

# Ensemble nonequivalence and Bose-Einstein condensation in weighted networks

Qi Zhang<sup>1,\*</sup> and Diego Garlaschelli<sup>1,2,†</sup>

<sup>1</sup>*Lorentz Institute for Theoretical Physics, Leiden University,  
Niels Bohrweg 2, 2333 CA Leiden, The Netherlands*

<sup>2</sup>*IMT School of Advanced Studies, Piazza San Francesco 19, 55100 Lucca, Italy*

(Dated: August 4, 2021)

The asymptotic (non)equivalence of canonical and microcanonical ensembles, describing systems with soft and hard constraints respectively, is a central concept in statistical physics. Traditionally, the breakdown of ensemble equivalence (EE) has been associated with nonvanishing relative canonical fluctuations of the constraints in the thermodynamic limit. Recently, it has been reformulated in terms of a nonvanishing relative entropy density between microcanonical and canonical probabilities. The earliest observations of EE violation required phase transitions or long-range interactions. More recent research on binary networks found that an extensive number of local constraints can also break EE, even in absence of phase transitions. Here we study for the first time ensemble nonequivalence in weighted networks with local constraints. Unlike their binary counterparts, these networks can undergo a form of Bose-Einstein condensation (BEC) producing a core-periphery structure where a finite fraction of the link weights concentrates in the core. This phenomenon creates a unique setting where local constraints coexist with a phase transition. We find surviving relative fluctuations only in the condensed phase, as in more traditional BEC settings. However, we also find a non-vanishing relative entropy density for all temperatures, signalling a breakdown of EE due to the presence of an extensive number of constraints, irrespective of BEC. Therefore, in presence of extensively many local constraints, vanishing relative fluctuations no longer guarantee EE.

## I. INTRODUCTION

Statistical ensembles were introduced by Gibbs [1] to mathematically describe systems at thermodynamic equilibrium, i.e. where certain conserved macroscopic properties (such as the total energy) are constant, while the microscopic state (i.e. the state of all the microscopic constituents) is subject to fluctuations. For a system with  $n$  units and discrete degrees of freedom, a statistical ensemble is a probability distribution  $P(\mathbf{W})$  over the collection  $\mathcal{W}_n = \{\mathbf{W}\}$  of all the possible (unobserved) microscopic states of the system, given a set of measurable macroscopic properties. Clearly,  $P(\mathbf{W}) \geq 0$  for all  $\mathbf{W} \in \mathcal{W}_n$  and  $\sum_{\mathbf{W} \in \mathcal{W}_n} P(\mathbf{W}) = 1$ . This distribution conceptualizes the fact that, ideally, repeated observations of the microscopic state would retrieve different (and independent) outcomes. It can be viewed as the probability distribution that maximizes the Gibbs-Shannon entropy functional

$$S[P] \equiv - \sum_{\mathbf{W} \in \mathcal{W}_n} P(\mathbf{W}) \ln P(\mathbf{W}), \quad (1)$$

under a set of (macroscopic) constraints, therefore being maximally noncommittal with respect to missing (microscopic) information [2].

Depending on the choice of the macroscopic properties being constrained, different statistical ensembles can be constructed. The *microcanonical ensemble*  $P_{\text{mic}}(\mathbf{W}|E)$  is used to describe systems with fixed total energy  $E$  (energetic isolation), while the *canonical ensemble*  $P_{\text{can}}(\mathbf{W}|\beta)$

is used to describe systems with fixed (inverse) temperature  $\beta$  (thermal equilibrium) [1]. For a physical system, the inverse temperature  $\beta$  equals  $1/kT$  where  $T$  is the absolute temperature and  $k$  is Boltzmann's constant. In both ensembles, the microscopic state  $\mathbf{W}$  is random, but the randomness is governed differently by the two distributions  $P_{\text{can}}(\mathbf{W}|\beta)$  and  $P_{\text{mic}}(\mathbf{W}|E)$ . In particular, while the microcanonical ensemble assigns each realized configuration  $\mathbf{W}$  a constant (deterministic) value  $E(\mathbf{W}) = E$  of the total energy (which can therefore be regarded as a 'hard' constraint corresponding to energetic isolation), the canonical ensemble assigns configurations a fluctuating (random) energy with a certain expected value  $\langle E \rangle_\beta = \sum_{\mathbf{W} \in \mathcal{W}_n} P_{\text{can}}(\mathbf{W}|\beta) E(\mathbf{W})$  and a positive standard deviation  $\sigma_\beta(E) > 0$  (i.e. the energy plays the role of a 'soft' constraint resulting from the contact with a heat bath at fixed inverse temperature  $\beta$ ).

### A. Conjugate ensembles

The two ensembles can be made conjugate to each other by choosing a specific value  $E^*$  and simultaneously setting the total energy  $E(\mathbf{W})$  of each realized configuration  $\mathbf{W}$  in the microcanonical ensemble equal to  $E(\mathbf{W}) \equiv E^*$  and the inverse temperature  $\beta$  in the canonical ensemble to the corresponding value  $\beta \equiv \beta^*$  such that the resulting average value  $\langle E \rangle_{\beta^*}$  of the fluctuating total energy under the canonical probability  $P_{\text{can}}(\mathbf{W}|\beta^*)$  equals  $E^*$ , i.e.

$$\langle E \rangle_{\beta^*} \equiv E^*. \quad (2)$$

For systems with finite size, the two conjugate ensembles are unavoidably different, because in the micro-

\* zhang@lorentz.leidenuniv.nl

† garlaschelli@lorentz.leidenuniv.nl, diego.garlaschelli@imtlucca.it

canonical ensemble the hardness of the constraint implies extra dependencies among the state of the microscopic constituents with respect to the canonical case. However, in the *thermodynamic limit* (i.e. when the number of units in the system goes to infinite) and under certain ‘natural’ circumstances, the two associated probabilistic descriptions are expected to become effectively equivalent (i.e. the canonical fluctuations and the microcanonical dependencies are both expected to play an asymptotically vanishing role) as a result of some form of the law of large numbers. This idea, which dates back to Gibbs himself [1] and has continued to attract a lot of interest until presently [3–6], goes under the name of *ensemble equivalence* (EE). When EE holds, one can treat the two ensembles as asymptotically interchangeable, and hence use any of them based on mathematical or computational convenience. For instance, analytical calculations are significantly easier in the canonical ensemble, while numerical randomizations of an initial configuration can be carried out more naturally in the microcanonical ensemble.

Most statistical physics textbooks still convey the message that EE is expected to hold in general as a sort of principle at the basis of ensemble theory. The possible breakdown of EE, also known as *ensemble nonequivalence* (EN), is still not discussed systematically in the literature. However, several observations of EN have been documented over the past decades [3, 5, 7–17]. These observations motivated various efforts aimed at elucidating both the possible physical mechanisms at the origin of EN and, in parallel, its proper mathematical definition(s).

### B. Physical mechanisms for ensemble (non)equivalence

Traditionally, the ‘natural’ circumstances generally invoked to ensure EE mainly concern the presence of (loosely speaking) ‘at most weak’ interactions between the constituents of the system. This condition is automatically realized when the system consists of independent units or units with short-range interactions and sufficiently high temperature (to stay away from possible low-temperature phases with broken symmetries, for which the canonical average value of the energy is no longer the typical value). Indeed, violations of EE have been documented in presence of long-range interactions (e.g. in gravitational systems) or phase transitions (e.g. in interacting spin systems) [3, 4, 7–15].

However, recent research on complex systems encountered beyond the usual realm of physics has found an additional mechanism that can break EE, even in presence of weak (or no) interactions: namely, the presence of *an extensive number of local constraints* [5, 6]. This situation is frequently found in networks with constraints on the number of links (*degree*) of each node. More specifically, the *binary configuration model* [18] is a widely used

null model of graphs with a given *degree sequence*, i.e. a given vector of node degrees. The model captures many properties found in real-world networks, because the local character of the degree constraint can accommodate the strong structural heterogeneity typically observed across nodes in real networks. Unlike the traditional thermodynamic example where the total energy (and possibly a small, finite number of additional macroscopic properties) is a global and unique constraint for the system, networks with given node degrees are systems with as many constraints as the number of fundamental units, i.e. where constraints are *extensive in number and local in nature*. This situation has been found to break the equivalence of the corresponding canonical and microcanonical ensembles, even without long-range interactions or phase transitions [5, 19, 20]. Notably, since systems with local constraints are generic models for virtually any heterogeneous system, the new mechanism significantly widens the range of real-world cases for which EE may break down. This novel result deserves further research.

### C. Mathematical definitions of ensemble (non)equivalence

Besides the aforementioned advances in the study of the physical mechanisms at the origin of EN, significant progress has been made in the mathematical characterization of EN as well. Traditionally, the informal criterion [1] used to test whether two conjugate ensembles are equivalent is checking whether the *relative fluctuations* of the constraint in the canonical ensemble, e.g. the ratio  $\sigma_{\beta^*}(E)/E^*$  of the canonical standard deviation to the average value of the energy, vanish in the thermodynamic limit. If this happens, then the canonical fluctuations of the total energy are negligible with respect to the total energy itself and, intuitively, the energy in the canonical ensemble can be thought of as an effectively deterministic quantity, very much like in the conjugated microcanonical ensemble. Similarly, the extra dependencies among the microscopic constituents in the microcanonical ensemble are expected to play a smaller and smaller role, thus coming closer to the canonical case.

More recent approaches have considered different rigorous definitions of EE, which can be beautifully summarized [4] as the following three notions: *thermodynamic* equivalence (convexity of the microcanonical entropy density), *macrostate* equivalence (equality of the expected values of macroscopic quantities under the two ensembles), and *measure* equivalence (vanishing of the relative entropy density between the microcanonical and canonical probability distributions). Under mild conditions, these notions have been shown to be equivalent [4]. In this paper, we use measure equivalence as it is more transparently related to the ensemble probabilities.

As a useful result, research on the relationship between statistical physics and combinatorics has revealed that

the relative entropy between the microcanonical and the canonical probability distributions is, under certain conditions, asymptotically proportional to the logarithm of the determinant of the covariance matrix of the *effective* constraints in the canonical ensemble [6]. The effective constraints are those that are neither redundant, i.e. trivially replicating other constraints, nor degenerate, i.e. deterministically restricting the canonical and microcanonical configurations in exactly the same way. For instance, formally imposing ‘two’ constraints where one is the total energy  $E$  and the other one is twice the total energy  $2E$  is clearly a redundant choice: the effective number of constraints is just one in this case. As another example, if in addition to the energy  $E$  we impose its square value  $E^2$ , then for those values of the Lagrange multipliers such that  $\langle E^2 \rangle = \langle E \rangle^2$  the variance of the energy will be zero also in the canonical ensemble:  $E$  will become degenerate and deterministically equal to its imposed value in both ensembles, so not contributing any difference between the two (by contrast, for parameter values such that  $\langle E^2 \rangle > \langle E \rangle^2$  there are no allowed configurations in the microcanonical ensemble because the hard values of  $E^2$  and  $E$  become conflicting, thereby breaking the equivalence with the canonical one). In general, if the problem is not ill-posed from the beginning, the number of effective constraints coincides with the number of enforced constraints. However, it may happen that some constraints become ineffective for certain degenerate values of the parameters. In any case, for a given parameter value the number of effective constraints coincides with the rank of the covariance matrix of all imposed constraints [6].

Since nonequivalence in the measure sense corresponds to the (super)extensivity of the relative entropy, studying the asymptotic behaviour of the determinant of the (effective) covariance matrix is enough in order to assess ensemble nonequivalence. It is worth noticing that, if there is a single constraint (say, the total energy  $E$ ), then the determinant of the covariance matrix coincides with the corresponding variance  $\sigma_{\beta^*}^2(E)$  and the relative entropy grows asymptotically (under the necessary hypotheses) as  $\ln \sigma_{\beta^*}^2(E) = 2 \ln \sigma_{\beta^*}(E)$ . On the other hand, since  $E$  is a global constraint, it is generally extensive in the number  $n$  of units of the system. Therefore the vanishing of the relative fluctuations, i.e.  $\sigma_{\beta^*}(E) = o(E)$ , implies that the relative entropy is subextensive, i.e. the relative entropy density vanishes in the thermodynamic limit. This suggests that, *in presence of a global constraint, the vanishing of the relative fluctuations implies ensemble equivalence*.<sup>1</sup> How this picture changes in presence of an extensive number of local constraints has not

been investigated yet. In particular, whether the vanishing of relative fluctuations still implies ensemble equivalence remains an open question.

#### D. The contribution of this paper

This paper connects to both lines of research described above (physical mechanisms and mathematical definitions for EN) and its aim is therefore twofold. On the one hand, we aim at investigating for the first (to the best of our knowledge) time the phenomenon of EN in a model system that combines the presence of an extensive number of local constraints with a phase transition. On the other hand, we aim at understanding whether the intuitive criterion of vanishing relative fluctuations of the constraints still ensures EE in this more general setting.

Concretely, we consider the *weighted configuration model* [18], namely a model of *weighted* (as opposed to binary) networks with given *strength sequence*, i.e. with given values of the *strength* (sum of the weights of incident links) of each node. The weighted character of the model allows for the emergence of a phase transition that is impossible to observe in the corresponding binary configuration model, namely Bose-Einstein Condensation (BEC) [21, 22]. For the sake of clarity, it is worth mentioning here that, although a form of BEC in networks was identified for the first time in growing binary graphs [23], the notion we refer to here refers to static networks and as such can only occur in weighted networks [21]. Indeed, while the configuration model for weighted networks obeys Bose-Einstein statistics, the configuration model for binary networks obeys Fermi-Dirac statistics [21, 22, 24]. BEC can arise in our model by appropriately tuning the strength sequence. In particular, we are going to show that we can make the strength sequence temperature-dependent and generate BEC by picking a sufficiently low temperature, below a certain critical value. The simplest such setting is one where the network has a ‘core-periphery’ structure, with BEC appearing in the core.

We find that, for all temperatures and irrespective of whether BEC emerges, the canonical and microcanonical ensembles are always nonequivalent as signalled by a non-vanishing relative entropy density. On the other hand, the relative fluctuations of all the constraints vanish when BEC is absent, while some of them do not vanish when BEC is present. This shows that the relative fluctuations cannot distinguish between equivalence and nonequivalence of the ensembles in this more general case where multiple constraints are present. In fact, what they do is detecting the presence of BEC. Therefore the traditional criterion for EE based on the vanishing of the relative fluctuations is no longer valid in presence of an extensive number of constraints, even when applied simultaneously to all constraints. These results enrich our understanding of the phenomenology of EN and shed more light on its relationship with both the extensivity of the constraints

<sup>1</sup> Note that the converse is not necessarily true. However, observing ensemble equivalence and non-vanishing relative fluctuations simultaneously requires some rather uncommon circumstances: for instance, if  $\sigma_{\beta^*}(E)$  grows like  $E^\alpha$  with  $\alpha \geq 1$  while the entropy still grows like  $E$ , then the relative entropy is still subextensive, while the relative fluctuations do not vanish.

and the presence of phase transitions.

The remainder of this paper is organized as follows. In Sec. II we rigorously define the canonical and microcanonical ensembles of weighted networks with given strength sequence. In Sec. III we introduce two criteria for the (non)equivalence of the ensembles, one based on the relative entropy between the corresponding probability distributions (measure equivalence) and one based on the relative fluctuations of the constraints. In Sec. IV we study in detail a model defined by the simplest family of strength sequences, driven by a temperature parameter, such that we can observe both a BEC and a non-BEC phase. In Sec. V we offer our conclusions. Finally, the Appendix contains useful calculations needed to establish the scaling of the relative entropy in all the regimes considered.

## II. CANONICAL AND MICROCANONICAL ENSEMBLES OF WEIGHTED NETWORKS

In this section, we introduce the definition of weighted networks and of canonical and microcanonical ensembles of weighted networks with given strength sequence.

### A. Weighted network ensembles

Weighted networks are widely used to describe systems with a large number of components and heterogeneous patterns of interaction [25]. We represent a possible configuration of a weighted network with  $n$  nodes as an  $n \times n$  weighted adjacency matrix  $\mathbf{W}$ . Each entry of the matrix  $w_{ij}$  ( $1 \leq i \leq n, 1 \leq j \leq n$ ) denotes the weight of the link between node  $i$  and node  $j$ , which is taken from the set  $\mathbb{N}$  of natural numbers (including zero, which corresponds to the absence of a link between  $i$  and  $j$ ). In this work, we only consider undirected networks without self-loops, thus the weighted matrix  $\mathbf{W}$  is a symmetric matrix ( $w_{ij} = w_{ji}$  for all  $i, j$ ) and its diagonal elements are zero ( $w_{ii} = 0$  for all  $i$ ). The number of independent entries of each such matrix is therefore  $\binom{n}{2} = n(n-1)/2$ .

An ensemble of weighted networks on  $n$  nodes is the discrete (infinite) set  $\mathcal{W}_n = \mathbb{N}^{n(n-1)/2}$  of all available configurations for the matrix  $\mathbf{W}$  and a probability distribution  $P(\mathbf{W})$  over  $\mathcal{W}_n$  that is specified by a given vector  $\vec{C}(\mathbf{W})$  of constraints, which can be enforced either as a soft constraint (canonical ensemble) or as a hard constraint (microcanonical ensemble) [18]. So the matrix  $\mathbf{W}$  is a possible outcome of a random variable. We will consider the weighted configuration model, for which the constraints are the strengths of all nodes, i.e. the strength sequence  $\vec{C}(\mathbf{W}) = \vec{s}(\mathbf{W})$ , where the strength  $s_i(\mathbf{W})$  of node  $i$  is a local sum of all the link weights that connect  $i$  to its neighbours in the particular network  $\mathbf{W}$ :

$$s_i(\mathbf{W}) = \sum_{j=1}^n w_{ij}, \quad i = 1, n. \quad (3)$$

Clearly, the number of scalar constraints is  $n$ , which coincides with the number of nodes, so this model is a perfect example of a system subject to an extensive number of local constraints. In fact, it is the weighted counterpart of the binary configuration model, where EN driven by local constraints was observed for the first time [5].

A crucial consequence of the presence of local constraints in weighted networks is illustrated in Fig.1, where we show an example of different networks with the same strength sequence. For a given choice of  $\vec{s}^*$ , in different realizations of the network each node  $i$  can have different neighbours and different distributions of weights on the links that connect it to those neighbours. In particular, the strength  $s_i^*$  can be more or less concentrated on specific neighbours (a property that is usually quantified by the so-called *disparity* [26]). However, more homogeneous choices of  $\vec{s}^*$  unavoidably result in less concentrated link weights, while more heterogeneous choices of  $\vec{s}^*$  impose more concentrated link weights. This fact will allow us to consider (in Sec. IV) different structural regimes ranging between two extreme limits: a constant (infinite-temperature) strength sequence implying that on average each node is connected to its neighbours in an equally strong way, and a ‘step-like’ (zero-temperature) strength sequence implying an extreme concentration of link weights among a small subset of the  $n$  nodes. In between these two limits, a certain critical temperature separates a ‘non-condensed’ (high-temperature) phase from a ‘condensed’ (low-temperature) phase featuring the properties of BEC.

