Nature of hypergraph *k*-core percolation problems

Ginestra Bianconi [1,2] and Sergey N. Dorogovtsev [3,4]
[1] *School of Mathematical Sciences, Queen Mary University of London, London, E1 4NS, United Kingdom*
[2] *Alan Turing Institute, 96 Euston Road, London NW1 2DB, United Kingdom*
[3] *Departamento de Física da Universidade de Aveiro & I3N, 3810-193 Aveiro, Portugal*
[4] *Ioffe Physico-Technical Institute, 194021 St. Petersburg, Russia*

(Received 28 July 2023; accepted 11 December 2023; published 17 January 2024)

Hypergraphs are higher-order networks that capture the interactions between two or more nodes. Hypergraphs can always be represented by factor graphs, i.e., bipartite networks between nodes and factor nodes (representing groups of nodes). Despite this universal representation, here we reveal that *k*-core percolation on hypergraphs can be significantly distinct from *k*-core percolation on factor graphs. We formulate the theory of hypergraph *k*-core percolation based on the assumption that a hyperedge can be intact only if all its nodes are intact. This scenario applies, for instance, to supply chains where the production of a product requires all raw materials and all processing steps; in biology it applies to protein-interaction networks where protein complexes can function only if all the proteins are present; and it applies as well to chemical reaction networks where a chemical reaction can take place only when all the reactants are present. Formulating a message-passing theory for hypergraph *k*-core percolation, and combining it with the theory of critical phenomena on networks, we demonstrate sharp differences with previously studied factor graph *k*-core percolation processes where it is allowed for hyperedges to have one or more damaged nodes and still be intact. To solve the dichotomy between *k*-core percolation on hypegraphs and on factor graphs, we define a set of pruning processes that act either exclusively on nodes or exclusively on hyperedges and depend on their second-neighborhood connectivity. We show that the resulting second-neighbor *k*-core percolation problems are significantly distinct from each other. Moreover we reveal that although these processes remain distinct from factor graphs *k*-core processes, when the pruning process acts exclusively on hyperedges the phase diagram is reduced to the one of factor graph *k*-cores.

DOI: 10.1103/PhysRevE.109.014307

I. INTRODUCTION

Higher-order networks—hypergraphs and simplicial complexes, representing multinode interactions—have recently gained significant attention [1–12]. Research interest is growing in both modeling higher-order network structures and investigating dynamical processes and cooperative systems on such networks [4,8,13–21]. Importantly, the features and characteristics of processes and cooperative phenomena on higher-order networks differ significantly from those on ordinary networks. In this work we focus on specific highly connected substructures in hypergraphs, namely, their *k*-cores.

Each hypergraph can be represented by an equivalent bipartite graph between nodes and hyperedges, factor nodes, which is called the factor graph of a hypergraph. Considering multiagent interactions, one can recognize two markedly distinct classes. In the first class [16,17,22], the removal (damaging) of one of the interacting agents (let the number of these agents exceed two) doesn't break interaction between the remaining agents. Typically this happens in networks of social interactions. For instance, an online social network group might still be working if one or more of its members are not participating on it actively. Such systems of multinode interactions are naturally described by factor graphs, i.e., bipartite networks. In the second class of multiagent interactions, the removal (damaging) of one of the interacting agents breaks interaction between the remaining agents. Supply chains and catalytic networks [23,24], protein-interaction networks [25], and networks of chemical reactions [26] provide an example for higher-order interactions of this kind. Such systems of multinode interactions are described by hypergraphs, in which the removal of a node results in the disappearance of all the adjacent hyperedges. For instance, the removal of a raw material will impede the production of a product, the absence of a protein will impede the formation of a protein complex, and the absence of a reactant will impede a chemical reaction to occur. Node percolation problems for these two classes of systems qualitatively differ from each other (compare Refs. [22] and [27]), although edge percolation on hypergraphs coincides with factor node percolation on corresponding factor graphs.

The issue of *k*-cores in ordinary networks has been extensively explored [28–33]. For a graph G, a k-core is the maximal subgraph G_k, in which each node has degree at least k, where k is a given threshold value. One can decompose a graph into the set of k-shells $G_k \backslash G_{k+1}$ and classify the nodes according to shells to which they belong [34,35]. The higher k-cores play a particular role in a graph in respect of rapid spreading phenomena, including fast disease spreading [36,37]. The highest k-core was observed to be the center localization in a number of network architectures [38]. The pruning process resulting in a k-core is quite efficient algorithmically: one must progressively remove each node with degree smaller than k until no such nodes remain. The result of this process in an infinite graph is a subgraph which can contain a single giant and many finite components. Sometimes it is this giant component that is referred to as the k-core. We shall focus on this component. In the infinite treelike networks, finite components in the k-cores are vanishingly rare if k exceeds 1.

The k-core problems for hypergraphs have received little attention thus far. The authors of Refs. [16,17] introduced the (k,n)-core in a factor graph as the maximal subgraph of the factor graph with nodes of degree equal to at least k and factor nodes of degree equal to at least n. This subgraph is the result of the progressive pruning of nodes with degrees smaller than k and factor nodes with degrees smaller than n. This definition and the pruning algorithm applied to a factor graph are relevant for the systems that are described by bipartite networks, like social interactions mentioned earlier. If we inspect, however, hypergraphs represented by the factor graphs emerging during the execution of this algorithm, we observe that during this pruning the cardinalities of hyperedges corresponding to factor nodes can decrease. Consequently, the (k,n)-cores introduced in this way are not subhypergraphs of the hypergraph in contrast to the k-cores of an ordinary graph.

In this work we adopt definitions of k-core hypergraph percolation that pertain to specific multinode interactions in which the removal or damage of one interacting agent disrupts the interaction among the remaining agents (e.g., supply chain, protein interaction networks, or networks of chemical reactions). In the present work we describe a set of k-core problems specific for hypergraphs and the corresponding pruning algorithms in which nodes and hyperedges are progressively removed (damaged) if their degrees and cardinalities, respectively, are smaller than given threshold values, k and n, and hence each step of these algorithms provides a subhypergraph of an initial hypergraph. The (k,n)-cores produced by these pruning algorithms are the maximal subhypergraphs whose vertices and hyperedges have degrees and cardinalities equal to at least k and n. These definitions specifically address multinode interactions that exhibit the following characteristic: if one of the interacting agents is removed, it disrupts the interaction among the remaining agents (as in chemical reactions).

For such (k,n)-cores in random hypergraphs, we explore phase transitions and obtain phase diagrams. These phase diagrams are significantly richer than for the k-cores in ordinary graphs, and in factor graphs [16,17] where the phase transition for the 2-core is continuous, while the phase transitions for $(k \geqslant 3)$-cores are hybrid. In particular, we observe a significant difference between the critical properties of the (2,2)-core on factor graphs and on hypergraphs. While in factor graphs the (2,2)-core percolation is always continuous [16,17], on hypergraphs we observe two transition lines on the phase diagram—the continuous transition line and the hybrid transition one. These lines converge at the tricritical point.

In order to solve the dichotomy between (k,n)-core percolation on factor graphs [16,17] and the k-core percolation on hypergraphs investigated here we introduce a class of k-core problems, in which the pruning process accounts not only for the closest neighborhood of a node (e.g., it accounts not only for hyperedges adjacent to a node but also for all their nodes). In this class of models, the pruning process can involve either exclusively nodes or exclusively hyperedges, and the pruning algorithm might depend on the nodes or hyperedges which are second neighbors within the factor graph.

We show that the second-neighbor k-core percolation involving exclusively pruning of the nodes or involving exclusively pruning of the hyperedges are distinct, and we relate these models to both hypergraph (k,n)-cores depending on the closest neighborhood and to factor graphs' (k,n)-core. In particular, we show that the percolation threshold for the second-neighbor k-core problems with pruning of the nodes coincides with the percolation threshold for the first-neighbor hypergraph k-core problems; however, in the limiting case in which only hyperedges are initially damaged, the second-neighbor k-core problems with pruning of the hyperedges coincides with the percolation threshold of the factor graph k-cores [16,17].

These results are obtained within a message-passing theory [39–42] exact for locally treelike hypergraphs and within the generating function theory of critical phenomena on networks, and it is here supported by simulations. The message-passing equations for the k-core percolation problems presented here are derived from their definition of the k-core problems using as starting point the message passing for percolation in hypergraphs [27].

Our approach is general, and the message-passing equations can be applied to arbitrary locally treelike hypergraphs. In particular, we apply these equations to random hypergraphs [12,43,44] with given cardinality and degree distributions. Possibly the proposed approach could be extended in the future in order to go beyond the locally treelike approximation due to recent advances on message passing on networks with loops [42,45].

The paper is structured as follows: In Sec. II we overview the k-core problem for ordinary graphs and develop the message-passing theory for it. Section III introduces the basic definitions and notations for hypergraphs, in particular, the definition of the subhypergraph of a hypergraph. In Sec. IV we derive the message-passing and the generating functions equations for the (k,n)-core problem on hypergraphs (the first-neighbor version in the sense of nodes and factor nodes in factor graphs). In Sec. V we derive the message-passing and the generating functions equations for the second-neighbor version of the (k,n)-core problem on hypergraphs. In Sec. VI we provide concluding remarks.

II. OVERVIEW OF k-CORE PERCOLATION ON SIMPLE NETWORKS

A. k-core and pruning algorithm

We consider a graph $G=(V,E)$. The k-core is the largest subgraph where intact vertices have at least k interconnections. We start from a configuration in which nodes are initially damaged with probability $1-p$. The k-core is obtained by the following algorithm:

(1) Damage iteratively all the nodes with fewer than k undamaged neighbors.

(2) The k-core is the giant component of the network induced by the undamaged nodes. This k-core is the giant subgraph induced by nodes left undamaged by the pruning process.

Note that in infinite locally treelike graphs finite k-core components are negligible.

B. Derivation of the message-passing algorithm for k-cores

Here we aim to derive the message-passing equations for the k-core starting directly from the pruning algorithm under the hypothesis that the network is locally treelike. To this end, let us assume that the initial damage of the nodes is exactly known and encoded in the indicator function $x_i \in \{0, 1\}$ specifying whether a node is initially damaged $x_i = 0$ or not $x_i = 1$.

At step (2) we assume to know whether each node i has been pruned or damaged ($s_i = 0$) or not ($s_i = 1$). The message-passing equations determining the giant component of the network formed by undamaged nodes are

$$\hat{\sigma}_{i\to j} = s_i\left[1 - \prod_{r\in N(i)\setminus j}(1-\hat{\sigma}_{r\to i})\right], \tag{1}$$

where $N(i)$ denotes the set of neighbors of node i. Note that the message $\hat{\sigma}_{i\to j}$ indicates whether node j connects node i to the giant component ($\hat{\sigma}_{i\to j} = 1$) or not ($\hat{\sigma}_{i\to j} = 0$), and it is defined assuming that node j is already in the giant component (see, for instance, discussion of the message-passing algorithm in multilayer networks with link overlap [41,46,47]). The indicator function determining whether node i belongs to the giant component or not is instead given by

$$\hat{\sigma}_{i} = s_i\left[1 - \prod_{r\in N(i)}(1-\hat{\sigma}_{r\to i})\right]. \tag{2}$$

Now at step (2) we have that a node that is undamaged and belongs to the k-core must receive at least k positive messages (i.e., it should be connected to at least k nodes in the giant component),

$$s_i = x_i \sum_{\substack{\Theta\subseteq N(i)\\ |\Theta|\geqslant k}} \prod_{r\in\Theta}\hat{\sigma}_{r\to i} \prod_{r\in N(i)\setminus\Theta}(1-\hat{\sigma}_{r\to i}), \tag{3}$$

where Θ is a subset of $N(i)$ including at least k nodes. Under the assumption that j is connected to the giant component, i.e., $\hat{\sigma}_{j\to i} = 1$, substituting the above expression for s_i into Eq. (1) for $\hat{\sigma}_{i\to j}$ we obtain then that at stationarity the messages $\sigma_{i\to j}$ satisfy

$$\hat{\sigma}_{i\to j} = x_i \sum_{\substack{\Theta\subseteq N(i)\setminus j\\ |\Theta|\geqslant k-1}} \prod_{r\in\Theta}\hat{\sigma}_{r\to i} \prod_{r\in N(i)\setminus(\Theta\cup j)}(1-\hat{\sigma}_{r\to i}), \tag{4}$$

while $\hat{\sigma}_i = s_i$ is given by Eq. (3).

