Recovery thresholds in the sparse planted matching problem

Guilhem Semerjian,[1] Gabriele Sicuro [ORCID],[1] and Lenka Zdeborová[2]

[1]*Laboratoire de Physique de l'École Normale Supérieure, ENS, Université PSL, CNRS, Sorbonne Université, Université de Paris, F-75005 Paris, France*

[2]*Université Paris-Saclay, CNRS, CEA, Institut de physique théorique, 91191, Gif-sur-Yvette, France*

We consider the statistical inference problem of recovering an unknown perfect matching, hidden in a weighted random graph, by exploiting the information arising from the use of two different distributions for the weights on the edges inside and outside the planted matching. A recent work has demonstrated the existence of a phase transition, in the large size limit, between a full and a partial-recovery phase for a specific form of the weights distribution on fully connected graphs. We generalize and extend this result in two directions: we obtain a criterion for the location of the phase transition for generic weights distributions and possibly sparse graphs, exploiting a technical connection with branching random walk processes, as well as a quantitatively more precise description of the critical regime around the phase transition.

I. INTRODUCTION

A matching of a graph is a subset of its edges such that each node belongs to at most one edge of the matching; in a perfect matching all the nodes are covered in this way [1]. In a weighted graph one defines the weight of a matching as the sum of the weights on its edges, and one can try to minimize or maximize this total weight under the (perfect) matching constraint [2]. This extremization is a problem of combinatorial optimization, widely studied in mathematics, computer science, and also in statistical physics. In this paper we study the planted matching problem, a statistical inference problem where one hides (plants) a perfect matching into a graph, the goal being to find it back. Planted matching problems arise in applications such as particle-tracking systems used in experimental physics [3]; the present paper concentrates on a more fundamental aspect, namely the mathematical description of the peculiar type of phase transition this inference problem exhibits.

Statistical inference on graphs and networks is an area of recent interest including problems such as community detection [4], group testing [5], planted Hamiltonian cycle recovery [6], certain types of error-correcting codes [7], and many others. All these problems share the common pattern of a signal being observed indirectly via the edges and weights of a graph, with the goal to infer the signal back from these observations. Interestingly as a function of the signal-to-noise ratio and in the limit of large system sizes one encounters sharp thresholds (phase transitions). A classical network-inference problem presenting a phase transition is the community detection in graphs created by the stochastic block model (SBM) [4]. In the SBM on sparse graphs a phase transition happens between a no-recovery phase where an estimation of the signal better than a random guess is impossible and a partial-recovery phase where a positive (but bounded away from one) correlation with the signal can be obtained. This phase transition can be of second or first order depending on the details of the model, an interesting connection was put forward between phase transitions of first order and existence of algorithmically hard phases [4]. A typology of phase transitions in problems where the detectability transition (from zero correlation to positive one) appears was recently presented in Ref. [8]. In the SBM, in order to obtain full recovery of the signal, i.e., a correlation with the signal converging to one, the average degree of the graph has to diverge as the logarithm of its size [9]. In low-density-parity-check error-correcting codes [7], another widely studied example of inference problem on sparse graphs, there is also a first order phase transition but this time from the partial-recovery phase to the full-recovery phase where the signal (codeword) can be reconstructed with an error that vanishes as the system size diverges.

Here we study the planted matching problem, where a perfect matching is hidden in a graph by adding edges to it. The information about which edges were added comes through the distribution of the weights, with different distributions on the planted and nonplanted edges. This problem was introduced in Ref. [3] as a toy model in a particle-tracking problem and was studied numerically by solving the corresponding recursive distributional equations for a particular case of the distribution of weights, suggesting a phase transition between full-recovery and partial-recovery phases. More recently, Ref. [10] rigorously analyzed another special case of the distribution of weights, and proved the existence of such a phase transition.

We generalize and extend these previous results in several directions (at the level of rigor of theoretical physics). We locate the recovery phase transition for generic weight distributions, considering also a sparse regime for the edges added to the planted configuration. As in Ref. [3] we use the standard cavity method related to the belief propagation approximation [11]. Our key contribution is an analytical insight into how the
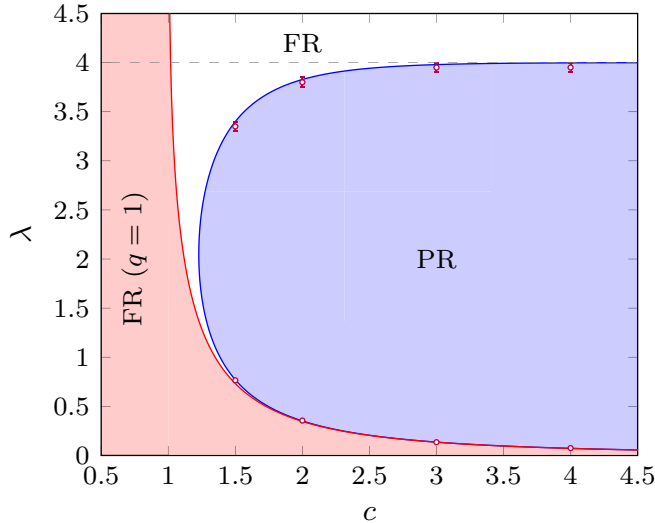
FIG. 1. Phase diagram of the planted matching problem with exponential planted weights on sparse graphs, exhibiting full-recovery (FR) and partial-recovery (PR) phases. The red area corresponds to the phase where FR is achievable by the simple leaf removal procedure and is delimited by the $\gamma = 1$ condition (red line). The blue area is the PR phase enclosed by the vanishing velocity criterion (blue line), the white domain corresponding to full recovery. The red dots have been obtained from the numerical resolution of the RDEs (29) by a population dynamics algorithm with 10^6 fields and mark the limit of existence of a nontrivial solution.

solution behaves, which allows us to derive a rather simple closed-form expression for the threshold, Eq. (45), that holds for generic distribution of weights and both sparse and fully connected graphs (the threshold for the sparse case converges to its fully connected limit exponentially fast in the average degree; see, for example, the phase diagram in Fig. 1). The results of Ref. [10] are also based on the cavity method, but apply only to fully connected graphs and only to the exponential distribution of weights on the planted edges. In this particular case the corresponding recursive distributional equations reduce into a closed system of differential equations. Instead we obtain the generic expression for the threshold by noticing a relation between the solution of the recursive distribution cavity equations and properties of branching random walk processes. The latter, and more generically the phenomenon of front propagation for reaction-diffusion equations, appear in a variety of context and have been extensively studied both in physics and in mathematics [12–22]; the precise way in which this connection arises here is nevertheless, as far as we know, original in the context of mean-field inference problems.

Both our work and Ref. [10] show that the recovery threshold in the planted matching problem is of a rather different nature than the thresholds known in the stochastic block model, error-correcting codes, or others discussed in the literature. Indeed it separates partial and full-recovery phases while occurring at a finite average degree, and there is no sign of a computational gap between the information theoretically optimal reconstruction accuracy and the one achievable by efficient algorithms. Another aspect in which this transition differs from more usual ones is its thermodynamic order:

for the specific case studied in Ref. [10] we provide a quantitatively more precise description of the critical regime around the phase transition, Eq. (88), showing that it is of infinite order in the usual thermodynamic classification (all the derivatives of the order parameter vanish at the transition point).

The rest of the paper is organized as follows. In Sec. II we define more explicitly the problem under study and introduce two statistical estimators [block and symbol, aximum a posteriori (MAP)] of the planted matching. In Sec. III we present the main equations (belief propagation and their probabilistic description) that governs the behavior of the problem. In Sec. IV we derive our first main result, namely, the location of the phase transition for arbitrary weight distributions (for the block MAP estimator), which is illustrated in Sec. V on several examples. Our second main result, i.e., a quantitatively more precise description of the critical regime around the phase transition, is explained in Sec. VI. We then study numerically in Sec. VII the threshold of the phase transition for the symbol MAP estimator, while conclusions and perspectives for future work are presented in Sec. VIII. Some more technical details are deferred to a pair of Appendices.

II. DEFINITIONS

A. Planted random weighted graphs

We shall consider weighted graphs denoted $\mathcal{G}_0 = (\mathcal{V}_0, \mathcal{E}_0, \underline{w})$, where $\mathcal{V}_0 = \{1, \ldots, N\}$ is the set of N vertices, N being an even integer, $\mathcal{E}_0$ the set of edges (unordered pairs of distinct vertices of $\mathcal{V}_0$), and $\underline{w} = \{w_e : e \in \mathcal{E}_0\}$ a collection of real weights assigned to each edge of the graph. We endow the set of weighted graphs with a probability distribution, the generation of $\mathcal{G}_0$ from this law corresponding to the following steps:

(1) One first chooses a perfect matching $\mathcal{M}$ of $\mathcal{V}_0$ uniformly at random among the $(N-1)!!$ possible ones, in other words $\mathcal{M}$ contains $N/2$ edges, each vertex of $\mathcal{V}_0$ belonging to exactly one edge of $\mathcal{M}$.

(2) The edge set $\mathcal{E}_0$ of $\mathcal{G}_0$ is made of the disjoint union of $\mathcal{M}$ and additional edges chosen at random: each of the $\binom{N}{2} - (N/2)$ possible edges not already included in $\mathcal{M}$ is added to $\mathcal{E}_0$ with probability c/N.

(3) The weights w_e are independent random variables, with an absolutely continuous distribution given by the density $\hat{p}$ if $e \in \mathcal{M}$ and p if $e \in \mathcal{E}_0 \setminus \mathcal{M}$.

We shall call planted (resp. nonplanted) edges those in $\mathcal{M}$ (resp. in $\mathcal{E}_0 \setminus \mathcal{M}$). The parameters of this random ensemble of weighted graphs are thus the even integer N, the parameter c controlling the density of nonplanted edges, and the two distributions $\hat{p}$ and p for the generation of the weights of the planted and nonplanted edges. In the formula the probability to generate a graph $\mathcal{G}_0$, given the choice of $\mathcal{M}$, translates from the above description as

$$\mathbb{P}(\mathcal{G}_0|\mathcal{M}) = \prod_{e \in \mathcal{M}} \hat{p}(w_e) \prod_{e \in \mathcal{E}_0 \setminus \mathcal{M}} p(w_e)$$
$$\times \left(\frac{c}{N}\right)^{|\mathcal{E}_0| - \frac{N}{2}} \left(1 - \frac{c}{N}\right)^{\binom{N}{2} - |\mathcal{E}_0|} \mathbb{I}(\mathcal{M} \subseteq \mathcal{E}_0),$$

$$(1)$$

where here and in the following $\mathbb{I}(A)$ denotes the indicator function of the event A. Note that the number of nonplanted edges concentrate in the large size (thermodynamic) limit $N \to \infty$ around its average value $cN/2$, assuming c remains fixed in this limit, and that these edges form essentially an Erdős-Rényi random graph of average degree c (modulo the exclusion of the planted edges).

The model studied in Refs. [3,10] corresponds to a dense, or fully connected, version of the model defined above, in which $\mathcal{G}_0$ is a complete weighted graph, $\mathcal{E}_0$, containing all the possible edges between the N vertices. In order to have a well-defined thermodynamic limit in this dense case it is necessary to rescale with N the weights on the nonplanted edges, i.e., to use an N-dependent distribution p. The simplest way to perform this rescaling is to use $p(w) = Q(w/N)/N$, where Q is a density with a support included in the nonnegative reals, and a positive density $Q(0) > 0$ at the origin [3,10,23]; without loss of generality we assume $Q(0) = 1$. The thermodynamic limit of this dense model is then equivalent to the large degree limit of the sparse one, $c \to \infty$ after $N \to \infty$, if one uses for the distribution p of the nonplanted edges the uniform distribution on the interval $[0, c]$. We shall thus study the richer sparse model, with finite c, and take the large degree limit when needed in order to compare our results with those of the dense case.

B. A statistical inference problem

The question we shall investigate in the following is whether the observation of a graph $\mathcal{G}_0$ generated according to the procedure above allows us to infer the hidden matching $\mathcal{M}$, assuming the observer knows the parameters c, p, and $\hat{p}$ of the model. In this setting all the information the observer can exploit to perform this task is contained in the posterior probability $\mathbb{P}(\mathcal{M}|\mathcal{G}_0)$. From the expression (1) of the graph generation probability, and the knowledge that the prior probability on $\mathcal{M}$ is uniform over the set of all perfect matchings, Bayes' theorem yields immediately the following expression for the posterior:

$$\mathbb{P}(\mathcal{M}|\mathcal{G}_0) \propto \prod_{e \in \mathcal{M}} \hat{p}(w_e) \prod_{e \in \mathcal{E}_0 \setminus \mathcal{M}} p(w_e) \mathbb{I}(\mathcal{M} \subseteq \mathcal{E}_0) \mathbb{I}_{\mathrm{pm}}(\mathcal{M}),$$

where the symbol $\propto$ hides a normalization constant independent of $\mathcal{M}$, and the last term is the indicator function of the event "$\mathcal{M}$ is a perfect matching." For notational simplicity it is convenient to encode a set of edges $\mathcal{M} \subseteq \mathcal{E}_0$ with binary variables, $\underline{m} = \{m_e : e \in \mathcal{E}_0\} \in \{0, 1\}^{\mathcal{E}_0}$, where $m_e = 1$ if and only $e \in \mathcal{M}$, and rewrite the posterior as

$$\mathbb{P}(\underline{m}|\mathcal{G}_0) \propto \prod_{e \in \mathcal{E}_0} \left[\frac{\hat{p}(w_e)}{p(w_e)}\right]^{m_e} \prod_{i=1}^{N} \mathbb{I}\left(\sum_{e \in \partial i} m_e = 1\right), \quad (2)$$

where ∂i denotes the set of edges incident to the vertex i. The observer will now compute an estimator $\widehat{\mathcal{M}}(\mathcal{G}_0)$, this function of the observations being "as close as possible" to the hidden matching $\mathcal{M}$. The optimal estimator actually depends on which notion of "closeness" between $\mathcal{M}$ and the estimator $\widehat{\mathcal{M}}(\mathcal{G}_0)$ is used.

