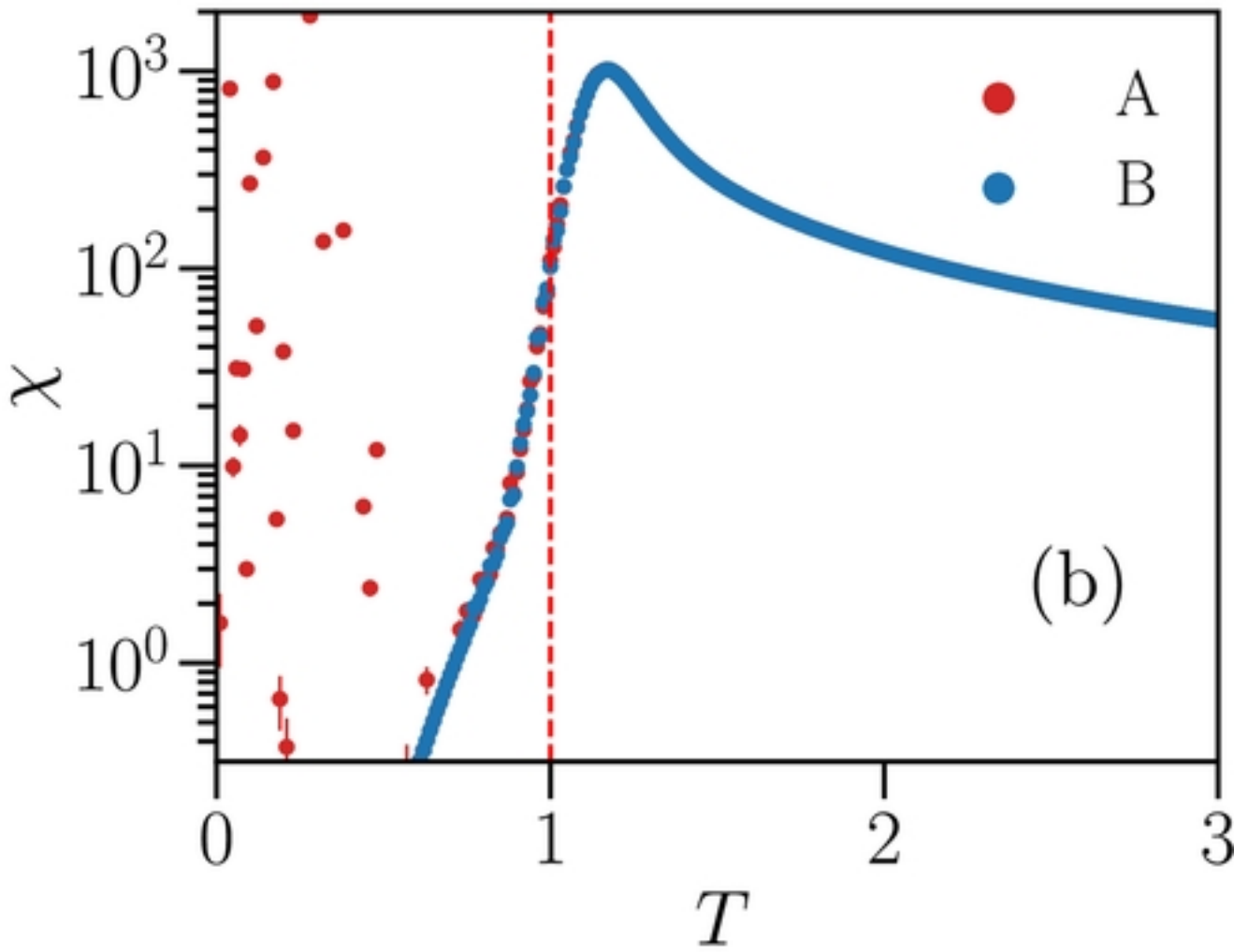

$p_{\rm in} = 10\%$
$p_{\rm in} = 20\%$
$\langle\sigma_i\rangle^{(\rm BM)}$
$\langle\sigma_i\rangle^{(\rm IF)}$
(a)
(b)
$C_{ij}^{(\rm BM)}$
$C_{ij}^{(\rm IF)}$
(c)
(d)