These message-passing equations can be averaged over the distribution of $\{x_i\}$ given by

$$P(\{x_i\}) = \prod_{i=1}^{N} p^{x_i}(1-p)^{1-x_i}, \tag{5}$$

and using the locally treelike approximation we get

$$\sigma_{i\to j} = p \sum_{\substack{\Theta\subseteq N(i)\setminus j\\ |\Theta|\geqslant k-1}} \prod_{r\in\Theta}\sigma_{r\to i} \prod_{r\in N(i)\setminus(\Theta\cup j)}(1-\sigma_{r\to i}), \tag{6}$$

$$\sigma_{i} = p \sum_{\substack{\Theta\subseteq N(i)\\ |\Theta|\geqslant k}} \prod_{r\in\Theta}\sigma_{r\to i} \prod_{r\in N(i)\setminus\Theta}(1-\sigma_{r\to i}). \tag{7}$$

Since these message-passing equations are more cumbersome to implement than the original pruning process, typically the message-passing algorithms are not applied to the k-core problems. However, their formulation can be used to derive the equations describing the pruning algorithm on a random (locally treelike) graph that we discuss next.

C. k-core transition on ordinary random networks

Having derived the message-passing algorithm describing the final outcome of the pruning process, we can now demonstrate how the known formulas for k-core percolation on a random network relate to the pruning algorithm. This exercise will help us clarify the correct equations determining (k, n)-core percolation on hypergraphs.

Let us consider the k-cores of networks provided by the configuration model with a given degree distribution $P(q)$. In the configuration model, to each network $G = (V, E)$ with $N = |V|$ nodes and the adjacency matrix $\mathbf{a}$, the following probability is assigned:

$$P(G) = \prod_{i<j} p_{ij}^{a_{ij}}(1-p_{ij})^{1-a_{ij}} \tag{8}$$

with

$$p_{ij} = \frac{q_i q_j}{\langle q\rangle N}, \tag{9}$$

where q_i is the degree of node i.

By averaging the message-passing equations over a network generated by the configuration model we get that Z indicating the average message $\sigma_{i\to j}$ in the configuration network ensemble is given by

$$Z = \sum_{q=k}^{\infty}\frac{qP(q)}{\langle q\rangle}\sum_{s=k-1}^{q-1}\binom{q-1}{s}Z^s(1-Z)^{q-1-s}, \tag{10}$$

and that $S_k = \langle\sigma_i\rangle$ indicating the fraction of nodes in the k-core is given by

$$S_k = \sum_{q=k}^{\infty}P(q)\sum_{s=k}^{q}\binom{q}{s}Z^s(1-Z)^{q-s}. \tag{11}$$

This latter quantity can be also written as

$$S_k = \sum_{s=k}^{\infty}S_k(s), \tag{12}$$

where $S_k(s)$ is the fraction of nodes that are in the k-core and have exactly degree $s \geqslant k$ within it. We have

$$S_k(s) = \sum_{q=s}^{\infty} P(q) \binom{q}{s} Z^s (1-Z)^{q-s}, \tag{13}$$

and Eq. (12) follows from the observation that

$$\begin{aligned} S_k &= \sum_{q=k}^{\infty} P(q) \sum_{s=k}^{q} \binom{q}{s} Z^s (1-Z)^{q-s} \\ &= \sum_{s=k}^{\infty} \sum_{q=s}^{\infty} P(q) \binom{q}{s} Z^s (1-Z)^{q-s}, \end{aligned} \tag{14}$$

where we have used the equality

$$\sum_{q=k}^{\infty} \sum_{s=k}^{q} = \sum_{s=k}^{\infty} \sum_{q=s}^{\infty}. \tag{15}$$

We conclude our overview of k-core percolation on an ordinary network by expressing Eq. (10) and Eq. (11) in terms of the generating functions of a degree distribution, $G(z) \equiv \sum_q P(q) z^q$, getting

$$Z = 1 - \frac{1}{\langle q \rangle} \sum_{s=0}^{k-2} \frac{Z^s}{s!} G^{(s+1)}(1-Z) \tag{16}$$

and

$$S_k = 1 - \sum_{s=0}^{k-1} \frac{Z^s}{s!} G^{(s)}(1-Z). \tag{17}$$

III. SUBHYPERGRAPHS OF HYPERGRAPHS

Denote a hypergraph by $\mathcal{H} = (V, H)$, where V and H are the sets of its vertices and hyperedges. We indicate with N the number of nodes, i.e., $|V| = N$, and with M the number of hyperedges, i.e., $|H| = M$, of the hypergraph. We indicate the nodes of the hypergraph with Latin letters $i, j, r, \ldots$ and the hyperedges of the hypergraph with Greek letters $\alpha, \beta, \gamma, \ldots$. Each hyperdege α determines a set of nodes

$$\alpha = [i_1, i_2, i_3, \ldots, i_{m_\alpha}] \tag{18}$$

with $m_\alpha \equiv |\alpha|$ indicating the cardinality (number of nodes) of the hyperdege α. Likewise, each node i has degree q_i indicating the number of hyperedges it belongs to.

A subhypergraph $\mathcal{S}$ of the hypergraph $\mathcal{H}$ is defined as $\mathcal{S} = (V_S, H_S)$, where $V_S \subset V$ and $H_S \subset H_{\max}$, where $H_{\max}$ is the full set of those hyperedges of H that have each of their end vertices belonging to V_S. In particular, $\mathcal{S}_{\text{induced}} = (V_S, H_{\max})$ is the vertex-induced subhypergraph of the hypergraph $\mathcal{H}$, induced by the set of vertices V_S. Importantly, any subhypergraph S of the hypergraph $\mathcal{H}$ cannot have hyperedges not belonging to $\mathcal{H}$, for example, hyperedges of smaller cardinalities.

IV. HYPERGRAPHS (k, n)-CORE PROBLEMS AND THEIR (FIRST-NEIGHBOR) PRUNING ALGORITHM

A. First-neighbor pruning algorithm

We consider a hypergraph $\mathcal{H} = (V, H)$. In this hypergraph we assume that a hyperedge is intact if it is not damaged and if none of its nodes are damaged. The hypergraph always can be represented as a factor graph, which is a bipartite networks having two types of nodes, namely, the nodes corresponding to the nodes of the hypergraphs and the factor nodes corresponding to the hyperedges of the hypergraph. Accordingly, we can choose to iteratively prune only nodes, only hyperedges, or both nodes and hyperedges. If the algorithm for the pruning only depends on the first neighbors of a node in the factor graph, we can treat all these variants simultaneously. On the other hand, when the pruning algorithm depends on the first and second neighbors of a node in the factor graph, we need to treat independently pruning of nodes and pruning of hyperedges as we will show in Sec. V. Moreover, for each of the pruning algorithms, the initial damage can target either nodes or hyperedges.

If we consider the pruning on the nodes and hyperedges depending only on the state of their neighbors in the factor graph we can define the hypergraph's (k, n)-core. The (k, n)-core is the maximal subhypergraph with the vertices whose (internal) degrees are at least k and the hyperedges have cardinalities equal or exceeding n. We start from a configuration in which nodes are initially damaged with probability $1 - p_N$ and/or hyperedges are initially damaged with probability $1 - p_H$. The (k, n)-core can be obtained using the following pruning algorithm:

(1) Damage iteratively all the hyperedges having fewer than n (undamaged) nodes and all the nodes with fewer than k undamaged hyperedges.

(2) The (k, n)-core is the giant subhypergraph induced by undamaged nodes and undamaged hyperedges.

Equivalently, this (k, n)-core can be obtained by, first, removing all hyperedges with cardinalities smaller than n and, second, progressively pruning all nodes with degrees smaller than k. The $(k, 2)$-core can be called the k-core for hypergraphs.

From the above definition of hypegraph k-core percolation we conclude that there are two major differences between hypegraph percolation and factor graph percolation [16,17]. First, and most importantly, in hypergraph percolation the damage of a node automatically disrupts all the hyperedges to which the node belongs while in factor graph percolation it reduces only by one the degree of the factor nodes to which the node is connected. Second, hypergraph k-cores are subhypegraphs of the original hypergraph, while this property is not preserved in factor graph percolation (see schematic representation of the difference in Fig. 1).

B. Message-passing algorithm for hypergraph (k, n)-core

In this subsection we derive the message-passing equations determining the (k, n)-core directly from the definition of problem given in the previous subsection. The obtained equations are very general and apply to every hypergraph under the assumption that the factor graph encoding for the hypergraph is locally treelike.

Since the definition of the (k, n)-core is given in terms of the giant subhypergraph induced by intact nodes and hyperedges, the message-passing equations for (k, n)-core percolation will be derived starting from the equations valid

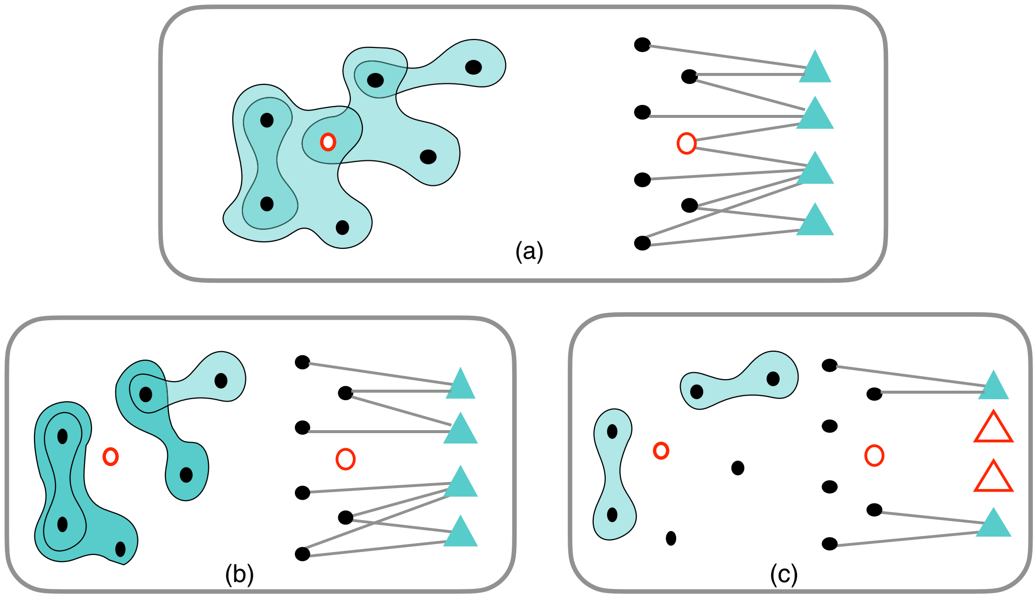

FIG. 1. Schematic representation of the difference between the factor graph [16,17] and hypergraph pruning algorithms. We consider the hypergraph in panel (a) and its factor graph representation with circles representing nodes and triangles representing factor nodes. In the factor graph pruning algorithm (b), when a node is damaged (empty circles), all the hyperedges including this node are reduced in size by one. In hypergraph percolation (c) all hyperedges containing the damaged node are damaged leading to damaged factor nodes (empty triangles) in the factor graph representation.

for hypergraph percolation [27], which will be used to identify this giant subhypergraph.

Here we recall that on hypergraph percolation [27] the node and hyperedge percolation are distinct and that an hyperedge is in the giant subhypergraph only if (1) it is not damaged, (2) none of its nodes are damaged, and (3) at least one of its nodes is connected to the giant subhypergraph.

If follows that the (k, n)-core algorithm described in the previous subsection is equivalent to the one in which the initial damage is modified as in the following: nodes are initially damaged with probability $1 - p_N$ and hyperedges α are initially damaged deterministically if their cardinality is smaller than n, i.e., $|\alpha| < n$, while if their cardinality is larger or equal to n, i.e., $|\alpha| \geqslant n$, they are damaged with probability $1 - p_H$. Given this initial damage, the (k, n)-core defined by the pruning algorithm is equivalent to the one obtained using the following pruning algorithm:

(1′) Damage iteratively all the nodes with fewer than k undamaged hyperedges.

(2′) The (k, n)-core is the giant subhypergraph induced by undamaged nodes and undamaged hyperedges.

In order to derive the message-passing algorithm for (k, n)-core percolation directly from this pruning algorithm, let us assume that the initial damage of the nodes is exactly known and encoded in the indicator function $x_i \in \{0, 1\}$ indicating whether a node is initially damaged $x_i = 0$ or not $x_i = 1$. Similarly, we assume that the initial damage of the hyperedges is exactly known and encoded by the product $y_\alpha\theta(|\alpha| - n) \in \{0, 1\}$ where $y_\alpha \in \{0, 1\}$ indicates whether the hyperedge α is randomly damaged while $\theta(|\alpha| - n)$ enforces the deterministic damage of hyperedges of cardinality less then n. [Note that here $\theta(z)$ indicates the Heaviside function $\theta(z) = 1$ if $z \geqslant 0$ and $\theta(z) = 0$ otherwise.]