If the measure of the distance between them is simply the indicator function $\mathbb{I}(\mathcal{M} \neq \widehat{\mathcal{M}}(\mathcal{G}_0))$, then the optimal

estimator, optimal in the sense that it minimizes this distance averaged over all realizations of the problem, is the one maximizing the posterior,

$$\widehat{\mathcal{M}}_{\mathrm{b}}(\mathcal{G}_0) = \underset{\underline{m}}{\mathrm{argmax}}\ \mathbb{P}(\underline{m}|\mathcal{G}_0), \quad (3)$$

where we slightly abused notations and used freely the equivalence between $\underline{m}$ and $\mathcal{M}$. Following the nomenclature of error-correcting codes [7] we shall call this the block maximal *a posteriori* (MAP) estimator, hence the subscript b. As we shall detail below the estimator $\widehat{\mathcal{M}}_{\mathrm{b}}$ is the perfect matching of $\mathcal{G}_0$ which minimizes the sum of some effective weights on the edges it contains.

If instead the distance to be minimized is the total number of misclassified edges, $|\mathcal{M} \triangle \widehat{\mathcal{M}}(\mathcal{G}_0)|$, with $\triangle$ the symmetric difference between sets, or equivalently the Hamming distance between the binary strings $\underline{m}$ and $\widehat{\underline{m}}(\mathcal{G}_0)$ encoding them, then the optimal estimator is the so-called symbol MAP one, denoted $\widehat{\mathcal{M}}_{\mathrm{s}}(\mathcal{G}_0)$, defined by the binary string $\widehat{\underline{m}}(\mathcal{G}_0)$ where, for all the edges $e \in \mathcal{E}_0$,

$$\widehat{m}_e(\mathcal{G}_0) = \underset{m_e}{\mathrm{argmax}}\ \mathbb{P}_e(m_e|\mathcal{G}_0), \quad (4)$$

with $\mathbb{P}_e$ the marginal of the posterior probability (2) for the edge e. Note that this estimator is not necessarily a perfect matching, nevertheless it is the one that minimizes the distance $|\mathcal{M} \triangle \widehat{\mathcal{M}}(\mathcal{G}_0)|$ on average over all realizations of the problem.

For future use let us define the (reduced) distance between the planted matching $\mathcal{M}$ and an arbitrary estimator $\widehat{\mathcal{M}}$ (i.e., a subset of the edge set $\mathcal{E}_0$) as

$$\varrho(\mathcal{M}, \widehat{\mathcal{M}}) = \frac{1}{N}|\mathcal{M} \triangle \widehat{\mathcal{M}}| \quad (5)$$

$$= \frac{1}{N} \sum_{e \in \mathcal{M}} \mathbb{I}(\widehat{m}_e = 0) + \frac{1}{N} \sum_{e \in \mathcal{E}_0 \setminus \mathcal{M}} \mathbb{I}(\widehat{m}_e = 1). \quad (6)$$

If $\widehat{\mathcal{M}}$ contains exactly $N/2$ edges, in particular if it is a perfect matching, this expression can be simplified as the two terms contribute in the same way (there are as many false positive as true negative errors in the identification of the edges), then

$$\varrho(\mathcal{M}, \widehat{\mathcal{M}}) = \frac{2}{N} \sum_{e \in \mathcal{M}} \mathbb{I}(\widehat{m}_e = 0). \quad (7)$$

Our goal in the rest of the article is to discuss the quality of the estimators defined above, in the thermodynamic limit $N \to \infty$, as a function of the parameters of the model. Following the studies [3,10] one expects to find phase transitions between full-recovery phases, in which all but a vanishing fraction of the edges of $\mathcal{M}$ can be recovered from the observation of $\mathcal{G}_0$, characterized by a vanishing average reconstruction error $\mathbb{E}[\varrho] = 0$, and partial-recovery phases where a positive fraction of the edges will be misclassified, $\mathbb{E}[\varrho] > 0$. Before entering the actual computations let us make two simple remarks in order to give the reader a first intuitive idea of the effect of the parameters on the inference difficulty. (1) The identification of the planted edges will be easier if the distributions p and $\hat{p}$ are less similar one to the other; in the extreme cases where $p = \hat{p}$ the weights contain absolutely no

information on $\mathcal{M}$, while if p and $\hat{p}$ have disjoint supports $\mathcal{M}$ can be identified by a simple inspection of the weights on the edges. (2) For a fixed choice of p and $\hat{p}$ the parameter c corresponds to a noise level: if c is very small $\mathcal{E}_0$ contains essentially only the sought-for edges of $\mathcal{M}$, increasing it the latter are hidden in the confusing nonplanted edges.

III. CAVITY METHOD EQUATIONS

A. A first pruning of the graph

Before proceeding further, let us observe that the inference problem can be in general reduced in size after some simple, preliminary observations. Following the remark (1) in a less drastic case, suppose that the supports of p and $\hat{p}$ are different (but not necessarily disjoint). Then an edge e bearing a weight w_e in the support of p but not in the one of $\hat{p}$ is, without doubt, nonplanted; conversely e is certainly planted if w_e is in the support of $\hat{p}$ but not in the one of p. All the edges identified in this way can be eliminated from $\mathcal{G}_0$; moreover the two vertices belonging to an edge identified as planted can also be eliminated, as well as the other edges incident to them, that cannot be planted by definition of a perfect matching.

To put these remarks on a quantitative ground let us denote $\text{supp}(\hat{p}) := \{w \in \mathbb{R} : \hat{p}(w) > 0\}$ (more precisely the closure of this set) the support of the distribution $\hat{p}$, and similarly $\text{supp}(p) := \{w \in \mathbb{R} : p(w) > 0\}$ the support of p. We define

$$\Gamma := \text{supp}(p) \cap \text{supp}(\hat{p}), \tag{8a}$$

$$\mu := \int_\Gamma p(w)\, dw, \tag{8b}$$

$$\hat{\mu} := \int_\Gamma \hat{p}(w)\, dw. \tag{8c}$$

A nonplanted edge e has weight $w_e \notin \Gamma$ and can thus be identified, with probability $1 - \mu$. Similarly, a planted edge e will have $w_e \notin \Gamma$ with probability $1 - \hat{\mu}$, and, in this case, it is surely an element of $\mathcal{M}$. We will denote

$$\mathcal{M}_0 := \{e \in \mathcal{E}_0 : w_e \in \text{supp}(\hat{p}) \setminus \text{supp}(p)\} \tag{9}$$

the set of planted edges immediately recognizable by means of these simple considerations. The edges in $\mathcal{M}_0$ can be removed from the graph, alongside with their endpoints and all edges incident to them. After this pruning process, the obtained graph has, on average and in the large N limit, $\hat{\mu}N$ surviving vertices, each of them with degree $1 + Z$, Z being a Poisson random variable of mean $\gamma := c\mu\hat{\mu}$ (each nonplanted edge is present with probability $\mu \frac{c}{N}$, but in a graph with $\hat{\mu}N$ vertices).

The distribution of the weights of the surviving edges is now conditioned to the fact that their values are in Γ. On the new pruned graph, therefore, the weight distributions are

$$P(w) := \frac{1}{\mu} p(w) \mathbb{I}(w \in \Gamma), \tag{10a}$$

$$\hat{P}(w) := \frac{1}{\hat{\mu}} \hat{p}(w) \mathbb{I}(w \in \Gamma) \tag{10b}$$

for the nonplanted and planted edges, respectively. We will denote $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \underline{w})$ the graph obtained after this pruning, with $\mathcal{V} \subseteq \mathcal{V}_0$ and $\mathcal{E} \subseteq \mathcal{E}_0$ the new vertex and edge sets.

B. The belief propagation equations

Here we present the belief propagation algorithm that was used in Ref. [3] for the planted matching problem. Let us introduce a positive parameter (fictitious inverse temperature) β, and consider the following probability distribution over the configurations $\underline{m} = \{m_e : e \in \mathcal{E}\} \in \{0, 1\}^{\mathcal{E}}$ of binary variables on the edges of a weighted graph $\mathcal{G}$,

$$\nu(\underline{m}) \propto e^{-\beta \sum_{e \in \mathcal{E}} m_e \omega_e} \prod_{i \in \mathcal{V}} \mathbb{I}\left(\sum_{e \in \partial i} m_e = 1\right), \tag{11}$$

where we introduced effective weights ω_e on the edges, that are computed from the observed weights w_e as $\omega_e = \omega(w_e)$, with

$$\omega(w) := -\ln \frac{\hat{P}(w)}{P(w)}. \tag{12}$$

To lighten the notation we kept implicit the dependency on β and $\mathcal{G}$ of the probability distribution ν; for $\beta = 1$ it coincides with the posterior defined in (2), when $\beta \to \infty$ it concentrates on the configurations maximizing the posterior, these two values of β allow us thus to deal with the symbol and block MAP estimators, respectively.

The probability distribution ν defined in Eq. (11) has the form of a Gibbs measure over all weighted perfect matchings of the graph $\mathcal{G}$. The exact computation of its marginals is an intractable task in general; we shall instead study it in an approximate way, using the belief propagation (BP) algorithm (see, for instance, Ref. [11] for a general introduction to BP as well as chapter 16 therein for its application to matching problems), that is conjectured to provide an asymptotically exact description in the large size limit for these sparse random graphs. One can indeed consider Eq. (11) as a graphical model and construct a factor graph starting from the graph $\mathcal{G}$,
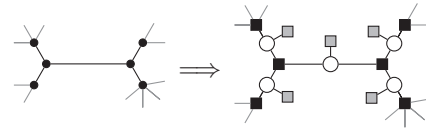

In this factor graph representation, each edge e is associated to a variable node $\bigcirc$, corresponding to the variable m_e. Each variable node is attached to two "function nodes" $\blacksquare$ located on the vertices of the original graph. Such function nodes express the constraint that the variables on the neighbor variable nodes have to sum to 1. Finally, a different kind of function node $\square$ is attached to each variable node, corresponding to the "local field" term $e^{-\beta \omega_e}$ in the measure (11).

The BP equations are then obtained by introducing "messages" on the edges of this factor graph, that mimic the marginal probabilities in amputated graphical models and would become exact if the factor graph were a tree. For the model at hand these messages are of the form $\nu_{i \to e}(m)$, from a vertex i to an edge $e = (i, j)$, and obey the following equations (one for each directed edge of the graph),

$$\nu_{i \to e}(m) \propto \sum_{\{m_{\tilde{e}}\}_{\tilde{e} \in \partial i \setminus e}} \mathbb{I}\left(m + \sum_{\tilde{e} \in \partial i \setminus e} m_{\tilde{e}} = 1\right) \prod_{\substack{\tilde{e} = (r, i) \\ \tilde{e} \in \partial i \setminus e}} \nu_{r \to \tilde{e}}(m_{\tilde{e}}) e^{-\beta m_{\tilde{e}} \omega_{\tilde{e}}}. \tag{13}$$

We adopt the convention $\sum_{a \in A} f(a) = 0$ and $\prod_{a \in A} f(a) = 1$ if $A = \emptyset$ for any function f. Pictorially, Eq. (13) can be rendered as
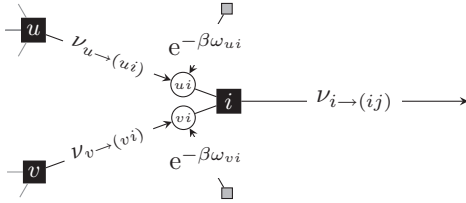
where the arrows indicate the directions of "propagation" of the messages.