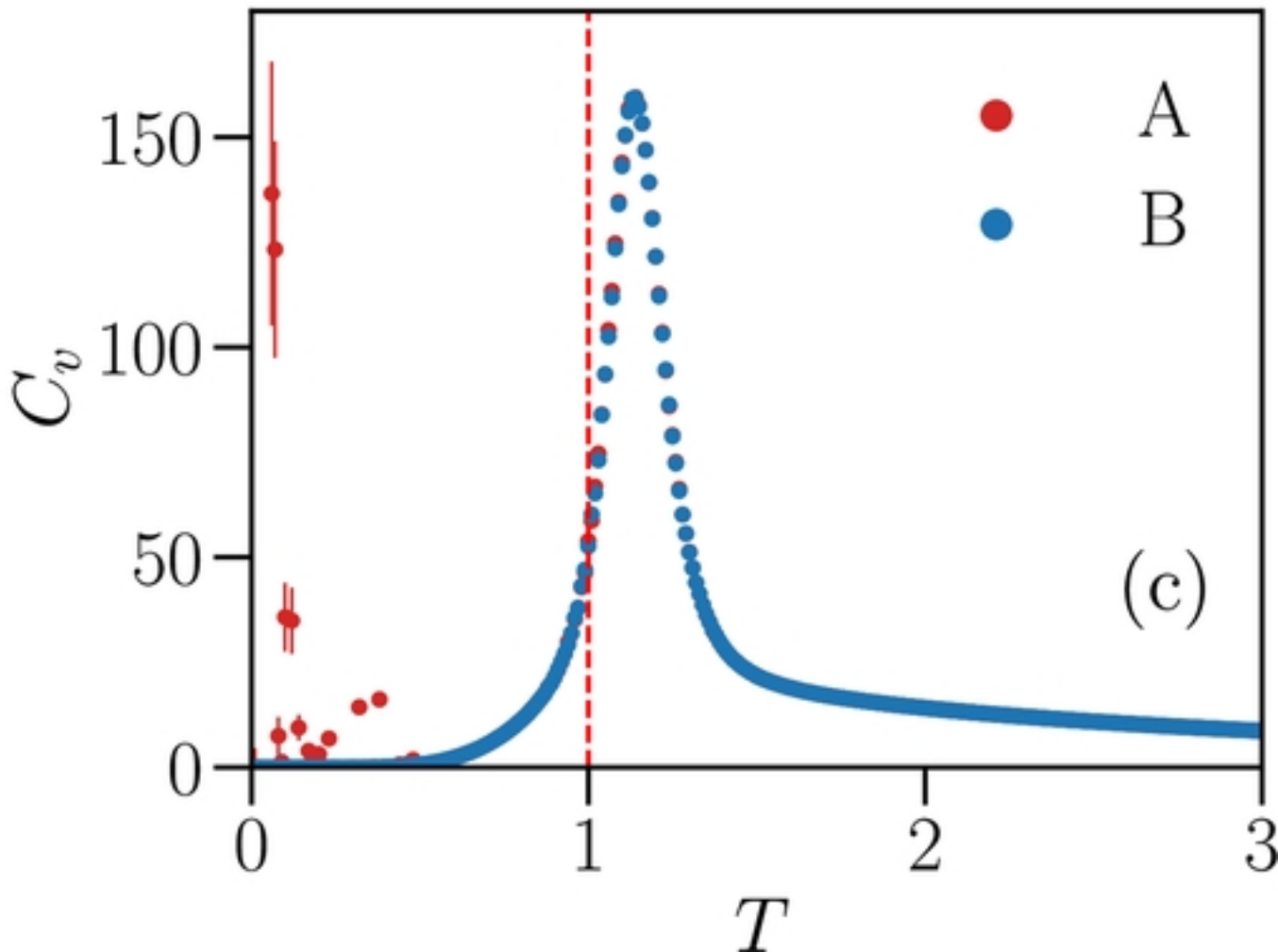

$N = 20$
$N = 40$
$N = 80$
$N = 120$
(a)
(b)
(c)
(d)
IF
BM
$P(K)$
K/N
10^0
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{-7}
0.00
0.25
0.50
0.75
1.00