The message-passing equations for (k, n)-core percolation are here derived starting from the definition of the (k, n)-core and the message-passing equation for hypergraph percolation [27].

At step (2′) of the pruning algorithm, assuming that we know the indicator functions $s_i \in \{0, 1\}$ and $s_\alpha \in \{0, 1\}$ indicating whether a node i or a hyperedge α is intact or not, the message-passing equations that determine the nodes in the giant subhypergraph induced by the intact nodes and hyperedges are the ones of hypergraph percolation [27]:

$$\hat{w}_{i\to\alpha} = s_i\left[1 - \prod_{\beta\in N(i)\setminus\alpha}(1 - \hat{v}_{\beta\to i})\right],$$
$$\hat{v}_{\alpha\to i} = s_\alpha\left(\prod_{j\in N(\alpha)\setminus i} s_j\right)\left[1 - \prod_{j\in N(\alpha)\setminus i}(1 - \hat{w}_{j\to\alpha})\right], \quad (19)$$

where $N(i)$ denotes the set of neighbors of node i and $N(\alpha)$ indicates the set of neighbors of factor node α. The indicator functions $\hat{\sigma}_i \in \{0, 1\}$ and $\hat{r}_\alpha \in \{0, 1\}$ indicating whether nodes and hyperedges are in the giant subhypergraph are expressed in terms of s_i and s_α and are given, respectively, by the equations [27]

$$\hat{\sigma}_i = s_i\left[1 - \prod_{\beta\in N(i)}(1 - \hat{v}_{\beta\to i})\right],$$
$$\hat{r}_\alpha = s_\alpha\left(\prod_{j\in N(\alpha)} s_j\right)\left[1 - \prod_{j\in N(\alpha)}(1 - \hat{w}_{j\to\alpha})\right]. \quad (20)$$

Note, however, that these equations assume that the indicator functions s_i and s_α are known, while here we want to obtain message-passing equations and also be able to determine their value as obtained by implementing the pruning algorithm (1′). According to the definition of the hypergraph (k, n)-core, the indicator function s_i obtained by the pruning algorithm is nonzero only if the node i receives at least k positive messages from its neighbors,

$$s_i = x_i \sum_{\substack{\Theta\subseteq N(i)\\ |\Theta|\geqslant k}} \prod_{\beta\in\Theta} \hat{v}_{\beta\to i} \prod_{\gamma\in N(i)\setminus\Theta} (1-\hat{v}_{\gamma\to i}). \quad (21)$$

The indicator function s_α obtained by the pruning algorithm is simply given by

$$s_\alpha = y_\alpha \theta(|\alpha| - n). \quad (22)$$

In order to get the message-passing equations for (k, n)-core percolation we need to insert these expressions for s_i and s_α into Eq. (19). Considering that the messages $\hat{v}_{\alpha\to i}$ is defined under the assumption that node i is in the giant subhypergraph and $\hat{w}_{i\to\alpha}$ is defined under the assumption that hyperedge α is in the giant subhypergraph, and exploiting the fact that the messages take only the 0,1 values, we get

$$\hat{w}_{i\to\alpha} = x_i \sum_{\substack{\Theta\subseteq N(i)\setminus\alpha\\ |\Theta|\geqslant k-1}} \prod_{\beta\in\Theta} \hat{v}_{\beta\to i} \prod_{\gamma\in N(i)\setminus(\Theta\cup\alpha)} (1-\hat{v}_{\gamma\to i}),$$

$$\hat{v}_{\alpha\to i} = y_\alpha\theta(|\alpha| - n) \prod_{j\in N(\alpha)\setminus i} \hat{w}_{j\to\alpha}. \quad (23)$$

Providing an intuitive explanation of these equations and their derivation might be instructive. The expression for $\hat{w}_{i\to\alpha}$ in (19) implies that the node sends a positive message to a neighbor if it is intact ($s_i = 1$) and if it receives at least a positive message from one of its hyperedges. The expression for s_i in Eq. (21) expresses that node i is intact if it is not initially damaged ($x_i = 1$) and if it receives at least k-positive messages from its neighbor hyperedges. It follows that under the assumption that α is in the giant subhypergraph, the message $\hat{w}_{i\to\alpha}$ is equal to one, if and only if $x_i = 1$, and node i receives at least $k - 1$ positive messages from neighbor hyperedges different from α as expressed by the first equation in (23). Similarly the equation for $\hat{v}_{\alpha\to i}$ in (19) implies that one hyperedge can send a positive message only if (1) it is not initially damaged, (2) all its nodes are intact, and (3) it receives at least a positive message from one of its nodes. The condition that all the nodes of the hyperedge must be intact, (i.e., must have $s_i = 1$) combined with the expression of s_i given by Eq. (21) implies that every node of the hyperedge should be connected to at least other $k - 1 \geqslant 1$ intact hyperedges. This happens if and only if each of these nodes sends a positive message to the hyperedge α leading to the second equation (23).

Following a similar line of thought, and exploiting the fact that both s_i, s_α and the messages $\hat{w}_{i\to\alpha}$, $\hat{v}_{\alpha\to i}$ are taking values 0,1, one can immediately show that

$$\hat{\sigma}_i = s_i = x_i \sum_{\substack{\Theta\subseteq N(i)\\ |\Theta|\geqslant k}} \prod_{\beta\in\Theta} \hat{v}_{\beta\to i} \prod_{\gamma\in N(i)\setminus\Theta} (1-\hat{v}_{\gamma\to i}),$$

$$\hat{r}_\alpha = y_\alpha\theta(|\alpha| - n) \prod_{j\in N(\alpha)} \hat{w}_{j\to\alpha}. \quad (24)$$

Hence Eqs. (23) and (24) are the message-passing equations for hypergraph (k, n)-core percolation when the initial random damage of nodes, i.e., $\{x_\alpha\}$, and of hyperedges, i.e., $\{y_\alpha\}$, is known.

Another set of message-passing equations hold when we do not have direct access to the configuration of the initial damage $\{x_i\}$, $\{y_\alpha\}$ but we know only the probability that nodes and hyperedges are initially intact, i.e., p_N and p_H, respectively. This second set of message-passing equations can be simply obtained by averaging the messages over the initial damage distribution

$$P(\{x_i\}, \{y_\alpha\}) = \prod_{i=1}^{N} p_N^{x_i}(1-p_N)^{1-x_i} \times \prod_{\alpha=1}^{M} p_H^{y_\alpha}(1-p_H)^{1-y_\alpha}. \quad (25)$$

In this way we obtain the following set of message-passing equations (note that the messages $w_{i\to\alpha}$ and $v_{\alpha\to i}$ now take real values between 0 and 1):

$$w_{i\to\alpha} = p_N \sum_{\substack{\Theta\subseteq N(i)\setminus\alpha\\ |\Theta|\geqslant k-1}} \prod_{\beta\in\Theta} v_{\beta\to i} \prod_{\gamma\in N(i)\setminus(\Theta\cup\alpha)} (1-v_{\gamma\to i}),$$

$$v_{\alpha\to i} = p_H\theta(|\alpha| - n) \prod_{j\in N(\alpha)\setminus i} w_{j\to\alpha}. \quad (26)$$

The probability σ_i that the node i belongs to the (k, n)-core and the probability r_α that the hyperedge α belongs to the (k, n)-core are given by

$$\sigma_i = p_N \sum_{\substack{\Theta\subseteq N(i)\\ |\Theta|\geqslant k}} \prod_{\beta\in\Theta} v_{\beta\to i} \prod_{\gamma\in N(i)\setminus\Theta} (1-v_{\gamma\to i}),$$

$$r_\alpha = p_H\theta(|\alpha| - n) \prod_{j\in N(\alpha)} w_{j\to\alpha}. \quad (27)$$

It follows that Eq. (26) and Eq. (27) uniquely determine the hypergraph (k, n)-core when we know only the probabilities p_N and p_H that nodes and hyperedges are initially undamaged, respectively. The fraction S_{kn} of nodes in the (k, n)-core can be expressed in terms of σ_i and r_α given by Eq. (27) as

$$S_{kn} = \frac{1}{N}\sum_{i=1}^{N} \sigma_i, \quad (28)$$

and the fraction R_{kn} of hyperedges in the (k, n)-core is given by

$$R_{kn} = \frac{1}{M}\sum_{\alpha=1}^{M} r_\alpha. \quad (29)$$

C. Hypergraph (k, n)-core percolation on random hypergraph

In many occasions the exact structure of the hypergraph might be unknown, and so we need to rely on predictions based on the hypergraph ensembles from which the hypergraph is drawn. Here we consider the ensembles of random hypergraphs $\mathcal{H} = (V, H)$ of $N = |V|$ nodes and $M = |H|$ hyperedges whose node degree distribution is $P(q)$ and whose distribution of hyperedge cardinalities is $Q(m)$. In this ensemble the probability of a hypergraph $\mathcal{H}$ of incidence matrix $\mathbf{b}$ is given by

$$P(\mathcal{H}) = \prod_{i=1}^{N}\prod_{\alpha=1}^{M} p_{i\alpha}^{b_{i\alpha}}(1-p_{i\alpha})^{1-b_{i\alpha}}, \tag{30}$$

with

$$p_{i\alpha} = \frac{q_i m_\alpha}{\langle q\rangle N}, \tag{31}$$

where q_i is the degree of node i and m_α is the cardinality of hyperedge α.

We can obtain the analytical equations determining the (k,n)-core percolation problem in this ensemble by averaging the message-passing equation over the probability $P(\mathcal{H})$.

Let us indicate with V and W, respectively, the averages of the messages $v_{i\to\alpha}$ and $w_{\alpha\to i}$ over the distribution $P(\mathcal{H})$. We obtain then the equations of V and W given by

$$\begin{aligned} V &= p_H \sum_{m\geqslant n} \frac{mQ(m)}{\langle m\rangle} W^{m-1},\\ W &= p_N \sum_{q=k}^{\infty} \frac{qP(q)}{\langle q\rangle} \sum_{s=k-1}^{q-1} \binom{q-1}{s} V^s(1-V)^{q-1-s}. \end{aligned} \tag{32}$$

Similarly the fractions of vertices, S_{kn}, and hyperedges, R_{kn}, belonging to the (k,n)-core can be obtained by Eq. (28) and Eq. (29) by averaging over $P(\mathcal{H})$ giving

$$R_{kn} = p_H \sum_{m\geqslant n} Q(m) W^m, \tag{33}$$

$$S_{kn} = p_N \sum_{q=k}^{\infty} P(q) \sum_{s=k}^{q} \binom{q}{s} V^s (1-V)^{q-s}. \tag{34}$$

In particular, setting $n=2$, we obtain the formulas for the k-cores in this problem.

Using the generating functions, we get

$$V = p_H\left[G_{Q1}(W) - \sum_{m<n}\frac{mQ(m)}{\langle m\rangle}W^{m-1}\right], \tag{35}$$

$$W = p_N - \frac{p_N}{\langle q\rangle}\sum_{s=0}^{k-2}\frac{V^s}{s!}G_P^{(s+1)}(1-V), \tag{36}$$

where $G_{Q1}(x) \equiv G'_Q(x)/G'_Q(1) = G'_Q(x)/\langle m\rangle$, and

$$R_{kn} = p_H\left[G_Q(W) - \sum_{m<n} Q(m)W^m\right], \tag{37}$$

$$S_{kn} = p_N - p_N\sum_{s=0}^{k-1}\frac{V^s}{s!}G_P^{(s)}(1-V). \tag{38}$$

D. Critical behavior of (k, n)-core percolation on random hypergraphs

The hypergraph (k,n)-core percolation process has a critical behavior that differs significantly from the k-core percolation on simple networks and the (k,n)-core percolation on factor graphs. One of the most striking properties of (k,n)-core percolation is the presence of discontinuous phase transitions also for $k=2$, while the k-core percolation on simple networks and the (k,n)-core percolation on factor graph are discontinuous only for $k\geqslant 3$. Here we will emphasize this significant difference showing that a continuous transition in hypergraph (k,n)-core percolation is possible only for $(k,n)=(2,2)$-core percolation also displaying the tricritical point at which the (2,2)-core percolation transition changes from continuous to hybrid transitions, i.e., discontinuous transitions displaying a singularity above the transition (see for a definition of hybrid transitions and background information [31,41,48]).