As the variables are binary, $m_e \in \{0, 1\}$ for each $e \in \mathcal{E}$, the messages can be conveniently parametrized in terms of "cavity fields" $h_{i \to e}$, one real number for each directed edge, as

$$\nu_{i \to e}(m) := \frac{e^{\beta m h_{i \to e}}}{1 + e^{\beta h_{i \to e}}}, \qquad (14)$$

so that the BP equations become in terms of the cavity fields:

$$h_{i \to e} = -\frac{1}{\beta} \ln \left(\sum_{\tilde{e} = (r,i) \in \partial i \setminus e} e^{-\beta(\omega_{\tilde{e}} - h_{r \to \tilde{e}})} \right). \qquad (15)$$

Once a solution of the set of BP equations has been found (for instance, by iterating them starting from a random or zero initial condition until convergence to a fixed point is reached), the BP approximation of the marginal probability of the variable m_e on the edge $e = (i, j)$ is given by

$$\nu_e(m) = \frac{e^{\beta m(h_{i \to e} + h_{j \to e} - \omega_e)}}{1 + e^{\beta(h_{i \to e} + h_{j \to e} - \omega_e)}}. \qquad (16)$$

The BP approximation to the symbol MAP estimator defined in (4) is thus obtained by solving the BP equations with $\beta = 1$, and estimating as a planted edge those for which $\nu_e(1) > 1/2$, namely,

$$\widehat{\mathcal{M}}_s(\mathcal{G}) := \left\{ e \in \mathcal{E} : \nu_e(1) > \tfrac{1}{2} \right\}$$
$$= \{ e = (i, j) \in \mathcal{E} : h_{i \to e} + h_{j \to e} > \omega_e \}. \qquad (17)$$

We will keep the same rule (17) for the conversion of a solution of the BP equations into an estimator of $\mathcal{M}$ for all values of β, and in particular for $\beta \to \infty$. If the block MAP configuration is unique, and if the marginal probabilities ν_e are computed exactly, then this is a legitimate way of determining the block MAP estimator (3). The BP algorithm is of course only an approximation here, but we conjecture it to be asymptotically exact, i.e., that the reduced Hamming distance between the block MAP estimator and its BP version vanishes in the thermodynamic limit. This relies on rigorous works that, even if they do not directly apply to the case considered here, have proven the exactness of the BP algorithm in similar settings. More precisely, Ref. [24] proved that for a given bipartite weighted graph, if the perfect matching with minimal weight is unique then the $\beta \to \infty$ version of the BP equations, associated to the inclusion rule (17), converges to the optimal configuration, in a number of iterations that scale with the gap between the optimal weight and the second minimum, and with the largest weight in the graph. Reference [25]

improved this convergence rate for typical bipartite graphs of a random ensemble, while Refs. [26,27] removed the bipartiteness assumption but added a hypothesis on the absence of fractional solutions for the linear programming relaxation of the problem.

C. Recursive distributional equations

The BP equations have been written in Eqs. (15) for a given instance of the graph $\mathcal{G}$; to obtain the average error on the ensemble of all possible instances of our problem we need to describe the statistics of the solutions of the BP equations. This step is known as density evolution in the context of error-correcting codes or as the cavity method in statistical mechanics [28]. We refer the reader to Refs. [29,30] for similar studies of the matchings in sparse, nonplanted random graphs, and to Refs. [31,32] which considered the weighted case (still without a planted structure).

Suppose that an instance is generated at random, that the BP equations are solved on it, and that a directed planted edge is chosen uniformly at random, say $i \to e$; let us call $\hat{\mathsf{H}}$ the random variable that has the law of the cavity field $h_{i \to e}$. We define similarly H as the random variable distributed as $h_{i \to e}$ when one chooses a nonplanted edge. Let us also introduce the random variables $\Omega = \omega(W)$ and $\hat{\Omega} = \omega(\hat{W})$, where W (resp. $\hat{W}$) is a random variable with density P (resp. $\hat{P}$). If one assumes that the typical realizations of ν have no long-range correlations (the so-called replica symmetric (RS) hypothesis), then (15) translates into recursive distributional equations (RDEs) between the random variables H and $\hat{\mathsf{H}}$. A vertex i in a directed planted edge $i \to e$ is incident to a Poissonian number of other nonplanted edges because of the Erdős-Rényi nature of the latter, and similarly if $i \to e$ is nonplanted there will be exactly one planted edge incident to i, and other nonplanted edges from the Erdős-Rényi part of the graph. With the RS assumption of independence of the incoming cavity fields one thus obtains

$$\hat{\mathsf{H}} \stackrel{\mathrm{d}}{=} -\frac{1}{\beta} \ln \left(\sum_{i=1}^{\mathsf{Z}} e^{-\beta(\Omega_i - \mathsf{H}_i)} \right), \qquad (18a)$$

$$\mathsf{H} \stackrel{\mathrm{d}}{=} -\frac{1}{\beta} \ln \left(e^{-\beta(\hat{\Omega} - \hat{\mathsf{H}})} + \sum_{i=1}^{\mathsf{Z}} e^{-\beta(\Omega_i - \mathsf{H}_i)} \right)$$

$$\stackrel{\mathrm{d}}{=} -\frac{1}{\beta} \ln(e^{-\beta(\hat{\Omega} - \hat{\mathsf{H}})} + e^{-\beta \hat{\mathsf{H}}'}). \qquad (18b)$$

In the equations above all random variables are independent, Z is Poisson distributed with mean γ, the Ω_i's have the same law as Ω, and similarly H_i's are independent copies of H, and $\hat{\mathsf{H}}'$ of $\hat{\mathsf{H}}$.

The average of the reconstruction error defined in (6) can be computed in this setting, recalling the inclusion rule (17), as

$$\mathbb{E}[\varrho] = \frac{\hat{\mu}}{2} \mathbb{P}[\hat{\mathsf{H}} + \hat{\mathsf{H}}' \leqslant \hat{\Omega}] + \frac{\hat{\mu}\gamma}{2} \mathbb{P}[\mathsf{H} + \mathsf{H}' > \Omega]. \qquad (19)$$

D. A second pruning of the graph

Our goal in the following will be to understand the properties of the random variables H and $\hat{\mathsf{H}}$ solutions of (18),

and their possible bifurcations when the parameters of the model are varied. As a first step in this direction we shall isolate the contribution of "hard fields," in other words the probabilities of the events $\hat{H} = +\infty$ and $H = -\infty$ for these random variables. Observe indeed that $\mathbb{P}[Z = 0] > 0$ in (18a), which leads to $\hat{H} = +\infty$, and that this event implies $H = -\infty$ in (18b). From both theoretical and practical point of views it is convenient to deal with these events explicitly, and we shall thus introduce the probabilities $\hat{q}$ and q of the events $\hat{H} = +\infty$ and $H = -\infty$, respectively, and two new random variables $\hat{H}$ and H that have the law of $\hat{H}$ and H conditional on being finite (we exclude the possibility of $\hat{H} = -\infty$ and $H = +\infty$, $\hat{H}$ and H are finite with probability one). In formulas these definitions amount to

$$H \stackrel{d}{=} \begin{cases} -\infty & \text{with prob. } q, \\ H & \text{with prob. } 1-q, \end{cases} \tag{20a}$$

$$\hat{H} \stackrel{d}{=} \begin{cases} +\infty & \text{with prob. } \hat{q}, \\ \hat{H} & \text{with prob. } 1-\hat{q}. \end{cases} \tag{20b}$$

Let us insert them in (18) in order to obtain the equations obeyed by q, $\hat{q}$, H, and $\hat{H}$. In the right-hand side of (18a) the number of infinite and finite H_i's are easily seen to be two independent Poisson random variables of parameters γq and $\gamma(1-q)$, respectively. $\hat{H}$ is infinite if and only if the second of this number vanishes, hence one has $\hat{q} = e^{-\gamma(1-q)}$ and

$$\hat{H} \stackrel{d}{=} -\frac{1}{\beta} \ln \left(\sum_{i=1}^{Z} e^{-\beta(\Omega_i - H_i)} \right), \tag{21}$$

where Z has the law of a Poisson random variable of parameter $\gamma(1-q)$ conditioned to be strictly positive, i.e., $\mathbb{P}[Z = k] = \frac{[\gamma(1-k)]^k}{k!(e^{\gamma(1-q)}-1)}$ for $k \geqslant 1$. In (18b) one sees that H is infinite if and only if $\hat{H}$ is infinite, hence $q = \hat{q}$ and

$$H \stackrel{d}{=} \begin{cases} \hat{\Omega} - \hat{H} & \text{with prob. } \hat{q}, \\ -\frac{1}{\beta} \ln(e^{-\beta(\hat{\Omega}-\hat{H})} + e^{-\beta \hat{H}'}) & \text{with prob. } 1-\hat{q}. \end{cases} \tag{22}$$

In summary the elimination of the hard fields amount to find the solution q of

$$q = e^{-\gamma(1-q)}, \tag{23}$$

and to study the finite random variables H and $\hat{H}$ solution of the RDEs

$$\hat{H} \stackrel{d}{=} -\frac{1}{\beta} \ln \left(\sum_{i=1}^{Z} e^{-\beta(\Omega_i - H_i)} \right), \tag{24a}$$

$$H \stackrel{d}{=} \begin{cases} \hat{\Omega} - \hat{H} & \text{with prob. } q, \\ -\frac{1}{\beta} \ln(e^{-\beta(\hat{\Omega}-\hat{H})} + e^{-\beta \hat{H}'}) & \text{with prob. } 1-q, \end{cases} \tag{24b}$$

where the variable Z in Eq. (24a) has distribution $\mathbb{P}[Z = k] = \pi_k$ with

$$\pi_k := \frac{q}{1-q} \frac{[(1-q)\gamma]^k}{k!} \qquad \text{for } k \geqslant 1. \tag{25}$$

The average reconstruction error (19) can be reexpressed in terms of the new random variables H and $\hat{H}$ as

$$\mathbb{E}[\varrho] = \frac{\hat{\mu}(1-q)^2}{2} \mathbb{P}[\hat{H} + \hat{H}' \leqslant \hat{\Omega}]$$
$$+ \frac{\hat{\mu}(1-q)^2 \gamma}{2} \mathbb{P}[H + H' > \Omega]. \tag{26}$$

This procedure of "hard fields" elimination that we explained on the RDEs admits also an interpretation on a single graph instance. As a matter of fact the presence of infinite fields on the planted edges can be traced back to the BP equation (15), which shows that $h_{i \to e} = +\infty$ if i is a leaf of the graph, i.e., i is of degree 1 and its only incident edge is e. But this fact allows us to unambiguously identify e as an edge of the planted matching, which by definition covers all the vertices of the graph. Then i, the edge $e = (i, j)$, the vertex j and all the edges incident to it can be removed, the latter being with certainty nonplanted edges. This leaf removal procedure can be iterated until either all the graph has been pruned, or stops when a nontrivial core without any leaf has been reached. The propagation of infinite fields in the BP equations is an equivalent way of describing this pruning algorithm. Note that such a leaf removal procedure has already been studied in the literature for standard Erdős-Rényi graphs [33,34], in this case a core percolation transition is found when the average degree of the graph crosses the Euler number value e: for sparser graphs the leaf removal procedure typically destroys the whole graph, while a nontrivial core survives for larger average degrees. In our case the core percolation transition happens at $\gamma = 1$ (which corresponds to the usual percolation transition of the Erdős-Rényi random graph superposed to the planted matching): if $0 < \gamma \leqslant 1$ the self-consistent equation (23) on q admits only the solution $q = 1$, which means that the leaf removal procedure allows us to recover completely the planted matching (up to a subextensive number of edges in the thermodynamic limit). On the contrary for $\gamma > 1$ the leaf removal stops with a nontrivial core [this explains why one should take the solution $q < 1$ of (23) when $\gamma > 1$]. We shall see in the following that full-recovery phases can exist also for $\gamma > 1$, but in that case the simple leaf removal procedure is not able to identify all the edges of the planted matching, the full recovery is due to a nontrivial amplification effect, by the iterations of the BP equations, of the information contained in the weights of the edges of the core.

IV. THE LOCATION OF THE PHASE TRANSITION FOR THE BLOCK MAP ESTIMATOR

A. RDEs for the block MAP

As explained in Sec. II B the block MAP estimator, that maximizes the probability of correct identification of the whole planted matching, is the configuration $\underline{m}$ that maximizes the posterior in Eq. (2), and can be obtained by taking the "zero-temperature" limit $\beta \to \infty$ in the probability distribution ν defined in (11). The BP equations that we wrote in (15) for a generic value of β become in this limit

$$h_{i \to e} = \min_{\tilde{e} = (r,i) \in \partial i \setminus e} [\omega_{\tilde{e}} - h_{r \to \tilde{e}}]; \tag{27}$$

the configuration maximizing the posterior (2) can be equivalently defined as the perfect matching of minimum cost on the weighted graph $(\mathcal{V}, \mathcal{E}, \underline{\omega})$, with the effective weights ω_e replacing the observed weights w_e. Hence these BP equations coincide with those written in Ref. [24] to study such minimum weight matching problems. With the inclusion criterion of (17) the BP approximation for the block MAP configuration is determined as

$$m_e = \mathbb{I}(h_{i \to e} + h_{j \to e} - \omega_e > 0). \tag{28}$$

The probabilistic treatment of the BP equations can also be specialized very easily to the limit case $\beta \to +\infty$, in particular the RDEs (24) yield

$$\hat{H} \overset{d}{=} \min_{1 \leqslant i \leqslant Z} [\Omega_i - H_i], \tag{29a}$$

$$H \overset{d}{=} \begin{cases} \hat{\Omega} - \hat{H} & \text{with prob. } q, \\ \min(\hat{\Omega} - \hat{H}, \hat{H}') & \text{with prob. } 1 - q, \end{cases} \tag{29b}$$

with the law of the random variable Z defined in (25). The average reconstruction error (26) can actually be simplified for this $\beta \to \infty$ situation into

$$\mathbb{E}[\varrho] = \hat{\mu}(1 - q)^2 \mathbb{P}[\hat{H} + \hat{H}' \leqslant \hat{\Omega}]; \tag{30}$$

as discussed in Sec. II B [see in particular (7)] there are as many misclassified planted and nonplanted edges when the estimator contains $N/2$ edges, which we argued to be asymptotically the case for $\beta \to \infty$, hence the two terms in (26) are equal. For completeness we show in Appendix A that this equality follows indeed from the RDE (29), modulo a hypothesis of continuity for the distributions of H and $\hat{H}$, that mimics the hypothesis of uniqueness of the block MAP assignment.