### B. Canonical weighted network ensemble

We first discuss how to implement the strength sequence constraint mathematically in the canonical ensemble (soft constraint) [18]. Recall that in the traditional canonical ensemble the inverse temperature  $\beta^*$  is the only (scalar) parameter of the canonical probability distribution  $P_{\text{can}}(\mathbf{W}|\beta^*)$ , conjugate to a certain (scalar) total energy  $E^*$  in the corresponding microcanonical ensemble. Explicitly,  $P_{\text{can}}(\mathbf{W}|\beta^*)$  is the Boltzmann distribution  $P_{\text{can}}(\mathbf{W}|\beta^*) = e^{-\beta^* E(\mathbf{W})}/Z(\beta^*)$  with inverse temperature  $\beta^* = 1/kT^*$ . By contrast, in our setting the canonical distribution  $P_{\text{can}}(\mathbf{W}|\vec{\beta}^*)$  has to depend on an  $n$ -dimensional vector  $\vec{\beta}^*$  of parameters, conjugate to the  $n$ -dimensional constraint  $\vec{s}^*$  which, in turn, defines the microcanonical ensemble. This distribution is found by maximizing the Gibbs-Shannon entropy functional defined in Eq. (1) under the soft constraint

$$\langle \vec{s} \rangle_{\vec{\beta}^*} = \sum_{\mathbf{W} \in \mathcal{W}_n} P_{\text{can}}(\mathbf{W}|\vec{\beta}^*) \vec{s}(\mathbf{W}) \equiv \vec{s}^* \quad (4)$$

which generalizes the conjugacy condition in Eq. (2). The solution to the maximization problem sees  $\vec{\beta}^*$  play the role of a vector of Lagrange multipliers coupled to the

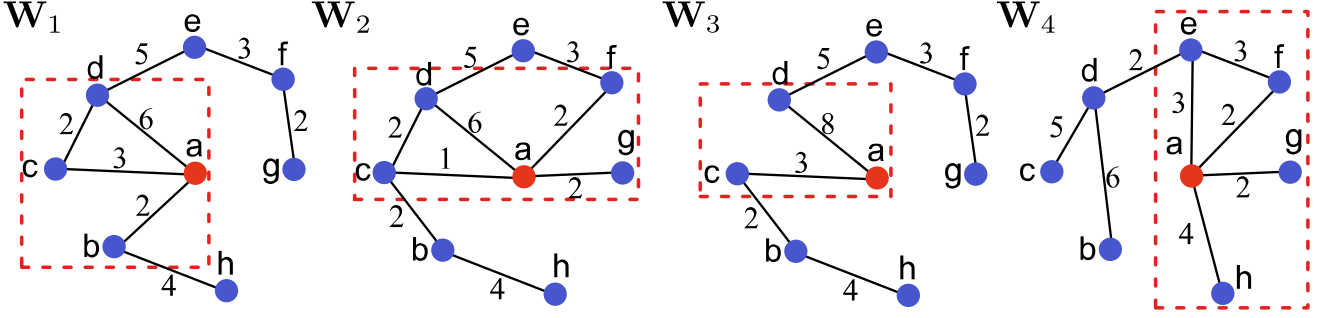


FIG. 1. Example of four weighted networks  $\mathbf{W}_1, \mathbf{W}_2, \mathbf{W}_3, \mathbf{W}_4$  having the same number  $n = 8$  of nodes (labelled from  $a$  to  $h$ ) and the same strength sequence  $\vec{s}(\mathbf{W}_i) = \vec{s}^* = (11, 6, 5, 13, 8, 5, 2, 4)$  ( $i = 1, 4$ ), but different structure ( $\mathbf{W}_i \neq \mathbf{W}_j$  for all  $i \neq j$ ). The dashed blocks highlight the links from node  $a$  to its neighbours: in different networks, node  $a$  can have different neighbours and, importantly, different distributions of link weights (more or less concentrated on specific neighbours). More homogeneous choices of  $\vec{s}^*$  would result in less concentrated link weights, while more heterogeneous choices of  $\vec{s}^*$  would result in more concentrated link weights.

strength sequence  $\vec{s}^*$ , and is given by [18]

$$P_{\text{can}}(\mathbf{W}|\vec{\beta}^*) = \frac{e^{-H(\mathbf{W}, \vec{\beta}^*)}}{Z(\vec{\beta}^*)}, \quad (5)$$

where the (network) *Hamiltonian*  $H(\mathbf{W}, \vec{\beta}) = \vec{\beta} \cdot \vec{s}(\mathbf{W})$  is the linear combination of the node strengths, the *partition function*  $Z(\vec{\beta}) = \sum_{\mathbf{W} \in \mathcal{W}_n} e^{-H(\mathbf{W}, \vec{\beta})}$  is a normalization constant, and  $\vec{\beta}^*$  is the unique parameter value realizing Eq. (4). Note that Eq. (5) has still the form of the Boltzmann distribution, with the important caution that *the inverse temperature has been reabsorbed into the Hamiltonian*. Therefore, to keep the parallel with the traditional physical situation, in our setting the Hamiltonian should be thought of as the inverse temperature times the energy, and  $\vec{\beta}^*$  as the inverse temperature times a vector of ‘fields’, each coupled to a different constraint. Clearly, since the probability in Eq. (5) must be dimensionless, the product  $\vec{\beta} \cdot \vec{s}(\mathbf{W})$  must be dimensionless as well. In Sec. IV, we will notice that the Hamiltonian can be further reinterpreted as also incorporating a ‘chemical potential’ governing the expected weight of the links in the network [24].

Notably,  $P_{\text{can}}(\mathbf{W}|\vec{\beta}^*)$  depends on  $\mathbf{W}$  only through  $\vec{s}(\mathbf{W})$ . In particular, it gives the same value  $P_{\text{can}}(\mathbf{W}^*|\vec{\beta}^*)$  to any network  $\mathbf{W}^*$  such that  $\vec{s}(\mathbf{W}^*) = \vec{s}^*$ . Explicitly, given the definition of node strength in Eq. (3), the network Hamiltonian can be written for a generic value of  $\vec{\beta}$  as

$$H(\mathbf{W}, \vec{\beta}) = \sum_{i=1}^n \sum_{i < j} (\beta_i + \beta_j) w_{ij}. \quad (6)$$

The partition function can be easily shown [18] to be

$$Z(\vec{\beta}) = \prod_{i=1}^n \prod_{i < j} \frac{1}{1 - e^{-(\beta_i + \beta_j)}} \quad (7)$$

provided that  $\beta_i + \beta_j > 0$  for all  $i, j$  (otherwise, the model admits no solution). The canonical probability distribution therefore factorizes over pairs of nodes as

$$P_{\text{can}}(\mathbf{W}|\vec{\beta}) = \prod_{i=1}^n \prod_{i < j} q_{ij}(w_{ij}|\vec{\beta}), \quad (8)$$

where

$$q_{ij}(w|\vec{\beta}) = \frac{e^{-(\beta_i + \beta_j)w}}{[1 - e^{-(\beta_i + \beta_j)}]^{-1}} \quad (9)$$

is the probability that the weight of the link between nodes  $i$  and  $j$  takes the particular value  $w$ . Therefore different pairs of nodes are statistically independent in the canonical ensemble (while they are not in the micro-canonical one).

Note that  $q_{ij}(w|\vec{\beta})$  is a geometric distribution [18, 27] with expected value

$$\begin{aligned} \langle w_{ij} \rangle_{\vec{\beta}} &= \sum_{w \in \mathbb{N}} w q_{ij}(w|\vec{\beta}) \\ &= \sum_{w \in \mathbb{N}} w \frac{e^{-(\beta_i + \beta_j)w}}{[1 - e^{-(\beta_i + \beta_j)}]^{-1}} \\ &= \frac{e^{-(\beta_i + \beta_j)}}{1 - e^{-(\beta_i + \beta_j)}} \end{aligned} \quad (10)$$

(representing the expected weight of the link connecting nodes  $i$  and  $j$ ) and variance

$$\begin{aligned} \text{Var}_{\vec{\beta}}(w_{ij}) &= \frac{e^{-(\beta_i + \beta_j)}}{[1 - e^{-(\beta_i + \beta_j)}]^2} \\ &= \langle w_{ij} \rangle_{\vec{\beta}} (1 + \langle w_{ij} \rangle_{\vec{\beta}}). \end{aligned} \quad (11)$$

As we will discuss in detail in Sec. IV, Eq. (10) has the form of Bose-Einstein statistics, where  $\langle w_{ij} \rangle_{\vec{\beta}}$  plays the role of an expected occupation number for the state labeled by nodes  $i$  and  $j$ . In an appropriate ‘low-temperature’ regime, BEC can emerge into the model

through the divergence of the occupation number  $\langle w_{ij} \rangle_{\vec{\beta}}$  for one (possibly degenerate) ‘ground state’ (corresponding to  $\beta_i + \beta_j \rightarrow 0^+$ ), while the occupation number for the other states remains finite [18, 21, 22]. This discussion requires a series of considerations that we leave for later. For the moment, we notice that Eq. (10) allows us to determine the special value  $\vec{\beta}^*$  corresponding to the given strength sequence  $\vec{s}^*$ . Summing over all nodes  $j \neq i$ , the average value of the strength of node  $i$  is

$$\langle s_i \rangle_{\vec{\beta}} = \sum_{j \neq i} \langle w_{ij} \rangle_{\vec{\beta}} = \sum_{j \neq i} \frac{e^{-(\beta_i + \beta_j)}}{1 - e^{-(\beta_i + \beta_j)}}, \quad (12)$$

whence we can reformulate Eq. (4) as

$$\sum_{j \neq i} \langle w_{ij} \rangle_{\vec{\beta}^*} = \sum_{j \neq i} \frac{e^{-(\beta_i^* + \beta_j^*)}}{1 - e^{-(\beta_i^* + \beta_j^*)}} \equiv s_i^* \quad i = 1, n \quad (13)$$

which fixes the unique parameter value  $\vec{\beta}^*$ . Notably, this value is also the one that maximizes the (log-)likelihood [18, 28], i.e.

$$\vec{\beta}^* = \operatorname{argmax}_{\vec{\beta}} \ln P_{\text{can}}(\mathbf{W}^* | \vec{\beta}), \quad (14)$$

where, again,  $\mathbf{W}^*$  is any configuration such that  $\vec{s}(\mathbf{W}^*) = \vec{s}^*$ . In general, it is not possible to write  $\vec{\beta}^*$  explicitly as a function of  $\vec{s}^*$ . However, Eq. (13) or equivalently Eq. (14) can be efficiently solved numerically [18, 27] using various algorithms that have been coded for this purpose [29, 30]. In any case, a general property (that we will use later) is that for any two nodes  $i$  and  $j$  with the same expected strength ( $s_i^* = s_j^*$ ), the corresponding parameters  $\beta_i^*$  and  $\beta_j^*$  obey the same equation in (13) and are therefore equal. In other words,  $s_i^* = s_j^*$  implies  $\beta_i^* = \beta_j^*$ .

Once  $\vec{\beta}^*$  is calculated, we can plug it back into Eq. (8) to finally obtain the probability distribution

$$P_{\text{can}}(\mathbf{W} | \vec{\beta}^*) = \prod_{i=1}^n \prod_{i < j} \frac{e^{-(\beta_i^* + \beta_j^*) w_{ij}}}{[1 - e^{-(\beta_i^* + \beta_j^*)}]^{-1}} \quad (15)$$

that characterizes the canonical ensemble entirely. For practical purposes,  $\vec{\beta}^*$  can be inserted into Eq. (9) to obtain the link weight probability  $q_{ij}(w | \vec{\beta}^*)$ , from which several expected network properties can be calculated very directly. For instance, besides the expected link weight  $\langle w_{ij} \rangle_{\vec{\beta}^*}$ , we can calculate the probability that nodes  $i$  and  $j$  are connected by a link, irrespective of the weight of the latter, as follows:

$$\begin{aligned} p_{ij}^* &\equiv \langle \Theta(w_{ij}) \rangle_{\vec{\beta}^*} \\ &= \sum_{w=1}^{\infty} w q_{ij}(w | \vec{\beta}^*) \\ &= 1 - q_{ij}(0 | \vec{\beta}^*) \\ &= e^{-(\beta_i^* + \beta_j^*)} \\ &= \frac{\langle w_{ij} \rangle_{\vec{\beta}^*}}{1 + \langle w_{ij} \rangle_{\vec{\beta}^*}}, \end{aligned} \quad (16)$$

where  $\Theta(x)$  denotes the Heaviside step function, defined as  $\Theta(x) = 1$  if  $x > 0$  and  $\Theta(x) = 0$  if  $x \leq 0$ . Note that, if  $i$  and  $j$  belong to the condensed state where the expected link weight  $\langle w_{ij} \rangle_{\vec{\beta}^*}$  diverges ( $\beta_i^* + \beta_j^* \rightarrow 0^+$ ), then they become deterministically connected, i.e.  $p_{ij}^* \rightarrow 1^-$ . By contrast, non-condensed states have  $\langle w_{ij} \rangle_{\vec{\beta}^*} < \infty$ ,  $\beta_i^* + \beta_j^* > 0$ , and  $p_{ij}^* < 1$ . The analogy with BEC will be discussed in much more detail in Sec. IV.

Besides the structural properties, one of the key quantities that we will need in order to determine EE (or the lack thereof) is the resulting canonical entropy  $S_{\text{can}}^*$ , obtained by inserting Eq. (15) into Eq. (1):

$$\begin{aligned} S_{\text{can}}^* &\equiv S[P_{\text{can}}(\mathbf{W} | \vec{\beta}^*)] \\ &= \langle H(\mathbf{W}, \vec{\beta}^*) \rangle + \ln Z(\vec{\beta}^*) \\ &= \vec{\beta}^* \cdot \langle \vec{s}(\mathbf{W}) \rangle + \ln Z(\vec{\beta}^*) \\ &= \vec{\beta}^* \cdot \vec{s}^* + \ln Z(\vec{\beta}^*) \\ &= -\ln P_{\text{can}}(\mathbf{W}^* | \vec{\beta}^*). \end{aligned} \quad (17)$$

Note that the calculation of the canonical entropy  $S_{\text{can}}^*$  of the entire weighted network ensemble only requires the knowledge of the probability of one generic network  $\mathbf{W}^*$  with strength sequence  $\vec{s}^*$ , which is in turn directly calculated through Eq. (15).

### C. Microcanonical weighted network ensemble

We now come to the microcanonical ensemble. Its governing probability distribution  $P_{\text{mic}}(\mathbf{W} | \vec{s}^*)$  can be obtained by maximizing the Gibbs-Shannon entropy functional in Eq. (1) under the hard constraint

$$\vec{s}(\mathbf{W}) = \vec{s}^* \quad (18)$$

that applies to each network  $\mathbf{W}$  realized (with positive probability) in the set  $\mathcal{W}_n$ . The solution is obviously the uniform probability distribution

$$P_{\text{mic}}(\mathbf{W} | \vec{s}^*) = \begin{cases} \Omega_{\vec{s}^*}^{-1} & \vec{s}(\mathbf{W}) = \vec{s}^* \\ 0 & \vec{s}(\mathbf{W}) \neq \vec{s}^* \end{cases} \quad (19)$$

where  $\Omega_{\vec{s}^*}$  is the number of networks for which the hard constraint in Eq. (18) is realized. An implicit assumption throughout this paper is that the particular strength sequence  $\vec{s}^*$  is *graphic*, i.e. it can be realized by at least one network, so that  $\Omega_{\vec{s}^*} > 0$ . In this case as well, the (microcanonical) entropy is obtained by inserting Eq. (19) into Eq. (1):

$$\begin{aligned} S_{\text{mic}}^* &\equiv S[P_{\text{mic}}(\mathbf{W} | \vec{s}^*)] \\ &= \ln \Omega_{\vec{s}^*} \\ &= -\ln P_{\text{mic}}(\mathbf{W}^* | \vec{s}^*), \end{aligned} \quad (20)$$

which is also known as Boltzmann entropy. Note that  $\mathbf{W}^*$  has the same meaning as in Eq. (17), therefore both the canonical and microcanonical entropies are equal to

minus the log of the corresponding probability, evaluated in any state  $\mathbf{W}^*$  realizing the hard constraint in Eq. (18).

Note that, although the derivation of  $P_{\text{mic}}(\mathbf{W}|\bar{s}^*)$  is formally much more direct than that of  $P_{\text{can}}(\mathbf{W}|\bar{\beta}^*)$  in the conjugate canonical ensemble, its explicit calculation is more challenging, as it requires the combinatorial enumeration of all the  $\Omega_{\bar{s}^*}$  weighted networks with strength sequence  $\bar{s}^*$  (as a side remark, it is precisely the local nature of  $\bar{s}^*$  that makes the calculation of  $\Omega_{\bar{s}^*}$  daunting). Here, we will employ a recently proposed saddle-point asymptotic formula, for a generic discrete system under a  $K$ -dimensional vector  $\vec{C}^*$  of *effective* (see Sec. I C) constraints, for the number  $\Omega_{\vec{C}^*}$  of microcanonical configurations [6]. The formula uses only conjugate canonical quantities, namely the canonical entropy  $S_{\text{can}}^*$  and the  $K \times K$  covariance matrix  $\Sigma^*$  among the  $K$  constraints in the canonical ensemble, and reads [6]

$$\Omega_{\vec{C}^*} = \frac{e^{S_{\text{can}}^*}}{\sqrt{\det(2\pi\Sigma^*)}} \prod_{k=1}^K [1 + O(1/\lambda_k^*)], \quad (21)$$

where  $\{\lambda_k^*\}_{k=1}^K$  are the eigenvalues of  $\Sigma^*$ . The symbol  $O(x)$  indicates a quantity with a finite limit when divided by  $x$  as  $n \rightarrow \infty$ , i.e.  $O(x)$  is asymptotically of the same order as  $x$ . Note that, since covariance matrices are positive semidefinite,  $\lambda_k^* \geq 0$  for all  $k$ . Moreover, since the constraints are assumed to be non-redundant, then  $\lambda_k^* > 0$  for all  $k$  [6] (if some of the constraints were redundant, there would be certain zero eigenvalues rendering the above equation inapplicable; that is why the formula should be applied to a maximal set of  $K$  non-redundant constraints). Finally, if these eigenvalues grow sufficiently fast as  $n \rightarrow \infty$ , then the product on the right hand side becomes irrelevant, in which case the knowledge of  $S_{\text{can}}^*$  and  $\det(2\pi\Sigma^*)$  is enough in order to characterize the asymptotics of  $\Omega_{\vec{C}^*}$ .

In our setting where  $\vec{C}^* = \bar{s}^*$  and  $K = n$  (node strengths are all mutually independent as it is not possible to guess any individual node strength from the knowledge of the other  $n - 1$  ones), we calculate the entries of  $\Sigma^*$  as

$$\begin{aligned} \Sigma_{ij}^* &\equiv \text{Cov}_{\bar{\beta}^*}(s_i, s_j) \\ &= \langle s_i s_j \rangle_{\bar{\beta}^*} - \langle s_i \rangle_{\bar{\beta}^*} \langle s_j \rangle_{\bar{\beta}^*} \\ &= \left. \frac{\partial^2 \ln Z(\vec{\beta})}{\partial \beta_i \partial \beta_j} \right|_{\vec{\beta}=\bar{\beta}^*} \end{aligned} \quad (22)$$

where  $Z(\vec{\beta})$  is given by Eq. (7). An explicit calculation gives

$$\begin{aligned} \Sigma_{ii}^* &= \text{Var}_{\bar{\beta}^*}(s_i) \\ &= \sum_{j \neq i} \frac{e^{-(\beta_i^* + \beta_j^*)}}{[1 - e^{-(\beta_i^* + \beta_j^*)}]^2} \\ &= \sum_{j \neq i} \langle w_{ij} \rangle_{\bar{\beta}^*} (1 + \langle w_{ij} \rangle_{\bar{\beta}^*}) \end{aligned} \quad (23)$$

for the diagonal entries (i.e. the variances of the constraints) and

$$\begin{aligned} \Sigma_{ij}^* &= \text{Cov}_{\bar{\beta}^*}(s_i, s_j) \\ &= \text{Var}_{\bar{\beta}^*}(w_{ij}) \\ &= \frac{e^{-(\beta_i^* + \beta_j^*)}}{[1 - e^{-(\beta_i^* + \beta_j^*)}]^2} \\ &= \langle w_{ij} \rangle_{\bar{\beta}^*} (1 + \langle w_{ij} \rangle_{\bar{\beta}^*}) \quad (i \neq j) \end{aligned} \quad (24)$$

for the off-diagonal entries (i.e. the covariances between distinct constraints).

We finally obtain

$$\begin{aligned} S_{\text{mic}}^* &= \ln \Omega_{\bar{s}^*} \\ &= S_{\text{can}}^* - \ln \sqrt{\det(2\pi\Sigma^*)} + \sum_{k=1}^n \ln[1 + O(1/\lambda_k^*)] \end{aligned} \quad (25)$$

where we have used  $\sqrt{\det(2\pi\Sigma^*)} = \prod_{k=1}^n \sqrt{2\pi\lambda_k^*}$ . Note that the eigenvalues  $\{\lambda_k^*\}_{k=1}^n$  are positive, the  $n$  node strengths being linearly independent constraints [6]. In principle, in order to compute (the leading term of) Eq. (25) explicitly, we need to specify a value for  $\bar{s}^*$ , calculate the resulting matrix  $\Sigma^*$ , and the eigenvalues of the latter. However, in Sec. III we show that the knowledge of the diagonal elements of the covariance matrix is enough for our purposes. This result is then used in Sec. IV when we consider specific choices of  $\bar{s}^*$ .

### III. EQUIVALENCE AND NONEQUIVALENCE OF WEIGHTED NETWORK ENSEMBLES

In this section we use the knowledge of the canonical and microcanonical probability distributions derived in the previous section in order to establish two criteria for the equivalence of ensembles of weighted networks, namely the one based on the vanishing of the relative entropy density between the two distributions [4] and the traditional one based on the vanishing of the canonical relative fluctuations of the constraints [1].

#### A. Relative entropy density

As we have anticipated, EE can be stated mathematically using three different notions, namely *thermodynamic*, *macrostate* and *measure* equivalence [4]. These definitions turn out to be, under mild assumptions, essentially equivalent [4]. Here, we use the definition in the measure sense, which has been recently formulated explicitly for binary network ensembles [5, 6, 19, 20] and is based on the vanishing of a suitable *relative entropy density* between the microcanonical and canonical probability distributions. Our calculations generalize those results to the case of weighted networks, for which measure equivalence has not been studied yet.