We consider the (k,n)-core percolation transitions on random hypergraphs captured by Eq. (32). By defining the functions

$$\begin{aligned} f_W(W) &= p_H \sum_{m\geqslant n}\frac{mQ(m)}{\langle m\rangle}W^{m-1},\\ f_V(V) &= p_N \sum_{q\geqslant k}\frac{qP(q)}{\langle q\rangle}\sum_{s\geqslant k-1}^{q-1}\binom{q-1}{s}V^s(1-V)^{q-1-s}, \end{aligned} \tag{39}$$

we write Eq. (32) as

$$V = f_W(W), \quad W = f_V(V). \tag{40}$$

These equations can be written as

$$h(V) = V - f_W(f_V(V)) = 0. \tag{41}$$

According to the theory of critical phenomena [41], we see that the lines of continuous (second-order) phase transitions are determined by the conditions

$$h(0) = 0, \quad h'(0) = 0; \tag{42}$$

the tricritical point is determined by

$$h(0) = 0, \quad h'(0) = 0, \quad h''(0) = 0; \tag{43}$$

while the lines of discontinuous (hybrid) phase transitions are determined by the equations

$$h(V^\star) = 0, \quad h'(V^\star) = 0 \tag{44}$$

with $V^\star > 0$. By direct inspection of these equations, it emerges immediately that the continuous (second-order) transitions lines and the tricritical point can occur only for $(k,n) = (2,2)$. In particular, a second order takes place for

$$1 = p_H p_N \frac{2Q(2)}{\langle m\rangle}\frac{\langle q(q-1)\rangle}{\langle q\rangle}, \tag{45}$$

while the tricritical point occurs when Eq. (45) is satisfied together with the following equation:

$$\frac{\langle q(q-1)(q-2)\rangle}{\langle q(q-1)\rangle} = p_H p_N^2 \frac{6Q(3)}{\langle m\rangle}\left(\frac{\langle q(q-1)\rangle}{\langle q\rangle}\right)^2. \tag{46}$$

Let us apply these equations to the uncorrelated hypergraph with a Poisson degree distribution $P(q)$ and a shifted Poisson cardinality distribution $Q(m)$. Note that we need to shift the Poisson $Q(m)$ distribution since cardinalities $m = 0, 1$ are

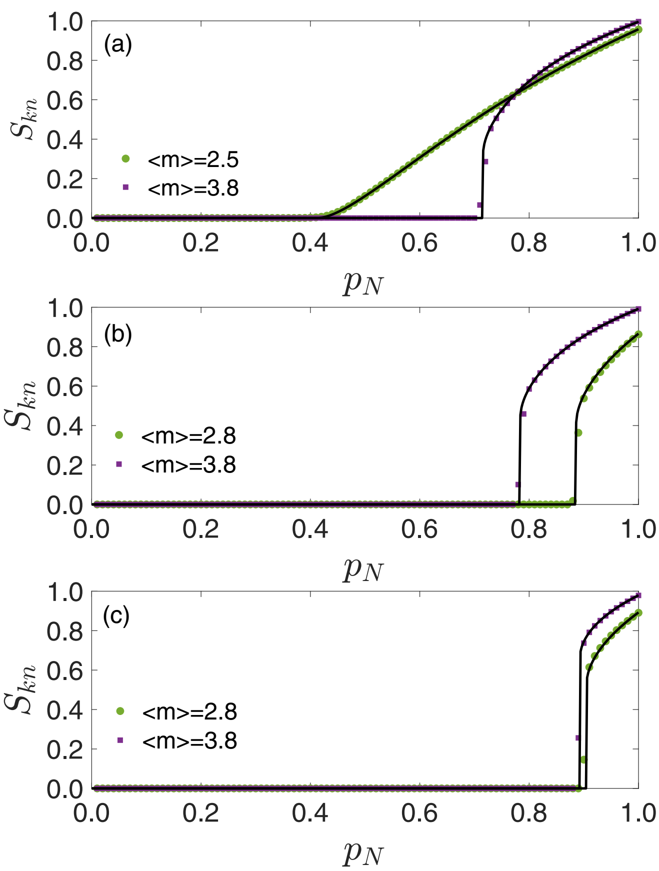

FIG. 2. Fraction of nodes in the (k, n)-core S_{kn} as a function of p_N for different (k, n)-cores and $p_H = 1$. The (k, n)-core percolation is displayed for the (2,2)-core (a), the (2,3)-core (b), and the (3,2)-core (c). The (2,2) percolation transition is continuous for $\langle m\rangle = 2.5$ and discontinuous for $\langle m\rangle = 3.5$. Note that for $\langle q\rangle = 2\langle m\rangle$ and $p_H = 1$ the tricritical point of the (2,2)-core occurs at $p_N = 0.492143\ldots$, $\langle m\rangle = 2.67731\ldots$. All hypergraphs have Poisson cardinality and degree distributions defined in Eq. (47) with average degree $\langle q\rangle = 2\langle m\rangle$, while $\langle m\rangle$ is indicated in the legend. Symbols indicate simulations obtained for $N = 10^4$ node hypergraphs averaged 100 times, and solid lines indicate our theoretical predictions.

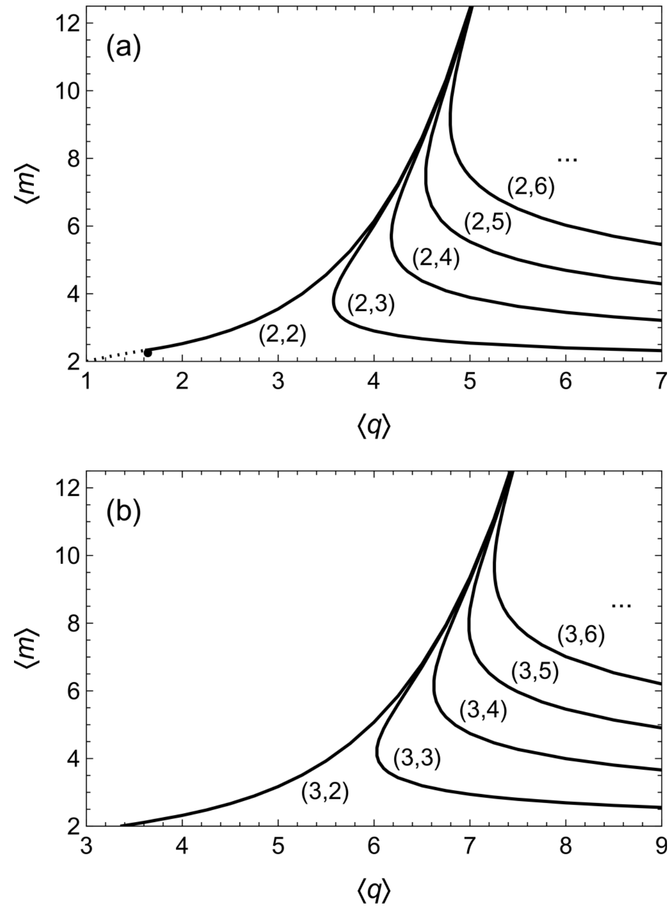

FIG. 3. Phase diagram of (k, n)-core percolation in the $\langle q\rangle$-$\langle m\rangle$ plane is shown for uncorrelated Poisson hypergraphs with $p_H = p_N = 1$. (a) Phase diagram for (k, n)-core percolation with (k, n) given by (2,2), (2,3), (2,4), (2,5), and (2,6); (b) core percolation phase diagram for (k, n) given by (3,2), (3,3), (3,4), (3,5), and (3,6). Each (k, n)-core exists in the whole region to the right of the corresponding boundary. All boundaries are discontinuous transitions with one exception, the leftmost dotted piece of the boundary for the (2,2)-core, which is a continuous phase transition. The tricritical point at the (2,2)-core's phase boundary is $\langle q\rangle_{\text{tricritical}} = 1.628\ldots = 7e^{1/3}/6$, $\langle m\rangle_{\text{tricritical}} = 2.333\ldots = 7/3$. For the (2,2)-core, the phase boundary ends at the point $\langle q\rangle = 1$, $\langle m\rangle = 2$. For the (3,2)-core, the phase boundary ends at the point $\langle q\rangle = 3.350919$, $\langle m\rangle = 2$.

impossible:

$$P(q) = e^{-\langle q\rangle}\frac{\langle q\rangle^q}{q!},$$
$$Q(m \geqslant 2) = e^{-(\langle m\rangle-2)}\frac{(\langle m\rangle - 2)^{m-2}}{(m-2)!}. \tag{47}$$

Note that $\langle m\rangle$ and $\langle q\rangle$ satisfy the equality $\langle m\rangle M = \langle q\rangle N$, so that $\langle m\rangle/\langle q\rangle = N/M$. The generating functions for these distributions are

$$G_P(z) = e^{\langle q\rangle(z-1)},$$
$$G_Q(z) = z^2 e^{(\langle m\rangle-2)(z-1)}. \tag{48}$$

In Fig. 2 we show the sizes of the (2,2), (2,3), and (3,2)-cores obtained from Monte Carlo simulations on random hypergraphs with the Poisson cardinality and degree distributions given by Eq. (47) with $\langle q\rangle = 2\langle m\rangle$ and $p_H = 1$. The simulations are in excellent agreement with our theoretical results and demonstrate that the (2,2)-core percolation can display both continuous and discontinuous transitions. Note that for $\langle q\rangle = 2\langle m\rangle$ and $p_H = 1$ the tricritical point of the (2,2)-core occurs at $p_N = 0.492143\ldots$, $\langle m\rangle = 2.67731\ldots$.

The phase diagram of the (k, n)-core percolation for $p_H = p_N = 1$ is shown in Fig. 3 in the $\langle q\rangle$-$\langle m\rangle$ phase space for $(2, n)$- and $(3, n)$-cores. The difference from the k-core problem for ordinary graphs, where the phase transition is continuous for $k = 2$ and discontinuous for $k \geqslant 3$ is apparent. Indeed, for random hypergraphs, the (2,2)-core phase boundary consists of two lines—a continuous transition (dotted) and a discontinuous one (solid)—converging at the tricritical point

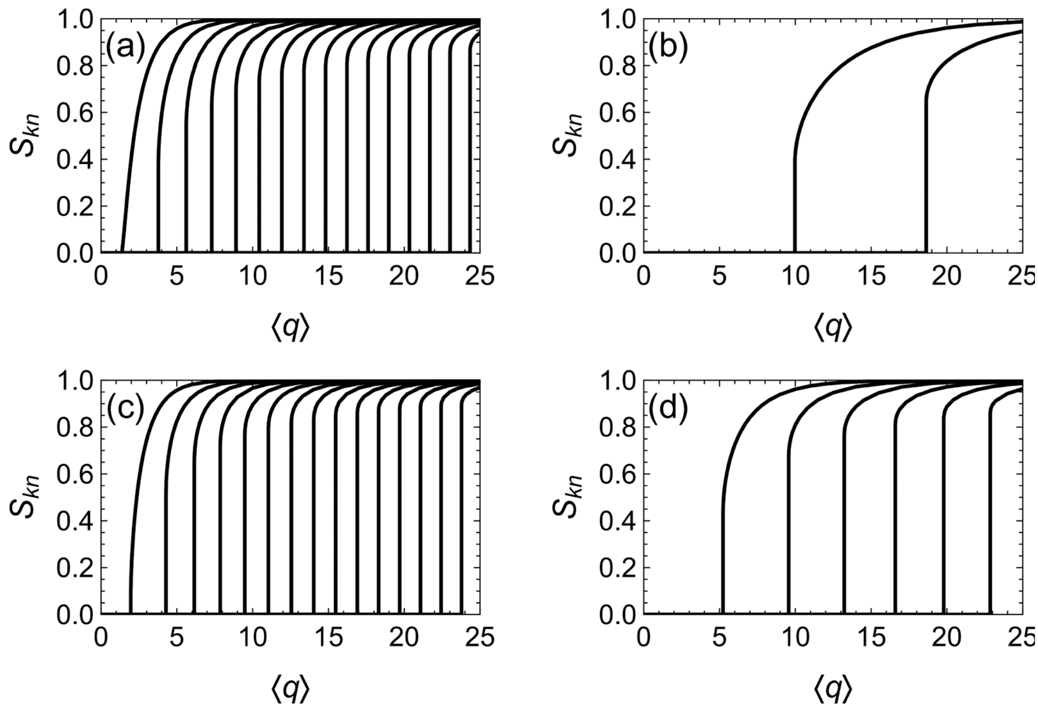

FIG. 4. Relative sizes S_{kn} of the (k, n)-cores vs $\langle q\rangle$ for $\langle m\rangle = 2.2$ (a, b) and $\langle m\rangle = 2.5$ (c, d) plotted for $p_N = p_H = 1$. (a) Curves from left to the right display the (k, n)-core percolation for (k, n) given by (2,2), (3, 2), (4, 2), (5, 2), (6, 2), (7, 2), (8, 2), (9, 2), (10, 2), (11, 2), (12, 2), (13, 2), (14, 2), (15, 2), (16, 2), and (17, 2); (b) curves from left to right display the (k, n)-core percolation for (k, n) given by (2, 3) and (3, 3). The phase transition for the (2,2)-core for $\langle m\rangle = 2.2$ is continuous, and for the other (k, n)-cores the transitions are hybrid. (c) Curves from left to the right display the (k, n)-core percolation with (k, n) given by (2,2), (3, 2), (4, 2), (5, 2), (6, 2), (7, 2), (8, 2), (9, 2), (10, 2), (11, 2), (12, 2), (13, 2), (14, 2), (15, 2), and (16, 2); (d) curves from left to right display (k, n)-core percolation with (k, n) given by (2, 3), (3, 3), (4, 3), (5, 3), (6, 3), and (7, 3).

with the following coordinates:

$$\langle q\rangle_{\text{tricritical}} = \frac{7}{6}\, e^{1/3} = 1.628\ldots,$$
$$\langle m\rangle_{\text{tricritical}} = \frac{7}{3} = 2.333\ldots. \tag{49}$$

The phase boundary for $(2, n)$- and $(3, n)$-core percolation can be obtained by imposing that $\langle m\rangle = 2$, which corresponds to the minimum possible value for the average cardinality of the hyperedges. Indeed, for $\langle m\rangle = 2$ the hypergraph reduces to an ordinary network.