The equalities in distribution between random variables stated in (29) can be equivalently rephrased as equations between the cumulative distribution functions of the variables H and $\hat{H}$. For a random variable X we shall define the c.d.f. F_X and its reciprocal $\bar{F}_X$ according to

$$F_X(x) = \mathbb{P}[X \leqslant x], \quad \bar{F}_X(x) = 1 - F_X(x) = \mathbb{P}[X > x].$$

One obtains from (29)

$$\bar{F}_H(h) = [q + (1 - q)\bar{F}_{\hat{H}}(h)] \int_\Gamma F_{\hat{H}}[\omega(w) - h]\hat{P}(w)\, dw \tag{31a}$$

and

$$\bar{F}_{\hat{H}}(h) = \frac{q}{1-q} \sum_{k=1}^\infty \frac{[\gamma(1-q)]^k}{k!} \left[\int_\Gamma F_H(\omega(w) - h) P(w)\, dw \right]^k$$

$$= \frac{\exp\left\{ -\gamma(1-q) \int_\Gamma \bar{F}_H[\omega(w) - h] P(w)\, dw \right\} - q}{1 - q}. \tag{31b}$$

These are integral nonlinear equations on the two functions F_H and $F_{\hat{H}}$, describing the thermodynamic limit of the planted matching problem. These recursive equations can also be understood as describing the optimal matching problem on an infinite tree. The authors of Ref. [10] show rigorously that for fully connected graphs the optimal configuration of the finite graph locally converges to the optimal matching of the infinite tree.

The above integral nonlinear equations are quite complicated to solve in general. It was shown in Ref. [10], that, for a rather specific case (in the large degree limit with an exponential distribution for the planted weights) these integral equations can be transformed into a system of ordinary differential equations (which we will detail in Sec. VI). Unfortunately such a simplification does not seem to hold besides this special case.

The question now is to understand the solution of (29) [or equivalently of (31)] as a function of the parameters of the model. It is easy to check that $\hat{H} = +\infty$, $H = -\infty$ [i.e., $F_{\hat{H}}(h) = 0$, $F_H(h) = 1$] is always a solution, for every choice of the parameters. If this is the correct solution then $\mathbb{E}[\varrho] = 0$, in other words one is in a full-recovery phase. The picture that emerges from the previous works [3,10] is that for some value of the parameters another solution of (29) exists and is attractive when running BP from an initial condition uncorrelated with the planted matching. One is then in a partial-recovery phase, with $\mathbb{E}[\varrho] > 0$. On the contrary if $\hat{H} = +\infty$, $H = -\infty$ is the only solution of (29) one is in a full-recovery phase, the hidden matching being sufficiently attractive to drive the iterations of the BP equations towards it. According to this description the phase transition between full and partial recovery corresponds to the disappearance of a nontrivial solution of the RDE (here and in the following nontrivial means distinct from $\hat{H} = +\infty$, $H = -\infty$). This can of course be studied numerically, and we shall display later on some results obtained in this way; in the next subsection we shall argue that, with some additional hypotheses on the nature of the transition one can compute its location analytically.

B. Locating the transition

Let us assume that the quantities p, $\hat{p}$, and c defining the model depend on some continuous parameter denoted λ, and that there exists a threshold value $\bar{\lambda}$ such that $\mathbb{E}[\varrho](\lambda) = 0$ for $\lambda > \bar{\lambda}$, and $\mathbb{E}[\varrho](\lambda) > 0$ for $\lambda < \bar{\lambda}$. We further assume that the transition at $\bar{\lambda}$ is continuous, i.e., $\mathbb{E}[\varrho](\lambda) \to 0$ as $\lambda \to \bar{\lambda}^-$. Under these hypotheses, the expression (30) of $\mathbb{E}[\varrho]$ reveals that $\mathbb{P}[\hat{H} + \hat{H}' > \hat{\Omega}] \to 1$ when λ reaches its threshold value from the partial-recovery phase. But if $\mathbb{P}[\hat{H} + \hat{H}' > \hat{\Omega}] = 1$ the minimum in (29b) is always realized by the first argument, which leads us to study the following, simplified form of the RDE (29):

$$\hat{K} \overset{d}{=} \min_{1 \leqslant r \leqslant Z} [\Omega_r - K_r], \tag{32a}$$

$$K \overset{d}{=} \hat{\Omega} - \hat{K}, \tag{32b}$$

which bear on a new couple of random variables $\hat{K}$ and K. The transition point will be characterized by the fact that the simplified RDE in Eqs. (32) has a nontrivial solution at $\bar{\lambda}$. To facilitate the discussion we define a new random variable,

$$\Xi \overset{d}{:=} \Omega - \hat{\Omega}, \tag{33}$$

in terms of which Eqs. (32) can be written as a distributional equation for $\hat{K}$ only,

$$\hat{K} \overset{d}{=} \min_{1 \leqslant r \leqslant Z} [\Xi_r + \hat{K}_r]. \tag{34}$$

This RDE is actually connected to the properties of branching random walk (BRW) processes, a subject that has generated a vast literature in both physics and mathematics [12–22], in the more general context of front propagation for reaction-diffusion equations. For the convenience of the reader we summarize here the definitions and the main properties we need about BRWs, more details can be found in Appendix B. A BRW describes the evolution of a population of particles that move along a continuous unidimensional spatial axis and multiply as time increases in discrete steps (the equivalent process in continuous time being the branching Brownian motion). More explicitly, at the initial generation $n = 0$ there is a single particle at the origin, $X_1^{(0)} = 0$. Each generation n is given by a set of particles in positions $\{X_k^{(n)}\}_k$ and is constructed iteratively. Each particle of the generation n, say, the ith one at position $X_i^{(n)}$, gives rise to a number (possibly infinite) of offsprings in the next generation, located at the positions $X_{i,r}^{(n+1)} = X_i^{(n)} + \Xi_{i,r}$ where the displacements $\{\Xi_{i,r}\}_r$ between the positions of a parent particle and its offsprings are, independently for each i, copies of an identical point process. In the simplest cases the number of offsprings is Z_i, an independent copy of the random variable Z for each parent i, and the displacements $\Xi_{i,r}$ are i.i.d. copies of a given random variable Ξ.

A realization of such a process is pictured on the example below:
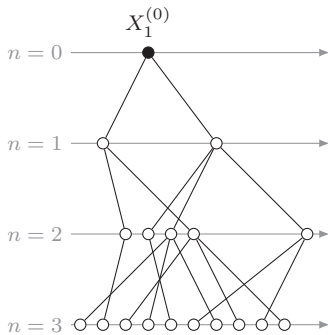
Note that BRWs combine and generalize Galton-Watson branching processes, that are recovered if one looks only at the number of particles in the BRW and discards their positions, and unidimensional random walks: the positions of the particles along a single branch of the BRW follow the law of a simple random walk.

Among several properties of BRWs one that has attracted significant research effort is the asymptotic behavior in the large n limit of the minimum of the process. Let us denote $\hat{K}^{(n)} = \min_k [X_k^{(n)}]$ the position of the leftmost particle in the nth generation. Decomposing a BRW of depth $n + 1$ into Z BRW of depth n attached to the root via Z displacements Ξ_r it is easy to convince oneself that $\hat{K}^{(n)}$ obey the following RDE:

$$\hat{K}^{(n+1)} \overset{d}{=} \min_{1 \leqslant r \leqslant Z} [\Xi_r + \hat{K}_r^{(n)}], \tag{35}$$

with the initial condition $\hat{K}^{(0)} = 0$ and the convention $\hat{K}^{(n)} = -\infty$ if the process is extinct before the nth generation. Such sequences of random variables have been extensively studied, and very precise mathematical results have been obtained. A first level of description [12–14] shows that, conditional on the nonextinction of the process, $\hat{K}^{(n)}$ has a ballistic behavior, namely, $\hat{K}^{(n)}/n \xrightarrow[n \to +\infty]{\text{a.s.}} v$, with an almost sure convergence towards a velocity v that can be computed in terms of the law of the point process of the displacements:

$$v = -\inf_{\theta > 0} \frac{1}{\theta} \ln \mathbb{E}\left[\sum_r e^{-\theta \Xi_r}\right]. \tag{36}$$

When the number of offspring is a random variable of law Z and the displacements i.i.d. copies of Ξ this simplifies into

$$v = -\inf_{\theta > 0} \frac{1}{\theta} \ln(\mathbb{E}[Z]\mathbb{E}[e^{-\theta \Xi}]); \tag{37}$$

see also Appendix B for a heuristic justification of this expression of the velocity. More recently a finer description of the limit has been obtained [18,20,22]: under some technical conditions there exists a constant C, such that, conditional on the nonextinction of the process,

$$\hat{K}^{(n)} - nv - C \log n \xrightarrow[n \to +\infty]{d} L, \tag{38}$$

where the convergence is in law and L is a finite random variable satisfying

$$L \overset{d}{=} -v + \min_{1 \leqslant r \leqslant Z} [\Xi_r + L_r]. \tag{39}$$

Note that this equation is invariant by translation: if L is a solution, then $L + a$ also is, for any constant a. Moreover the left tail behavior of the limit random variable L was established in Refs. [20,22] to be

$$\mathbb{P}[L \leqslant z] \sim \alpha z \, e^{\theta_* z} \quad \text{as } z \to -\infty, \tag{40}$$

where θ_* is the minimizer of (37), and $\alpha < 0$ a constant.

Let us now come back to the planted matching problem and specialize these results taking for the law Z of the offspring of the BRW the expression (25), and for Ξ the random variable defined in (33), where we recall that Ω (resp. $\hat{\Omega}$) is the random variable $\omega(W)$ with the function ω of (12) and W drawn with the distribution P (resp. $\hat{P}$). Note that in our case $Z \geqslant 1$ with probability 1, hence the probability of extinction of the BRW process is equal to zero. The expression of the velocity (37) becomes then

$$v = -\inf_{\theta > 0} \frac{\ln [I(\theta)I(1 - \theta)]}{\theta}, \tag{41}$$

with

$$I(\theta) = \sqrt{\gamma} \int_\Gamma \hat{P}(w)^\theta P(w)^{1-\theta} \, dw, \tag{42}$$

where we remind readers of the definition of Γ in Eq. (8a). One realizes at this point that if the parameters of the problem are such that $v = 0$, then the random variable L solution of (39) and constructed through the large generation limit of the BRW is a nontrivial solution of Eq. (34): this is precisely the condition we argued to be satisfied at the continuous

phase transition between full and partial-recovery phases. The vanishing velocity criterion

$$\inf_{\theta > 0} \frac{\ln [I(\theta)I(1-\theta)]}{\theta} = 0 \tag{43}$$

is actually equivalent to $I(1/2) = 1$ because the argument of the logarithm is a convex function symmetric around $\theta = 1/2$, and therefore this condition becomes

$$\int_\Gamma \sqrt{\hat{P}(w)P(w)} \, dw = \frac{1}{\sqrt{\gamma}}, \tag{44}$$

or equivalently in terms of the original parameters:

$$\int_\Gamma \sqrt{\hat{p}(w)p(w)} \, dw = \frac{1}{\sqrt{c}}. \tag{45}$$

Note that the Cauchy-Schwarz inequality implies

$$\int_\Gamma \sqrt{\hat{P}(w)P(w)} \, dw \leqslant 1, \tag{46}$$

hence (44) cannot be satisfied if $\gamma < 1$. This is perfectly consistent with what we found in Sec. III D, if $\gamma < 1$ the only solution to (23) is $q = 1$, signaling a phase where full recovery can be achieved by the leaf removal procedure.

Equation (45) is our first main result; it provides a prediction for the locus of the continuous phase transition in the parameter space $(p, \hat{p}, c)$ of the model. We shall simplify it in the large degree limit in Sec. IV C, and test it numerically on several examples in Sec. V. Before that we shall make a series of remarks on the reasoning which led to it and on its consequences.

(1) We have implicitly assumed that $v = 0$ is a necessary condition for (34) to have a nontrivial solution, in other words that the solution of (39) is unique (modulo the invariance under translations) and can thus be realized as the (properly shifted) large n limit of the BRW construction. This uniqueness is actually an open question in mathematics, stated as open problem 46 in Ref. [19].

(2) We justified the introduction of the simplified RDE (32) by an assumption on the continuity of $\mathbb{E}[\varrho]$. We can be more precise in some cases; suppose that the random variable $\hat{\Omega}$ is not bounded from above, i.e., that $\omega(w)$ diverges to $+\infty$ at some point in Γ (as we will see later there are nontrivial examples where this property can be true or false). Then a continuously vanishing $\mathbb{E}[\varrho]$ in Eq. (30) can occur only if $\hat{H}$ diverges to $+\infty$ as $\lambda \to \bar{\lambda}^-$. In that case we can restate more precisely our hypothesis as the existence of a function $m(\lambda)$ that diverges to $+\infty$ as $\lambda \to \bar{\lambda}^-$, such that

$$\hat{H} - m(\lambda) \xrightarrow[\lambda \to \bar{\lambda}^-]{d} \hat{K}, \tag{47}$$

with $\hat{K}$ solution of (34). The case studied in Ref. [10] falls in this category, and Sec. VI will be devoted to the determination of the divergence of $m(\lambda)$.