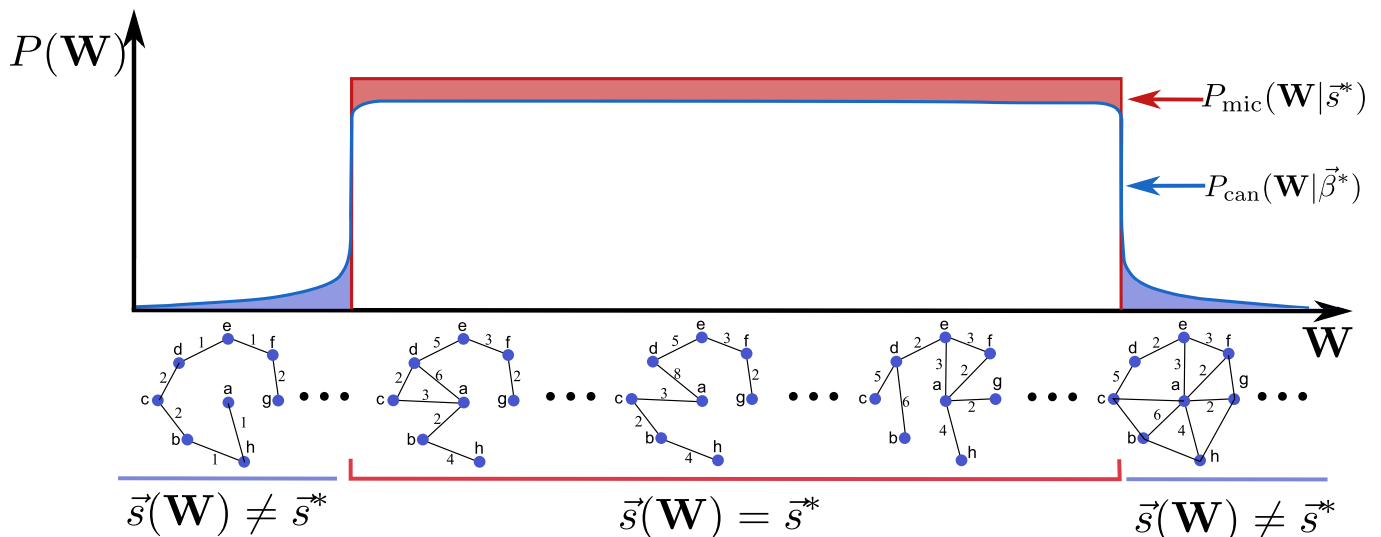


FIG. 2. Illustration of the nonequivalence of ensembles of weighted networks with given strength sequence  $\vec{s}^*$  in the measure sense. Schematically, the  $x$  axis represents all weighted networks  $\mathbf{W} \in \mathcal{W}_n$ . Here,  $n = 8$  and  $\vec{s}^* = (11, 6, 5, 13, 8, 5, 2, 4)$ . The  $\Omega_{\vec{s}^*}$  networks matching the particular strength sequence  $\vec{s}^*$ , i.e. those for which  $\vec{s}(\mathbf{W}) = \vec{s}^*$ , are represented in the middle. The  $y$  axis represents the canonical and microcanonical probabilities for each network. The microcanonical distribution  $P_{\text{mic}}(\mathbf{W}|\vec{s}^*)$  assigns zero probability to the networks for which  $\vec{s}(\mathbf{W}) \neq \vec{s}^*$ , and uniform probability  $P_{\text{mic}}(\mathbf{W}|\vec{s}^*) = \Omega_{\vec{s}^*}^{-1}$  to each network  $\mathbf{W}^*$  for which  $\vec{s}(\mathbf{W}^*) = \vec{s}^*$ . By contrast, the conjugate canonical distribution  $P_{\text{can}}(\mathbf{W}|\vec{\beta}^*)$  assigns positive probability to all the networks in  $\mathcal{W}_n$  and therefore has ‘tails’ that extend all over the  $x$  axis. Normalization implies that, while also the canonical probability gives a constant value  $P_{\text{can}}(\mathbf{W}^*|\vec{\beta}^*)$  to each network  $\mathbf{W}^*$  with the strength sequence  $\vec{s}^*$ , this value is (for  $n$  finite) smaller (lower plateau) than the corresponding microcanonical one (upper plateau):  $P_{\text{can}}(\mathbf{W}^*|\vec{\beta}^*) < P_{\text{mic}}(\mathbf{W}^*|\vec{s}^*)$ . Indeed, the blue and red areas should be equal because of normalization. Intuitively, the two ensembles become equivalent in the thermodynamic limit if, sufficiently fast as  $n \rightarrow \infty$ , the canonical tails vanish (the blue areas disappear) and the canonical plateau ‘catches up’ with the microcanonical one (the red area disappears). Measure equivalence formalizes this ‘sufficiently fast’ rigorously, finding that EE corresponds to the condition  $\lim_{n \rightarrow \infty} [\ln P_{\text{mic}}(\mathbf{W}^*|\vec{s}^*) - \ln P_{\text{can}}(\mathbf{W}^*|\vec{\beta}^*)]/n = 0$ . It turns out that this is not the case for the model discussed here: canonical and microcanonical ensembles of weighted networks with given strength sequence are not equivalent.

In general, the relative entropy (or Kullback-Leibler divergence) between two distributions  $P$  and  $Q$ , both having support over a discrete set  $\mathcal{W}_n = \{\mathbf{W}\}$  of configurations in analogy with Eq. (1), is defined as

$$D[P||Q] \equiv \sum_{\mathbf{W} \in \mathcal{W}_n} P(\mathbf{W}) \ln \frac{P(\mathbf{W})}{Q(\mathbf{W})} \quad (26)$$

and quantifies ‘how far’ the distribution  $P$  is from the reference distribution  $Q$  [31]. When  $P$  and  $Q$  represent the microcanonical and canonical distributions respectively, it can be shown [5, 6] that  $D[P||Q]$  reduces to the difference between the canonical and microcanonical entropy, which can be both estimated on a single configuration realizing the hard constraint (as we have indeed shown in the previous section for our weighted network model). Moreover, its calculation asymptotically requires only the canonical covariance matrix  $\Sigma^*$  between the constraints.

We are now going to see how these general results apply to our specific case. A visual illustration of the idea behind measure equivalence for our ensembles of weighted networks with given strength sequence is provided in Fig. 2. Following Eq. (26), the relative entropy

between  $P_{\text{mic}}(\mathbf{W}|\vec{s}^*)$  and  $P_{\text{can}}(\mathbf{W}|\vec{\beta}^*)$  is defined as

$$\begin{aligned} D^* &\equiv D[P_{\text{mic}}(\mathbf{W}|\vec{s}^*)||P_{\text{can}}(\mathbf{W}|\vec{\beta}^*)] \\ &= \sum_{\mathbf{W} \in \mathcal{W}_n} P_{\text{mic}}(\mathbf{W}|\vec{s}^*) \ln \frac{P_{\text{mic}}(\mathbf{W}|\vec{s}^*)}{P_{\text{can}}(\mathbf{W}|\vec{\beta}^*)}. \end{aligned} \quad (27)$$

By inserting Eq. (19) into Eq. (27) and using the fact that  $P_{\text{can}}(\mathbf{W}|\vec{\beta}^*)$  has the same value for any network  $\mathbf{W}^*$  matching the hard constraint  $\vec{s}(\mathbf{W}^*) = \vec{s}^*$ , we confirm that  $D^*$  can be estimated pointwise on  $\mathbf{W}^*$  as

$$D^* = \ln P_{\text{mic}}(\mathbf{W}^*|\vec{s}^*) - \ln P_{\text{can}}(\mathbf{W}^*|\vec{\beta}^*). \quad (28)$$

Moreover, using Eqs. (17) and (20), we also confirm that it reduces to the entropy difference

$$D^* = S_{\text{can}}^* - S_{\text{mic}}^*. \quad (29)$$

Now, Eq. (25) immediately allows us to obtain

$$D^* = \ln \sqrt{\det(2\pi\Sigma^*)} + \sum_{k=1}^n \ln[1 + O(1/\lambda_k^*)] \quad (30)$$

which depends only on the eigenvalues of the canonical covariance matrix  $\Sigma^*$ , whose diagonal and off-diagonal entries are given in Eqs. (23) and (24) respectively.



The definition of measure equivalence is the vanishing of the relative entropy *density*, i.e. of the ratio  $D^*/n$ , in the thermodynamic limit [4]. Explicitly, EE in the measure sense corresponds to

$$d^* \equiv \lim_{n \rightarrow \infty} \frac{D^*}{n} = 0 \quad (31)$$

or equivalently [6]

$$D^* = o(n), \quad (32)$$

where  $o(n)$  denotes a quantity that, if divided by  $n$ , vanishes as  $n \rightarrow \infty$ . Equation (29) allows us to understand the above definition of macrostate EE as follows. The microcanonical entropy  $S_{\text{mic}}^*$  is the logarithm of the number of accessible configurations under hard constraints, while the canonical entropy  $S_{\text{can}}^*$  is the logarithm of the corresponding ‘effective’ number of configurations under soft constraints. EE requires that, as  $n$  increases, the typical configurations of the system under the two ensembles become the same, i.e. that the (effective) numbers of configurations in the two ensembles become closer to each other. This cannot happen if, as we keep adding one more unit to the system, the difference between the two entropies (i.e. the relative entropy  $D^*$ ) keeps increasing by an arbitrary amount. The criterion in Eq. (31) establishes that if the entropy difference per node, i.e.  $D^*/n$ , does not vanish as  $n$  diverges, then the two ensembles cannot be equivalent.

Equation (32) implies that, in order to assess whether the system is under EE, we do not need the exact value of  $D^*$ , but only its leading order with respect to  $n$ . Then from Eq. (30) we see that, since the term  $1 + O(1/\lambda_k^*)$  is at most of the same order as  $\sqrt{2\pi\lambda_k^*}$ , the presence of  $O(1/\lambda_k^*)$  does not affect ensemble (non)equivalence:

$$D^* = O\left(\sum_{k=1}^n \ln \lambda_k^*\right) = O(\ln \det \Sigma^*). \quad (33)$$

So, in order to check whether Eq. (32) holds, it is ultimately enough to check whether

$$\ln \det \Sigma^* = o(n). \quad (34)$$

On the other hand, since our hypothesis of non-redundant constraints implies  $\lambda_k^* > 0$  for all  $k$  (as discussed in Sec. II C), we see that the contribution of the term  $\sum_{k=1}^n \ln[1 + O(1/\lambda_k^*)]$  to the relative entropy in Eq. (30) is at most  $O(n)$ . So if  $\ln \sqrt{\det(2\pi\Sigma^*)}$  grows faster than  $O(n)$  then Eq. (30) reduces to the stronger result

$$D^* \approx \ln \sqrt{\det(2\pi\Sigma^*)} \approx \frac{1}{2} \ln \det \Sigma^*, \quad (35)$$

i.e. the leading term of the relative entropy (not only its leading order) can be calculated exactly from  $\ln \det \Sigma^*$  (throughout this paper, the symbol “ $x \approx y$ ” indicates that the leading term of  $x$  and  $y$  is asymptotically the

same, i.e. the two quantities differ by a quantity that vanishes if divided by either  $x$  or  $y$  as  $n \rightarrow \infty$ ). Depending on the regimes considered later on in the paper, different techniques for calculating (the leading order of) the determinant of the covariance matrix  $\Sigma^*$  can be used. We will discuss these techniques when needed, and refer to the Appendix for explicit calculations. We will show that, except in a certain zero-temperature limit, the conditions ensuring Eq. (35) are met and the leading order of the relative entropy can be calculated exactly.

## B. Relative fluctuations of the constraints

We now consider the relative fluctuations of the constraints, whose behaviour in the thermodynamic limit is, historically, the traditional criterion used to check whether statistical ensembles are equivalent [1]. In the standard situation, where the canonical and microcanonical ensembles are defined through a single scalar constraint on the total energy  $E$ , the relative fluctuations are captured by a single scalar quantity  $r^* \equiv \sigma_{\beta^*}(E)/E^*$  representing the ratio of the canonical standard deviation of the energy to the mean energy itself. EE is then associated with the vanishing of  $r^*$  as  $n \rightarrow \infty$ . In statistics,  $r^*$  is called the *coefficient of variation* of the random variable  $E$  with expected value  $E^*$  and variance  $\sigma_{\beta^*}^2(E)$ .

In the case of networks with local constraints, there are  $n$  coefficients of variation to consider. They have been calculated for both the binary and the weighted versions of the configuration model [27]. In extreme summary, those results show that, in the binary case (where there is a constraint on the degree  $k_i^*$  for each node  $i = 1, n$ ), an *upper bound*  $1/\sqrt{k_i^*}$  for the relative fluctuation  $r_i^* = \sigma_{\beta^*}(k_i)/k_i^*$  can be established. By contrast, in the weighted case (corresponding to the model considered here with a constraint on the strength  $s_i^*$  for each node  $i = 1, n$ ), the value  $1/\sqrt{s_i^*}$  becomes a *lower bound* for the relative fluctuation  $r_i^* = \sigma_{\beta^*}(s_i)/s_i^*$  [27]. Those results have two implications.

First, in the binary case the only regime for which general conclusions can be drawn about the vanishing of the relative fluctuations is the so-called *dense* regime where the expected degree of all nodes diverges, hence  $r_i^* \rightarrow 0$ . In the opposite *sparse* regime where the average degree of all nodes is finite, we only have a finite upper bound for the relative fluctuations, but their actual value depends on the specific network. In general, however, the decreasing behaviour of the upper bound for the relative fluctuations in binary networks with increasing degrees is opposite to that of the relative entropy density, which increases as the expected degree increases [5, 19, 20].

Second, in the weighted case we have a somewhat opposite situation where we can only conclude that, in the sparse regime where the expected strengths are finite (apart from possible hubs), the relative fluctuations do not vanish. By contrast, in the dense case where the expected strengths diverge, the relative fluctuations can

in principle pick any value. Confusingly, even if in the weighted case the lower bound for the relative fluctuations of the strengths goes to zero for nodes with diverging strength  $s_i^*$ , previous results seem to indicate that the realized value of  $r_i^*$  in networks with heterogeneous strength sequence actually *increases* for higher  $s_i^*$  [27]. This behaviour suggests that weighted networks, due to the many possible ways in which weight can accumulate on links, can behave very differently from binary networks. This observation requires further research and strengthens our motivation for studying the relative fluctuations in weighted networks in different scenarios (ranging from homogeneous to heterogenous concentrations of link weights), in conjunction with the relative entropy and, in general, ensemble (non)equivalence.

Using Eq. (23), we can immediately calculate the standard deviation of each constraint  $s_i$  around its expected value  $s_i^*$  as

$$\begin{aligned}\sigma_{\bar{\beta}^*}(s_i) &= \sqrt{\text{Var}_{\bar{\beta}^*}(s_i)} \\ &= \sqrt{\sum_{i \neq j} \frac{e^{-(\beta_i^* + \beta_j^*)}}{[1 - e^{-(\beta_i^* + \beta_j^*)}]^2}} \\ &= \sqrt{\sum_{i \neq j} \langle w_{ij} \rangle_{\bar{\beta}^*} (1 + \langle w_{ij} \rangle_{\bar{\beta}^*})} \\ &= \sqrt{s_i^* + \sum_{i \neq j} \langle w_{ij} \rangle_{\bar{\beta}^*}^2}.\end{aligned}\quad (36)$$

The relative fluctuation of the strength is therefore

$$r_i^* = \frac{\sigma_{\bar{\beta}^*}(s_i)}{s_i^*} = \sqrt{\frac{1}{s_i^*} + \frac{\sum_{i \neq j} \langle w_{ij} \rangle_{\bar{\beta}^*}^2}{(\sum_{i \neq j} \langle w_{ij} \rangle_{\bar{\beta}^*})^2}},\quad (37)$$

Since  $\langle w_{ij} \rangle_{\bar{\beta}^*} \geq 0$  for all  $i, j$ , we have  $(\sum_{i \neq j} \langle w_{ij} \rangle_{\bar{\beta}^*})^2 \geq \sum_{i \neq j} (\langle w_{ij} \rangle_{\bar{\beta}^*})^2$ , showing that  $1/\sqrt{s_i^*}$  is indeed a lower bound for  $r_i^*$ . When studying the asymptotic behaviour of the relative fluctuations in the thermodynamic limit, we will be interested in whether the limit

$$\rho_i^* \equiv \lim_{n \rightarrow \infty} r_i^* \quad (38)$$

is zero (vanishing relative fluctuations) or positive (non-vanishing relative fluctuations) for each node  $i = 1, n$ .

#### IV. BEC IN WEIGHTED NETWORKS

In physical systems composed of bosons, i.e. particles obeying Bose-Einstein statistics, BEC is a phase transition whereby, below a certain critical temperature, a finite fraction of the total number of particles condenses in the ground state, i.e. the state with lowest energy (or more generally in a finite number of states with lowest energy). BEC was theoretically predicted by Satyendra Nath Bose and Albert Einstein in 1924 [32], and

it has since then been observed in various physical systems. Models of BEC have been studied in different statistical ensembles in the standard case with only global constraints (total energy and/or total number of particles) [33–38]. Although the detailed phenomenology exhibited by these models depends on the choice of the energy and the structure of the interactions, it is generally found that EE breaks down in the condensed (low-temperature) phase, as signalled by nonvanishing relative fluctuations of the constraints.

In this Section, we are going to show that a form of BEC, even if quite different from that found in more traditional physical settings, can also appear in our ensembles of weighted networks. The possible onset of BEC in our system creates an ideal situation where an EE-breaking phase transition can be studied in combination with an additional and unrelated mechanism for the breakdown of EE, i.e. the presence of local constraints, which is always active in both the condensed and the non-condensed phases. To illustrate our results, we first make some important clarifications in order to establish a rigorous link from weighted network ensembles to Bose-Einstein statistics and then study the different phases of the model.

##### A. Bose-Einstein statistics in weighted networks

As we have already recalled, weighted networks with a constraint on the strength sequence obey Bose-Einstein statistics, as opposed to binary networks that obey Fermi-Dirac statistics [21, 22, 24]. Indeed, inserting Eq. (6) into Eq. (5) we get the probability of a configuration for a gas of free particles in the *grandcanonical ensemble*<sup>2</sup>, where the pair  $i, j$  labels an energy state, the weight  $w_{ij}$  is the number of particles in that state (occupation number), and the sum  $\beta_i + \beta_j$  can be interpreted as

$$\beta_i + \beta_j = \frac{\epsilon_{ij} - \mu(T)}{kT}.\quad (39)$$

In the latter expression,  $\epsilon_{ij}$  represents the energy of the state,  $1/kT$  is the inverse temperature, and  $\mu(T)$  is the *chemical potential* (required to fix the same expected overall number of particles for all values of  $T$ ) [24]. Indeed, as we discussed in Sec. II B, in our setting the energy and temperature (and in this case, the chemical potential as well) are all reabsorbed into  $\vec{\beta}$ . Therefore we

<sup>2</sup> In the grandcanonical ensemble, both the total energy and the total number of particles are treated as soft constraints. With respect to the canonical ensemble, the appearance of the number of particles as an additional soft constraint requires the introduction of an extra Lagrange multiplier, the chemical potential. Interestingly, in the context of BEC a fourth (so-called ‘Maxwell’s Demon’) ensemble has also been introduced where the total number of particles is soft while the total energy is hard [33].

can interpret the link weight  $w_{ij}$  as the number of ‘elementary particles’ of weight, i.e. the number of quanta of unit weight, populating the link between nodes  $i$  and  $j$  [21, 22]. The total number of such particles in the system is the total weight  $W$  of all links in the network:

$$W(\mathbf{W}) = \sum_{i=1}^n \sum_{j<i} w_{ij} = \frac{1}{2} \sum_{i=1}^n s_i(\mathbf{W}). \quad (40)$$

The ‘weighted’ property  $w_{ij} \in \mathbb{N}$ , which leads to Eqs. (7), (8) and (9), corresponds to the possibility that the same state (pair of nodes) is occupied by indefinitely many particles (subject to the average number dictated by the chemical potential), which is a property of bosons. By contrast, in binary networks one has to impose  $w_{ij} \in (0, 1)$ , which is a property of fermions [21]. An extensive treatment of the role of chemical potential and temperature in binary networks can be found in [24]. Here, to properly interpret what the weighted model is doing, we should give a series of clarifications.

First, we should make a clear distinction between the  $n$  ‘units’ of our system (i.e. the nodes of the network) and the  $W$  ‘particles’ of weight that, as a formal analogy, can be interpreted as populating the links of the network. The former are the real constituents of our physical system, while the latter are a mathematical abstraction used to represent the nature of the interactions (links) between such constituents. If we imagine doubling the size of our network, we should imagine doubling the number  $n$  of nodes: indeed, we can imagine the network ‘growing in size’ by adding one single node at a time, but we cannot imagine adding one single *pair* of nodes at a time, without actually adding  $n$  new pairs. One should also not be tempted to regard node pairs as the fundamental units by the fact that, mathematically, the  $n(n-1)/2$  variables  $\{w_{ij}\}$  involving different pairs of nodes are independent random variables: actually, this only occurs in the canonical ensemble and would in any case not be true for more general choices of the constraints. Moreover, even the  $n(n-1)/2$  independent node pairs in the canonical ensemble cannot be assigned independent values of the parameters, since there are only  $n$  parameters corresponding to the Lagrange multipliers attached to each node. Explicit (and strong) consequences of this fact will be illustrated precisely in the context of BEC. Therefore the physical size of our system is  $n$ , and this is why in Eq. (31) we defined the relative entropy density as the relative entropy divided by  $n$  in the first place. How the total weight  $W$  varies with the system size  $n$  depends on a specific property, i.e. on how we make the entries of  $\vec{s}^*$  (and the resulting value of  $W^* = \sum_{i=1}^n s_i^*/2$ ) scale with  $n$ . For instance, we may choose to be in the sparse regime where  $\vec{s}^*$  remains finite as  $n \rightarrow \infty$ , or in the dense regime where  $\vec{s}^*$  diverges as  $n \rightarrow \infty$ . As we show below, the latter is the relevant case for BEC to emerge.