For the (2,2)-core, the phase boundary ends at the point $\langle q\rangle = 1$, $\langle m\rangle = 2$. One can see in Fig. 8(b) that for the (3,2)-core, the phase boundary ends at a point on a line $\langle m\rangle = 2$. The coordinate $\langle q\rangle$ of this point can be obtained exactly. We substitute the degree and cardinality distributions, Eq. (47), into the equations (44) for $k = 3$, $n = 2$, and $\langle m\rangle = 2$ [one can conveniently use the generating functions of the distributions, Eq. (48)]. This results in the following equations:

$$V^\star = 1 - (1 + \langle q\rangle V^\star)e^{-\langle q\rangle V^\star},$$
$$1 = \langle q\rangle^2 V^\star e^{-\langle q\rangle V^\star} \tag{50}$$

for $\langle q\rangle$ and $V^\star$. Excluding $V^\star$ from Eq. (50),

$$\langle q\rangle V^\star = \tfrac{1}{2}[\langle q\rangle - 1 + \sqrt{(\langle q\rangle - 1)^2 - 4}], \tag{51}$$

we get the equation for $\langle q\rangle$:

$$e^{\langle q\rangle - 1 + \sqrt{(\langle q\rangle - 1)^2 - 4}} = \langle q\rangle^2 \frac{\langle q\rangle - 1 + \sqrt{(\langle q\rangle - 1)^2 - 4}}{\langle q\rangle - 1 - \sqrt{(\langle q\rangle - 1)^2 - 4}}, \tag{52}$$

whose root $\langle q\rangle = 3.350919$. Thus the phase boundary of the (3,2)-core ends at the point $\langle q\rangle = 3.350919$, $\langle m\rangle = 2$.

Figure 4 shows the dependencies of the relative sizes S_{kn} of the (k, n)-cores on $\langle q\rangle$ for different values of mean cardinality $\langle m\rangle$ for the Poisson hypergraph with $p_N = p_H = 1$.

V. HYPERGRAPH (k, n)-CORE SECOND-NEIGHBOR PROBLEMS AND THEIR PRUNING ALGORITHM

A. Second-neighbor pruning algorithm

Until now we have defined the (k, n)-core of a the hypergraph based on a pruning algorithm that prunes nodes and hyperedges according to their connectivity. However, there is another possibility, i.e., pruning nodes or hyperedges depending on the state of their second neighbors in the factor graph. This implies a set of algorithms pruning nodes considering the connectivity of the hyperedges they belong to pruning hyperedges according to the connectivity of the nodes belonging to it.

To this end we distinguish two types of second-neighbor (k, n)-core problems: in the first nodes are iteratively pruned, and in the second hyperedges are iteratively pruned. We note that there is no symmetry between these two pruning algorithms. This is due to the fact that in hypergraph percolation hyperedges in order to be belong to the giant component must have all their nodes undamaged, while no corresponding constraint holds for the nodes. Interestingly we will observe important differences between the second-neighbor

(k, n)-cores obtained pruning only nodes and the ones obtained pruning only hyperedges.

Let us consider these two algorithms and their corresponding message-passing equations separately.

B. Message-passing equations for second-neighbor node-pruning algorithm

Let us start from a configuration in which we initially damage nodes with probability $1 - p_N$ and/or hyperedges with probability $1 - p_H$. If we consider the pruning on the nodes the second-neighbor hypergraph (k, n)-core can be obtained using the following pruning algorithm:

(1) Damage iteratively all nodes belonging to fewer than k hyperedges each connected to at least n (undamaged) nodes.

(2) Define the (k, n)-core as the giant component of the network induced by the undamaged nodes and their connected hyperedges.

As we will see in the following this algorithm is very closely related to the algorithm defined in Sec. IV A. Note that also in this case, as in the algorithm defined in Sec. IV A, due to the definition of the hypergraph giant component, every hyperedge of cardinality $m \geqslant n$ belonging to the (k, n)-core will be connected to exactly $m \geqslant n$ undamaged nodes. Therefore the result of the algorithm is unchanged if only hyperedges of cardinality less than n are pruned at stage (1).

In order to derive the corresponding message-passing equation we start with the message-passing equations [27] implementing hypergraph percolation at step (2). Using the same notation used in Sec. IV A we see therefore that the messages $\hat{w}_{i\to\alpha}$, $\hat{v}_{\alpha\to i}$ obey

$$\hat{w}_{i\to\alpha} = s_i\left[1 - \prod_{\beta\in N(i)\setminus\alpha}(1 - \hat{v}_{\beta\to i})\right],$$
$$\hat{v}_{\alpha\to i} = y_\alpha\left(\prod_{j\in N(\alpha)\setminus i} s_j\right)\left[1 - \prod_{j\in N(\alpha)\setminus i}(1 - \hat{w}_{j\to\alpha})\right], \qquad (53)$$

where s_i indicates whether node i has been damaged or pruned, $s_i = 0$, or not, $s_i = 1$. The indicator functions $\hat{\sigma}_i \in \{0, 1\}$ and $\hat{r}_\alpha \in \{0, 1\}$, indicating whether nodes and hyperedges are in the giant component and hence in the (k, n)-core, are expressed in terms of s_i, x_i, and y_α, and given, respectively, by the equations

$$\hat{\sigma}_i = s_i\left[1 - \prod_{\beta\in N(i)}(1 - \hat{v}_{\beta\to i})\right],$$
$$\hat{r}_\alpha = y_\alpha\left(\prod_{j\in N(\alpha)} s_j\right)\left[1 - \prod_{j\in N(\alpha)}(1 - \hat{w}_{j\to\alpha})\right]. \qquad (54)$$

The pruning of the nodes determines the indicator function s_i, which is nonzero only if the node i receives at least k positive messages from its hyperedge neighbors of cardinality at least n,

$$s_i = x_i \sum_{\substack{\Theta\subseteq N(i)\\ |\Theta|\geqslant k}} \prod_{\beta\in\Theta} \tilde{v}_{\beta\to i} \prod_{\gamma\in N(i)\setminus\Theta}(1 - \tilde{v}_{\gamma\to i}), \qquad (55)$$

where we have defined

$$\tilde{v}_{\alpha\to i} = \theta(|\alpha| - n)\hat{v}_{\alpha\to i} \qquad (56)$$

with $\theta(x) = 1$ if $x \geqslant 0$ and $\theta(x) = 0$ otherwise. Inserting this expression of s_i into Eq. (53), and taking into account that the messages $\hat{v}_{\alpha\to i}$ are defined under the assumption that node i is in the giant hypergraph component and the messages $\hat{w}_{i\to\alpha}$ are defined under the assumption that hyperedge α is in the giant component, exploiting the fact that the messages take only 0, 1 values we get

$$\hat{w}_{i\to\alpha} = x_i \sum_{\substack{\Theta\subseteq N(i)\setminus\alpha\\ |\Theta|\geqslant k-1}} \prod_{\beta\in\Theta} \tilde{v}_{\beta\to i} \prod_{\gamma\in N(i)\setminus(\Theta\cup\alpha)}(1 - \tilde{v}_{\gamma\to i}),$$
$$\tilde{v}_{\alpha\to i} = y_\alpha\theta(|\alpha| - n)\prod_{j\in N(\alpha)\setminus i}\hat{w}_{j\to\alpha}. \qquad (57)$$

Exploiting furthermore the fact that both s_i and the messages $\hat{w}_{i\to\alpha}$, $\hat{v}_{\alpha\to i}$ take values 0, 1, it is also immediate to show that

$$\hat{\sigma}_i = s_i = x_i \sum_{\substack{\Theta\subseteq N(i)\\ |\Theta|\geqslant k}} \prod_{\beta\in\Theta} \tilde{v}_{\beta\to i} \prod_{\gamma\in N(i)\setminus\Theta}(1 - \tilde{v}_{\gamma\to i}),$$
$$\hat{r}_\alpha = y_\alpha \prod_{j\in N(\alpha)}\hat{w}_{j\to\alpha}. \qquad (58)$$

Another set of message-passing equations holds when we do not know direct access to the configuration of the initial damage $\{x_i\}$, $\{y_\alpha\}$ but we know only the probability that nodes and hyperedges are initially intact, i.e., p_N and p_H, respectively. This second set of message-passing equations can be simply obtained by averaging the messages over the initial damage distribution $P(\{x_i\}, \{y_\alpha\})$ given by Eq. (25). In this way we obtain the following set of message-passing equations (note that the messages $w_{i\to\alpha}$ and $v_{\alpha\to i}$ now take real values between 0 and 1):

$$w_{i\to\alpha} = p_N \sum_{\substack{\Theta\subseteq N(i)\setminus\alpha\\ |\Theta|\geqslant k-1}} \prod_{\beta\in\Theta} v_{\beta\to i} \prod_{\gamma\in N(i)\setminus(\Theta\cup\alpha)}(1 - v_{\gamma\to i}),$$
$$v_{\alpha\to i} = p_H\theta(|\alpha| - n)\prod_{j\in N(\alpha)\setminus i} w_{j\to\alpha}. \qquad (59)$$

The probability σ_i that the node i belongs to the (k, n)-core and the probability r_α that the hyperedge α belongs to the (k, n)-core are given by

$$\sigma_i = p_N \sum_{\substack{\Theta\subseteq N(i)\\ |\Theta|\geqslant k}} \prod_{\beta\in\Theta} v_{\beta\to i} \prod_{\gamma\in N(i)\setminus\Theta}(1 - v_{\gamma\to i}),$$
$$r_\alpha = p_H \prod_{j\in N(\alpha)} w_{j\to\alpha}. \qquad (60)$$

The fraction S_{kn} of nodes in the (k, n)-core is given by

$$S_{kn} = \frac{1}{N}\sum_{i=1}^{N}\sigma_i, \qquad (61)$$

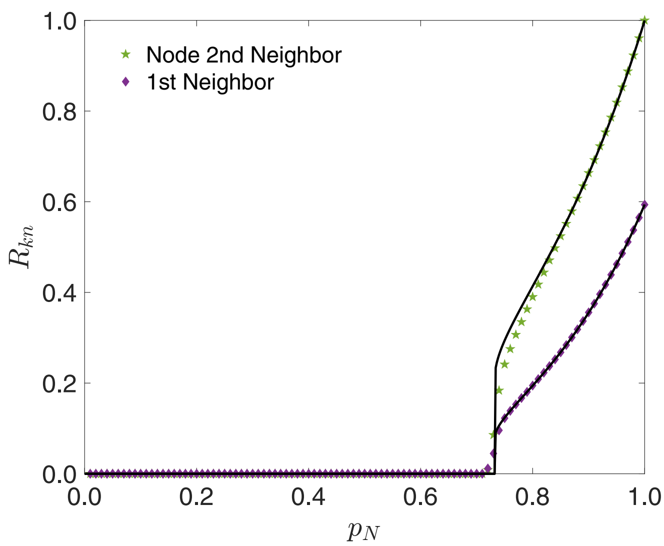

FIG. 5. Fraction of hyperedges R_{kn} in the (2,4)-core is plotted vs p_N for the second-neighbor node (2,4)-core and for the first-neighbor (2,4)-core percolation. The percolation threshold is the same but R_{kn} differ. The hypergraph has Poisson cardinality and degree distribution with $\langle m\rangle = 4$, $\langle q\rangle = 16$ and number of nodes $N = 5000$. The symbols correspond to Monte Carlo simulations averaged over 100 iterations for $p_H = 1$. The solid lines are the theoretical predictions.

and the fraction R_{kn} of hyperedges in the (k, n)-core is given by

$$R_{kn} = \frac{1}{M}\sum_{\alpha=1}^{M} r_\alpha. \tag{62}$$

Therefore this algorithm is essentially reduced to the first-neighbor (k, n)-core studied in Sec. IV A. Indeed, the percolation threshold and the nature of the transition are the same, and the fractions of nodes within these (k, n)-cores, S_{kn}, also coincide, while only the fractions of hyperedges within the (k, n)-cores, R_{kn}, can in general differ; see Fig. 5.