(3) Independently of the continuity assumption, a point in parameter space with $v > 0$ is most certainly in a full-recovery phase, according to the following reasoning. Instead of the fixed point condition (29) consider an iterative version of these equations,

$$\hat{H}^{(n+1)} \stackrel{d}{=} \min_{1 \leqslant i \leqslant Z} \left[\Omega_i - H_i^{(n)} \right], \tag{48a}$$

$$H^{(n)} \stackrel{d}{=} \begin{cases} \hat{\Omega} - \hat{H}_1^{(n)} & \text{with prob. } q, \\ \min \left(\hat{\Omega} - \hat{H}_1^{(n)}, \, \hat{H}_2^{(n)} \right) & \text{with prob. } 1-q, \end{cases} \tag{48b}$$

that defines a sequence of random variables $\hat{H}^{(n)}$, with the initial condition $\hat{H}^{(0)} = 0$. Comparing these equations with (35) one can show by induction on n that $\hat{K}^{(n)}$ is stochastically smaller [35] than $\hat{H}^{(n)}$; we detail this proof in Appendix B 2. Here we shall only recall that given two random variables X and Y one says that X is stochastically smaller than Y, to be denoted $X \preceq Y$, if and only if $\mathbb{P}[X > x] \leqslant \mathbb{P}[Y > x]$ for all x. This condition is equivalent to the existence of a coupling $(\hat{X}, \hat{Y})$, i.e., a random vector with marginal laws equal to those of X and Y, respectively, such that $\mathbb{P}[\hat{X} \leqslant \hat{Y}] = 1$. In our case if $v > 0$ we have seen that $\hat{K}^{(n)}$ diverges to $+\infty$ in the large n limit, hence by this stochastic comparison argument this will also be the case of $\hat{H}^{(n)}$. With the assumption that a nontrivial solution of the fixed point equation (29), if it exists, will be reached as the large n limit of the sequence $\hat{H}^{(n)}$, this allows us to conclude that $v > 0$ rules out such a nontrivial fixed point, hence is a criterion for a full-recovery phase.

C. The large degree limit

As explained in Sec. II the large degree limit, $c \to \infty$ taken here after the thermodynamic limit $N \to \infty$, allows us to recover the dense models defined on complete graphs in Refs. [3,10], if one performs an appropriate rescaling of the distribution of the weights on the nonplanted edges. Consider indeed the condition (45): if p and $\hat{p}$ are kept constant as $c \to \infty$ this becomes $\int_\Gamma \sqrt{\hat{p}(w)p(w)} \, dw = 0$, which is never satisfied unless Γ is empty (and in this case the planted edges can be identified by inspection of the weight on the edges, as p and $\hat{p}$ have disjoint supports). Indeed, if the weights on both types of edges are of the same order of magnitude, around one vertex the planted weight will be hidden among the $O(c)$ nonplanted ones and impossible to distinguish in the $c \to \infty$ limit. To have a nontrivial partial–full-recovery transition for $c \to \infty$ it is therefore necessary to scale the nonplanted edges weights, the simplest way being to take the nonplanted weight distribution uniform on the interval $[0, c]$. The condition (45) becomes then

$$\int \sqrt{\hat{p}(w)} \, dw = 1, \tag{49}$$

where we assumed for simplicity of notation that the support of $\hat{p}$ is included in the positive real axis.

V. EXAMPLES

We shall now confront our analytical prediction for the location of the phase transition with numerical results, obtained both for finite N by solving the BP equations on single samples of the problem, and in the thermodynamic limit by solving numerically the RDEs. Some details on these numerical procedures are given in Sec. V D.

For concreteness we will always take an uniform distribution for the nonplanted weights,

$$p(w) = \frac{1}{c} \mathbb{I}(0 \leqslant w \leqslant c), \tag{50}$$

for which Eq. (45) further simplifies as

$$\int_\Gamma \sqrt{\hat{p}(w)}\, dw = 1, \quad \Gamma = \text{supp}(\hat{p}) \cap [0, c]. \tag{51}$$

We will present our results for different choices of $\hat{p}$, some of them partially investigated in the literature.

A. The exponential distribution

Let us start by considering the exponential distribution

$$\hat{p}(w) = \lambda\, e^{-\lambda w}\, \mathbb{I}(w \geqslant 0), \tag{52}$$

for which, in the large degree limit, Ref. [10] proved that $\lambda < 4$ is a partial-recovery phase, while $\lambda > 4$ corresponds to full recovery.

Our predictions are summarized in the phase diagram in the (c, λ) plane displayed in Fig. 1. The red line corresponds to the condition $\gamma = 1$ below which the leaf removal procedure described in Sec. III D recovers completely the hidden matching; here $\mu = 1$ and $\hat{\mu} = 1 - e^{-\lambda c}$, and the equation of this line is thus $\lambda = -\log[1 - (1/c)]/c$. The blue line is instead the vanishing velocity criterion (51), which becomes for this choice of $\hat{p}$:

$$1 = \int_0^c \sqrt{\lambda}\, e^{-\frac{\lambda w}{2}}\, dw = 2 \frac{1 - e^{-\frac{c\lambda}{2}}}{\sqrt{\lambda}}, \tag{53}$$

a relation that can be inverted in

$$c = -\frac{2}{\lambda} \ln\left(1 - \frac{\sqrt{\lambda}}{2}\right). \tag{54}$$

This blue line separates a domain, in blue in Fig. 1, where $v < 0$, corresponding to a partial-recovery phase, from a full-recovery phase (in white) with $v > 0$. Note that the curve has a minimal abscissa of $c \approx 1.2277$ below which there is full recovery for all λ. In the large degree limit the two branches of the blue line converge to $\lambda = 0$ and $\lambda = 4$, we thus recover the results of [10].

The red dots on this phase diagram have been obtained from a numerical resolution of Eqs. (29), using a population dynamics algorithm, and correspond to the limit values for which we found a nontrivial solution of the equations. They are in agreement, within numerical accuracy, with the analytical prediction. In Fig. 2 we compare our prediction for the average reconstruction error $\mathbb{E}[\varrho]$ (nonzero in the partial-recovery phase) obtained by Eqs. (29) and (30) (in the thermodynamic limit) with the numerical results obtained running a belief propagation based on Eq. (27) (on finite graphs). The agreement between these two procedures is very good, except for smoothing finite-size effects close to the phase transitions.
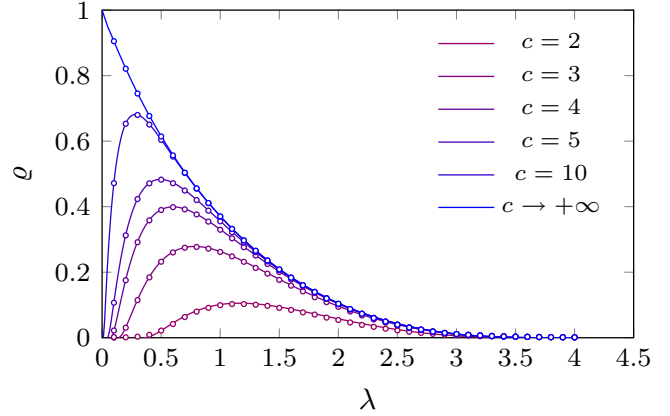

FIG. 2. Reconstruction error for the planted matching problem with exponential planted weights and different values of c. The lines have been obtained numerically solving the RDEs (29) using a population dynamics algorithm with 10^6 fields. The dots are the results of the resolution of the BP equations (27) on graphs of $N = 10^3$ vertices, averaged over 10^4 instances, the $c \to +\infty$ case corresponding to the complete graph.

B. The folded Gaussian case

We have also considered a planted weight distribution of the folded Gaussian form,

$$\hat{p}(w) = \sqrt{\frac{2}{\pi \lambda}}\, e^{-\frac{w^2}{2\lambda}}\, \mathbb{I}(w \geqslant 0), \tag{55}$$

that had been investigated previously in Ref. [3] in the large degree limit. Our prediction for this case, easily obtained by plugging this expression in (51) and taking the $c \to \infty$ limit, is of a phase transition at

$$\bar{\lambda} = \frac{1}{2\pi} \approx 0.1591 \tag{56}$$

between a partial-recovery phase for $\lambda > \bar{\lambda}$ and a full-recovery phase for $0 < \lambda < \bar{\lambda}$. This agrees qualitatively with the numerical investigations of Ref. [3], but not with the value of the threshold that was estimated in Ref. [3] to be $\bar{\lambda} = 0.174(4)$. We believe this discrepancy is due to finite-population size effects (that are particularly severe in this kind of problems, as discussed in Sec. V D) in the study of Ref. [3]. To check this point we solved the equations with a careful numerical integration (see Sec. V D) of the recursive equations, which convincingly suggests that $\bar{\lambda} < 0.160(1)$, as shown in Fig. 3.

C. The truncated power-law case

Let us consider as a final example the case where the density of the weights on the planted edges varies as a power law on a finite interval,

$$\hat{p}(w) = \frac{\alpha w^{\alpha-1}}{\lambda^\alpha} \mathbb{I}(0 \leqslant w \leqslant \lambda), \tag{57}$$

with $\alpha > 0$. For the sake of simplicity, we will restrict our analysis to the case $c > \lambda$. Then one finds immediately that $\Gamma = [0, \lambda]$, $\hat{P} = \hat{p}$, P is the uniform distribution on $[0, \lambda]$ and
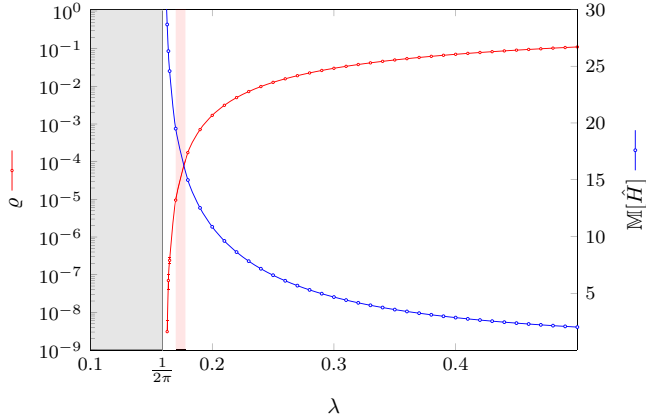
FIG. 3. Planted matching problem with (folded) Gaussian weights, in the large degree limit; the curves have been obtained from a numerical resolution of Eqs. (31), the red line corresponding to the reconstruction error ϱ, the blue line to the median cavity fields on planted edges $\mathbb{M}[\hat{H}]$. The light-red interval corresponds to the estimate for the transition point given in Ref. [3].

$\gamma = \lambda$, in such a way that as soon as $c > \lambda$ the problem on the pruned graph is completely independent of c.

The transition condition $\gamma = 1$ for the full recoverability of the planted matching by the leaf removal algorithm is thus $\lambda = 1$, which yields the red domain in the phase diagram of Fig. 4. The vanishing velocity condition (51) becomes here

$$\bar{\lambda} = \frac{(1 + \alpha)^2}{4\alpha}, \tag{58}$$

which is plotted as a blue line in Fig. 4, the full-recovery phase corresponding to the domain $\lambda < \bar{\lambda}$. This prediction is in agreement with the numerical results obtained by a population dynamics resolution of the RDEs.
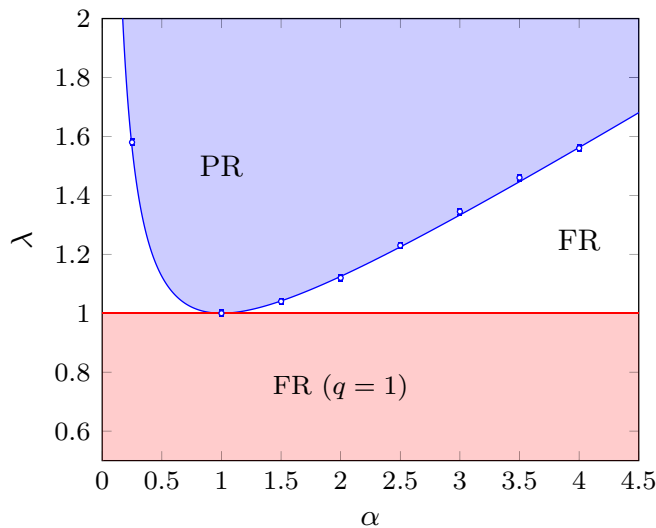

FIG. 4. Phase diagram of the recovery transition for the planted matching problem with a truncated power-law distribution for the planted weights. The red line corresponds to the $\gamma = 1$ bound for the full-recovery transition. The circles are the transition points obtained running a population dynamics algorithm with 10^6 fields.

For $\alpha = 1$, that corresponds to the planted weights uniformly distributed on $[0, \lambda]$, one can actually solve the RDEs explicitly. Indeed in this case the function $\omega(w)$ in Eq. (12) vanishes, hence Eqs. (29) reduce to

$$\hat{H} \stackrel{d}{=} -\max_{1 \leqslant i \leqslant Z} [H_i], \tag{59a}$$

$$H \stackrel{d}{=} \begin{cases} -\hat{H} & \text{with prob. } q, \\ \min(-\hat{H}, \hat{H}') & \text{with prob. } 1 - q, \end{cases} \tag{59b}$$

which obviously admit the solution $H \stackrel{d}{=} \hat{H} \stackrel{d}{=} 0$. Indeed the effective weights on the pruned graph are all equal, the planted matching is one of the many perfect matchings of this reduced graph, but there is no information contained in the weights to decide which one. In a simple-minded application of the inclusion rule (17) one would include in the estimator the edges of the pruned graph independently with probability $1/2$, leading to an average estimation error $\mathbb{E}[\varrho] = (1 - q)^2(1 + \lambda)/4$ for $\lambda > 1$, and of course $\mathbb{E}[\varrho] = 0$ for $\lambda \leqslant 1$.