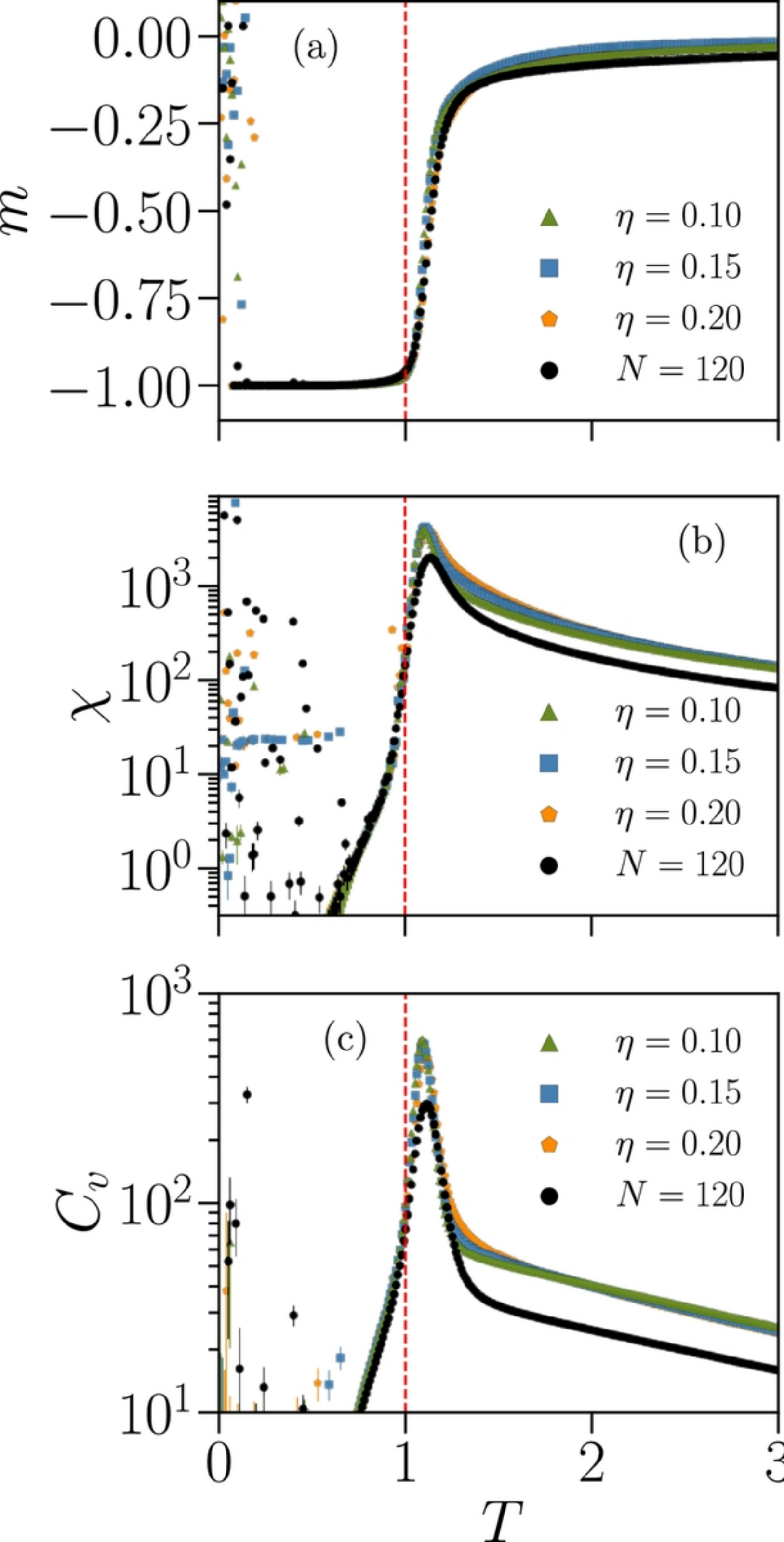

Local
Random
(a)
(b)
(c)
(d)
$\langle\sigma_i\rangle^{(\mathrm{BM})}$
$\langle\sigma_i\rangle^{(\mathrm{IF})}$
$C_{ij}^{(\mathrm{BM})}$
$C_{ij}^{(\mathrm{IF})}$
−0.94
−0.95
−0.96
−0.97
−0.98
−0.99
−1.00
0.023
0.018
0.013
0.008
0.003
−0.002