C. Message-passing equations for second-neighbor hyperedge pruning algorithm

We start from the configuration in which we initially damage either nodes with probability $1 - p_N$ and/or hyperedges with probability $1 - p_H$. If we consider the pruning on the hyperedges of the hypergraph, the (k, n)-core can be obtained using the following pruning algorithm:

(1) Damage iteratively all hyperedges having fewer than n nodes each connected to at least k undamaged hyperedges.

(2) Define the (k, n)-core as the giant component of the network induced by the undamaged hyperedges.

This (k, n)-core is the maximal connected subhypergraph, each of whose hyperedges has at least n nodes with degrees at least k. Note that this pruning algorithm doesn't change the number of nodes in the network. It stays equal to N.

Our starting point is always the set of message-passing equations for hypergraph percolation [27] where now at step (2) of the pruning process each hyperedge α is either damaged ($s_\alpha = 0$) or not damaged ($s_\alpha = 1$). Using always the same notations we have been using so far we get

$$\hat{w}_{i\to\alpha} = x_i\left[1 - \prod_{\beta\in N(i)\setminus\alpha}(1 - \hat{v}_{\beta\to i})\right],$$

$$\hat{v}_{\alpha\to i} = s_\alpha\left(\prod_{j\in N(\alpha)\setminus i} x_j\right)\left[1 - \prod_{j\in N(\alpha)\setminus i}(1 - \hat{w}_{j\to\alpha})\right]. \tag{63}$$

The indicator functions $\hat{\sigma}_i \in \{0, 1\}$ and $\hat{r}_\alpha \in \{0, 1\}$ indicating whether nodes and hyperedges are in the giant component at step (2) and hence in the (k, n)-core are expressed in terms of x_i and s_α, and they are given, respectively, by the equations

$$\hat{\sigma}_i = x_i\left[1 - \prod_{\beta\in N(i)}(1 - \hat{v}_{\beta\to i})\right],$$

$$\hat{r}_\alpha = s_\alpha\left(\prod_{j\in N(\alpha)} x_j\right)\left[1 - \prod_{j\in N(\alpha)}(1 - \hat{w}_{j\to\alpha})\right]. \tag{64}$$

The pruning of the hyperedge determines the indicator function s_α. The indicator function s_α is nonzero only if the hyperedge is connected to al least n nodes, each connected to at least k undamaged hyperedges [belonging to the (k, n)-core or giant component]. Let us define a node i to be active if it receives at least k positive messages from its neighbors,

$$a_i = x_i \sum_{\substack{\Theta\subseteq N(i)\\ |\Theta|\geqslant k}} \prod_{\beta\in\Theta} \hat{v}_{\beta\to i} \prod_{\gamma\in N(i)\setminus\Theta}(1 - \hat{v}_{\gamma\to i}). \tag{65}$$

The damage of the hyperedges is therefore determined by the activity of their nodes by

$$s_\alpha = y_\alpha \sum_{\substack{\Theta\subseteq N(\alpha)\\ |\Theta|\geqslant n}} \prod_{r\in\Theta} a_r\hat{w}_{r\to\alpha} \prod_{r\in N(\alpha)\setminus\Theta}(1 - a_r\hat{w}_{r\to\alpha}). \tag{66}$$

Let us define $\tilde{w}_{i\to\alpha}$ as

$$\tilde{w}_{i\to\alpha} = a_i\hat{w}_{i\to\alpha}. \tag{67}$$

In order to express $\hat{v}_{\alpha\to i}$ we need to distinguish the case in which i is active, i.e., $a_i = 1$, and the case in which a_i is not active. Moreover, using the fact that s_α, a_i and the message all take values 0, 1 we obtain that

$$\begin{aligned}\hat{v}_{\alpha\to i} =& \tilde{y}_{\alpha,i}a_i \sum_{\substack{\Theta\subseteq N(\alpha)\setminus i\\ |\Theta|\geqslant n-1}} \prod_{r\in\Theta}\tilde{w}_{r\to\alpha} \prod_{r\in N(\alpha)\setminus(\Theta\cup i)}(1 - \tilde{w}_{r\to\alpha})\\ &+ \tilde{y}_{\alpha,i}(1 - a_i) \sum_{\substack{\Theta\subseteq N(\alpha)\setminus i\\ |\Theta|\geqslant n}} \prod_{r\in\Theta}\tilde{w}_{r\to\alpha} \prod_{r\in N(\alpha)\setminus(\Theta\cup i)}(1 - \tilde{w}_{r\to\alpha}),\end{aligned}$$

where $\tilde{y}_{\alpha,i}$ is given by

$$\tilde{y}_{\alpha,i} = y_\alpha\left(\prod_{j\in N(\alpha)\setminus i} x_j\right), \tag{68}$$

and $\tilde{w}_{i\to\alpha}$ is given by

$$\tilde{w}_{i\to\alpha} = a_i \sum_{\substack{\Theta\subseteq N(i)\setminus\alpha\\ |\Theta|\geqslant k-1}} \prod_{\beta\in\Theta}\hat{v}_{\beta\to i} \prod_{\gamma\in N(i)\setminus(\Theta\cup\alpha)}(1 - \hat{v}_{\gamma\to i}). \tag{69}$$

Similarly one can show that the indicator functions $\hat{\sigma}_i$ and $\hat{r}_\alpha$ are given by

$$\hat{\sigma}_i = x_i\left[1 - \prod_{\alpha\in N(i)} \hat{v}_{\alpha\to i}\right],$$
$$\hat{r}_\alpha = \hat{y}_\alpha \sum_{\substack{\Theta\subseteq N(\alpha)\\ |\Theta|\geqslant n}} \prod_{r\in\Theta} \tilde{w}_{r\to\alpha} \prod_{r\in N(i)\setminus\Theta} (1-\tilde{w}_{r\to\alpha}), \tag{70}$$

where

$$\hat{y}_\alpha = y_\alpha\left(\prod_{j\in N(\alpha)} x_j\right). \tag{71}$$

We now derive the second set of message-passing equations that hold when we do not have direct access to the configuration of the initial damage $\{x_i\}$, $\{y_\alpha\}$ but we know only the probability that nodes and hyperedges are initially intact, i.e., p_N and p_H, respectively, by averaging the messages over the initial damage distribution $P(\{x_i\}, \{y_\alpha\})$. First, we observe that $\tilde{w}_{i\to\alpha}$ are nonzero only if the node i is active. Therefore we consider only the average message $v_{\alpha\to i} = \langle a_i \hat{v}_{\alpha\to i}\rangle$ and the average message $w_{i\to\alpha} = \langle \tilde{w}_{i\to\alpha}\rangle$, which constitute the closed form equations determining the percolation threshold. In this way, paying attention to the fact that the messages take real values between 0 and 1, we obtain the following set of message-passing equations:

$$v_{\alpha\to i} = p_{HN} \sum_{\substack{\Theta\subseteq N(\alpha)\setminus i\\ |\Theta|\geqslant n-1}} \prod_{r\in\Theta} w_{r\to\alpha} \prod_{r\in N(\alpha)\setminus(\Theta\cup i)} (1-w_{r\to\alpha}),$$
$$w_{i\to\alpha} = \sum_{\substack{\Theta\subseteq N(i)\setminus\alpha\\ |\Theta|\geqslant k-1}} \prod_{\beta\in\Theta} v_{\beta\to i} \prod_{\gamma\in N(i)\setminus(\Theta\cup\alpha)} (1-v_{\gamma\to i}), \tag{72}$$

where $p_{HN} = p_H p_N^{m-1}$.

The probability r_α that the hyperedge α belongs to the (k, n)-core is given by

$$r_\alpha = p_{HN} \sum_{\substack{\Theta\subseteq N(\alpha)\\ |\Theta|\geqslant n}} \prod_{r\in\Theta} w_{r\to\alpha} \prod_{r\in N(\alpha)\setminus\Theta} (1-w_{r\to\alpha}). \tag{73}$$

However, we need additional care to express the probability σ_i that node i belongs to the (k, n)-core. In particular, the giant component will include all active nodes and the inactive nodes that are intact and are connected to at least one undamaged hyperedge. Note that an hyperedge including an inactive node can be active only if at least n of its other nodes are active. This implies that the all intact nodes will belong to the giant component unless both (1) and (2) are satisfied. These two conditions are (1) the node is not connected to any hyperedge including at least to n other active nodes and (2) the nodes belongs to fewer than k hyperedges linked to $n-1$ other active nodes. It follows from this that the probability that a node belongs to the (k, n)-core is

$$\sigma_i = p_N \times\left[1 - \left(\sum_{\substack{\Theta\subseteq N(i)\\ |\Theta|\leqslant k-1}} \prod_{r\in\Theta} \theta_{r\to i} \prod_{r\in N(\alpha)\setminus\Theta} (1-\rho_{\alpha\to i})\right)\right], \tag{74}$$

where $\theta_{\alpha\to i}$ and $\rho_{\alpha\to i}$ are given by

$$\theta_{\alpha\to i} = p_{HN} \sum_{\substack{\Theta\subseteq N(\alpha)\setminus i\\ |\Theta|= n-1}} \prod_{r\in\Theta} w_{r\to\alpha} \prod_{r\in N(\alpha)\setminus(\Theta\cup i)} (1-w_{r\to\alpha}),$$
$$\rho_{\alpha\to i} = \theta_{\alpha\to i} + p_{HN} \sum_{\substack{\Theta\subseteq N(\alpha)\setminus i\\ |\Theta|\geqslant n}} \prod_{r\in\Theta} w_{r\to\alpha} \prod_{r\in N(\alpha)\setminus(\Theta\cup i)} (1-w_{r\to\alpha}). \tag{75}$$

The fraction S_{kn} of nodes in this (k, n)-core is given by

$$S_{kn} = \frac{1}{N}\sum_{i=1}^{N} \sigma_i, \tag{76}$$

and the fraction R_{kn} of hyperedges in the (k, n)-core is given by

$$R_{kn} = \frac{1}{M}\sum_{\alpha=1}^{M} r_\alpha. \tag{77}$$

D. Discussion of differences between first-neighbor and second-neighbor pruning algorithm

When we consider the second-neighbor pruning algorithms, node pruning and hyperedges pruning give rise to very different definition of (k, n)-cores.

For the random hypergraphs belonging to the configuration model ensembles, the equations determining the average messages $W = \langle w_{i\to\alpha}\rangle$ and $V = \langle v_{\alpha\to i}\rangle$ of the second-neighbor (k, n)-core with pruning of nodes are

$$V = p_H \sum_{m\geqslant n} \frac{mQ(m)}{\langle m\rangle} W^{m-1},$$
$$W = p_N \sum_{q=k}^{\infty} \frac{qP(q)}{\langle q\rangle} \sum_{s=k-1}^{q-1} \binom{q-1}{s} V^s (1-V)^{q-1-s}. \tag{78}$$

These are the same equations determining the average messages of the first-neighbor (k, n)-core algorithm. It follows that the phase diagram of the first-neighbor (k, n)-core percolation coincides with the phase diagram for the second-neighbor (k, n)-core percolation with pruning of the nodes. However, the order parameters might differ. Indeed, the order parameter $S_{kn} = \langle \sigma_i\rangle$ and $R_{kn} = \langle r_\alpha\rangle$, where σ_i and r_α are given by Eqs. (74) and (73), are

$$R_{kn} = p_H \sum_m Q(m) W^m, \tag{79}$$
$$S_{kn} = p_N \sum_{q=k}^{\infty} P(q) \sum_{s=k}^{q} \binom{q}{s} V^s (1-V)^{q-s}. \tag{80}$$

Note that only Eq. (79) differs from the corresponding equation determining the first-neighbor (k, n)-core percolation, Eq. (33), while Eqs. (80) and (34) coincide. Indeed, Eq. (79) includes a sum extended to hyperedges of arbitrary cardinality m, while in Eq. (33) the sum is extended only to hyperedges of cardinality $m \geqslant n$. It follows that the order parameter S_{kn} is unchanged if one considers first-neighbor (k, n)-core percolation or second-neighbor (k, n)-core percolation with node pruning, but the order parameter R_{kn} can change for $n \geqslant 3$. In order to demonstrate this, we show in Fig. 5 Monte Carlo results for first-neighbor (k, n)-core percolation and for second-neighbor (k, n)-core percolation with node pruning. The results are in very good agreement with our theoretical predictions.