D. A note on the numerical procedures

Most of the thermodynamic limit results presented above have been obtained by a numerical resolution of Eqs. (29) via a population dynamics algorithm [28]. The idea of this method, which is very commonly used to solve RDEs, is to represent the law of a random variable X as the empirical distribution of a sample $\{X_1, \ldots, X_{\mathcal{N}}\}$ of its representants, with $\mathcal{N} \gg 1$ to improve the accuracy of the method. In terms of cumulative distributions this corresponds to the approximation

$$F_X(x) \approx \frac{1}{\mathcal{N}} \sum_{i=1}^{\mathcal{N}} \mathbb{I}(X_i \leqslant x). \tag{60}$$

One considers then the iterative version of the RDE written in (48), and update the population according to these rules. For instance each representant $\hat{H}_i$ at the iteration $n + 1$ is generated, independently, by drawing an integer Z with the law (25), Z copies of the random variable Ω, and Z representants of H at the iteration n, by a uniform choice over the $\mathcal{N}$ ones. These quantities are then combined according to the right-hand side of (48a) to compute $\hat{H}_i$. The sample of representants of H at the iteration $n + 1$ is then generated similarly according to (48b). These steps are repeated a large number n of times, the type of phase (partial or full recovery) is then decided according to the convergence or divergence to $+\infty$ of the population representing $\hat{H}^{(n)}$ in the large n limit. The accurate determination of such a phase transition suffers from finite population size effects that are much more severe than in usual applications of the population dynamics algorithm. Indeed, the transition is governed by an instability that manifests itself as a front propagation in the cumulative distribution function; such front propagations are generically driven by the behavior in the exponentially small tail far away from the front [15–17]. As the finite population size implies a cutoff of $1/\mathcal{N}$ on the smallest representable value of the cumulative distribution function, this translates into logarithmic finite population size effects on the velocity of the front and the location of the phase transition, at variance with the usual $1/\mathcal{N}$ corrections for the computation of observables

as empirical averages. We refer the reader to Ref. [15] for a quantitative study of these logarithmic corrections in the velocity of a front in presence of a threshold in its tail.

We thus believe that the discrepancy in the folded Gaussian case between our analytical prediction $\bar{\lambda} = \frac{1}{2\pi} \approx 0.159$ and the numerical estimate $\bar{\lambda} = 0.174(4)$ of Ref. [3] can be ascribed to these strong finite-$\mathcal{N}$ effects. The results presented in Fig. 3 that supports this thesis have been obtained with another numerical procedure: instead of the population representation (60) of the cumulative distribution functions $F_H(h)$ and $F_{\hat{H}}(h)$ we stored their values in M points $h_1 < h_2 < \cdots < h_M$ over a given interval $[h_1, h_M]$ and updated them using Eqs. (31) until a certain convergence criterion was satisfied (until the L^2 distance between the solution at step n and the solution at step $n - 1$ was smaller than a given, prefixed tolerance ϵ). The advantage of this method is that the cutoff h_M can be taken arbitrarily large, in such a way that $\bar{F}_{\hat{H}}(h_M)$ is very small, hence bypassing the threshold at $1/\mathcal{N}$ of the population dynamics algorithm. After convergence the function $F_{\hat{H}}(h)$ can be used to estimate $\mathbb{E}[\varrho]$, e.g., by a Monte Carlo integration.

VI. A MORE PRECISE DESCRIPTION OF THE CRITICAL REGIME

Once the threshold value of a parameter has been determined it is natural to aim at a more quantitative description of the transition in its critical regime. In the case considered in this paper of planted models that undergo a continuous transition from partial recovery for $\lambda < \bar{\lambda}$ to full recovery for $\lambda > \bar{\lambda}$ this point amounts to describe how the average reconstruction error $\mathbb{E}[\varrho]$ vanishes as $\lambda \to \bar{\lambda}^-$. This was raised as open question 2 in Ref. [10], and we shall study it in the model defined therein, i.e., with an exponential distribution for the planted weights, in the large degree limit. This case allows for some technical simplifications; as shown in Ref. [10] the RDEs can then be reduced to a system of ordinary differential equations (ODEs), that we first recall in the next subsection before studying their solution in the critical regime.

A. The ODEs for the exponential model

Let us specialize our formalism with the following choices of weight distributions: $\hat{p}(w) = \lambda\,e^{-\lambda w}$ for $w \geqslant 0$, and $p(w) = 1/c$ for $w \in [0, c]$. The intersection of their supports is thus $\Gamma = [0, c]$, and one finds $\mu = 1$, $\hat{\mu} = 1 - e^{-\lambda c}$. The reduced distributions $\hat{P}$ and P are then, on their common support Γ, $\hat{P}(w) = \lambda\,e^{-\lambda w}/\hat{\mu}$ and $P(w) = 1/c$, which gives an effective weight function $\omega(w) = \lambda w - \ln(\lambda c/\hat{\mu})$. The parameter q is the solution of $q = e^{-\gamma(1-q)}$ with $\gamma = c\hat{\mu}$.

To simplify the notations, and to get closer to the conventions used in Ref. [10], we shall define random variables X and Y that are affine transformations of $\hat{H}$ and H, respectively. More precisely we define their cumulative functions as

$$F_X(x) := F_{\hat{H}}\left[\lambda x - \frac{1}{2}\ln\left(\frac{\lambda c}{\hat{\mu}}\right)\right], \tag{61}$$

$$F_Y(x) := F_H\left[\lambda x - \frac{1}{2}\ln\left(\frac{\lambda c}{\hat{\mu}}\right)\right], \tag{62}$$

and keep the convention $\bar{F} = 1 - F$ for reciprocal cumulative distributions. Equations (31) become

$$\bar{F}_Y(x) = [q + (1 - q)\bar{F}_X(x)]\frac{\lambda}{\hat{\mu}}\int_0^c e^{-\lambda w}F_X(w - x)\,dw, \tag{63a}$$

$$\bar{F}_X(x) = \frac{\exp\left[-\hat{\mu}(1 - q)\int_{-x}^{c-x}\bar{F}_Y(w)\,dw\right] - q}{1 - q}. \tag{63b}$$

Taking the limit $c \to +\infty$, in which $\hat{\mu} \to 1$ and $q \to 0$, Eqs. (63) become

$$\bar{F}_Y(x) = \bar{F}_X(x)\int_0^{+\infty}\lambda\,e^{-\lambda w}F_X(w - x)\,dw, \tag{64a}$$

$$\bar{F}_X(x) = \exp\left[-\int_{-x}^{+\infty}\bar{F}_Y(w)\,dw\right]. \tag{64b}$$

These equations between cumulative distribution functions correspond to the following RDEs on X and Y:

$$X \overset{d}{=} \min\{\xi_i - Y_i\}, \tag{65a}$$

$$Y \overset{d}{=} \min(\eta - X, X'), \tag{65b}$$

where in the first line the ξ_i's are the points of a Poisson point process of intensity 1 on the positive real axis, and in the second line η has an exponential distribution of parameter λ. It is convenient to introduce the auxiliary function $V(x)$ defined as the cumulative distribution of the random variable $X - \eta$,

$$V(x) := \mathbb{P}[X - \eta \leqslant x] = \int_0^{+\infty}\lambda\,e^{-\lambda w}F_X(w + x)\,dw, \tag{66}$$

in such a way that (64a) can be rewritten $\bar{F}_Y(x) = \bar{F}_X(x)V(-x)$. Taking derivatives with respect to x in (64b) and (66), and denoting for simplicity $F = F_X$, one obtains

$$F'(x) = [1 - F(x)][1 - F(-x)]V(x), \tag{67a}$$

$$V'(x) = \lambda(V(x) - F(x)), \tag{67b}$$

where the form of the equation on V crucially depends on the exponential character of the distribution of the planted weight η. These two equations on F and V are not yet ODEs because one of the arguments in the right-hand side of (67a) is $-x$ instead of x; to bypass this difficulty one introduces two additional functions, $G(x) := F(-x)$ and $W(x) := V(-x)$, in such a way that the four-dimensional vector $(F, G, V, W)(x)$ obeys an autonomous first-order ODE, from the solution of which the average reconstruction error is computed as

$$\mathbb{E}[\varrho] = \mathbb{P}[X + X' \leqslant \eta]$$
$$= 2\int_0^\infty [1 - F(x)][1 - G(x)]V(x)W(x)\,dx; \tag{68}$$

see Ref. [10] for the details of the derivation of the integral expression of $\mathbb{E}[\varrho]$.

The dimensionality of the problem can be reduced by exploiting the conservation law $F(x)W(x) + G(x)V(x) - V(x)W(x) = 0$ for all x. Introducing finally $U(x) = F(x)/V(x)$ it is shown in Ref. [10] that the problem reduces to
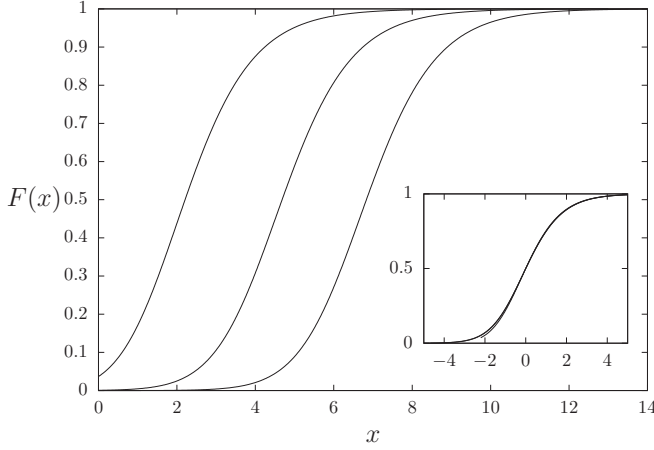
FIG. 5. Cumulative distribution $F(x)$ for, from left to right, $\lambda = 3$, $\lambda = 3.8$, $\lambda = 3.92$. Inset: the three curves have been shifted horizontally by their medians $\mathbb{M}(\lambda) = F^{-1}(1/2)$, the collapse confirms the hypothesis stated in (71) of a convergence in distribution of the shifted random variables.
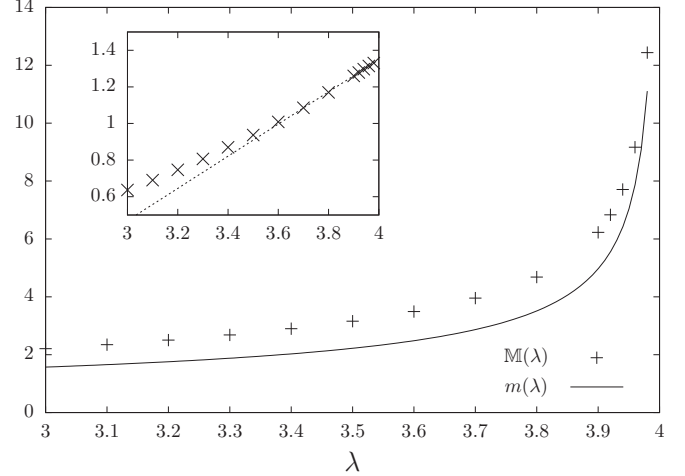

FIG. 6. Median $\mathbb{M}(\lambda)$ of the cumulative distribution in the exponential case, compared to the analytical prediction (78) for $m(\lambda)$. Inset: the dots show $\mathbb{M}(\lambda) - m(\lambda)$, the dashed line is a linear fit that confirms the convergence of $\mathbb{M}(\lambda) - m(\lambda)$ to a finite constant in the limit $\lambda \to 4$.

solve, for $x \geqslant 0$, the following ODE on the three-dimensional vector $(U, V, W)(x)$:

$$U'(x) = -\lambda U(x)[1 - U(x)]$$
$$+ [1 - U(x)V(x)]\{1 - [1 - U(x)]W(x)\}, \quad (69\text{a})$$

$$V'(x) = \lambda V(x)[1 - U(x)], \quad (69\text{b})$$

$$W'(x) = -\lambda W(x)U(x), \quad (69\text{c})$$

with the initial conditions

$$U(0) = \tfrac{1}{2}, \quad V(0) = W(0). \quad (70)$$

Even if the notation does not show it explicitly the solution of these ODEs depends of course on λ, both directly as λ appears in (69) and indirectly through the initial condition $V(0) = W(0)$. It is indeed shown in Ref. [10] that for a given $\lambda < 4$ there is a unique choice of this initial condition that yields a proper solution, i.e., one in which F and V have the properties of cumulative distribution functions (nondecreasing and bounded between 0 and 1).