Second, we stress that, irrespective of the above, we always consider a hard number  $n$  of nodes, and this is why we compare (only) the canonical (soft value of  $\vec{s}$ )

and microcanonical (hard value of  $\vec{s}$ ) ensembles of networks, both for fixed  $n$  (which sets the dimension of  $\vec{s}$ ). We do *not* consider the grandcanonical ensemble of *network configurations* where  $n$  is soft. The grandcanonical ensemble introduced in the aforementioned analogy with systems of bosons is a different one; it may be denoted as an ensemble of *weight quanta* in a network with fixed  $n$  and originates from the fact that the Hamiltonian in Eq. (6), and consequently the total link weight (not  $n$ ) in Eq. (40), is a fluctuating quantity in the canonical ensemble of network configurations. The fluctuations in the (rescaled) energy  $H$  (canonical ensemble of network configurations) are seen as fluctuations in the particle number  $W$  (grandcanonical ensemble of weight quanta) in the Bose-Einstein analogy. Fluctuations in the particle number have been the subject of many studies in the literature on BEC [33–38]. Note that, in both canonical and microcanonical ensembles, the individual link weights  $\{w_{ij}\}$  are fluctuating quantities, despite the fact that the total link weight  $W^*$  is constant in the microcanonical ensemble. Therefore the numbers of ‘weight particles’ of individual links fluctuate in both ensembles.

Third, while we necessarily discuss the (non)equivalence of the canonical and microcanonical ensembles in the thermodynamic limit  $n \rightarrow \infty$ , the total weight  $W^*$  can (and, across the canonical ensemble, will in any case) be arbitrarily large even for finite  $n$ . Indeed, the phase transition that we are about to discuss (namely, BEC) does not *per se* require the limit  $n \rightarrow \infty$ , while it definitely requires the limit  $W^* \rightarrow \infty$ . Abstractly, these two limits (and the associated phenomena of EN and BEC respectively) may appear as mathematically unrelated. However, in practice they are physically related once the scaling of  $\vec{s}^*$  with  $n$  is specified. In particular, we are going to show that, in order to observe BEC, we need be in a dense regime where  $W^* = O(n^2)$ . This ensures that, when taking the thermodynamic limit  $n \rightarrow \infty$  in order to study ensemble (non)equivalence, we are automatically implying  $W^* \rightarrow \infty$  so that we can check for BEC at the same time.

Last, we recall that  $\vec{\beta} \cdot \vec{s}(\mathbf{W})$  has to be dimensionless in order to ensure that the probability is a number. Therefore, since  $w_{ij}$  is dimensionless, so are  $s_i^*$  and  $\beta_i^*$ . In turn, this implies that both sides of Eq. (39) must be dimensionless. On the other hand, when modelling a real system, the ‘energy’  $\epsilon_{ij}$  may represent any physical ‘cost’ associated to the link between nodes  $i$  and  $j$  (more precisely, it represents the cost of reinforcing  $w_{ij}$  by a unit of weight) and may therefore carry its own unit of measure (e.g. it may depend on some distance between nodes  $i$  and  $j$ ). Necessarily, the chemical potential  $\mu(T)$  carries the same units as the energy. As for the ‘temperature’  $T$ , it may be chosen to be dimensionless as it merely represents a control parameter (this is the choice that we will make later); alternatively, it may carry the same units of the energy if it is useful that temperature and energy live on the same scale. Irrespective of this

choice, in our setting the ‘Boltzmann constant’  $k$  is simply a constant that takes care of all dimensional units of measure and makes the ratio on the right hand side of Eq. (39) dimensionless.

## B. Core-periphery networks

With the above clarifications, we can finally go back to our model. In the traditional physical situation, in the canonical ensemble the energy  $\epsilon_{ij}$  of each state  $i, j$  is a constant and the temperature  $T$  can be varied. Clearly,  $\epsilon_{ij}$  is independent of  $T$ , while the chemical potential  $\mu(T)$  is chosen, as a function of temperature, in order to realize the correct ( $T$ -independent) expected total number  $\langle W \rangle^* \equiv W^*$  of particles for all values of  $T$ . In this ‘direct problem’, every state will therefore have an expected occupation number governed by  $\epsilon_{ij}$ ,  $T$  and  $\mu(T)$ . In our ‘inverse’ setting,  $T^*$  and  $\mu^*$  are instead reabsorbed into  $\vec{\beta}^*$ , which in turn is induced by the chosen value of the strength sequence (rather than the other way around). We should therefore regard  $\vec{s}^* = \vec{s}^*(T)$  and  $\vec{\beta}^* = \vec{\beta}^*(T)$  as  $T$ -dependent, while  $W^*$  remains  $T$ -independent. This means that the chemical potential  $\mu^*(T)$  should be such that

$$\sum_{i=1}^n s_i^*(T) = 2W^* \quad \forall T \geq 0. \quad (41)$$

BEC emerges when, below a certain critical temperature  $T_c$ , the occupation number of the state with minimum energy  $\epsilon_{\min} = \min_{i,j} \{\epsilon_{ij}\}$  (ground state), or of a finite number of states with lowest energy, becomes so large that it reaches a finite fraction of the total number  $W^*$  of particles. Clearly, this requires the existence of at least two different energy levels (the ground state and at least one ‘excited’ state). Therefore the simplest way to obtain BEC in our model is by considering a strength sequence of the form

$$s_i^*(T) = \begin{cases} s_+^*(T) & i = 1, n_+ \\ s_-^*(T) & i = n_+ + 1, n \end{cases} \quad s_+^*(T) \geq s_-^*(T), \quad (42)$$

i.e. by partitioning the  $n$  nodes into two classes, which we call *core* and *periphery*: the core has a finite number

$$n_+ = O(1) \quad (43)$$

of nodes, each having a ‘large’ strength  $s_+^*(T)$ , while the periphery has an extensive number

$$n_- = n - n_+ = O(n) \quad (44)$$

of nodes, each having a ‘small’ strength  $s_-^*(T)$ . What we mean precisely by ‘small’ and ‘large’ will be clarified below. For the moment, we notice that the BEC phase ( $T < T_c$ ) corresponds to picking a ‘condensed’ value of  $\vec{s}^*(T < T_c)$  such that, in the thermodynamic limit, the core takes up a finite fraction of the total weight  $W^*$

of all links in the network, despite having a finite size. In particular, in the zero-temperature limit *all* the total weight  $W^*$  is in the core. By contrast, the non-condensed phase  $T > T_c$  is one where  $\vec{s}^*(T > T_c)$  is such that no individual link receives a finite fraction of  $W^*$ . In particular, the infinite-temperature limit should be such that the energy difference between ground and excited states becomes ineffective, i.e.  $s_+^*(T \rightarrow \infty) = s_-^*(T \rightarrow \infty)$ . The different phases can be efficiently monitored by introducing a temperature-dependent order parameter  $Q^*(T)$ , as we show below.

We stress that, since we are ultimately interested in the relative fluctuations of the canonical constraints and in the relative entropy that can be asymptotically calculated purely from canonical quantities according to Eq. (33), practically we only need to study the canonical ensemble. The only check we need to make is that, whenever we speak of the system being in a certain ‘phase’, this statement does not depend on the particular ensemble. In other words, we need to show that the order parameter has always the same value in the canonical and microcanonical ensembles.

Before studying the individual phases, let us make some general considerations, valid for all values of  $T$ . We first find the value of  $\vec{\beta}^*(T)$  corresponding to the value of  $\vec{s}^*(T)$  in Eq. (42). As we mentioned,  $s_i^*(T) = s_j^*(T)$  implies  $\beta_i^*(T) = \beta_j^*(T)$ , therefore the entries of  $\vec{\beta}^*(T)$  take only two values  $\beta_+^*(T)$  and  $\beta_-^*(T)$  such that

$$\beta_i^*(T) = \begin{cases} \beta_+^*(T) & i = 1, n_+ \\ \beta_-^*(T) & i = n_+ + 1, n \end{cases} \quad \beta_+^*(T) \leq \beta_-^*(T). \quad (45)$$

These values solve the  $n$  equations in (13), which here reduce to the two independent equations

$$(n_+ - 1)w_+^*(T) + n_-w_0^*(T) \equiv s_+^*(T) \quad (46)$$

$$(n_- - 1)w_-^*(T) + n_+w_0^*(T) \equiv s_-^*(T) \quad (47)$$

where we have defined

$$w_+^*(T) = \frac{e^{-2\beta_+^*(T)}}{1 - e^{-2\beta_+^*(T)}} \quad (48)$$

as the expected link weight  $\langle w_{ij} \rangle_{\vec{\beta}^*(T)}$  between any two nodes in the core ( $i, j = 1, n_+$ ),

$$w_-^*(T) = \frac{e^{-2\beta_-^*(T)}}{1 - e^{-2\beta_-^*(T)}} \quad (49)$$

as the expected link weight  $\langle w_{ij} \rangle_{\vec{\beta}^*(T)}$  between any two nodes in the periphery ( $i, j = n_+ + 1, n$ ), and

$$\begin{aligned} w_0^*(T) &= \frac{e^{-\beta_-^*(T) - \beta_+^*(T)}}{1 - e^{-\beta_-^*(T) - \beta_+^*(T)}} \\ &= \frac{\sqrt{w_+^*(T)w_-^*(T)}}{\sqrt{1 + w_+^*(T)}\sqrt{1 + w_-^*(T)} - \sqrt{w_+^*(T)w_-^*(T)}} \\ &= \frac{1}{\sqrt{1 + 1/w_+^*(T)}\sqrt{1 + 1/w_-^*(T)} - 1} \end{aligned} \quad (50)$$

as the expected link weight  $\langle w_{ij} \rangle_{\beta^*(T)}$  between any node in the core and any node in the periphery ( $i = 1, n_+$  and  $j = n_+ + 1, n$  or  $j = 1, n_+$  and  $i = n_+ + 1, n$ ). Note that  $\beta_+^*(T) \leq \beta_-^*(T)$  implies  $w_-^*(T) \leq w_0^*(T) \leq w_+^*(T)$ .

Now, solving Eqs. (46) and (47), we obtain the explicit values of  $\beta_+^*(T)$  and  $\beta_-^*(T)$  appearing in Eq. (45):

$$\beta_{\pm}^*(T) = \frac{1}{2} \ln \frac{1 + w_{\pm}^*(T)}{w_{\pm}^*(T)} = \frac{1}{2} \ln \left( 1 + \frac{1}{w_{\pm}^*(T)} \right). \quad (51)$$

Also note from Eq. (50) that

$$\beta_+^*(T) + \beta_-^*(T) = \ln \left( 1 + \frac{1}{w_0^*(T)} \right). \quad (52)$$

Also, using Eq. (16) we can define

$$p_{\pm}^*(T) \equiv e^{-2\beta_{\pm}^*(T)} = \frac{w_{\pm}^*(T)}{1 + w_{\pm}^*(T)} \quad (53)$$

as the probability that a link exists (irrespective of its weight) between any two core-core (+) or any two periphery-periphery (-) nodes, and

$$p_0^*(T) \equiv e^{-\beta_+^*(T) - \beta_-^*(T)} = \frac{w_0^*(T)}{1 + w_0^*(T)} \quad (54)$$

as the probability that a link exists (irrespective of its weight) between a core node and a periphery node.

From Eq. (39), we notice that the existence of the two values above for the entries of  $\vec{\beta}^*(T)$  implies that there are three energy levels, associated with the energies

$$\epsilon_+^* = \mu^*(T) + 2kT\beta_+^*(T), \quad (55)$$

$$\epsilon_-^* = \mu^*(T) + 2kT\beta_-^*(T), \quad (56)$$

$$\epsilon_0^* = \mu^*(T) + kT[\beta_+^*(T) + \beta_-^*(T)] = \frac{\epsilon_+^* + \epsilon_-^*}{2}, \quad (57)$$

where  $\epsilon_+^* \leq \epsilon_0^* \leq \epsilon_-^*$  (we recall that all energy values are finite and independent of both  $T$  and  $n$ ). The appearance of *three* distinct energy levels out of just *two* values of the fundamental Lagrange multipliers confirms the interpretation that the true units of the system are the nodes and not the node pairs: it would indeed be impossible for our system to exhibit exactly two energy states, or in general to engineer an arbitrary number of energy states for the node pairs, since the only arbitrary values are those that can be attached to nodes, not to node pairs. Also note that all the three levels above are degenerate: the  $n_+(n_+ - 1)/2$  pairs of nodes in the core have the same expected link weight  $w_+^*(T)$  and energy  $\epsilon_+^*$ , the  $n_-(n_- - 1)/2$  pairs of nodes in the periphery have the same expected link weight  $w_-^*(T)$  and energy  $\epsilon_-^*$ , and the  $n_+n_-$  pairs of nodes across core and periphery have the same expected link weight  $w_0^*(T)$  and energy  $\epsilon_0^*$ . Therefore the ground state has energy  $\epsilon_{\min}^* = \epsilon_+^*$  and degeneracy  $n_+(n_+ - 1)/2$ . These degeneracies are dictated by the numbers of nodes in the two sets and cannot be assigned arbitrarily.

The occupation number of the ground state (with energy  $\epsilon_+^*$ ) coincides with the expected weight of all links between core nodes (total ‘core-core’ weight):

$$W_+^*(T) = \frac{n_+(n_+ - 1)}{2} w_+^*(T). \quad (58)$$

Similarly, the occupation number of the first excited state (with energy  $\epsilon_0$ ) coincides with the expected weight of all links between nodes across core and periphery (total ‘core-periphery’ weight):

$$W_0^*(T) = n_+n_-w_0^*(T). \quad (59)$$

Finally, the occupation number of the second excited state (with energy  $\epsilon_-$ ) coincides with the expected weight of all links between periphery nodes (total ‘periphery-periphery’ weight):

$$W_-^*(T) = \frac{n_-(n_- - 1)}{2} w_-^*(T). \quad (60)$$

By writing  $W^*$  as the sum of its core-core, core-periphery and periphery-periphery components, we get

$$\begin{aligned} W^* &= W_+^*(T) + W_0^*(T) + W_-^*(T) \\ &= \frac{n_+(n_+ - 1)}{2} w_+^*(T) + n_+n_-w_0^*(T) \\ &\quad + \frac{n_-(n_- - 1)}{2} w_-^*(T). \end{aligned} \quad (61)$$

Using Eq. (41), the total weight can also be expressed as

$$W^* = \frac{n_+s_+^*(T) + n_-s_-^*(T)}{2} \quad (62)$$

which, through Eqs. (46) and (47), indeed reduces to Eq. (61).

We stress again that the chemical potential  $\mu^*(T)$  appearing in Eqs. (55), (56) and (57) plays the role of a global Lagrange multiplier ensuring that, for all values of  $T$ , the total expected weight is  $W^*$ . Note that the  $T$ -independence of  $W^*$  allows us to conclude immediately that its value should be of order

$$W^* = O(n^2) \quad (63)$$

because, in particular, in the non-condensed phase all the  $n(n - 1)/2$  individual link weights  $w_{\pm}^*$ ,  $w_0^*$  must be finite by definition. As we have anticipated, this result ensures that in the thermodynamic limit ( $n \rightarrow \infty$ ) we automatically have  $W^* \rightarrow \infty$ , so that we can study ensemble (non)equivalence and BEC simultaneously, thereby ‘physically’ connecting two otherwise mathematically unrelated limits. We also note that, irrespective of temperature, the network is always in the dense regime. We can therefore introduce the average expected link weight

$$w^* \equiv \frac{2W^*}{n(n - 1)} = O(1), \quad (64)$$

which is a  $T$ -independent, finite parameter of our model, controlling the overall link weight in the network. Clearly,

$$w_-^*(T) \leq w^* \leq w_+^*(T) \quad \forall T. \quad (65)$$

It is good to remark again that, in our ‘inverse’ problem (construction of the conjugate canonical and micro-canonical ensembles), the parameters of the model are the values of the constraints, which here reduce to the two (diverging when  $n \rightarrow \infty$ ) numbers  $s_{\pm}^*(T)$ . However, to allow consistent comparisons for different temperatures, not all strength sequences are allowed, but only those that can be obtained from one another by varying  $T$ . In particular the values  $s_{\pm}^*(T)$  have to be specified for each value of  $T$  and be such that the total weight is always  $W^*$ . Indeed Eq. (62) shows that only two of the three quantities  $s_{\pm}^*(T)$ ,  $W^*$  are independent. By contrast, the traditional ‘direct’ problem in physics sees the three energies  $\epsilon_{\pm}^*$  and  $\epsilon_0^*$  (which do not depend on  $T$ ) as the parameters of the model, plus either  $w^*$  or the chemical potential  $\mu^*(T)$ . However, Eq. (57) shows that  $\epsilon_0^* = (\epsilon_+^* + \epsilon_-^*)/2$ , indicating only two independent values of the energy (say  $\epsilon_{\pm}^*$ ), as a result of the fact that there are only two types of nodes. Moreover, we may set the minimum energy  $\epsilon_+^* \equiv 0$  without loss of generality, because any overall energy shift can be reabsorbed into the chemical potential. We can therefore rename the only remaining independent value of the energy as  $\epsilon^* \equiv \epsilon^* > 0$ , and similarly  $\epsilon_0^* = \epsilon^*/2 > 0$ . Using these replacements into Eqs. (55) and (56), and combining the two equations, we get

$$\mu^*(T) = -2kT\beta_+^*(T) = \epsilon^* - 2kT\beta_-^*(T) \quad (66)$$

which is a convenient expression for solving the ‘direct’ problem. Rearranging, we obtain

$$\epsilon^* = 2kT [\beta_-^*(T) - \beta_+^*(T)]. \quad (67)$$

and, using Eqs. (51) and (53),

$$\begin{aligned} e^{-\epsilon^*/kT} &= e^{2[\beta_+^*(T) - \beta_-^*(T)]} \\ &= \frac{p_-^*(T)}{p_+^*(T)} \\ &= \frac{w_-^*(T)[1 + w_+^*(T)]}{w_+^*(T)[1 + w_-^*(T)]}, \end{aligned} \quad (68)$$

which shows how the energy difference  $\epsilon^*$  between periphery-periphery (+) and core-core (−) states is related to the corresponding connection probabilities  $p_{\pm}^*(T)$  and expected link weights  $w_{\pm}^*(T)$ . Therefore the most compact way of parametrizing the direct problem is by specifying only the two finite, positive and  $T$ -independent numbers  $\epsilon^*$  and  $w^*$ , and explore the resulting network properties by finding  $\mu(T)$  (as a function of  $\epsilon^*$ ,  $w^*$  and  $T$ ) and varying  $T$  as a control parameter. This will indeed allow us to easily explore the different (high- and low-temperature) phases consistently.

In our model, BEC occurs below a critical temperature  $T_c$  such that a finite fraction of the total weight  $W^*$  condenses in the core, which remains of finite size (i.e. of a finite number  $n_+$  of nodes) even when the size of the whole network diverges. This corresponds to requiring that, as  $n \rightarrow \infty$ ,  $n_+$  remains finite as dictated by Eq. (43),  $W^*$  diverges, and  $W_+^*(T)$  takes up a finite fraction of  $W^*$ . Rigorously, we can define this fraction (for finite  $n$ ) as

$$Q_n^*(T) \equiv \frac{W_+^*(T)}{W^*} \quad (69)$$

and use it to introduce the *order parameter* as

$$Q^*(T) \equiv \lim_{n \rightarrow \infty} Q_n^*(T) = \lim_{n \rightarrow \infty} \frac{W_+^*(T)}{W^*}. \quad (70)$$

We can then define the BEC phase as a phase emerging below a certain critical temperature  $T_c$  such that

$$Q^*(T < T_c) > 0. \quad (71)$$

By contrast, the non-BEC phase is such that

$$Q^*(T > T_c) = 0. \quad (72)$$

A visual anticipation of the qualitative behaviour that our system will exhibit is provided in Fig. 3.