The critical behavior in second-neighbor (k, n)-core percolation with pruning of the hyperedges is distinct from these results. Indeed, not only the order parameters can differ from the first-neighbor (k, n)-core transition, but also the nature of the transition and its critical points. Indeed, the system of equations determining the nature of the phase transition reads in this case as

$$V = p_H \sum_{m\geqslant n} p_N^{m-1} \frac{mQ(m)}{\langle m\rangle} \sum_{s=n-1}^{m-1} \binom{m-1}{s} W^s (1-W)^{m-1-s},$$
$$W = \sum_{q\geqslant k} \frac{qP(q)}{\langle q\rangle} \sum_{s=k-1}^{q-1} \binom{q-1}{s} V^s (1-V)^{q-1-s}, \quad (81)$$

where V and W are the average messages. The equation determining the fraction of hyperedges, R_{kn}, in the second-neighbor (k, n)-core with pruning of hyperedges is given by

$$R_{kn} = p_H \sum_{m\geqslant n} p_N^m Q(m) \sum_{s=n}^{m} \binom{m}{s} W^s (1-W)^{m-s}. \quad (82)$$

Moreover, the equation determining the fraction of nodes S_{kn} in the second-neighbor (k, n)-core with pruning of hyperedges is more subtle. This equation is

$$S_{kn} = p_N \sum_q P(q) \left[1 - \sum_{s\leqslant \min(k-1,q)} \binom{q}{s} \tilde{V}^s (1-\tilde{V}-\hat{V})^{m-s} \right], \quad (83)$$

where $\hat{V}$ and $\tilde{V}$ are given by

$$\hat{V} = p_H \sum_{m\geqslant n+1} p_N^{m-1} \frac{mQ(m)}{\langle m\rangle} \sum_{s=n}^{m-1} \binom{m-1}{s} W^s (1-W)^{m-1-s},$$
$$\tilde{V} = p_H \sum_{m\geqslant n} p_N^{m-1} \frac{mQ(m)}{\langle m\rangle} \binom{m-1}{n-1} W^{n-1} (1-W)^{m-s}. \quad (84)$$

The rationale behind Eqs. (83) and (84) was explained while deriving the message-passing Eqs. (74) and (75), from which these equations directly follow.

Intuitively an active node will be always part of the second-neighbor node (k, n)-core. An intact inactive node will be part of the (k, n)-core only if it belongs to at least one hyperedge that belongs to the (k, n)-core. It follows that a node will be always in second-neighbor node (k, n)-core unless (1) none of its hyperedges is connected to at least n other active nodes, which occurs with probability $\hat{V}$, and (2) there are fewer than k hyperedges connected to $n-1$ active nodes, which occurs with probability $\tilde{V}$. Note that condition (2) together with (1) ensures that the node is not active.

The phase diagram of second-neighbor (k, n)-core percolation with pruning of the hyperdeges is very different from the phase diagram for second-neighbor (k, n)-core percolation with pruning of the nodes. In particular, the phase transition is continuous if and only if $(k, n) = (2, 2)$ with the second-order phase transition line obtained for

$$1 = p_H \frac{\langle m(m-1)p_N^{m-1}\rangle}{\langle m\rangle} \frac{\langle q(q-1)\rangle}{\langle q\rangle}, \quad (85)$$

and the phase transition is hybrid for any other (k, n). In Fig. 6 we show the order parameter S_{kn} as a function of p_H for different (k, n)-cores for a random Poisson hypergraph with the hyperedge cardinality and node degree distributions given by Eq. (47). The figure demonstrates excellent agreement with our theoretical predictions.

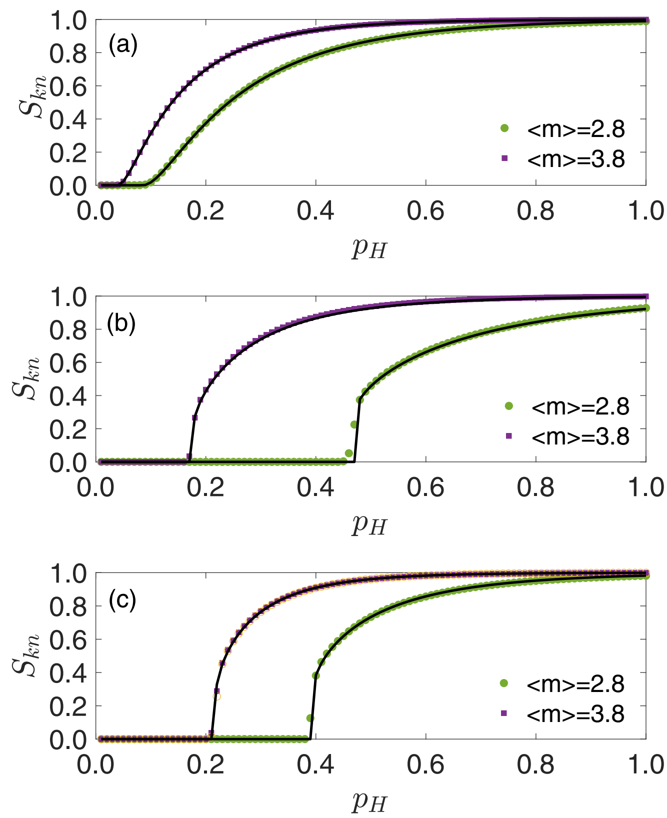

FIG. 6. Fraction S_{kn} of nodes in the second-neighbor (k, n)-core with hyperedge pruning is plotted vs p_H for different values of k and n: $(k, n) = (2, 2)$ (a); $(k, n) = (2, 3)$ (b); $(k, n) = (3, 2)$ (c). The hypergraphs have Poisson cardinality and degree distributions given by Eq. (47), $\langle q\rangle = 2\langle m\rangle$, $\langle m\rangle$ indicated in the legend, and $N = 10^4$ nodes.

Let us now compare the equations determining the second-neighbor (k, n)-core percolation with pruning of the hyperedges to the factor graph (k, n)-core equations [16,17] characterizing the sub-factor graph induced by the nodes of at least degree k and the factor nodes of at least cardinality n. In

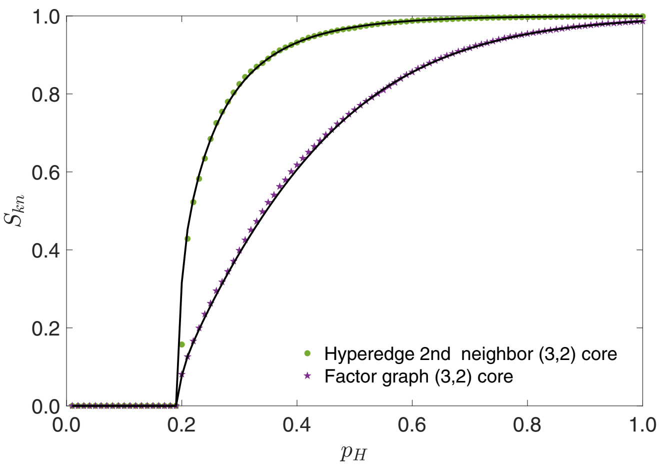

FIG. 7. Fraction S_{kn} of nodes in the (3,2)-core is plotted vs p_H for the second-neighbor (k, n)-core with pruning of the hyperedges and for the factor graph (k, n)-core. The latter displays the same critical threshold of the first but a smaller fraction of nodes in the care. The hypergraph has Poisson cardinality and degree distribution with $\langle m\rangle = 4.0$ and $\langle q\rangle = 2\langle m\rangle$.

this model the fraction of nodes S_{kn} and the fraction of factor nodes (hyperedges) in the (k, n)-core R_{kn} are given by

$$R_{kn} = p_H \sum_{m\geqslant n} Q(m) \sum_{s=n}^{k} \binom{k}{s} W^s(1-W)^{m-s},$$
$$S_{kn} = p_N \sum_{q\geqslant k} P(q) \sum_{s=k}^{n} V^s(1-V)^{q-s}, \tag{86}$$

with W and V obeying

$$V = p_H \sum_{m\geqslant n} \frac{mQ(m)}{\langle m\rangle} \sum_{s=n-1}^{m-1} \binom{m-1}{s} W^s(1-W)^{m-1-s},$$
$$W = p_N \sum_{q\geqslant k} \frac{qP(q)}{\langle q\rangle} \sum_{s=k-1}^{q-1} \binom{q-1}{s} V^s(1-V)^{q-1-s}. \tag{87}$$

We note that when $p_N = 1$ the nature of the phase transition of the second-neighbor hypergraph (k, n)-core percolation and its percolation threshold coincides with the one of the (k, n)-core on factor graphs. Moreover, R_{kn} coincides for the two models while S_{kn} differs (see Fig. 7).

It follows that the phase diagram of the second-neighbor hypergraph (k, n)-core percolation reduces to the phase diagram of the (k, n)-cores of a factor graph for $p_N = 1$. Figure 8 shows this phase diagram for $p_H = p_N = 1$. Comparing this figure with the corresponding one for the first-neighbor (k, n)-core percolation problem (Fig. 3), we notice the absence of the tricritical point in Fig. 8 with hybrid transitions only present for $k \geqslant 3$.

The explicit equation for the continuous transition line for (2,2)-core in this phase diagram is given by

$$\langle m\rangle = \frac{1 + \sqrt{1 + 8\langle q\rangle^2}}{2\langle q\rangle}. \tag{88}$$

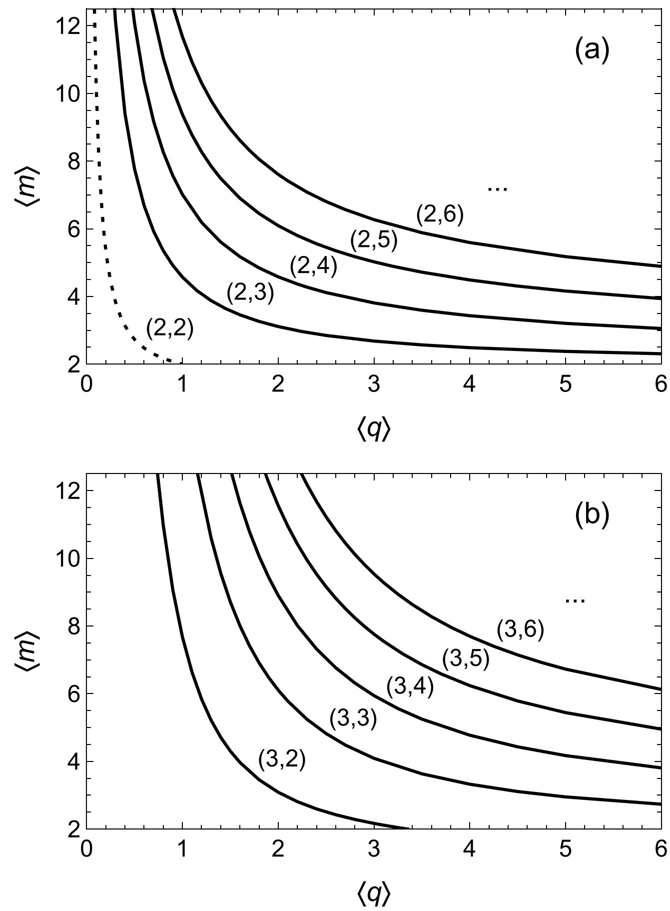

FIG. 8. The phase diagram of the second-neighbor hyperedge (k, n)-core percolation in the $\langle q\rangle$-$\langle m\rangle$ plane is shown for uncorrelated Poisson hypergraphs with $p_H = p_N = 1$. (a) Phase diagram for (k, n)-core percolation with (k, n) given by (2,2), (2,3), (2,4), (2,5), and (2,6); (b) the (k, n)-core percolation phase diagram for (k, n) given by (3,2), (3,3), (3,4), (3,5), and (3,6). Each core exists in the whole region to the right of the corresponding boundary. All boundaries are discontinuous (hybrid) transitions with one exception, namely, the (2,2)-core, which is always a continuous phase transition. For the (2,2)-core, the phase boundary ends at the point $\langle q\rangle = 1$, $\langle m\rangle = 2$. For the (3,2)-core, the phase boundary ends at the point $\langle q\rangle = 3.3509\ldots$, $\langle m\rangle = 2$.