B. The divergence of X in the limit $\lambda \to 4$

We present in Fig. 5 the cumulative distribution F of the random variable X, for three values of λ increasing towards the critical value $\lambda = 4$, obtained by a numerical resolution of the ODE (69). This plot suggests that F drifts without deformation when approaching the transition; this impression is confirmed by the inset of the figure, which shows a very good collapse of the curves once shifted by the median $\mathbb{M}(\lambda) = F^{-1}(1/2)$ of X (any other quantile would have led to the same collapse, with an additional constant shift of the horizontal axis). This observation has two equivalent translations: from the probabilistic point of view it corresponds to the existence of a function $m(\lambda)$ and a random variable $\hat{X}$ such that

$$X - m(\lambda) \xrightarrow[\lambda \to 4^-]{d} \hat{X}, \quad (71)$$

as was stated in (47) for generic weight distributions, $m(\lambda)$ differing from $\mathbb{M}(\lambda)$ by an arbitrary constant. From the analytic point of view it means that the solution of the ODEs admits a scaling regime $x = z + m(\lambda)$ when z is kept fixed while $\lambda \to 4^-$, described by functions $\hat{F}(z), \hat{V}(z), \hat{U}(z)$, defined as

$$\hat{F}(z) = \lim_{\lambda \to 4^-} F[z + m(\lambda)], \quad (72)$$

similar definitions holding for $\hat{V}$ and $\hat{U}$. The dots in the main panel of Fig. 6 represent the numerically determined value of $\mathbb{M}(\lambda)$, and suggest a divergence of this quantity as $\lambda \to 4^-$. Supposing that $m(\lambda)$ indeed diverges one can simplify the ODEs (67) in this scaling regime, with $F(-x) \to 0$; this yields

$$\hat{F}'(x) = [1 - \hat{F}(x)]\hat{V}(x), \quad (73\text{a})$$

$$\hat{V}'(x) = 4[\hat{V}(x) - \hat{F}(x)]. \quad (73\text{b})$$

$\hat{F}$ is the cumulative distribution of the limit variable $\hat{X}$, solution of the simplified RDE

$$\hat{X} \stackrel{d}{=} \min\{\xi_i - \hat{Y}_i\}, \quad (74\text{a})$$

$$\hat{Y} \stackrel{d}{=} \eta - \hat{X}. \quad (74\text{b})$$

Studying this simplified ODE in the $z \to -\infty$ limit where it can be linearized, or appealing to the theorems explained in (40) for the left tail behavior of the limit random variable in the BRW interpretation, one finds

$$\hat{F}(z) \underset{z \to -\infty}{\sim} -A z\, e^{2z}, \quad \hat{U}(z) \underset{z \to -\infty}{\sim} \frac{1}{2} - \frac{1}{4z}, \quad (75)$$

where $A > 0$ is a constant that cannot be fixed because of the invariance by translation of the equation.

In order to determine the sought-for divergence of $m(\lambda)$ as $\lambda \to 4^-$ we need now to study the solution of the ODE (69) on another scaling regime, $x = t m(\lambda)$ with $t \in [0, 1)$ kept fixed in the limit, that will allow us to take into account the initial condition (70). The derivation will then conclude by

a matching argument at the common boundary of the two scaling regimes, $t \to 1^-$ and $z \to -\infty$.

In order to study the scaling regime $x = tm(\lambda)$ we first notice that $V(0) \to 0$ as $\lambda \to 4^-$, because V is the cumulative distribution function of the random variable $X - \eta$ that diverges in this limit. It is thus instructive to solve first Eqs. (69) with $V(0) = W(0) = 0$, denoting U_0, V_0, W_0 its solution, even if this cannot be exactly a proper solution. One finds $V_0(x) = W_0(x) = 0$ for all $x \geqslant 0$, and (69a) simplifies into

$$U_0'(x) = -\lambda U_0(x)[1 - U_0(x)] + 1, \tag{76}$$

an equation that can be solved exactly for any $\lambda < 4$ with the initial condition $U_0(0) = 1/2$, yielding

$$U_0(x) = \frac{1}{2} + \frac{1}{2}\sqrt{\frac{4-\lambda}{\lambda}} \tan\left[\frac{x}{2}\sqrt{\lambda(4-\lambda)}\right]. \tag{77}$$

This expression diverges when the argument of the tangent reaches $\pi/2$, which gives us a natural candidate for the scale $m(\lambda)$ at which the first regime ends, namely,

$$m(\lambda) = \frac{\pi}{2\sqrt{4-\lambda}}, \tag{78}$$

and a conjecture for the behavior of the solution $U(x)$ of the full ODE in the first scaling regime, namely,

$$\lim_{\lambda \to 4^-} \frac{1}{\sqrt{4-\lambda}} \left\{ U[tm(\lambda)] - \frac{1}{2} \right\} = \frac{1}{4} \tan\left(t\frac{\pi}{2}\right). \tag{79}$$

We have assumed here that $V(0)$, even if strictly nonzero for all $\lambda < 4$, is sufficiently small for the second line in (69a) to be negligible, and hence for U to coincide with U_0 at the dominant order in this scaling regime. To check the self-consistency of this hypothesis we first give an exact expression of V and W in terms of U obtained by integration of (69b) and (69c):

$$V(x) = V(0) \exp\left\{ \lambda \int_0^x [1 - U(y)]\, dy \right\}, \tag{80}$$

$$W(x) = e^{-\lambda x} V(x). \tag{81}$$

Inserting the scaling ansatz (79) into (80) we obtain

$$\lim_{\lambda \to 4^-} \frac{1}{V(0)} V(tm(\lambda))\, e^{-2tm(\lambda)} = \cos\left(t\frac{\pi}{2}\right), \tag{82}$$

the behavior of W being easy to deduce from the one of V because of (81). We fix now the initial condition $V(0)$ by matching the behavior $t \to 1^-$ of this expression with the limit $z \to -\infty$ of the other regime, which from (75) is $\hat{V}(z) \sim -2Az\, e^{2z}$, with the correspondence $t \sim 1 + \frac{z}{m(\lambda)}$. This yields

$$V(0) = 2A\, e^{-2m(\lambda)} \frac{1}{\sqrt{4-\lambda}} \tag{83}$$

and allows us to check that indeed the first line of (69a) is dominant in the scaling regime $t = m(\lambda)$ as long as $t < 1$, confirming the self-consistency of our hypothesis. In addition the behavior of U in (75) and (79) matches at the boundary of the two scaling regimes. Note that the indeterminacy of the constant A, because of the invariance by translation of the equations on $\hat{F}$ and $\hat{V}$, is related in (83) to the additive arbitrary constant that can be added to $m(\lambda)$.

The inset of Fig. 6 presents a numerical confirmation of this reasoning: the difference between the numerically determined median of X and our formula (78) is seen to converge to a finite constant when $\lambda \to 4^-$ (with corrections that seem polynomial in $4 - \lambda$). We have also checked that the numerical results for $U(x)$ and $V(x)$ are compatible with the scaling ansatz of (79) and (82).

C. The critical behavior of $\mathbb{E}[\varrho]$

We would like now to use our prediction (78) for the divergence of X in order to describe the way in which the average reconstruction error $\mathbb{E}[\varrho]$ vanishes at the transition. The expression of the latter, given in (68), can be rewritten as

$$\mathbb{E}[\varrho] = \mathbb{E}[e^{-\lambda X}]^2. \tag{84}$$

Indeed, the exponential distribution of η is such that $\mathbb{P}[\eta \geqslant x] = e^{-\lambda x}$, and X and X' in (68) are independent random variables with the same law. Recalling the convergence in distribution stated in (71), and the prediction (78) of $m(\lambda)$, it would be tempting to write

$$\mathbb{E}[\varrho] \underset{\lambda \to 4^-}{\propto} e^{-\frac{4\pi}{\sqrt{4-\lambda}}} \mathbb{E}[e^{-4\hat{X}}]^2. \tag{85}$$

Unfortunately this result has to be amended: as the tail of $\hat{X}$ varies as e^{2z} for $z \to -\infty$ [cf. Eq. (75)] the expectation value $\mathbb{E}[e^{-4\hat{X}}]$ is infinite. Using the integral expression of $\mathbb{E}[\varrho]$ given in (68), one finds that the leading contribution is given by the scaling regime $x = tm(\lambda)$ and is of the form

$$2 \int_0^{m(\lambda)} V(x) W(x)\, dx = 2 \int_0^{m(\lambda)} V(x)^2\, e^{-\lambda x}\, dx \tag{86}$$

$$\propto V(0)^2 \int_0^1 \cos^2\left(t\frac{\pi}{2}\right) m(\lambda)\, dt. \tag{87}$$

The asymptotic form of the initial condition stated in (83), combined with the expression of $m(\lambda)$ given in (78), yields finally the prediction

$$\mathbb{E}[\varrho] \underset{\lambda \to 4^-}{\propto} e^{-\frac{2\pi}{\sqrt{4-\lambda}}} (4-\lambda)^{-3/2}. \tag{88}$$

Note that all the derivatives of $\mathbb{E}[\varrho]$ vanish as $\lambda \to 4^-$ because of the essential singularity of the exponential term, the transition is thus of infinite order in the usual thermodynamic classification.

We present in Fig. 7 our numerical results for $\mathbb{E}[\varrho]$, computed by a numerical integration of the ODE and the integral expression in (68). The main panel shows qualitatively that $\mathbb{E}[\varrho]$ is indeed very flat close to the transition; the rescaling performed in the inset is in agreement with the asymptotic form (88).

VII. THE SYMBOL MAP CASE

We have discussed above the phase diagram of the problem, and distinguished in particular full and partial-recovery phases, considering the block MAP estimator, i.e., the $\beta \to \infty$ version of the BP equations. The phases were thus defined according to whether the average reconstruction error $\mathbb{E}[\varrho_b]$ vanished in the thermodynamic limit or not, the subscript b specifying the use of the block MAP estimator in the
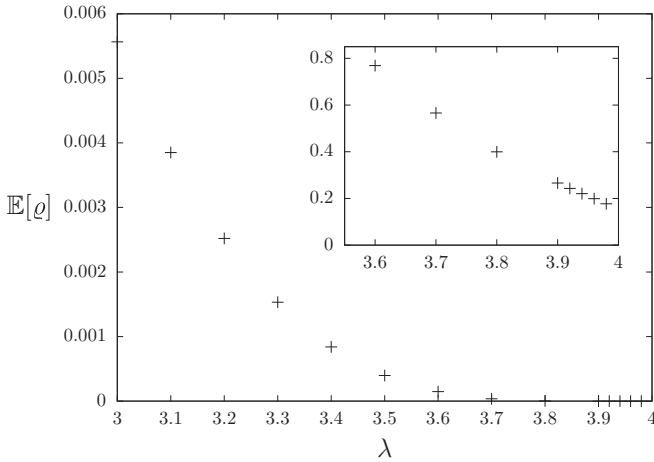
FIG. 7. The average reconstruction error $\mathbb{E}[\varrho]$ that vanishes continuously as $\lambda \to 4^-$. Inset: $\mathbb{E}[\varrho]\, e^{\frac{2\pi}{\sqrt{4-\lambda}}} (4-\lambda)^{3/2}$ as a function of λ, the convergence to a positive constant as $\lambda \to 4^-$ confirms (88).

computation. However, we explained in Sec. II B that the estimator that minimizes the average reconstruction error is the symbol MAP one, obtained with $\beta = 1$, with an average reconstruction error denoted $\mathbb{E}[\varrho_s]$. As $\mathbb{E}[\varrho_s] \leqslant \mathbb{E}[\varrho_b]$ the phases shown to be of the full-recovery type for $\beta \to \infty$ are certainly so also for the symbol MAP estimator, one can nevertheless wonder if the converse is true, namely if some choices of parameters yield $0 = \mathbb{E}[\varrho_s] < \mathbb{E}[\varrho_b]$.

We have investigated this question numerically, by solving with a population dynamics algorithm the RDEs (24) with $\beta = 1$, and computed $\mathbb{E}[\varrho_s]$ from (26). Our results are presented in Fig. 8; for concreteness we have used the exponential distribution of Eq. (52) for the planted weights, and
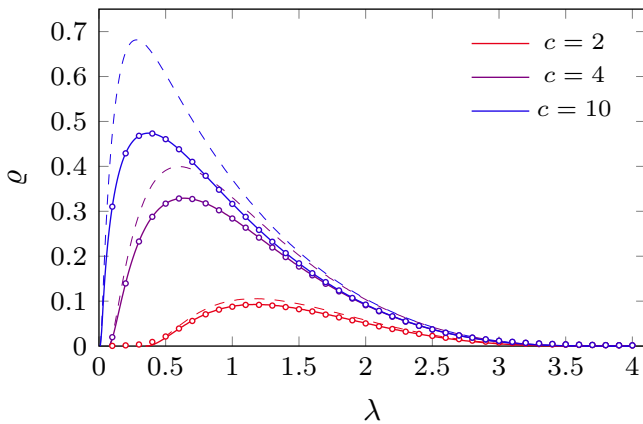

FIG. 8. The average reconstruction error for the symbol MAP estimator ($\beta = 1$) with exponentially distributed planted weights, at different values of c. The solid lines have been obtained numerically solving the RDEs in Eq. (24) with $\beta = 1$ using a population dynamics algorithm with 10^6 fields. The dots corresponds to the same error rate estimated running BP over 10^4 instances of the problem, on graphs with $N = 10^3$ vertices. The dashed lines corresponds to the reconstruction error of the block MAP estimator presented in Fig. 2, which are larger than the symbol MAP ones for the same value of c (encoded by the color of the curve).

several values of the average degree c (the nonplanted weight distribution being uniform on $[0, c]$). We found indeed that $\mathbb{E}[\varrho_s] \leqslant \mathbb{E}[\varrho_b]$ (the block MAP results, previously presented on Fig. 2, are drawn with dashed lines). Within our numerical accuracy the transition to the full-recovery phases occur for the same values of the parameters in the symbol and block MAP cases; this is in agreement with a conjecture of Ref. [10]; see open question 1 therein.