In conjunction with BEC, we will investigate ensemble (non)equivalence. Therefore, in each phase of the model we will consider the relative entropy between the microcanonical and canonical ensembles and the relative fluctuations of the constraints. The criterion for measure equivalence is based on the relative entropy in Eq. (33), and useful techniques for the calculation of the determinant of the covariance matrix  $\Sigma^*$  in each phase are provided in the Appendix. Clearly, from Eqs. (23) and (45) we see that the diagonal entries  $\Sigma_{ii}^*$  of  $\Sigma^*$  take two possible values:

$$\Sigma_{ii}^*(T) = \sigma_{\beta^*(T)}^2(s_i) = \begin{cases} \Sigma_+^*(T) & i = 1, n_+ \\ \Sigma_-^*(T) & i = n_+ + 1, n \end{cases} \quad (73)$$

where

$$\begin{aligned} \Sigma_{\pm}^*(T) &= \frac{(n_{\pm} - 1)e^{-2\beta_{\pm}^*(T)}}{\left[1 - e^{-2\beta_{\pm}^*(T)}\right]^2} + \frac{n_{\mp}e^{-\beta_{\mp}^*(T) - \beta_{\pm}^*(T)}}{\left[1 - e^{-\beta_{\mp}^*(T) - \beta_{\pm}^*(T)}\right]^2} \\ &= (n_{\pm} - 1)w_{\pm}^*(T) [1 + w_{\pm}^*(T)] \\ &\quad + n_{\mp}w_0^*(T) [1 + w_0^*(T)]. \end{aligned} \quad (74)$$

Recalling Eq. (7), we remark that the canonical entropy  $S_{\text{can}}^*(T)$  can be easily calculated from Eq. (17) as

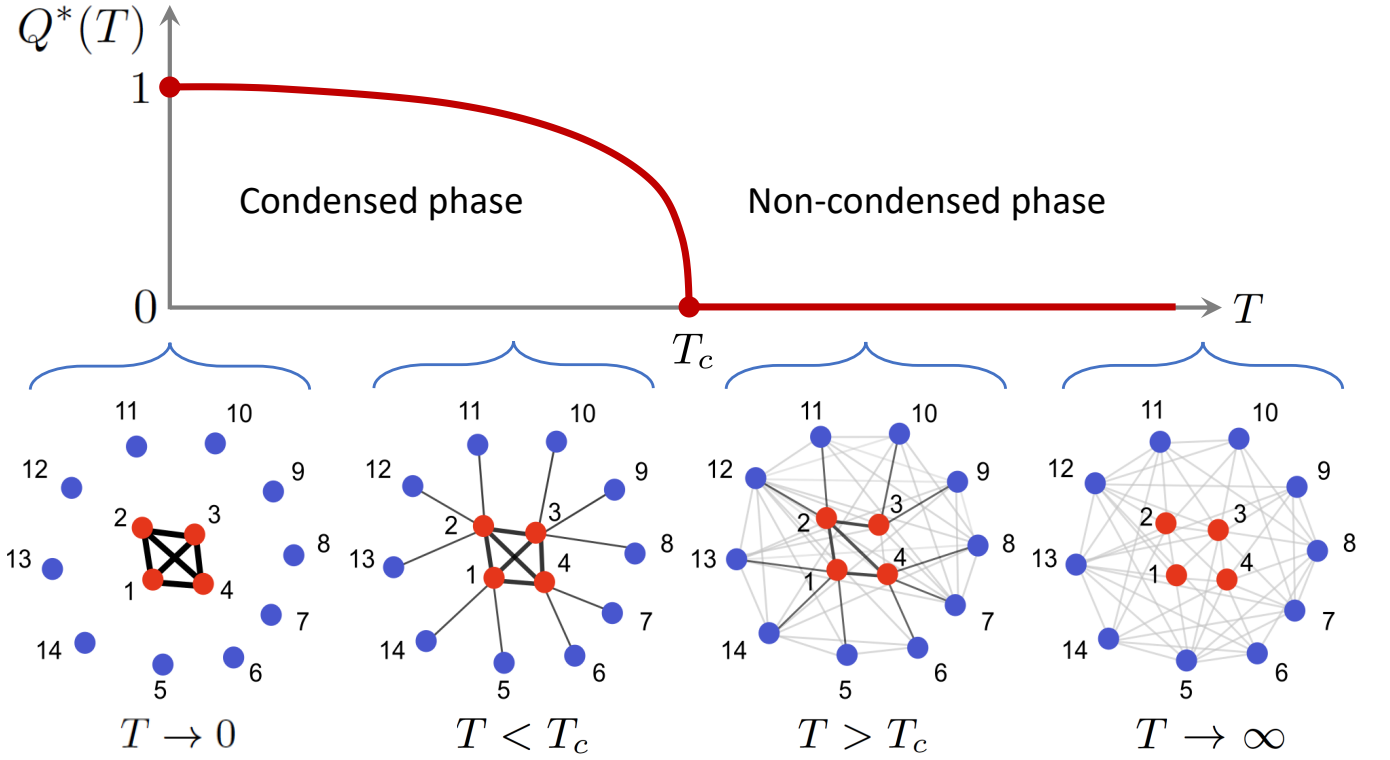


FIG. 3. Illustration of possible realizations of a network as a function of temperature, from higher (right) to lower (left) values of  $T$ . As a schematic example, a network with  $n = 14$  nodes, of which  $n_+$  are core nodes and  $n_- = 10$  are peripheral nodes, is considered. The order parameter  $Q^*(T)$  is zero for temperatures above the critical temperature  $T_c$ , while it is positive for temperatures below  $T_c$ , increasing towards 1 at zero temperature. At infinite temperature ( $T \rightarrow \infty$ ), there is no distinction between core and periphery: all links have the same probability of existing and the same expected weight. At lower but supercritical temperature ( $T > T_c$ ), a quantitative (but not yet qualitative) distinction between core and periphery appears: core-core links have higher probability and expected weight than core-periphery links, which in turn have higher probability and expected weight than periphery-periphery links, however all these probabilities and expected link weights are of the same (finite) order  $O(1)$ . Below a certain critical temperature ( $T < T_c$ ), the distinction between core and periphery becomes qualitative and more dramatic (the core forms a ‘condensate’): the expected link weights are of order  $O(n^2)$  for core-core links (with the corresponding connection probabilities approaching one) and  $O(1)$  for core-periphery and periphery-periphery links. Finally, at zero temperature ( $T \rightarrow 0$ ) the condensate decouples from the rest: peripheral nodes are completely disconnected and all links end up in the core, with an expected weight still of order  $O(n^2)$ .

the following sum of five terms:

$$\begin{aligned}
 S_{\text{can}}^*(T) &= \vec{\beta}^*(T) \cdot \vec{s}^*(T) + \ln Z[\vec{\beta}^*(T)] \\
 &= n_+ \beta_+^*(T) s_+^*(T) \\
 &\quad + n_- \beta_-^*(T) s_-^*(T) \\
 &\quad + \frac{n_+(n_+ - 1)}{2} \ln \frac{1}{1 - e^{-2\beta_+^*(T)}} \\
 &\quad + \frac{n_-(n_- - 1)}{2} \ln \frac{1}{1 - e^{-2\beta_-^*(T)}} \\
 &\quad + n_+ n_- \ln \frac{1}{1 - e^{-\beta_+^*(T) - \beta_-^*(T)}}, \quad (75)
 \end{aligned}$$

while the microcanonical entropy  $S_{\text{mic}}^*(T)$  is in general hard to compute, as it requires an explicit enumeration. However, its leading order can be obtained combining Eqs. (25) and (75).

The relative fluctuations of the constraints take the

form

$$r_i^*(T) = \begin{cases} r_+^*(T) & i = 1, n_+ \\ r_-^*(T) & i = n_+ + 1, n \end{cases} \quad (76)$$

where, from Eq. (37),

$$\begin{aligned}
 r_{\pm}^*(T) &= \frac{\sqrt{\Sigma_{\pm}^*(T)}}{s_{\pm}^*(T)} \quad (77) \\
 &= \sqrt{\frac{1}{s_{\pm}^*(T)} + \frac{(n_{\pm} - 1)[w_{\pm}^*(T)]^2 + n_{\mp}[w_0^*(T)]^2}{[s_{\pm}^*(T)]^2}} \\
 &= \sqrt{\frac{1}{s_{\pm}^*(T)} + \frac{(n_{\pm} - 1)[w_{\pm}^*(T)]^2 + n_{\mp}[w_0^*(T)]^2}{[(n_{\pm} - 1)w_{\pm}^*(T) + n_{\mp}w_0^*(T)]^2}}.
 \end{aligned}$$

Therefore in the thermodynamic limit the relative fluctuations of the constraints, as defined in Eq. (38), take

only the two possible limiting values

$$\rho_i^*(T) = \begin{cases} \rho_+^*(T) & i = 1, n_+ \\ \rho_-^*(T) & i = n_+ + 1, n \end{cases} \quad (78)$$

where

$$\rho_{\pm}^*(T) = \lim_{n \rightarrow \infty} r_{\pm}^*(T). \quad (79)$$

Armed with the above general results, we can now study each phase in detail.

### C. Non-condensed phase

Let us start from the non-BEC phase ( $T > T_c$ ). We first consider the finite-temperature case ( $T_c < T < \infty$ ) and then the infinite-temperature limit ( $T \rightarrow \infty$ ). As all the interesting phenomenology (in terms of both BEC and EN) occurs in the thermodynamic limit  $n \rightarrow \infty$ , we look for the asymptotic behaviour of all quantities in that limit.

#### 1. Finite (supercritical) temperature: $T_c < T < \infty$

Since, by definition, when  $T > T_c$  there is no concentration of ‘particles’ of weight on any of the links, all the expected link weights must be separately finite, i.e.  $w_+(T > T_c)$ ,  $w_-(T > T_c)$  and  $w_0^*(T > T_c)$  are all  $O(1)$ . Consequently, from Eqs. (53) and (54) we see that all the connection probabilities  $p_+^*(T > T_c)$ ,  $p_-^*(T > T_c)$  and  $p_0^*(T > T_c)$  are strictly smaller than one, i.e. missing links can occur anywhere in the network. Using this fact into Eqs. (46), (47), (58), (59) and (60), and using Eqs. (43) and (44), we immediately get the strength of nodes in the core, i.e.

$$\begin{aligned} s_+^*(T > T_c) &= (n_+ - 1)w_+^*(T > T_c) + n_-w_0^*(T > T_c) \\ &\approx n w_0^*(T > T_c) \\ &= O(n), \end{aligned} \quad (80)$$

and in the periphery, i.e.

$$\begin{aligned} s_-^*(T > T_c) &= (n_- - 1)w_-^*(T > T_c) + n_+w_0^*(T > T_c) \\ &\approx n w_-^*(T > T_c) \\ &= O(n). \end{aligned} \quad (81)$$

Similarly, for  $W_{\pm}^*(T > T_c)$ ,  $W_0^*(T > T_c)$  we get

$$W_+^*(T > T_c) = n_+(n_+ - 1)w_+^*(T > T_c)/2 = O(1), \quad (82)$$

$$W_-^*(T > T_c) \approx n^2 w_-^*(T > T_c)/2 = O(n^2), \quad (83)$$

$$W_0^*(T > T_c) \approx n_+ n w_0^*(T > T_c) = O(n), \quad (84)$$

from which we see that in this phase the total weight  $W^*$  is essentially all in the periphery, i.e.

$$W_-^*(T > T_c) = W^* - o(n^2) \approx W^*, \quad (85)$$

$$w_-^*(T > T_c) = w^* - o(1) \approx w^*. \quad (86)$$

We stress that the above result does not mean that the core is empty or that there are no connections between core and periphery. Rather, it indicates that the total weight of all core-core and core-periphery connections is asymptotically negligible with respect to the total weight located inside the periphery, simply because the number of periphery-periphery node pairs dominates the number of core-periphery and core-core pairs. In particular, the finite parameter  $w_+^*(T > T_c)$  can take an arbitrarily large value, without ‘moving’ the (finite and positive) value of the average link weight  $w^*$ . All positive values of  $w_+^*(T > T_c)$  are therefore allowed. By contrast,  $w_-^*(T > T_c)$  is forced to take (to leading order) only the value  $w^*$ .

Using Eqs. (81), (83) and (84), we write the order parameter as

$$\begin{aligned} Q^*(T > T_c) &\equiv \lim_{n \rightarrow \infty} \frac{W_+^*(T > T_c)}{W^*} \\ &= \lim_{n \rightarrow \infty} \frac{W^* - W_-^*(T > T_c) - W_0^*(T > T_c)}{W^*} \\ &= 1 - \lim_{n \rightarrow \infty} \frac{n^2 w_-^*(T > T_c)}{2W^*} \\ &= 1 - \lim_{n \rightarrow \infty} \frac{n s_-^*(T > T_c)}{2W^*}, \\ &= 0, \end{aligned} \quad (87)$$

confirming the definition of non-condensed phase in Eq. (72) and showing that, since both  $s_-^*(T > T_c)$  and  $W^*$  have by construction the same value in the canonical and microcanonical ensemble, the order parameter must be zero in both ensembles, for all values of  $T > T_c$ . Therefore, whenever one ensemble is in the non-condensed phase, the other ensemble is the non-condensed phase as well. Importantly, this allows us to refer to the conjugate canonical and microcanonical ensembles ‘in the non-condensed phase’ consistently.

To solve the ‘inverse’ problem, we use Eqs. (51) and (52) and invert Eqs. (80) and (81) to get

$$\begin{aligned} \beta_-^*(T > T_c) &= \frac{1}{2} \ln \left( 1 + \frac{1}{w_-^*(T > T_c)} \right) \\ &\approx \frac{1}{2} \ln \left( 1 + \frac{n}{s_-^*(T > T_c)} \right) \end{aligned} \quad (88)$$

and

$$\begin{aligned} \beta_+^*(T > T_c) + \beta_-^*(T > T_c) &= \ln \left( 1 + \frac{1}{w_0^*(T > T_c)} \right) \\ &\approx \ln \left( 1 + \frac{n}{s_+^*(T > T_c)} \right). \end{aligned} \quad (89)$$

Then, subtracting Eq. (88) from Eq. (89), we get

$$\begin{aligned} \beta_+^*(T > T_c) &= \frac{1}{2} \ln \left( 1 + \frac{1}{w_+^*(T > T_c)} \right) \\ &\approx \ln \frac{1 + \frac{n}{s_+^*(T > T_c)}}{\sqrt{1 + \frac{n}{s_-^*(T > T_c)}}}. \end{aligned} \quad (90)$$



Equations (88) and (90) express  $\beta_{\pm}^*(T > T_c)$  as a function of the two (diverging) constraints  $s_{\pm}^*(T > T_c)$ , or equivalently as a function of the finite parameters  $w_{\pm}^*(T > T_c)$ , which have to be specified for all values of  $T$ .

To solve the ‘direct’ problem, we first note an important consequence of Eq. (86): in the large  $n$  limit,  $w^*(T > T_c)$  and  $\beta^*(T > T_c)$  are independent of temperature. Indeed, using Eq. (81) and (86) into Eq. (88), we get asymptotically

$$\beta_{-}^*(T > T_c) \approx \frac{1}{2} \ln \left( 1 + \frac{1}{w^*} \right) \quad (91)$$

and, using Eq. (67),

$$\beta_{+}^*(T > T_c) \approx \frac{1}{2} \ln \left( 1 + \frac{1}{w^*} \right) - \frac{\epsilon^*}{2kT}. \quad (92)$$

When inserted into Eq. (66), the above expressions allow us to directly obtain the chemical potential as

$$\begin{aligned} \mu(T > T_c) &= \epsilon^* - 2kT\beta_{-}^*(T > T_c) \\ &\approx \epsilon^* - kT \ln \left( 1 + \frac{1}{w^*} \right) + o(T). \end{aligned} \quad (93)$$

As anticipated, the above result provides the solution to the direct problem in terms of the two finite constants  $\epsilon^*$  and  $w^*$ , and allows us to explore the model by varying  $T$  throughout the non-condensed phase  $T > T_c$ .

We now consider ensemble (non)equivalence. Inserting Eqs. (80) and (81) into Eq. (74), we obtain the variance of the strength of nodes in the core, i.e.

$$\begin{aligned} \Sigma_{+}^*(T > T_c) &\approx n w_0^*(T > T_c) [1 + w_0^*(T > T_c)] \\ &\approx s_{+}^*(T > T_c) \left[ 1 + \frac{s_{+}^*(T > T_c)}{n} \right] \\ &= O(n), \end{aligned} \quad (94)$$

and in the periphery, i.e.

$$\begin{aligned} \Sigma_{-}^*(T > T_c) &\approx n w_{-}^*(T > T_c) [1 + w_{-}^*(T > T_c)] \\ &\approx s_{-}^*(T > T_c) \left[ 1 + \frac{s_{-}^*(T > T_c)}{n} \right] \\ &= O(n). \end{aligned} \quad (95)$$

As we show in Appendix A 1, it is possible to show that the leading term of the determinant of the covariance matrix  $\Sigma^*(T > T_c)$  in this non-condensed phase is

$$\det[\Sigma^*(T > T_c)] = \prod_{i=1}^n \Sigma_{ii}^*(T > T_c) + O(n^{n-2}). \quad (96)$$

Using Eqs. (74), (94) and (95) we obtain

$$\begin{aligned} \prod_{i=1}^n \Sigma_{ii}^*(T > T_c) &= [\Sigma_{+}^*(T > T_c)]^{n+} [\Sigma_{-}^*(T > T_c)]^{n-} \\ &= O(n^n). \end{aligned} \quad (97)$$

Inserting this result into Eq. (96), we find

$$\det[\Sigma^*(T > T_c)] = O(n^n) + O(n^{n-2}) = O(n^n), \quad (98)$$

showing that the determinant is dominated by the diagonal entries of  $\Sigma^*(T > T_c)$ . Taking the logarithm, we obtain

$$\ln \det \Sigma^*(T > T_c) = O(n \ln n) \quad (99)$$

which, when compared with Eq. (34), shows that the system is under ensemble nonequivalence. We note that the  $O(n \ln n)$  scaling of  $\ln \det \Sigma^*(T > T_c)$  ensures that Eq. (35) holds, so the leading order of the relative entropy can be calculated exactly as

$$D^*(T > T_c) \approx \frac{1}{2} \ln \det \Sigma^*(T > T_c) \approx \frac{1}{2} n \ln n \quad (100)$$

where we have used Eqs. (94) and (95) into Eq. (96). The above result is in line with the scaling of the relative entropy found in the case of binary networks with a constraint on the node degrees in the dense regime [5, 19, 20]. Another similarity between the two models is the order of the canonical entropy:

$$S_{\text{can}}^*(T > T_c) = O(n^2), \quad (101)$$

which can be easily seen from Eq. (75) using  $\beta_{\pm}^*(T > T_c) = O(1)$  and  $s_{\pm}^*(T > T_c) = O(n)$ , as found in Eqs. (80), (81), (88) and (90). Then Eq. (25) also implies

$$S_{\text{mic}}^*(T > T_c) = O(n^2). \quad (102)$$

Note that, even if the relative entropy is subleading with respect to the canonical and microcanonical entropies, it is still superextensive in the number  $n$  of units of the system, thereby breaking ensemble equivalence as in binary networks with fixed degrees. Therefore the result in Eq. (99) is another observation, for the first time in weighted networks, of the fact that ensemble equivalence can be broken by the presence of an extensive number of local constraints, even away from phase transitions.

Coming to the relative fluctuations of the constraints, we see from Eqs. (77) and (79) that

$$\begin{aligned} r_{\pm}^*(T > T_c) &= \frac{\sqrt{\Sigma_{\pm}^*(T > T_c)}}{s_{\pm}^*(T > T_c)} \\ &\approx \sqrt{\frac{n + s_{\pm}^*(T > T_c)}{n s_{\pm}^*(T > T_c)}} \\ &= O\left(\frac{1}{\sqrt{n}}\right) \end{aligned} \quad (103)$$

and

$$\rho_{\pm}^*(T > T_c) = 0. \quad (104)$$

We therefore observe that in the non-condensed phase the decay of the relative fluctuations of each constraint is of

the same order  $O(1/\sqrt{n})$  as generally observed for the global constraint (total energy) in a system with short-range interactions away from phase transitions. However, while in the traditional situation the vanishing of the relative fluctuations implies that the relative entropy is subextensive and that the relative entropy density vanishes in the thermodynamic limit (as discussed in Sec. I C), here the relative entropy density does not vanish and the ensembles are not equivalent. Therefore we find that, *in systems with an extensive number of local constraints, the vanishing of even all the relative fluctuations does not ensure ensemble equivalence.*

## 2. Infinite temperature: $T \rightarrow \infty$

The extreme regime of the non-condensed phase is the infinite-temperature case, which can be explored by taking the limit  $T \rightarrow \infty$  in the solution to the ‘direct’ problem provided by Eq. (93). In such a limit, Eq. (39) implies that  $\beta_+^*(T > T_c)$  and  $\beta_-^*(T > T_c)$  converge to the same value  $\beta_\infty^*$  given by

$$\beta_+^*(T \rightarrow \infty) = \beta_-^*(T \rightarrow \infty) = \beta_\infty^* \equiv \frac{1}{2} \ln \left( 1 + \frac{1}{w^*} \right). \quad (105)$$

Then, through Eqs. (48), (49), (50), (53) and (54), we get

$$w_+^*(T \rightarrow \infty) = w_-^*(T \rightarrow \infty) = w_0^*(T \rightarrow \infty) = w^*, \quad (106)$$

$$p_+^*(T \rightarrow \infty) = p_-^*(T \rightarrow \infty) = p_0^*(T \rightarrow \infty) = p^*, \quad (107)$$

i.e. all node pairs have the same expected link weight  $w^*$  and connection probability  $p^*$  given by

$$p^* = \frac{w^*}{1 + w^*}. \quad (108)$$

This is the characteristic situation in the infinite-temperature limit of Bose-Einstein statistics, where each particle is equally likely distributed across all energy levels. Here, this situation translates in the graph becoming completely homogeneous: the distinction between core and periphery disappears as the finite difference between energy levels becomes entirely dominated by the diverging temperature. The expected strength of every node has the same value  $s^* \equiv (n-1)w^*$ :

$$s_+^*(T \rightarrow \infty) = s_-^*(T \rightarrow \infty) = s^* = (n-1)w^*, \quad (109)$$

i.e. the strength sequence becomes a constant vector.

Clearly, the above result does not change the value of the order parameter in Eq. (87):

$$Q^*(T \rightarrow \infty) = 0. \quad (110)$$

Similarly, the final results in eqs. (99) and (104) about the simultaneous breakdown of ensemble equivalence and the vanishing of all relative fluctuations carry over to the infinite-temperature limit, so in principle we do not have

to further discuss this case. However, the fact that the strength sequence becomes a constant vector allows us to calculate many of the properties of the model exactly, so this example is a very transparent and instructive one. It is therefore worth considering it in some more detail, also because some of the following results will be useful in the (much less trivial) zero-temperature limit as well.