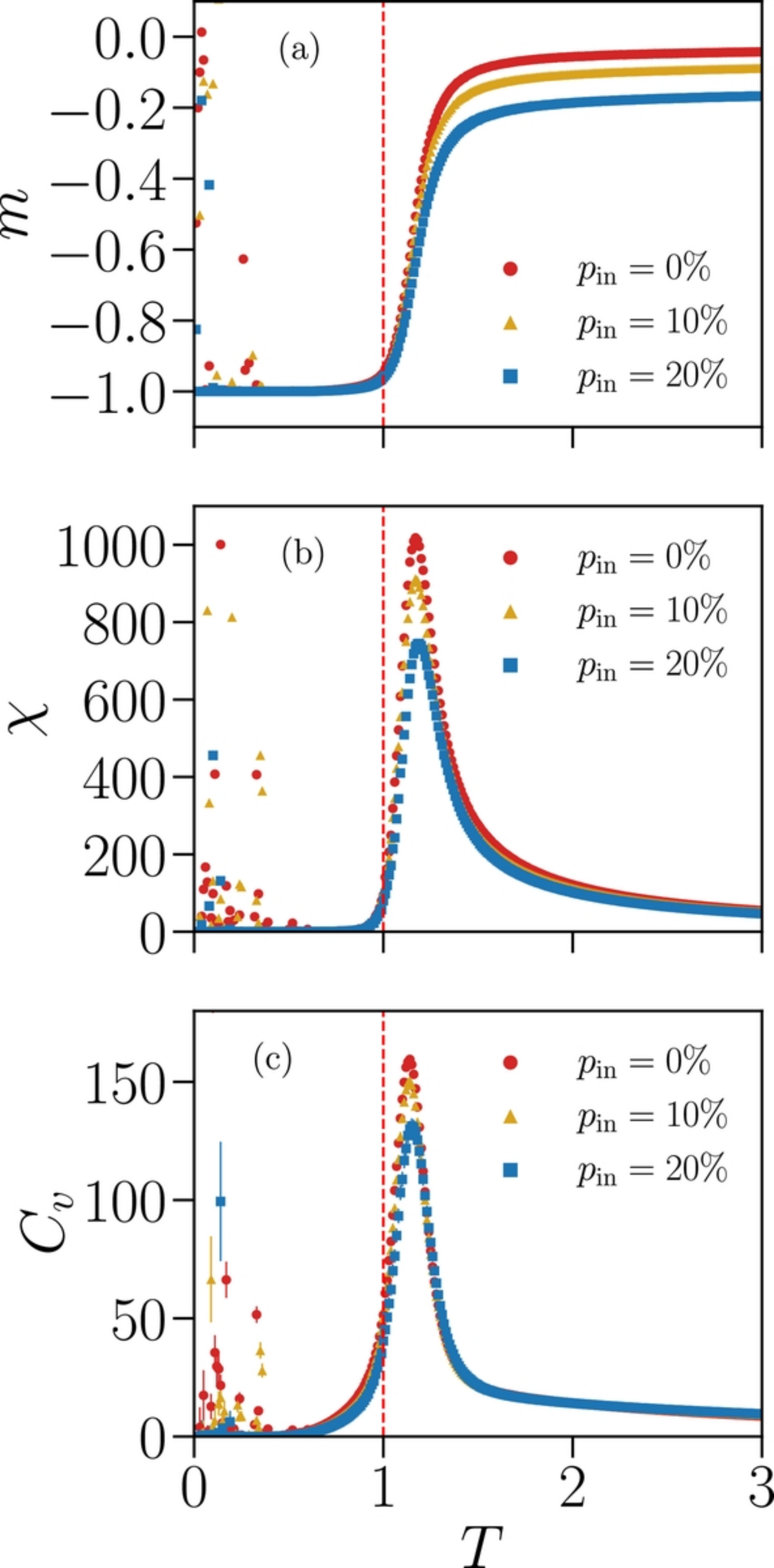

Local
Random
IF
BM
(a)
(b)
(c)
(d)
$P(K)$
K/n
$T_{ijk}^{(\mathrm{BM})}$
$T_{ijk}^{(\mathrm{IF})}$