Furthermore, the end point of the phase boundary (hybrid transition line) for the (3,2)-core is given by

$$\langle q\rangle = 3.3509\ldots, \ \langle m\rangle = 2. \tag{89}$$

Here the number $3.3509\ldots = (1 + x + x^2)/x$, where x is the nonzero root of the equation

$$1 + x + x^2 = e^x. \tag{90}$$

VI. CONCLUSIONS

In this work we have developed a message-passing theory for hypergraph (k, n)-core percolation assuming that hyperedges can be intact only if all their nodes are undamaged. This simple hypothesis is relevant for a wide variety of real scenarios, including supply networks, protein interactions networks, and networks of chemical reactions. The k-core decomposition is a widely used tool for the discovery of highly connected substructures within complex networks, which essentially determine the character of cooperative and spreading phenomena in networks. We demonstrate that k-core problems for hypergraphs are significantly different from the k-core problem on ordinary graphs. While the hypergraph structure is represented by an equivalent bipartite graph between nodes and hyperedges—factor nodes (factor graph)—here we reveal that the set of k-cores on hypergraphs is distinct from this set for their factor graphs [16,17].

The reason for this difference is that the deletion of a node in a hypergraph also removes all the adjacent hyperedges, while the deletion of a node in a factor graph doesn't lead to the removal of factor nodes, only the connections of the neighboring factor nodes to the removed node disappear. Accounting for this difference, we describe a set of k-core problems [also called first-neighbor (k, n)-core problems] for hypergraphs and the corresponding pruning algorithms in which nodes and hyperedges are progressively removed (damaged) if their degrees and cardinalities, respectively, fall behind given threshold values, k and n. We obtain phase diagrams for such (k, n)-cores in random hypergraphs. In contrast to ordinary graphs, where the phase transition for the 2-cores is continuous, while the phase transitions for ($k \geqslant 3$)-cores are hybrid, for the (2,2)-core we observe two transition lines on the phase diagram—the continuous transition line and the hybrid transition one. These lines converge at the tricritical point.

In order to bridge the gap between the k-core problems defined on hypegraphs and on factor graphs, we introduce a class of hypergraph k-core problems in which the pruning process involves only nodes or only hyperedges and accounts for the connectivity of their neighbors in the factor graph. We call these latter problems second-neighbor (k, n)-core percolation processes. We show that the second-neighbor (k, n)-core percolation process where only nodes are pruned is rather distinct from the one where only hyperedges are pruned. In particular the nature of the (k, n)-core percolation transition and the percolation threshold of the two variants of the second-neighbor (k, n)-core percolation process is different. The second-neighbor (k, n)-core percolation process with node pruning has a phase diagram that coincides with the first-neighbor (k, n)-core process. The second-neighbor (k, n)-core percolation process with hyperedge pruning has a phase diagram that for $p_N = 1$ coincides with the factor graph (k, n)-core percolation problems [16,17]. Note, however, that the order parameters for second-neighbor (k, n)-cores with pruning of nodes or hyperedges [fractions of nodes and hyperedges within these (k, n)-cores] do not all reduce to the ones for the hypergraph or factor graph (k, n)-cores.

We suggest that this work will highlight the important differences between hypergraphs and factor graphs and will contribute to a better understanding of specific critical phenomena in higher-order networks. It is our hope and trust that the first-neighbor and second-neighbor (k, n)-core hypergraph problems defined here might find wide applications in the study or real-world higher-order networks.

[1] G. Bianconi, *Higher-Order Networks* (Cambridge University Press, Cambridge, 2021).

[2] F. Battiston, E. Amico, A. Barrat, G. Bianconi, G. Ferraz de Arruda, B. Franceschiello, I. Iacopini, S. Kéfi, V. Latora, Y. Moreno *et al.*, The physics of higher-order interactions in complex systems, Nat. Phys. **17**, 1093 (2021).

[3] A. R. Benson, D. F. Gleich, and J. Leskovec, Higher-order organization of complex networks, Science **353**, 163 (2016).

[4] G. Bianconi and R. M. Ziff, Topological percolation on hyperbolic simplicial complexes, Phys. Rev. E **98**, 052308 (2018).

[5] F. Battiston, G. Cencetti, I. Iacopini, V. Latora, M. Lucas, A. Patania, J.-G. Young, and G. Petri, Networks beyond pairwise interactions: Structure and dynamics, Phys. Rep. **874**, 1 (2020).

[6] S. Majhi, M. Perc, and D. Ghosh, Dynamics on higher-order networks: A review, J. R. Soc. Interface **19**, 20220043 (2022).

[7] S. Boccaletti, P. De Lellis, C. I. del Genio, K. Alfaro-Bittner, R. Criado, S. Jalan, and M. Romance, The structure and dynamics of networks with higher order interactions, Phys. Rep. **1018**, 1 (2023).

[8] G. Ferraz de Arruda, M. Tizzani, and Y. Moreno, Phase transitions and stability of dynamical processes on hypergraphs, Commun. Phys. **4**, 24 (2021).

[9] M. Barthelemy, Class of models for random hypergraphs, Phys. Rev. E **106**, 064310 (2022).

[10] V. Salnikov, D. Cassese, and R. Lambiotte, Simplicial complexes and complex systems, Eur. J. Phys. **40**, 014001 (2019).

[11] B. C. Coutinho, A.-K. Wu, H.-J. Zhou, and Y.-Y. Liu, Covering problems and core percolations on hypergraphs, Phys. Rev. Lett. **124**, 248301 (2020).

[12] G. Bianconi, Statistical physics of exchangeable sparse simple networks, multiplex networks, and simplicial complexes, Phys. Rev. E **105**, 034310 (2022).

[13] H. Sun, F. Radicchi, J. Kurths, and G. Bianconi, The dynamic nature of percolation on networks with triadic interactions, Nat. Commun. **14**, 1308 (2023).

[14] G. F. de Arruda, G. Petri, and Y. Moreno, Social contagion models on hypergraphs, Phys. Rev. Res. **2**, 023032 (2020).

[15] I. Iacopini, G. Petri, A. Barrat, and V. Latora, Simplicial models of social contagion, Nat. Commun. **10**, 2485 (2019).

[16] J. Lee, K.-I. Goh, D.-S. Lee, and B. Kahng, (k, q)-core decomposition of hypergraphs, Chaos Solitons Fractals **173**, 113645 (2023).

[17] M. Mancastroppa, I. Iacopini, G. Petri, and A. Barrat, Hyper-cores promote localization and efficient seeding in higher-order processes, Nat. Commun. **14**, 6223 (2023).

[18] A. P. Millán, J. J. Torres, and G. Bianconi, Explosive higher-order Kuramoto dynamics on simplicial complexes, Phys. Rev. Lett. **124**, 218301 (2020).

[19] G. St-Onge, H. Sun, A. Allard, L. Hébert-Dufresne, and G. Bianconi, Universal nonlinear infection kernel from heterogeneous exposure on higher-order networks, Phys. Rev. Lett. **127**, 158301 (2021).

[20] H. Peng, C. Qian, D. Zhao, M. Zhong, J. Han, R. Li, and W. Wang, Message passing approach to analyze the robustness of hypergraph, arXiv:2302.14594.

[21] E. López, Weighted projected networks: Mapping hypergraphs to networks, Phys. Rev. E **87**, 052813 (2013).

[22] H. Sun and G. Bianconi, Higher-order percolation processes on multiplex hypergraphs, Phys. Rev. E **104**, 034306 (2021).

[23] S. Thurner, P. Klimek, and R. Hanel, Schumpeterian economic dynamics as a quantifiable model of evolution, New J. Phys. **12**, 075029 (2010).

[24] R. Hanel, S. A. Kauffman, and S. Thurner, Phase transition in random catalytic networks, Phys. Rev. E **72**, 036117 (2005).

[25] F. Klimm, C. M. Deane, and G. Reinert, Hypergraphs for predicting essential genes using multiprotein complex data, J. Complex Netw. **9**, cnaa028 (2021).

[26] J. Jost and R. Mulas, Hypergraph Laplace operators for chemical reaction networks, Adv. Math. **351**, 870 (2019).

[27] G. Bianconi and S. N. Dorogovtsev, preceding paper, Theory of percolation on hypergraphs, Phys. Rev. E **109**, 014306 (2024).

[28] S. B. Seidman, Network structure and minimum degree, Soc. Netw. **5**, 269 (1983).

[29] B. Bollobás, The evolution of sparse graphs, in *Graph Theory and Combinatorics: Proc. Cambridge Combinatorial Conference in Honour of Paul Erdős*, edited by B. Bollobás (Academic Press, San Diego, 1984), pp. 335–357.

[30] J. Chalupa, P. L. Leath, and G. R. Reich, Bootstrap percolation on a Bethe lattice, J. Phys. C: Solid State Phys. **12**, L31 (1979).

[31] S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes, k-core organization of complex networks, Phys. Rev. Lett. **96**, 040601 (2006).

[32] A. V. Goltsev, S. N. Dorogovtsev, and J. F. F. Mendes, k-core (bootstrap) percolation on complex networks: Critical phenomena and nonlocal effects, Phys. Rev. E **73**, 056101 (2006).

[33] S. N. Dorogovtsev and J. F. F. Mendes, *The Nature of Complex Networks* (Oxford University Press, Oxford, 2022).

[34] J. I. Alvarez-Hamelin, L. Dall'Asta, A. Barrat, and A. Vespignani, Large-scale networks fingerprinting and visualization using the k-core decomposition, in *NIPS'05: Proceedings of the 18th International Conference on Neural Information Processing Systems 18*, edited by Y. Weiss, B. Schölkopf, and J. C. Platt (MIT Press, Cambridge, Massachusetts, 2006), pp. 41–50.

[35] S. Carmi, S. Havlin, S. Kirkpatrick, Y. Shavitt, and E. Shir, A model of Internet topology using k-shell decomposition, Proc. Natl. Acad. Sci. USA **104**, 11150 (2007).

[36] M. Kitsak, L. K. Gallos, S. Havlin, F. Liljeros, L. Muchnik, H. E. Stanley, and H. A. Makse, Identification of influential spreaders in complex networks, Nat. Phys. **6**, 888 (2010).

[37] M. Serafino, H. S. Monteiro, S. Luo, S. D. S. Reis, C. Igual, A. S. L. Neto, M. Travizano, J. S. Andrade, Jr., and H. A. Makse, Digital contact tracing and network theory to stop the spread of COVID-19 using big-data on human mobility geolocalization, PLoS Comput. Biol. **18**, e1009865 (2022).

[38] R. Pastor-Satorras and C. Castellano, Distinct types of eigenvector localization in networks, Sci. Rep. **6**, 18847 (2016).

[39] M. E. J. Newman, Message passing methods on complex networks, Proc. R. Soc. A **479**, 20220774 (2023).

[40] M. Weigt and H. Zhou, Message passing for vertex covers, Phys. Rev. E **74**, 046110 (2006).

[41] G. Bianconi, *Multilayer Networks: Structure and Function* (Oxford University Press, Oxford, 2018).

[42] G. T. Cantwell, A. Kirkley, and F. Radicchi, Heterogeneous message passing for heterogeneous networks, Phys. Rev. E **108**, 034310 (2023).

[43] M. E. J. Newman, Random graphs with clustering, Phys. Rev. Lett. **103**, 058701 (2009).

[44] S. Yoon, A. V. Goltsev, S. N. Dorogovtsev, and J. F. F. Mendes, Belief-propagation algorithm and the Ising model on networks with arbitrary distributions of motifs, Phys. Rev. E **84**, 041144 (2011).

[45] A. Kirkley, G. T. Cantwell, and M. E. J. Newman, Belief propagation for networks with loops, Sci. Adv. **7**, eabf1211 (2021).

[46] D. Cellai, S. N. Dorogovtsev, and G. Bianconi, Message passing theory for percolation models on multiplex networks with link overlap, Phys. Rev. E **94**, 032301 (2016).

[47] F. Radicchi and G. Bianconi, Redundant interdependencies boost the robustness of multiplex networks, Phys. Rev. X **7**, 011013 (2017).

[48] S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes, Critical phenomena in complex networks, Rev. Mod. Phys. **80**, 1275 (2008).