In particular, Eqs. (105) and (109) imply that Eqs. (23) and (24) can be rewritten as

$$\Sigma_{ii}^*(T \rightarrow \infty) = \frac{(n-1)e^{-2\beta_\infty^*}}{(1 - e^{-2\beta_\infty^*})^2} = w^*(1+w^*)(n-1) \quad (111)$$

for all  $i = 1, n$  and

$$\Sigma_{ij}^*(T \rightarrow \infty) = \frac{e^{-2\beta_\infty^*}}{(1 - e^{-2\beta_\infty^*})^2} = w^*(1+w^*) \quad (112)$$

for all  $i \neq j$  respectively. In Appendix A 2 we show that the above expressions can be used to calculate the determinant of  $\Sigma^*(T \rightarrow \infty)$  exactly as

$$\det \Sigma^*(T \rightarrow \infty) = 2(n-1)(n-2)^{n-1} [w^*(1+w^*)]^n, \quad (113)$$

from which we confirm, without having made the approximation in eq. (96), that

$$\ln \det \Sigma^*(T \rightarrow \infty) = O(n \ln n) \quad (114)$$

and that

$$D^*(T \rightarrow \infty) \approx \frac{1}{2} \ln \det \Sigma^*(T \rightarrow \infty) \approx \frac{1}{2} n \ln n. \quad (115)$$

Clearly, Eqs. (101) and (102) hold in this limit as well:

$$S_{\text{can}}^*(T \rightarrow \infty) = O(n^2), \quad S_{\text{mic}}^*(T \rightarrow \infty) = O(n^2). \quad (116)$$

Finally, from eqs. (109) and (111) we see that eq. (77) leads in this case to a unique value for the coefficient of variation of all the strengths:

$$r_i^*(T \rightarrow \infty) = \frac{\sqrt{\Sigma_{ii}^*(T \rightarrow \infty)}}{s_i^*(T \rightarrow \infty)} = \sqrt{\frac{1+w^*}{(n-1)w^*}} \quad \forall i, \quad (117)$$

so that

$$\rho_i^*(T \rightarrow \infty) = 0 \quad \forall i, \quad (118)$$

in accordance with eq. (104).

## D. Condensed phase

We now consider the BEC phase ( $T < T_c$ ). We first derive general results and then discuss the finite-temperature case and the zero-temperature limit separately.

By the definition in Eq. (71), the condensed phase must be such that a positive fraction  $Q^*(T < T_c) > 0$  of the total weight lies in the core, i.e. (to leading order)

$$W_+^*(T < T_c) \approx Q_n^*(T < T_c) W^* \approx Q^*(T < T_c) W^* \quad (119)$$

which necessarily means

$$w_+^*(T < T_c) = O(n^2), \quad W_+^*(T < T_c) = O(n^2) \quad (120)$$

and  $p_+^*(T < T_c) \approx 1$ , i.e. the core does not have missing links (the presence of core-core links is no longer a random event, while the weight of such links is still a random variable). As expected, BEC corresponds to the divergence of  $w_+^*(T < T_c)$ , and we now see that the speed of this divergence is of order  $n^2$  in our model. For convenience, we may define

$$\psi_+^*(T < T_c) = \lim_{n \rightarrow \infty} \frac{w_+^*(T < T_c)}{n^2} \quad (121)$$

which is finite and positive, so that

$$w_+^*(T < T_c) \approx \psi_+^*(T < T_c) n^2. \quad (122)$$

Combining Eqs. (50) and (120) we see that

$$w_0^*(T < T_c) \approx \frac{1}{\sqrt{1 + 1/w_-^*(T < T_c)} - 1}, \quad (123)$$

which inserted into Eq. (61) shows that, to leading order,

$$w^* \approx n_+(n_+ - 1)\psi_+^*(T < T_c) + w_-^*(T < T_c), \quad (124)$$

an expression that relates the finite parameters of the model with each other in the condensed phase. Therefore we see from Eq. (119) that

$$w_-^*(T < T_c) \approx [1 - Q_n^*(T < T_c)]w^* \quad (125)$$

and

$$W_-^*(T < T_c) \approx [1 - Q_n^*(T < T_c)]W^*. \quad (126)$$

Inserting Eq. (125) into Eq. (123) yields

$$w_0^*(T < T_c) \approx \frac{1}{\sqrt{1 + \frac{1}{[1 - Q_n^*(T < T_c)]w^*}} - 1} \quad (127)$$

and

$$W_0^*(T < T_c) \approx \frac{n_+n_-}{\sqrt{1 + \frac{1}{[1 - Q_n^*(T < T_c)]w^*}} - 1}. \quad (128)$$

The above expressions show that neither  $w_-^*(T < T_c)$  nor  $w_0^*(T < T_c)$  diverge, indicating that BEC occurs only in the ground state and that  $p_-^*(T < T_c) < 1$  and  $p_0^*(T < T_c) < 1$ , i.e. there can be missing links in the periphery and between core and periphery. Moreover, we see that  $W_0^*(T < T_c)$  is subleading with respect to both  $W_+^*(T < T_c)$  and  $W_-^*(T < T_c)$ : although individual core-periphery links have an expected weight  $w_0^*(T < T_c)$  larger than the expected weight  $w_-^*(T < T_c)$  of individual periphery-periphery links, the number  $n_+n_-$  of core-periphery links is of smaller order with respect to the number  $n_-(n_- - 1)/2$  of periphery-periphery links, and as a result the total weight of all core-periphery links is of smaller order as well.

To obtain  $s_{\pm}^*(T < T_c)$  to leading order, we use Eqs. (46) and (47):

$$\begin{aligned} s_+^*(T < T_c) &= (n_+ - 1)w_+^*(T < T_c) + n_-w_0^*(T < T_c) \\ &\approx (n_+ - 1)\psi_+^*(T < T_c)n^2, \end{aligned} \quad (129)$$

$$\begin{aligned} s_-^*(T < T_c) &= (n_- - 1)w_-^*(T < T_c) + n_+w_0^*(T < T_c) \\ &\approx n w^*[1 - Q_n^*(T < T_c)]. \end{aligned} \quad (130)$$

Now, combining the above expressions, we see that the order parameter defined in Eq. (70) can be written as

$$\begin{aligned} Q^*(T < T_c) &= \lim_{n \rightarrow \infty} \frac{W_+^*(T < T_c)}{W_+^*(T < T_c) + W_-^*(T < T_c)} \\ &= \lim_{n \rightarrow \infty} \frac{1}{1 + \frac{n_-(n_- - 1)w_-^*(T < T_c)}{n_+(n_+ - 1)w_+^*(T < T_c)}} \\ &= \lim_{n \rightarrow \infty} \frac{1}{1 + \frac{n_-s_-^*(T < T_c)}{n_+s_+^*(T < T_c)}} \\ &= \frac{1}{1 + \frac{w_-^*(T < T_c)}{n_+(n_+ - 1)\psi_+^*(T < T_c)}} \\ &= \frac{n_+(n_+ - 1)\psi_+^*(T < T_c)}{w^*} > 0. \end{aligned} \quad (131)$$

Besides quantifying the order parameter, the above calculation shows that, since the value of  $Q^*(T < T_c)$  only depends on the values of  $s_+^*(T < T_c)$  and  $s_-^*(T < T_c)$  (which by construction are the same in the canonical and microcanonical ensembles), whenever one ensemble is in the BEC phase, the other ensemble is the BEC phase as well, for all temperatures  $T < T_c$ . As for the non-condensed, this allows us to refer to the conjugate canonical and microcanonical ensembles being ‘in the same phase’ consistently. Inverting Eq. (131), we can also express the parameter  $\psi_+^*(T < T_c)$  in terms of the order parameter as follows:

$$\psi_+^*(T < T_c) = \frac{w^* Q^*(T < T_c)}{n_+(n_+ - 1)}. \quad (132)$$

The ‘inverse’ problem is solved by inverting Eqs. (129) and (130) and using them into Eq. (51) to get

$$\begin{aligned} \beta_+^*(T < T_c) &\approx \frac{1}{2} \ln \left( 1 + \frac{1}{n^2 \psi_+^*(T < T_c)} \right) \\ &\approx \frac{1}{2n^2 \psi_+^*(T < T_c)} \\ &\approx \frac{n_+ - 1}{2s_+^*(T < T_c)} \\ &\approx \frac{n_+(n_+ - 1)}{2n^2 w^* Q_n^*(T < T_c)}, \end{aligned} \quad (133)$$

$$\begin{aligned} \beta_-^*(T < T_c) &= \frac{1}{2} \ln \left( 1 + \frac{1}{w_-^*(T < T_c)} \right) \\ &\approx \frac{1}{2} \ln \left( 1 + \frac{n}{s_-^*(T < T_c)} \right) \\ &\approx \frac{1}{2} \ln \left( 1 + \frac{1}{w^*[1 - Q_n^*(T < T_c)]} \right). \end{aligned} \quad (134)$$

The above equations solve the inverse problem by expressing  $\beta_{\pm}^*(T < T_c)$  as a function of the constraints  $s_{\pm}^*(T < T_c)$ , which can in turn be expressed either in terms of the finite parameters  $\psi_{\pm}^*(T < T_c)$  and  $w_{\pm}^*(T < T_c)$  or in terms of  $Q_n^*(T < T_c)$  and the temperature-independent parameter  $w^*$ .

Again, the ‘direct’ problem requires finding the chemical potential  $\mu^*(T)$  as a function of  $\epsilon^*$ ,  $w^*$  and  $T$ . From Eq. (66) we get

$$\begin{aligned} \mu^*(T < T_c) &= -2kT\beta_{\pm}^*(T < T_c) \\ &\approx -\frac{kT}{n^2\psi_{\pm}^*(T < T_c)} \\ &\approx -kT\frac{n_+(n_+ - 1)}{n^2w^*Q_n^*(T < T_c)}. \end{aligned} \quad (135)$$

We now consider the variance of the constraints. Inserting Eqs. (120), (125) and (127) into Eqs. (74), we immediately see that

$$\Sigma_{+}^*(T < T_c) \approx (n_+ - 1)[\psi_{+}^*(T < T_c)]^2 n^4, \quad (136)$$

$$\Sigma_{-}^*(T < T_c) \approx w_{-}^*(T < T_c)[1 + w_{-}^*(T < T_c)]n. \quad (137)$$

### 1. Finite (subcritical) temperature: $0 < T < T_c$

In this regime we have

$$0 < Q^*(0 < T < T_c) < 1, \quad (138)$$

which, as clear from Eqs. (125) and (127), implies

$$w_{-}^*(0 < T < T_c) \approx [1 - Q^*(0 < T < T_c)]w^*, \quad (139)$$

$$w_0^*(0 < T < T_c) \approx \frac{1}{\sqrt{1 + \frac{1}{[1 - Q^*(0 < T < T_c)]w^*} - 1}}}, \quad (140)$$

where both quantities are  $O(1)$ . Using these results, it is possible to show (see Appendix A3) that the leading term of the determinant of the covariance matrix between the constraints is

$$\det[\Sigma^*(0 < T < T_c)] = O(n^{n+3n_+}), \quad (141)$$

implying

$$\ln \det[\Sigma^*(0 < T < T_c)] = O(n \ln n), \quad (142)$$

which is the same scaling found in Eq. (99) for the non-condensed phase. The criterion for measure equivalence in Eq. (34) is again violated, showing that ensemble equivalence does not hold in the condensed case as well. The leading term of the relative entropy can still be calculated exactly from Eq. (35) and is the same as the one found in Eq. (100) for the non-condensed phase:

$$D^*(0 < T < T_c) \approx \frac{1}{2} \ln \det \Sigma^*(0 < T < T_c) \approx \frac{1}{2} n \ln n. \quad (143)$$

Similarly, the canonical entropy is still of order  $O(n^2)$ , as can be seen by inserting Eqs. (129), (130), (133) and (134) into Eq. (75). We therefore retrieve

$$S_{\text{can}}^*(0 < T < T_c) = O(n^2), \quad (144)$$

$$S_{\text{mic}}^*(0 < T < T_c) = O(n^2). \quad (145)$$

Coming to the relative fluctuations, from Eqs. (77), (79), (136) and (137) we obtain

$$\rho_{+}^*(0 < T < T_c) = \frac{1}{\sqrt{n_+ - 1}}, \quad (146)$$

$$\rho_{-}^*(0 < T < T_c) = 0. \quad (147)$$

The above result can be interpreted as follows. The term  $\sum_{i \neq j} (\langle w_{ij} \rangle_{\beta^*})^2 / (\sum_{i \neq j} \langle w_{ij} \rangle_{\beta^*})^2$  in Eq. (37) is an inverse participation ratio, taking values in the range  $[(n-1)^{-1}, 1]$  and quantifying the inverse of the effective number of link weights contributing to the strength of node  $i$  [27]. Here, for a node in the core, there is a finite number  $n_+ - 1$  of dominant link weights, each equal to  $w_{+}^* = O(n^2)$ , while the remaining  $n_-$  weights are of smaller order. Taking the thermodynamic limit, these  $n_+ - 1$  dominant weights lead to the value for  $\rho_{+}^*$  in Eq. (146). By contrast, for a node in the periphery, all the expected link weights are of the same order, so the inverse participation ratio, and consequently the value of  $\rho_{-}^*$  in Eq. (147), vanishes. It should be noted that, even if the expected strength of the core nodes is much bigger than that of the periphery nodes, the relative fluctuations of the core nodes do not vanish, while those of the peripheral nodes do.

The fact that BEC occurs necessarily among the core nodes confirms that the units of the system are the nodes, and not the node pairs: the ‘ground state pairs’ are necessarily all and only the pairs of ‘ground state nodes’. Indeed, one cannot decide arbitrarily which node pairs form the degenerate ground state where condensation occurs. This would have been possible only if node pairs were the fundamental units, by assigning the same degenerate ground state energy value to any set of node pairs (including pairs not necessarily involving the same set of nodes). For instance, it would have been possible to include the pairs  $(i, j)$  and  $(i, k)$ , without necessarily including the pair  $(j, k)$ , in the degenerate ground state (which is instead unavoidable in our system).

### 2. Zero temperature: $T \rightarrow 0$

We finally consider the zero-temperature limit as the extreme case of the condensed phase. Importantly, we have to be careful how we approach the two limits  $T \rightarrow 0$  and  $n \rightarrow \infty$ . Indeed, we are going to show that taking the limit  $T \rightarrow 0$  while  $n$  is kept fixed leads to results that cannot be subsequently carried over to the thermodynamic limit by taking the limit  $n \rightarrow \infty$ . Since we are interested precisely in the thermodynamic limit, we have

to take a different route. To show the difference, we consider the zero-temperature limit first in the case of finite  $n$  and then in the case of growing  $n$ .

If  $n$  is finite, the zero-temperature limit simply represents the situation where the only populated state is the degenerate ground state corresponding to the links in the core, i.e.

$$Q_n^*(T \rightarrow 0) = 1 \quad (148)$$

by construction. All other links are not present. As usual, if the ground state is not degenerate, then the microcanonical entropy is zero, while if the ground state is degenerate, then the microcanonical entropy approaches a value called *residual entropy* which, for a system of fixed size, is a constant that depends only on the degeneracy (these statements usually go under the names of *Third Law of Thermodynamics* or, somehow improperly, *Nernst Theorem*) [39]. In our setting the ground state is non-degenerate only if  $n_+ = 2$ , in which case the link between the two core nodes is the one with minimum energy. In the general case  $n_+ > 2$ , the ground state is degenerate and both the microcanonical and canonical entropies are strictly positive. In any case, for finite  $n$  the zero-temperature limit is characterized by the fact that, in both the canonical and microcanonical ensembles, all nodes in the periphery are *deterministically* isolated, i.e. necessarily isolated in all realizations of the network. The periphery becomes completely disconnected, both internally and from the core. Note that this is one of the degenerate situations (mentioned in Sec. IC) where, even if the constraints are in principle all mutually independent, for certain degenerate parameter value(s) some of them become ‘hard’ in both ensembles, thereby not contributing any difference between the two ensembles. Note that if we simply take the ideal limit  $n \rightarrow \infty$  starting from this zero-temperature state, we would be considering the degenerate situation where an infinite number of isolated ‘peripheral’ nodes are added to the fully connected core. These nodes are unavoidably disconnected in both ensembles, so their contribution to the system is purely formal. The only variability (hence the only possible source of nonequivalence) comes from the core, which keeps having a finite number  $n_+$  of nodes: as an extreme signature of BEC, the condensate behaves as an effectively lower-dimensional object.

In order to access the thermodynamic limit, we therefore have to consider from the beginning the case where  $n$  can grow indefinitely. We are going to show that the main difference arises from the fact that the temperature can only correspond to graphical strength sequences, which on the other hand depend on  $n$ . Therefore one should expect a certain  $n$ -dependent temperature  $T_n$ . At that point, a temperature value  $T_n > 0$  that is small but finite when  $n$  is finite (allowing for certain populated excited states besides the ground state) may actually approach zero as  $n$  diverges, i.e.  $\lim_{n \rightarrow \infty} T_n = 0$ . The corresponding excited states will effectively become part of the accessible configurations in the zero-temperature limit and

contribute an extra residual entropy.

To study this scenario, we start from the consideration that if the two limits  $T \rightarrow 0$  and  $n \rightarrow \infty$  were taken simultaneously for quantities that depend on both  $Q_n^*$  and  $n$ , e.g. terms such as  $n(1 - Q_n^*)$ , we would encounter indeterminate expressions. We therefore need to understand how, as  $T$  goes to zero,  $Q_n^*$  goes to one as a function of  $n$ , for  $n$  large. We recall that our starting point is always the value  $\bar{s}^*(T)$  of the constraints. The temperature  $T$  is a parameter that allows us to vary  $\bar{s}^*(T)$ , while keeping it graphic, i.e. realizable in at least one configuration of the network. For large  $n$ , we therefore have to identify the possible states of the network, hence the values of  $\bar{s}^*(T)$ , closest to zero temperature, i.e. when  $T \simeq 0$  (we will use the symbol ‘ $\simeq$ ’ to denote this near-zero-temperature behaviour of any quantity, thereby keeping the notation distinct from the symbol ‘ $\approx$ ’ that will still denote the leading order of any quantity for large  $n$ ). This is easily done by realizing that, if we start from some  $\bar{s}^*(T > 0)$  and decrease the temperature towards zero, the lowest excited state accessible to the network (before all links condense in the core) is one where only the smallest possible number  $\Delta W^*$  of the  $W^*$  particles of weight remain out of the core, while keeping the strength sequence  $\bar{s}^*(T \simeq 0)$  in the form given by Eq. (42). This state is necessarily such that  $s_-^*(T \simeq 0) = 1$  (which is the minimum non-zero value of the strength, recalling that the strength is a non-negative integer by construction) and can be realized in multiple ways: either by connecting the  $n_-$  peripheral nodes in pairs, thus creating  $n_-/2$  periphery-periphery links of unit weight and energy  $\epsilon^*$  (in which case  $\Delta W^* = n_-/2$ ), or by connecting each peripheral node to a core node, thus creating  $n_-$  core-periphery links of unit weight and energy  $\epsilon^*/2$  (in which case  $\Delta W^* = n_-$ ), or finally by combining both types of situations. Recalling the discussion in Sec. IV B, in all cases the  $\Delta W^*$  links outside of the core have collectively the same energy  $\epsilon^* n_-/2$  while the links in the core have zero energy; indeed, all such configurations are equiprobable. If we also consider the next excited states with  $s_-^*(T \simeq 0) = 2, 3, \dots$ , in general we will have  $\Delta W^*(T \simeq 0) = n_- \ell/2$  where  $\ell$  is a small (in a sense that will be clear in a moment) integer.

The above considerations imply that the possible values of  $Q_n^*$  close to zero temperature are of the form

$$\begin{aligned} Q_n^*(T \simeq 0) &= \frac{W^* - \Delta W^*(T \simeq 0)}{W^*} \\ &= 1 - \frac{n_- \ell/2}{W^*} \\ &\approx 1 - \frac{\ell}{n w^*}. \end{aligned} \quad (149)$$

Basically, the above expression makes it explicit that, since the strength is a discrete quantity, technically the temperature can only take discrete values in order to keep the strength sequence graphic, so the role of  $T$  is taken up by  $\ell$  (which is an integer) and a low temperature corresponds to a ‘small’, i.e. finite or at most  $o(n)$ , value of

$\ell$ . Indeed, in the thermodynamic limit we recover

$$Q^*(T \simeq 0) = \lim_{n \rightarrow \infty} Q_n^*(T \simeq 0) = 1, \quad (150)$$

confirming that  $\ell = o(n)$  leads to the correct zero-temperature limit. At the same time, Eq. (149) shows that, to recover any finite-temperature value  $Q^*(T > 0) < 1$ , we would need  $\ell$  to grow linearly in  $n$  in the thermodynamic limit (note that  $\ell$  cannot grow faster than  $n$ , because  $\Delta W^*$  cannot grow faster than  $n^2$ , which is the order of  $W^*$ ). The  $\ell = o(n)$  regime considered here is therefore genuinely different from the positive-temperature cases discussed so far.