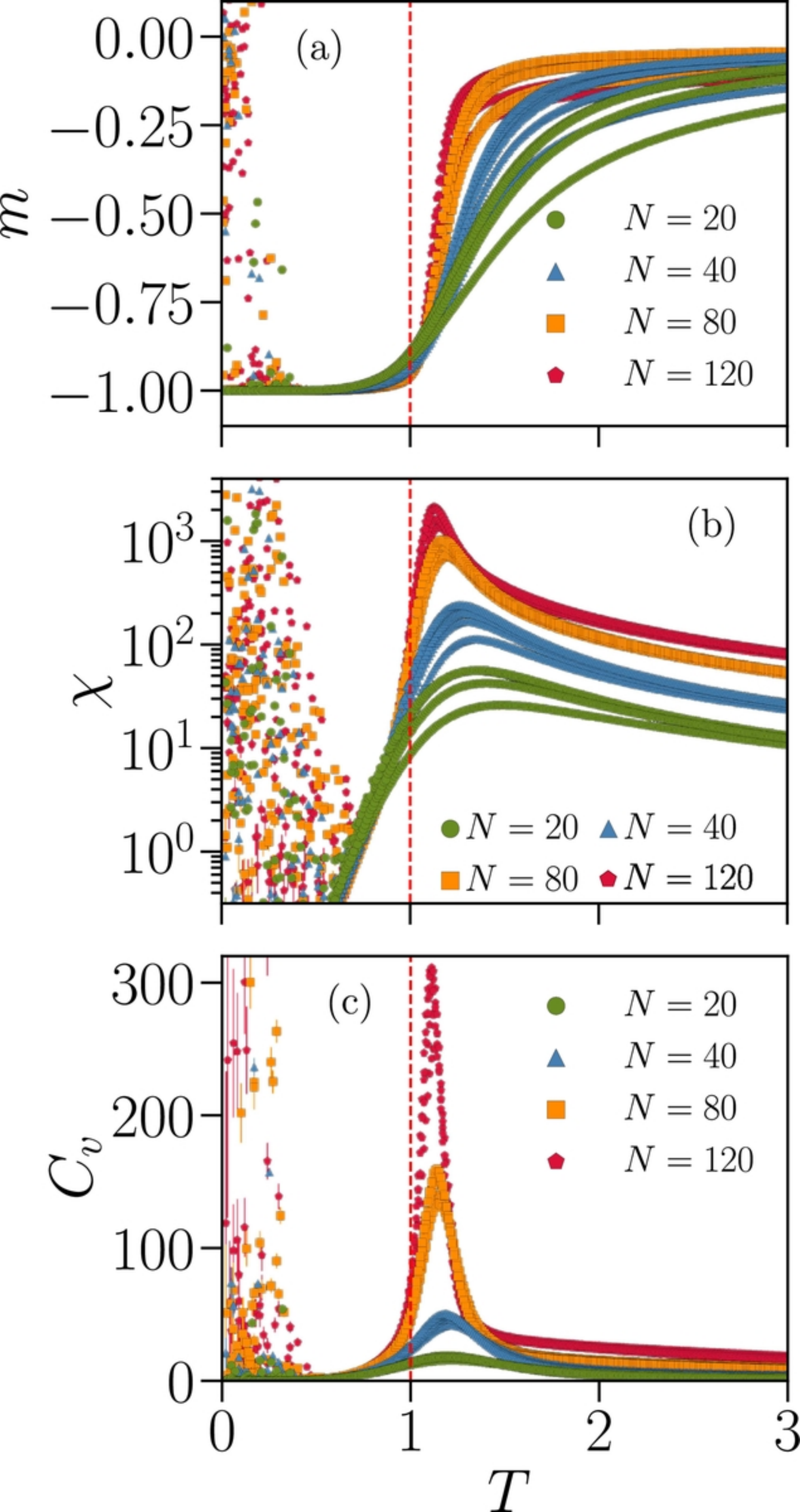

0.00
−0.25
−0.50
−0.75
−1.00
m
A
B
(a)
0
1
2
3
T


10^3
10^2
10^1
10^0
χ
A
B
(b)
0
1
2
3
T


150
100
50
0
C_v
A
B
(c)
0
1
2
3
T


(a)
m
0.00
−0.25
−0.50
−0.75
−1.00
$\eta = 0.10$
$\eta = 0.15$
$\eta = 0.20$
$N = 120$
(b)
χ
10^3
10^2
10^1
10^0
$\eta = 0.10$
$\eta = 0.15$
$\eta = 0.20$
$N = 120$
(c)
C_v
10^3
10^2
10^1
$\eta = 0.10$
$\eta = 0.15$
$\eta = 0.20$
$N = 120$
0
1
2
3
T


(a)
(b)
(c)
m
χ
C_v
T
$p_{\rm in} = 0\%$
$p_{\rm in} = 10\%$
$p_{\rm in} = 20\%$
0.0
−0.2
−0.4
−0.6
−0.8
−1.0
1000
800
600
400
200
0
150
100
50
0
1
2
3


(a)
0.00
−0.25
−0.50
−0.75
−1.00
m
N = 20
N = 40
N = 80
N = 120
(b)
10^3
10^2
10^1
10^0
χ
N = 20
N = 40
N = 80
N = 120
(c)
300
200
100
0
C_v
N = 20
N = 40
N = 80
N = 120
0
1
2
3
T