Having characterized the zero-temperature limit in this way, we can calculate

$$\lim_{n \rightarrow \infty} n[1 - Q_n^*(T \simeq 0)] = \frac{\ell}{w^*}, \quad (151)$$

from which we can obtain various asymptotic expressions. Indeed, from Eqs. (125) and (126) we obtain

$$w_-^*(T \simeq 0) \approx [1 - Q_n^*(T \simeq 0)]w^* \approx \frac{\ell}{n} \quad (152)$$

and

$$W_-^*(T \simeq 0) \approx \frac{\ell n}{2}. \quad (153)$$

Similarly, expanding Eqs. (127) and (128) for  $Q_n^*$  close to 1, we get

$$w_0^*(T \simeq 0) \approx \sqrt{[1 - Q_n^*(T \simeq 0)]w^*} \approx \sqrt{\frac{\ell}{n}} \quad (154)$$

and

$$W_0^*(T \simeq 0) \approx n_+ \sqrt{\ell n}. \quad (155)$$

Since both  $W_-^*(T \simeq 0)$  and  $W_0^*(T \simeq 0)$  are subleading with respect to  $W_+^*(T \simeq 0)$ , we have  $W_+^*(T \simeq 0) \approx W^*$  which can be rewritten as  $w_+^*(T \simeq 0)n_+(n_+ - 1)/2 \approx w^*n(n - 1)/2$ . This implies

$$w_+^*(T \simeq 0) \approx \frac{w^*n^2}{n_+(n_+ - 1)} \quad (156)$$

and, from Eq. (121),

$$\psi_+^*(T \simeq 0) = \frac{w^*}{n_+(n_+ - 1)}, \quad (157)$$

consistently with Eq. (132).

We now note that, inserting Eqs. (152) and (156) into Eq. (68), we obtain the anticipated dependence of  $T_n$  (for  $T_n \simeq 0$ ) on  $n$ :

$$e^{-\epsilon^*/kT_n} = \frac{w_-^*(T_n \simeq 0)[1 + w_+^*(T_n \simeq 0)]}{w_+^*(T_n \simeq 0)[1 + w_-^*(T_n \simeq 0)]} \approx \frac{\ell}{n}. \quad (158)$$

Inverting, we find how the temperature approaches zero as  $n$  grows:

$$T_n \approx \frac{\epsilon^*}{k \ln(n/\ell)} \approx \frac{\epsilon^*}{k \ln n}. \quad (159)$$

The above result, which is independent of  $\ell$ , connects the thermodynamic limit  $n \rightarrow \infty$  with the zero-temperature limit  $T \rightarrow 0$  in our setting and confirms that it would be inappropriate to first identify the ground state are the core links by taking the limit  $T \rightarrow 0$  and subsequently let  $n$  grow. On the contrary, the zero-temperature state turns out to be the entire set of configurations obtained displacing a certain number of units of weight out of the core and such that  $\ell = o(n)$ . Inserting Eqs. (152), (154) and (156) into Eqs. (53) and (54) we can characterize these configurations through the connection probabilities

$$p_+^*(T \simeq 0) \approx 1 - \frac{n_+(n_+ - 1)}{w^*n^2}, \quad (160)$$

$$p_-^*(T \simeq 0) \approx \frac{\ell}{n}, \quad (161)$$

$$p_0^*(T \simeq 0) \approx \sqrt{\frac{\ell}{n}}, \quad (162)$$

which in the thermodynamic limit behave as expected for the ground state, i.e.  $p_+^*(T \rightarrow 0) = 1$ ,  $p_-^*(T \rightarrow 0) = 0$  and  $p_0^*(T \rightarrow 0) = 0$ .

Using Eqs. (152) and (154), we can calculate the strengths from Eqs. (129) and (130) as

$$s_+^*(T \simeq 0) \approx (n_+ - 1)\psi_+^*(T \simeq 0)n^2 \approx \frac{w^*}{n_+}n^2 \quad (163)$$

$$s_-^*(T \simeq 0) \approx nw_-^*(T \simeq 0) + n_+w_0^*(T \simeq 0) \approx \ell. \quad (164)$$

The ‘inverse’ problem is solved by Eqs. (133) and (134), which now become

$$\beta_+^*(T \simeq 0) \approx \frac{n_+(n_+ - 1)}{2n^2w^*}, \quad (165)$$

$$\beta_-^*(T \simeq 0) \approx \frac{1}{2} \ln \frac{n}{\ell} \approx \frac{1}{2} \ln n. \quad (166)$$

By contrast, the solution to the ‘direct’ problem is given through the chemical potential, obtained inserting Eqs. (159) and (165) into Eq. (135):

$$\begin{aligned} \mu^*(T \simeq 0) &= -2kT\beta_+^*(T \simeq 0) \\ &\approx -\frac{\epsilon^*n_+(n_+ - 1)}{w^*n^2 \ln n}. \end{aligned} \quad (167)$$

The variances of the constraints can be calculated inserting Eqs. (152) and (157) into Eqs. (136) and (137). This yields

$$\Sigma_+^*(T \simeq 0) \approx \frac{(w^*)^2}{n_+^2(n_+ - 1)}n^4, \quad (168)$$

$$\Sigma_-^*(T \simeq 0) \approx \ell. \quad (169)$$

Using the above relationships, it is possible to show (see Appendix A 4) that the leading order of the determinant of the covariance matrix is

$$\det[\Sigma^*(T \simeq 0)] = O(n^{4n_+} \ell^{n-n_+}). \quad (170)$$

This leads to

$$\ln \det[\Sigma^*(T \simeq 0)] = O(n \ln \ell), \quad (171)$$

which is of smaller order compared to the scaling  $O(n \ln n)$  found in Eqs. (99), (114) and (142) for all positive temperatures, but still signalling the breakdown of ensemble equivalence. Equation (171) implies that the requirements ensuring the validity of Eq. (35) are not met, therefore in this case the leading term of the relative entropy cannot be calculated exactly. However, the leading order is still given by Eq. (33)

$$D^*(T \simeq 0) = O(\ln \det \Sigma^*(T \simeq 0)) = O(n \ln \ell). \quad (172)$$

The canonical entropy  $S_{\text{can}}^*(T \simeq 0)$  can be calculated by using Eqs. (165) and (166) to rewrite the sum of the five terms appearing into Eq. (75) as the following sum:

$$S_{\text{can}}^*(T \simeq 0) = \frac{n_+(n_+ - 1)}{2} + \frac{\ell}{2} n \ln n + n_+(n_+ - 1) \ln n + \frac{\ell}{2} n + n_+ \sqrt{n\ell} = O(\ell n \ln n), \quad (173)$$

which, unless  $\ell = O(n/\ln n)$  or bigger, is different from the scaling  $O(n^2)$  obtained for finite temperatures in Eqs. (101) and (144). This slower increase of the canonical entropy with  $n$  confirms that in the zero-temperature limit the system behaves as a Bose-Einstein condensate, a phenomenon that determines a strong reduction in the dimensionality of the space of allowed configurations. Note that since the relative entropy is of order given by Eq. (171), which is smaller than the order of  $S_{\text{can}}^*(T \simeq 0)$ , the leading term of the microcanonical entropy must be

$$S_{\text{mic}}^*(T \simeq 0) \approx \frac{\ell}{2} n \ln n = O(\ell n \ln n). \quad (174)$$

Combining Eqs. (163)-(169), we finally obtain the relative canonical fluctuations

$$\rho_+^*(T \simeq 0) = \frac{\sqrt{n_+}}{n_+ - 1}, \quad (175)$$

$$\rho_-^*(T \simeq 0) = \frac{1}{\sqrt{\ell}}, \quad (176)$$

which differ from the results in Eqs. (146) and (147) obtained in the (subcritical) finite-temperature case. In particular, now both  $\rho_-^*(T \simeq 0)$  and  $\rho_+^*(T \simeq 0)$  (if  $\ell$  is finite) are non-zero. Note that, while Eq. (147) can be formally retrieved from Eq. (176) by letting  $\ell$  grow linearly in  $n$ , Eq. (146) cannot be retrieved from Eq. (175), because the assumption  $\ell = o(n)$  has already been exploited in the derivation.

### E. Critical temperature: $T = T_c$

Having characterized the model in all regimes, we can now discuss more easily what happens right at the critical temperature  $T = T_c$ . Clearly, a very interesting question

is whether the phase transition is of first or second order. A first-order phase transition is obtained when the order parameter  $Q^*(T)$  jumps discontinuously from zero to a strictly positive value as  $T$  is lowered through  $T_c$ . In such a case, the left and right limits of  $Q^*(T)$  at  $T = T_c$  are different:

$$Q^*(T \rightarrow T_c^-) > Q^*(T \rightarrow T_c^+) = 0. \quad (177)$$

By contrast, the phase transition is of second order if the order parameter increases continuously from zero to positive values as  $T$  is lowered through  $T_c$ :

$$Q^*(T \rightarrow T_c^-) = Q^*(T \rightarrow T_c^+) = 0. \quad (178)$$

In principle, in our setting we can engineer the order of the phase transition as we like: as clear from Eq. (131), the value of the order parameter for values slight below the critical temperature is governed by the value of  $\psi_+^*(T \lesssim T_c)$  defined in Eq. (121). So, if we choose  $\psi_+^*(T \rightarrow T_c^-) > 0$  the transition will be first-order, while if we choose  $\psi_+^*(T \rightarrow T_c^-) = 0$  the transition will be second-order. While both choices are possible, the case  $\psi_+^*(T \rightarrow T_c^-) > 0$  is somewhat unnatural, since it would ‘forbid’ all those strength sequences  $\bar{s}^*(T < T_c)$  that, while being both graphic and perfectly consistent with the definition of ‘condensed’ given in Sec. IV D 1, are such that asymptotically  $w_+^*(T < T_c) < n^2 \psi_+^*(T \rightarrow T_c^-)$  or equivalently, by virtue of Eq. (124), such that  $w_-^*(T < T_c) < w^* - n_+(n_+ - 1) \psi_+^*(T \rightarrow T_c^-)$ . Note that the latter inequality implies that  $w_-^*(T)$  would experience a finite jump from  $w^* - n_+(n_+ - 1) \psi_+^*(T \rightarrow T_c^-) < w^*$  to  $w^*$  as  $T$  is raised from a value just below  $T_c$  to a value just above  $T_c$ : as discussed in Sec. IV C 1,  $w^*$  is (to leading order) the only allowed value for  $w_-^*(T)$  above the critical temperature.

Therefore we find more appropriate to choose  $\psi_+^*(T)$  such that  $\psi_+^*(T \rightarrow T_c^-) = 0$ . In this way, all values of  $w_-^*(T < T_c)$  in the range  $[0, w^*]$  are allowed and there is no discontinuity for  $w_-^*(T)$  at  $T_c$ : Eq. (124) implies that its left limit is  $w_-^*(T \rightarrow T_c^-) \approx w^*$ , which coincides with its right limit  $w_-^*(T \rightarrow T_c^+) \approx w^*$  implied by Eq. (86). With this choice, we can locate the critical temperature  $T_c$  by equating the right and left limits of Eq. (68): since  $\lim_{T \rightarrow T_c^-} e^{-\epsilon^*/kT} = \lim_{T \rightarrow T_c^+} e^{-\epsilon^*/kT} = e^{-\epsilon^*/kT_c}$ , the right and left limits of

$$\frac{w_-^*(T)[1 + w_+^*(T)]}{w_+^*(T)[1 + w_-^*(T)]} \quad (179)$$

must coincide. This implies

$$\frac{1 + w_+^*(T \rightarrow T_c^+)}{w_+^*(T \rightarrow T_c^+)} = \frac{1 + w_+^*(T \rightarrow T_c^-)}{w_+^*(T \rightarrow T_c^-)} = 1 \quad (180)$$

since in the thermodynamic limit  $w_+^*(T \rightarrow T_c^-) = \infty$ . The above expression in turn implies  $w_+^*(T \rightarrow T_c^+) = \infty$ . Note that this is consistent with the fact that, as discussed in Sec. IV C 1,  $w_+^*(T > T_c)$  is finite but can be

arbitrarily large, while not altering (to leading order) the average weight  $w^*$ . So,  $w_+^*(T > T_c)$  can grow indefinitely as  $T$  decreases towards  $T_c$ , and take an infinite limit as  $T \rightarrow T_c^-$ , consistently with the fact that, for even lower temperatures,  $w_+^*(T)$  diverges with speed  $n^2$  as dictated by Eq. (120). Inserting Eq. (180) into Eq. (68), we get

$$e^{-\epsilon^*/kT_c} = \frac{w^*}{1+w^*} \quad (181)$$

as  $n \rightarrow \infty$ . We therefore obtain

$$T_c = \frac{\epsilon^*}{k \ln \left(1 + \frac{1}{w^*}\right)}, \quad (182)$$

finally showing how the critical temperature depends on the expected link weight  $w^*$ , on the energy difference  $\epsilon^*$  between periphery-periphery and core-core links, and on the constant  $k$  converting the units of the ‘cost of links’ (energy) to those of the temperature. These results explain much of the information anticipated previously in Fig. 3 and summarized therein.

Adopting the view that the order parameter is continuous through the critical value  $T_c$ , we notice that the ‘direct’ solutions in Eqs. (93) and (135), as well as the behaviour at the critical point  $T = T_c$ , can be combined into the general ‘phenomenological’ expression

$$\mu^*(T) \approx \epsilon^* - kT \ln \left(1 + \frac{1}{w^* [1 - Q_n^*(T)]}\right) + o(T) \quad (183)$$

which is valid for all values of the temperature. Indeed, when  $T \geq T_c$  the order parameter is zero and the above expression reduces to Eq. (93), while when  $T < T_c$  the order parameter takes the positive value in Eq. (131) and the above expression reduces to Eq. (135). The extreme limits  $T \rightarrow \infty$  and  $T \rightarrow 0$  can be retrieved from Eq. (183) as well.

## V. CONCLUSIONS

We have investigated the breakdown of equivalence between canonical and microcanonical ensembles of weighted networks with local constraints on the strength of each node (*weighted configuration model* [18]). While ensemble nonequivalence in the corresponding *binary configuration model* (i.e. binary networks with given node degrees) had already been studied in detail [5, 19, 20, 40, 41], a similar analysis for weighted networks had not been carried out so far. As a unique and novel ingredient in the case considered here, weighted networks can undergo BEC, a phase transition that is impossible to observe in the unweighted case. BEC emerges when a finite fraction of the total weight of all links condenses in a finite number of links. We constructed the simplest model exhibiting such behaviour: a network with a finite core, an infinite periphery, and a temperature-dependent strength sequence. This setting allows us to

combine for the first time, in a single model, two completely different mechanisms that can potentially destroy the equivalence of ensembles: a phase transition (a condition exhibited in the earliest observations of ensemble nonequivalence [3, 7–17]) and an extensive number of local constraints (an ingredient found in more recent investigations on network ensembles [5, 6, 19, 20]).

We have considered two criteria for ensemble equivalence: the traditional and intuitive one based on the vanishing of the relative canonical fluctuations of the constraints in the thermodynamic limit [1] and the more recent and rigorous one based on the vanishing of the relative entropy density between microcanonical and canonical probability distributions (*measure equivalence*) [4]. While in the standard situation (i.e. under only one or a finite number of global constraints) the vanishing of the relative fluctuations implies measure equivalence, the relationship between the two criteria had not been investigated in presence of an extensive number of local constraints yet. Technically, while the relative fluctuations can be calculated exactly (as they are purely canonical quantities), the relative entropy requires in principle unfeasible microcanonical calculations but it can still be calculated asymptotically via a recently proposed saddle-point technique showing that its leading term is the logarithm of the determinant of the matrix of canonical covariances between the constraints.

We found that, for all positive temperatures, the relative entropy is  $O(n \ln n)$  while the canonical and microcanonical entropies are  $O(n^2)$ . These behaviours mimic the corresponding ones found for the binary configuration model in the dense regime [5, 19, 20]. This result shows that, for all  $T > 0$  (including  $T \rightarrow \infty$ ), the relative entropy is subleading with respect to the canonical and microcanonical entropies, but is still superextensive in the number of nodes  $n$ , which in all network models represents the number of units (physical size) of the system. In the zero-temperature limit, we found slower scalings for the canonical, microcanonical and relative entropies. This is due to the fact that, in both canonical and microcanonical ensembles, the peripheral nodes are asymptotically disconnected from all other nodes in each possible realization of the network. In this zero-temperature limit, the condensate effectively behaves as a lower-dimensional system, as commonly observed in the physics of BEC. Its entropy is the residual entropy resulting from the degeneracy of the ground state. The scaling of the relative entropy still indicates ensemble nonequivalence. We note that in the binary configuration model (which obeys Fermi-Dirac rather than Bose-Einstein statistics) the zero-temperature phase is one where the canonical and microcanonical ensembles are instead *identical*, because in both ensembles the pairs of nodes below a certain ‘Fermi energy’ (whose value coincides with the chemical potential) are surely connected, while those above it are surely disconnected [24]. Therefore we can conclude that, irrespective of BEC, at all temperatures ensemble equivalence is broken by the pres-



ence of an extensive number of local constraints, as in the binary configuration model (for which, however, BEC cannot occur). So the condensation phase transition (occurring at some critical temperature  $T_c > 0$ ) appears to have no effect on ensemble equivalence.

On the other hand, the calculation of the canonical relative fluctuations of the constraints shows that they are sensitive to the phase transition, while they cannot be used to characterize ensemble (non)equivalence as traditionally expected. Indeed, we found that in the non-condensed phase ( $T > T_c$ ) the relative fluctuations of the strength of all the  $n$  nodes vanish in the thermodynamic limit. By contrast, in the condensed phase ( $T < T_c$ ) the relative fluctuations of the strength of the  $n_+$  nodes in the core do not vanish, while those for the  $n_-$  nodes in the periphery still do (except in the zero-temperature limit, for which not even the relative fluctuations for the peripheral nodes vanish). Therefore, as the temperature is lowered below the critical temperature, there is a sudden change in the relative fluctuations but no change in the scaling of the relative entropy. Conversely, as the temperature is further lowered to zero, there is a sudden change in the scaling of the relative entropy, while the relative fluctuations for the core nodes remain non-zero, albeit with a different value. These results show that, at least in the dense case studied here, the relative entropy and the relative fluctuations capture different aspects of the phenomenology of the proposed model, the former being sensitive to the presence of local constraints and the latter being sensitive to the phase transition. In any case, in presence of an extensive number of local constraints the vanishing of (even all) the canonical relative fluctuations does not guarantee measure equivalence and is therefore no longer a valid criterion for ensemble equivalence as intuitively expected.

We stress that, while the network model presented here is deliberately simple from the structural point of view (a core-periphery network with local, but homogeneous, constraints), it could certainly serve as a reference for more complicated models (e.g. a core-periphery network with local and heterogeneous constraints). Indeed, ensemble nonequivalence will still be manifest in such a generalized model for all positive temperatures, because research on binary networks with given degrees has shown that nonequivalence is due to the locality of the constraints, and not to their specific value [5, 6, 19, 20]. Additionally, since a more heterogeneous choice of the constraints can only increase the number of states with different energy in the network, we expect that BEC will still emerge below some critical temperature. In general, we expect a qualitatively similar behaviour to the one found here, with only quantitative differences.

The concept of ensemble equivalence is central for the foundations of statistical physics, irrespective of the particular system being considered. The findings documented here shed new light on the breakdown of EE, on the (possibly misleading) criteria used to detect it, and on the (so far undocumented) interplay between differ-

ent mechanisms producing it. We hope they can inspire future research on these subjects.

## ACKNOWLEDGMENTS

This work is supported by the Chinese Scholarship Council (No. 201606990001), the Dutch Econophysics Foundation (Stichting Econophysics, Leiden, the Netherlands) and the Netherlands Organization for Scientific Research (NWO/OCW).

### Appendix A: Determinant of the covariance matrix

$\Sigma^*(T)$  is the  $n \times n$  canonical covariance matrix between the strengths of all nodes, with entries

$$\begin{aligned} \Sigma_{ij}^*(T) &= \begin{cases} \text{Var}_{\beta^*(T)}(s_i) & i = j \\ \text{Cov}_{\beta^*(T)}(w_{ij}) & i \neq j \end{cases} \\ &= \begin{cases} \Sigma_{ii}^*(T) & i = j \\ \langle w_{ij} \rangle_{\beta^*(T)} [1 + \langle w_{ij} \rangle_{\beta^*(T)}] & i \neq j \end{cases}. \end{aligned} \quad (\text{A1})$$

Combining Eqs. (23) and (24) with the results discussed in Sec. IV B for our core-periphery model, it is easy to see that, for all values of temperature,  $\Sigma^*(T)$  is a combination of four blocks

$$\Sigma^*(T) = \begin{bmatrix} \mathbf{A}(T) & \mathbf{B}(T) \\ \mathbf{C}(T) & \mathbf{D}(T) \end{bmatrix}, \quad (\text{A2})$$

where  $\mathbf{A}(T)$  is the  $n_+ \times n_+$  submatrix of covariances between the strengths of nodes in the core, with entries

$$A_{ij}(T) = \begin{cases} \Sigma_+^*(T) & i = j \\ w_+^*(T)[1 + w_+^*(T)] & i \neq j \end{cases}, \quad (\text{A3})$$

$\mathbf{B}(T)$  is the  $n_+ \times n_-$  submatrix of covariances between the strengths of nodes across core and periphery, with entries

$$B_{ij}(T) = w_0^*(T)[1 + w_0^*(T)] \quad \forall i, j, \quad (\text{A4})$$

$\mathbf{C}(T)$  is a  $n_- \times n_+$  matrix equal to the transpose of  $\mathbf{B}(T)$ , and  $\mathbf{D}(T)$  is the  $n_- \times n_-$  submatrix of covariances between the strengths of nodes in the periphery, with entries

$$D_{ij}(T) = \begin{cases} \Sigma_-^*(T) & i = j \\ w_-^*(T)[1 + w_-^*(T)] & i \neq j \end{cases}. \quad (\text{A5})$$

Depending on the range of temperature values of interest, different techniques become useful in order to calculate the determinant of  $\Sigma^*(T)$ . We therefore consider each regime separately below.

### 1. Non-condensed phase

In the regime of finite supercritical temperature ( $T > T_c$ ) discussed in Sec. IV C 1, it is possible to show that the asymptotic behaviour of  $\det \Sigma^*(T > T_c)$  can be decomposed as the product of the diagonal elements of  $\Sigma^*(T > T_c)$ , plus a correction. Our rationale for this decomposition comes from the fact that, as noted in Sec. IV C 1,  $w_+(T > T_c)$ ,  $w_-(T > T_c)$  and  $w_0^*(T > T_c)$  are all  $O(1)$ , i.e. the expected link weights are all of the same, finite order. Consequently, the block structure depicted in Eq. (A2) does not identify any particular difference in the order of magnitude of the entries of  $\Sigma^*(T > T_c)$ . Rather, an important property of  $\Sigma^*(T > T_c)$  in this regime is that its diagonal entries are, on average,  $n$  times bigger than its off-diagonal ones. Indeed, the off-diagonal entries are  $O(1)$ , while the diagonal ones are  $O(n)$ . Then the asymptotic behaviour of  $\det \Sigma^*(T > T_c)$  must be essentially dictated by the product of the diagonal entries of  $\Sigma^*(T > T_c)$ .

To make this intuition more rigorous, we recall that if a  $k \times k$  matrix  $\mathbf{L}$  can be decomposed as

$$\mathbf{L} = \mathbf{M} + \varepsilon \mathbf{N}, \quad (\text{A6})$$

where  $\mathbf{M}$  is a diagonal matrix with entries of finite order and  $\varepsilon \mathbf{N}$  is a perturbation, then Jacobi's formula applies as follows:

$$\begin{aligned} \det \mathbf{L} &= \det(\mathbf{M} + \varepsilon \mathbf{N}) \\ &= \det \mathbf{M} + \varepsilon (\det \mathbf{M}) \text{tr}(\mathbf{M}^{-1} \mathbf{N}) + O(\varepsilon^2). \end{aligned} \quad (\text{A7})$$

Moreover, if the diagonal elements of  $\mathbf{N}$  are equal to 0, then the product  $\mathbf{M}^{-1} \mathbf{N}$  is a  $k \times k$  zero matrix and therefore

$$\text{tr}(\mathbf{M}^{-1} \mathbf{N}) = 0. \quad (\text{A8})$$

Equation (A8) then becomes

$$\begin{aligned} \det \mathbf{L} &= \det \mathbf{M} + O(\varepsilon^2) \\ &= \prod_{i=1}^k L_{ii} + O(\varepsilon^2). \end{aligned} \quad (\text{A9})$$

Turning to the matrix  $\Sigma^*(T > T_c)$ , we note that the above hypotheses apply by setting

$$k \equiv n, \quad \varepsilon \equiv \frac{1}{n}, \quad \mathbf{L} \equiv \frac{\Sigma^*(T > T_c)}{n} \quad (\text{A10})$$

and defining the entries of  $\mathbf{M}$  and  $\mathbf{N}$  as

$$M_{ij} = \delta_{ij} \Sigma_{ij}^*(T > T_c) / n, \quad (\text{A11})$$

$$N_{ij} = (1 - \delta_{ij}) \Sigma_{ij}^*(T > T_c), \quad (\text{A12})$$

where  $\delta_{ij}$  is the Kronecker delta symbol. Equation (A9) then becomes

$$\det \left( \frac{\Sigma^*(T > T_c)}{n} \right) = \frac{1}{n^n} \prod_{i=1}^n \Sigma_{ii}^*(T > T_c) + O(n^{-2})$$

and, finally,

$$\begin{aligned} \det \Sigma^*(T > T_c) &= n^n \det \left( \frac{\Sigma^*(T > T_c)}{n} \right) \\ &= \prod_{i=1}^n \Sigma_{ii}^*(T > T_c) + O(n^{n-2}), \end{aligned} \quad (\text{A13})$$

proving Eq. (96) used in the main text.

### 2. Infinite-temperature limit

In the infinite-temperature limit discussed in Sec. (IV C 2), the determinant can be calculated exactly as follows. From Eqs. (111) and (112) we see that, if we introduce a  $k \times k$  matrix  $\mathbf{Z}_k$  defined as

$$\mathbf{Z}_k = \begin{pmatrix} k-1 & 1 & \cdots & 1 & 1 \\ 1 & k-1 & 1 & \cdots & 1 \\ \vdots & & \ddots & & \vdots \\ 1 & \cdots & 1 & k-1 & 1 \\ 1 & 1 & \cdots & 1 & k-1 \end{pmatrix}, \quad (\text{A14})$$

then we can rewrite the covariance matrix as

$$\Sigma^*(T \rightarrow \infty) = w^*(1 + w^*) \mathbf{Z}_n. \quad (\text{A15})$$

Clearly, the calculation of  $\det \Sigma^*(T \rightarrow \infty)$  reduces to the calculation of  $\det \mathbf{Z}_n$ :

$$\det \Sigma^*(T \rightarrow \infty) = [w^*(1 + w^*)]^n \det \mathbf{Z}_n. \quad (\text{A16})$$

To compute  $\det \mathbf{Z}_k$  for arbitrary  $k$ , we note that

$$\mathbf{Z}_k = (k-2) \mathbf{I}_k + \mathbf{u}_k^T \mathbf{u}_k = (k-2) \left( \mathbf{I}_k + \frac{\mathbf{u}_k^T \mathbf{u}_k}{k-2} \right), \quad (\text{A17})$$

where  $\mathbf{I}_k$  is the  $k \times k$  identity matrix and

$$\mathbf{u}_k = (1, \dots, 1) \quad (\text{A18})$$

is the  $k$ -dimensional row vector with all unit entries. Then, using Sylvester's identity  $\det(\mathbf{I}_k + \mathbf{X}\mathbf{Y}) = \det(\mathbf{I}_l + \mathbf{Y}\mathbf{X})$  (where  $\mathbf{X}$  is a  $k \times l$  matrix,  $\mathbf{Y}$  is an  $l \times k$  matrix, and  $\mathbf{I}_k$  and  $\mathbf{I}_l$  are  $k \times k$  and  $l \times l$  identity matrices respectively) with  $l = 1$ ,  $\mathbf{X} = \mathbf{u}_k^T$ ,  $\mathbf{Y} = \mathbf{u}_k$  and  $\mathbf{I}_l = 1$ , we get

$$\begin{aligned} \det \mathbf{Z}_k &= (k-2)^k \det \left( \mathbf{I}_k + \frac{\mathbf{u}_k^T}{\sqrt{k-2}} \frac{\mathbf{u}_k}{\sqrt{k-2}} \right) \\ &= (k-2)^k \det \left( 1 + \frac{\mathbf{u}_k}{\sqrt{k-2}} \frac{\mathbf{u}_k^T}{\sqrt{k-2}} \right) \\ &= (k-2)^k \left( 1 + \frac{k}{k-2} \right) \\ &= 2(k-1)(k-2)^{k-1}. \end{aligned} \quad (\text{A19})$$

Combining Eqs. (A16) and (A19), and setting  $k = n$ , we obtain exactly Eq. (113) used in the main text.

### 3. Condensed phase

In the regime of subcritical temperature ( $T < T_c$ ) discussed in Sec. IV D, the block structure indicated in Eq. (A2) becomes particularly relevant, as it captures the important differences in the order of magnitude of both diagonal and off-diagonal entries of  $\Sigma^*(T < T_c)$  calculated using Eqs. (122), (125), (127), (136) and (137). We first express each block conveniently and then proceed to the calculation of the determinant. Inserting Eqs. (122) and (136) into Eq. (A3), we obtain

$$\mathbf{A}(T < T_c) \approx [\psi_+^*(T < T_c)]^2 n^4 \mathbf{Z}_{n_+}, \quad (\text{A20})$$

where  $\mathbf{Z}_k$  is still the matrix defined in Eq. (A14). Next, we note from Eq. (A4) that

$$\mathbf{B}(T < T_c) = w_0^*(T < T_c)[1 + w_0^*(T < T_c)] \mathbf{u}_{n_+}^T \mathbf{u}_{n_-}, \quad (\text{A21})$$

where  $\mathbf{u}_k$  is still given by Eq. (A18). Similarly,

$$\mathbf{C}(T < T_c) = w_0^*(T < T_c)[1 + w_0^*(T < T_c)] \mathbf{u}_{n_-}^T \mathbf{u}_{n_+}. \quad (\text{A22})$$

Finally, inserting Eq. (137) into Eq. (A5), we obtain

$$\mathbf{D}(T < T_c) \approx w_-^*(T < T_c)[1 + w_-^*(T < T_c)] \mathbf{Z}_{n_-} \quad (\text{A23})$$

Now, since  $\mathbf{A}(T < T_c)$  is invertible, it is useful to exploit the block structure of  $\Sigma^*(T < T_c)$  by expressing its determinant as

$$\det \Sigma^*(T < T_c) = \det \mathbf{A}(T < T_c) \det \bar{\mathbf{A}}(T < T_c), \quad (\text{A24})$$

where

$$\bar{\mathbf{A}}(T < T_c) \equiv \mathbf{D}(T < T_c) - \mathbf{C}(T < T_c) \mathbf{A}^{-1}(T < T_c) \mathbf{B}(T < T_c) \quad (\text{A25})$$

is the so-called Shur complement of  $\mathbf{A}(T < T_c)$ . To calculate  $\det \mathbf{A}(T < T_c)$ , we use Eq. (A20) and immediately obtain

$$\begin{aligned} \det \mathbf{A}(T < T_c) &\approx [\psi_+^*(T < T_c)]^{2n_+} n^{4n_+} \det \mathbf{Z}_{n_+} \\ &= O(n^{4n_+}) \end{aligned} \quad (\text{A26})$$

where, using Eq. (A19),

$$\det \mathbf{Z}_{n_+} = 2(n_+ - 1)(n_+ - 2)^{n_+ - 1}. \quad (\text{A27})$$

To calculate  $\det \bar{\mathbf{A}}(T < T_c)$ , we first use Eq. (A20) and obtain

$$\mathbf{A}^{-1}(T < T_c) \approx [\psi_+^*(T < T_c)]^{-2} n^{-4} \mathbf{Z}_{n_+}^{-1}, \quad (\text{A28})$$

where, using Eq. (A14),  $\mathbf{Z}_k^{-1}$  is easily calculated by direct inversion of  $\mathbf{Z}_k$  as

$$\begin{aligned} \mathbf{Z}_k^{-1} &= c_k \begin{pmatrix} 2k-3 & -1 & \cdots & -1 & -1 \\ -1 & 2k-3 & -1 & \cdots & -1 \\ \vdots & & \ddots & & \vdots \\ -1 & \cdots & -1 & 2k-3 & -1 \\ -1 & -1 & \cdots & -1 & 2k-3 \end{pmatrix} \\ &= \frac{\mathbf{I}_k}{k-2} - c_k \mathbf{u}_k^T \mathbf{u}_k \\ &= \frac{1}{k-2} \left[ \mathbf{I}_k - \frac{\mathbf{u}_k^T \mathbf{u}_k}{2(k-1)} \right] \end{aligned} \quad (\text{A29})$$

with

$$c_k = \frac{1}{2(k-1)(k-2)}. \quad (\text{A30})$$

Inserting Eqs. (A21),(A22), (A23) and (A28) into Eq. (A25), and noticing that

$$\begin{aligned} \mathbf{u}_{n_-}^T \mathbf{u}_{n_+} \mathbf{Z}_{n_+}^{-1} \mathbf{u}_{n_+}^T \mathbf{u}_{n_-} &= c_{n_+} n_+ (n_+ - 2) \mathbf{u}_{n_-}^T \mathbf{u}_{n_-} \\ &= \frac{n_+}{2(n_+ - 1)} \mathbf{u}_{n_-}^T \mathbf{u}_{n_-}, \end{aligned} \quad (\text{A31})$$

we can obtain the Shur complement of  $\mathbf{A}(T < T_c)$  as

$$\begin{aligned} \bar{\mathbf{A}}(T < T_c) &\approx w_-^*(T < T_c)[1 + w_-^*(T < T_c)] \mathbf{Z}_{n_-} \quad (\text{A32}) \\ &\quad - \frac{n_+[w_0^*(T < T_c)]^2 [1 + w_0^*(T < T_c)]^2}{2(n_+ - 1)[\psi_+^*(T < T_c)]^2 n^4} \mathbf{u}_{n_-}^T \mathbf{u}_{n_-} \end{aligned}$$

from which we can calculate  $\det \bar{\mathbf{A}}(T < T_c)$ . We have to distinguish the cases  $0 < T < T_c$  and  $T \simeq 0$ , as they are characterized by different scalings of  $w_-^*(T < T_c)$  and  $w_0^*(T < T_c)$ . In the rest of this section we consider the case of finite temperature, while the zero-temperature limit is considered in the next section.

When  $0 < T < T_c$ , we recall from Eqs.(139) and (140) that both  $w_-^*(0 < T < T_c)$  and  $w_0^*(0 < T < T_c)$  are  $O(1)$ . From Eq. (A32) we therefore see that all the off-diagonal entries of  $\bar{\mathbf{A}}(0 < T < T_c)$  are  $O(1)$ , while all the diagonal ones are  $O(n_-)$ . This implies that we can use the decomposition in Eq. (A6) where

$$k \equiv n_-, \quad \varepsilon \equiv \frac{1}{n_-}, \quad \mathbf{L} \equiv \frac{\bar{\mathbf{A}}(0 < T < T_c)}{n_-}. \quad (\text{A33})$$

Equation (A9) then implies

$$\begin{aligned} \det \bar{\mathbf{A}}(0 < T < T_c) &= \prod_{i=1}^{n_-} \bar{A}_{ii}(0 < T < T_c) + O(n_-^{n_- - 2}) \\ &= O(n_-^{n_-}). \end{aligned} \quad (\text{A34})$$

Combining Eqs. (A26) and (A34) into Eq. (A24), we finally obtain the full determinant of  $\Sigma^*(0 < T < T_c)$ . We are interested only in its scaling with  $n$ , which is

$$\det \Sigma^*(0 < T < T_c) = O(n^{4n_+ + n_-}) = O(n^{n^2 + 3n_+}), \quad (\text{A35})$$

proving Eq. (141) used in the main text.

### 4. Zero-temperature limit

In the zero-temperature limit, all calculations of the previous section remain valid until and including Eq. (A32). The scaling of the entries of  $\bar{\mathbf{A}}(T \simeq 0)$  will however be different. Indeed, we recall from Eqs.(152) and (154) that  $w_-^*(T \simeq 0) \approx \ell/n$  and  $w_0^*(T \simeq 0) \approx \sqrt{\ell/n}$ . Inserted into Eq. (A32), these expressions imply that all the diagonal entries of  $\bar{\mathbf{A}}(T \simeq 0)$  are asymptotically equal

to  $\ell$ , while all the off-diagonal ones are  $O(\ell/n)$ . We can therefore use the decomposition in Eq. (A6) where

$$k \equiv n_-, \quad \varepsilon \equiv \frac{1}{n}, \quad \mathbf{L} \equiv \overline{\mathbf{A}}(T \simeq 0). \quad (\text{A36})$$

Equation (A9) then implies

$$\begin{aligned} \det \overline{\mathbf{A}}(T \simeq 0) &= \prod_{i=1}^{n_-} \overline{A}_{ii}(T \simeq 0) + O(n^{-2}) \\ &= \ell^{n_-} + O(n^{-2}). \end{aligned} \quad (\text{A37})$$

Combined with Eq. (A26) into Eq. (A24), the above result leads to the full determinant of  $\Sigma^*(T < T_c)$ , whose scaling with  $n$  is

$$\det \Sigma^*(T \simeq 0) = O(n^{4n + \ell^m}), \quad (\text{A38})$$

proving Eq. (170) used in the main text.

- 
- [1] J. Gibbs, *The Collected Works of JW Gibbs* (Yale University, New Haven, CT, 1957) **2** (1902).
- [2] E. T. Jaynes, *Physical review* **106**, 620 (1957).
- [3] A. Campa, T. Dauxois, and S. Ruffo, *Physics Reports* **480**, 57 (2009).
- [4] H. Touchette, *Journal of Statistical Physics* **159**, 987 (2015).
- [5] T. Squartini, J. de Mol, F. den Hollander, and D. Garlaschelli, *Physical review letters* **115**, 268701 (2015).
- [6] T. Squartini and D. Garlaschelli, *arXiv preprint arXiv:1710.11422* (2017).
- [7] R. S. Ellis, K. Haven, and B. Turkington, *Journal of Statistical Physics* **101**, 999 (2000).
- [8] M. Blume, V. Emery, and R. B. Griffiths, *Physical review A* **4**, 1071 (1971).
- [9] J. Barré, D. Mukamel, and S. Ruffo, *Physical Review Letters* **87**, 030601 (2001).
- [10] R. S. Ellis, H. Touchette, and B. Turkington, *Physica A: Statistical Mechanics and its Applications* **335**, 518 (2004).
- [11] D. Lynden-Bell, *Physica A: Statistical Mechanics and its Applications* **263**, 293 (1999).
- [12] P.-H. Chavanis, *Astronomy & Astrophysics* **401**, 15 (2003).
- [13] M. d'Agostino, F. Gulminelli, P. Chomaz, M. Bruno, F. Cannata, R. Bougault, F. Gramigna, I. Iori, N. Le Neindre, G. Margagliotti, *et al.*, *Physics Letters B* **473**, 219 (2000).
- [14] J. Barré and B. Gonçalves, *Physica A: Statistical Mechanics and its Applications* **386**, 212 (2007).
- [15] C. Radin and L. Sadun, *Journal of Physics A: Mathematical and Theoretical* **46**, 305002 (2013).
- [16] R. S. Ellis, K. Haven, and B. Turkington, *Nonlinearity* **15**, 239 (2002).
- [17] M. Kastner, *Physical review letters* **104**, 240403 (2010).
- [18] T. Squartini and D. Garlaschelli, *Maximum-Entropy Networks: Pattern Detection, Network Reconstruction and Graph Combinatorics* (Springer, 2017).
- [19] D. Garlaschelli, F. Den Hollander, and A. Roccaverde, *Journal of Physics A: Mathematical and Theoretical* **50**, 015001 (2016).
- [20] A. Roccaverde, *Indagationes Mathematicae* **30**, 7 (2019).
- [21] J. Park and M. E. Newman, *Physical Review E* **70**, 066117 (2004).
- [22] D. Garlaschelli and M. I. Loffredo, *Physical review letters* **102**, 038701 (2009).
- [23] G. Bianconi and A.-L. Barabási, *Physical review letters* **86**, 5632 (2001).
- [24] D. Garlaschelli, S. E. Ahnert, T. Fink, and G. Caldarelli, *Entropy* **15**, 3148 (2013).
- [25] M. E. Newman, A.-L. E. Barabási, and D. J. Watts, *The structure and dynamics of networks*. (Princeton university press, 2006).
- [26] M. Á. Serrano, M. Boguñá, and R. Pastor-Satorras, *Physical Review E* **74**, 055101 (2006).
- [27] T. Squartini, R. Mastrandrea, and D. Garlaschelli, *New Journal of Physics* **17**, 023052 (2015).
- [28] D. Garlaschelli and M. I. Loffredo, *Physical Review E* **78**, 015101 (2008).
- [29] "<https://it.mathworks.com/matlabcentral/fileexchange/46912-max-sam-package-zip>".
- [30] "<https://meh.imtlucca.it>".
- [31] S. Kullback and R. A. Leibler, *The annals of mathematical statistics* **22**, 79 (1951).
- [32] A. Einstein, *Königliche Preußische Akademie der Wissenschaften*. **261-267** (1924).
- [33] P. Navez, D. Bitouk, M. Gajda, Z. Idziaszek, and K. Rzażewski, *Physical review letters* **79**, 1789 (1997).
- [34] M. Holthaus, E. Kalinowski, and K. Kirsten, *Annals of Physics* **270**, 198 (1998).
- [35] W. Mullin and J. Fernandez, *American Journal of Physics* **71**, 661 (2003).
- [36] S. Chatterjee and P. Diaconis, *Journal of Physics A: Mathematical and Theoretical* **47**, 085201 (2014).
- [37] S. Tarasov, V. V. Kocharovsky, and V. Kocharovsky, *Journal of Statistical Physics* **161**, 942 (2015).
- [38] A. Crisanti, A. Sarracino, and M. Zannetti, *Physical Review Research* **1**, 023022 (2019).
- [39] C. Kittel and H. Kroemer, *Thermal Physics* (W.H. Freeman, San Francisco, 1980).
- [40] D. Garlaschelli, F. den Hollander, and A. Roccaverde, *Journal of Statistical Physics* **173**, 644 (2018).
- [41] P. Dionigi, D. Garlaschelli, F. den Hollander, and M. Mandjes, "A spectral signature of breaking of ensemble equivalence for constrained random graphs," <https://arxiv.org/abs/2009.05155> (2020).