VIII. FUTURE WORK

Let us conclude by giving some thoughts on how our study could be extended. One could try to study the critical regime for generic distributions, i.e., extend the results of Sec. VI that were obtained only for the exponential distribution and in the large degree limit. We expect the exponent $-1/2$ for the divergence of the median of the fields to be rather universal, but the form of the vanishing of $\mathbb{E}[\varrho]$ should be much more dependent on the details of the models. One motivation for this direction of research is the difficulty of an accurate numerical determination of the location of the phase transition, as discussed in Sec. V D. The numerical accuracy problems should be less stringent further away from $\bar{\lambda}$ inside the partial-recovery phase, hence an extrapolation of $\mathbb{M}[\hat{H}]$, if one has a prediction for its functional form, should lead to more precise determinations of threshold parameters.

It would also be interesting to further investigate the possibility of discontinuous recovery phase transitions, for which the derivation presented in Sec. IV B would fail. We did not find evidence for their occurrence, but we cannot exclude this possibility because of the limited accuracy of our numerical results. Such situations might occur for contorted weight distributions, or if instead of Erdős-Rényi random graphs one hides the planted matching in a configuration model with some well-chosen degree distributions, for which Ref. [30] unveiled the existence of multiple BP fixed points.

The coincidence of the thresholds for full recovery of the symbol and block MAP estimators observed numerically in Sec. VII also calls for further investigation and for an analytical argument supporting (or disproving) it. This point is also connected to the apparent absence of statistical to computational gaps in this problem: the block MAP estimator, being a minimal weight perfect matching, can be determined in polynomial time [2], and the results of Refs. [24–27] strongly suggest that it can be asymptotically (in the large size limit) obtained by the $\beta \to \infty$ BP equations. An exact computation of the symbol MAP estimator is instead a computationally hard problem, but it is tempting to conjecture that the BP algorithm with $\beta = 1$ reaches asymptotically the information theoretically optimal reconstruction error $\mathbb{E}[\varrho_s]$.

A k factor of a graph is a set of edges such that each node belongs to exactly k edges of the factor; a perfect matching is thus a special case of this definition with $k = 1$. It would therefore be interesting to study the planted k-factor problem for generic values of k. For $k = 2$ the problem is related to the planted Hamiltonian cycle that was considered in Ref. [6]. The planted k factor could also be studied using the cavity approach and the associated belief propagation equations. At variance with the matching case there is, for generic k, no efficient algorithm even for the block MAP estimator; this

opens the possibility for computationally hard phases in such a generalization.

Another natural direction for future work is a rigorous proof of our results, notably of the threshold given in Eq. (45) and the critical behavior stated in Eq. (88). While the local-weak-convergence proof of Ref. [10] can likely be extended to generic weights distribution and to the sparse graph settings, it is not clear how to control rigorously the solution of the recursive distributional equations, in particular the reasoning at the beginning of section IV B. The stochastic comparison argument explained in remark (3) at the end of Sec. IV B and expanded upon in Appendix B 2, should provide a scheme for a rigorous proof of full recovery when $v > 0$, the much more challenging question is to prove partial recovery when $v < 0$.

ACKNOWLEDGMENTS

APPENDIX A: THE RECONSTRUCTION ERROR FOR THE BLOCK MAP ESTIMATOR

We prove in this Appendix that the equality of the two expressions (26) and (30) of the average reconstruction error when $\beta \to \infty$ follows from the RDE (29). We have thus to prove that $\mathbb{P}[\hat{H} + \hat{H}' \leqslant \hat{\Omega}] = \gamma \, \mathbb{P}[H + H' > \Omega]$. We first notice that (29b) implies that for any real x one has $\mathbb{P}[H \geqslant x] = \mathbb{P}[\hat{\Omega} - \hat{H} \geqslant x](q + (1-q)\mathbb{P}[\hat{H} \geqslant x])$, hence

$$\mathbb{P}[\hat{\Omega} - \hat{H} \geqslant x] = \frac{\mathbb{P}[H \geqslant x]}{q + (1-q)\mathbb{P}[\hat{H} \geqslant x]}. \qquad (A1)$$

Multiplying this expression by $-\frac{d}{dx}\mathbb{P}[\hat{H} \geqslant x]$, which is the density of the random variable $\hat{H}$, we obtain

$$\begin{aligned}
&\mathbb{P}[\hat{H} + \hat{H}' \leqslant \hat{\Omega}] \\
&= \int_{-\infty}^{+\infty} \left(-\frac{d}{dx}\mathbb{P}[\hat{H} \geqslant x] \right) \mathbb{P}[\hat{\Omega} - \hat{H} \geqslant x] \\
&= \frac{1}{1-q} \int_{-\infty}^{+\infty} \left(-\frac{d}{dx}\ln\{q + (1-q)\mathbb{P}[\hat{H} \geqslant x]\} \right) \mathbb{P}[H \geqslant x] \\
&= \frac{1}{1-q} \int_{-\infty}^{+\infty} \left(\frac{d}{dx}\mathbb{P}[H \geqslant x] \right) \ln\{q + (1-q)\mathbb{P}[\hat{H} \geqslant x]\},
\end{aligned}$$
(A2)

where we performed an integration by part; the integrated term vanishes because there is no mass at infinity in the law of H and $\hat{H}$.

We exploit now the other RDE (29a), that gives

$$\begin{aligned}
\mathbb{P}[\hat{H} \geqslant x] &= \sum_{k=1}^{\infty} \frac{q}{1-q} \frac{[(1-q)\gamma]^k}{k!} \mathbb{P}[\Omega - H \geqslant x]^k \\
&= \frac{q}{1-q}(e^{(1-q)\gamma \mathbb{P}[\Omega - H \geqslant x]} - 1).
\end{aligned}$$

This yields

$$\ln(q + (1-q)\mathbb{P}[\hat{H} \geqslant x]) = -(1-q)\gamma \mathbb{P}[\Omega - H < x], \tag{A3}$$

where we used the equation $q = e^{-\gamma(1-q)}$ to simplify the expression. Inserting (A3) in (A2) gives

$$\begin{aligned}
\mathbb{P}[\hat{H} + \hat{H}' \leqslant \hat{\Omega}] &= \gamma \int_{-\infty}^{+\infty} \left(-\frac{d}{dx}\mathbb{P}[H > x] \right) \mathbb{P}[\Omega - H < x] \\
&= \gamma \, \mathbb{P}[H + H' > \Omega],
\end{aligned}$$

which proves our claim. Note that this derivation relies crucially on the hypothesis that H and $\hat{H}$ have a continuous distribution, which allowed us to introduce their density and to perform integration by parts to connect the two terms of (26). We expect this to be the case when the effective weight distribution is continuous, in such a way that the minimal weight perfect matching is unique (on a finite graph); a counterexample is discussed in Sec. V C.

APPENDIX B: DETAILS ABOUT THE SIMPLIFIED RDE IN SEC. IV B

We provide in this Appendix some additional details about the simplified RDE defined in Sec. IV B; we first give a heuristic justification of the velocity (37) of the leftmost particle of a BRW, then we detail the stochastic ordering argument that leads to the divergence of $\hat{H}^{(n)}$ when $v > 0$.

1. Heuristic derivation of the velocity in the BRW process

We will present a reasoning typical of the physics literature on front propagation in reaction-diffusion systems and equations of the FKPP type (see, for instance, Refs. [15–17]) that leads to the expression (37) for the velocity of the leftmost particle of the BRW.

We define the cumulative distribution function of $\hat{K}^{(n)}$ as $F(x, n) = \mathbb{P}[\hat{K}^{(n)} \leqslant x]$. For a given time n this is an increasing function of x, from 0 to 1 as x increases from $-\infty$ to $+\infty$. The RDE (35) translates into an evolution equation for F as the discrete time increases,

$$F(x, n+1) = 1 - \sum_{k=1}^{\infty} \pi_k \left[1 - \int F(x - \Xi, n)\chi(\Xi)\,d\Xi \right]^k,$$

where π_k is the probability law of the random variable Z, and χ the density of Ξ. We assume that at large times F exhibits a front propagating at a velocity v, and denote F_v the shape of the front in the reference frame moving at this velocity: $F(z + vn, n) \to F_v(z)$ as $n \to \infty$. This gives the following equation on F_v:

$$F_v(z - v) = 1 - \sum_{k=1}^{\infty} \pi_k \left[1 - \int F_v(z - \Xi)\chi(\Xi)\,d\Xi \right]^k,$$

which is equivalent to the RDE (39) on the limit random variable L. When $z \to -\infty$ the distribution function vanishes, in this limit we can thus linearize the equation on F_v, which yields

$$F_v(z - v) = \left(\sum_{k=1}^{\infty} \pi_k k\right) \int F_v(z - \Xi) \chi(\Xi) \, d\Xi. \quad (B1)$$

This linear (integral) equation admits solutions of the form $F_v(z) = e^{\theta z}$, with $\theta > 0$ to respect the increasing character of distribution functions, if θ and v obey the condition

$$e^{-\theta v} = \left(\sum_{k=1}^{\infty} \pi_k k\right) \int e^{-\theta \Xi} \chi(\Xi) \, d\Xi, \quad (B2)$$

which gives a relation $v = v(\theta)$ corresponding to (37). The linearized equation thus admits a family of solutions parametrized by the tail exponent $\theta > 0$, corresponding to velocities $v(\theta)$. The delicate point in this reasoning, for which we refer the reader to the literature, is the justification of the minimum velocity selection principle, namely the fact that the relevant solution of the full nonlinear equation on F_v is the one minimizing $v(\theta)$, as stated in (37).

Note that the minimizer θ_* of $v(\theta)$ corresponds to a double root of the characteristic equation of the linearized equation on F_v, which thus admits as solutions the linear combinations of $e^{\theta_* z}$ and $z \, e^{\theta_* z}$. This enlightens the statement made in (40) for the left tail behavior of the limit random variable L, obtained rigorously in Refs. [20,22].

2. Stochastic ordering argument

Let us prove here the claim made in remark (3) of Sec. IV B, namely, that the sequence of random variables $\hat{K}^{(n)}$ defined in (35) by the simplified RDE provides a stochastic lower-bound for the sequence $\hat{H}^{(n)}$ of the complete RDE (48). We recall that a random variable X is said to be stochastically smaller than a random variable Y if and only if $\mathbb{P}[X > x]$ $\leqslant \mathbb{P}[Y > x]$ for all x, which we denote $X \preceq Y$. A very useful equivalent characterization of this property [35] is the existence of a coupling $(\hat{X}, \hat{Y})$, i.e., a random vector with marginal laws equal to those of X and Y respectively, such that $\mathbb{P}[\hat{X} \leqslant \hat{Y}] = 1$. In other words if $X \preceq Y$ one can consider that X and Y are defined on the same probability space and that $X \leqslant Y$ with probability one.

To simplify the comparison between (35) and (48) we break the iteration (35) in two steps and define another sequence of random variables $K^{(n)}$, with

$$\hat{K}^{(n+1)} \overset{d}{=} \min_{1 \leqslant i \leqslant Z} \left[\Omega_i - K_i^{(n)} \right], \quad (B3a)$$

$$K^{(n)} \overset{d}{=} \hat{\Omega} - \hat{K}^{(n)}. \quad (B3b)$$

We claim that if the initial condition for $\hat{H}^{(n)}$ and $\hat{K}^{(n)}$ is the same, namely, $\hat{H}^{(0)} = \hat{K}^{(0)} = 0$, then for all $n \geqslant 0$ one has $\hat{K}^{(n)} \preceq \hat{H}^{(n)}$ and $H^{(n)} \preceq K^{(n)}$. The proof is obtained by two induction steps on n.

Suppose that $H^{(n)} \preceq K^{(n)}$; then one can couple (48a) and (B3a) by taking the same random variables Z and Ω_i in both, and by using the existence of the coupling of $H^{(n)}$ and $K^{(n)}$ to ensure that $H_i^{(n)} \leqslant K_i^{(n)}$ with probability one, for all $i \in \{1, \dots, Z\}$. This yields a coupling of $\hat{H}^{(n+1)}$ and $\hat{K}^{(n+1)}$ such that $\hat{K}^{(n+1)} \leqslant \hat{H}^{(n+1)}$ with probability one, which proves that $\hat{K}^{(n+1)} \preceq \hat{H}^{(n+1)}$.

Assume now that $\hat{K}^{(n)} \preceq \hat{H}^{(n)}$, and couple (48b) and (B3b) by taking the same random variable $\hat{\Omega}$ in both, the same $\hat{H}_1^{(n)}$ in the two alternatives of (48b), and by ensuring that $\hat{K}^{(n)} \leqslant \hat{H}_1^{(n)}$ with probability one. We have thus coupled $H^{(n)}$ and $K^{(n)}$ in such a way that $H^{(n)} \leqslant K^{(n)}$ with probability one, hence yielding $H^{(n)} \preceq K^{(n)}$.

As the initial condition obviously satisfies the induction hypothesis $\hat{K}^{(0)} \preceq \hat{H}^{(0)}$ this proves our claim: for all n one has $\hat{K}^{(n)} \preceq \hat{H}^{(n)}$, and in particular if $v > 0$ the divergence to $+\infty$ of the sequence $\hat{K}^{(n)}$ as $n \to \infty$ implies the one of $\hat{H}^{(n)}$.

[1] L. Lovász and D. Plummer, *Matching Theory*, AMS Chelsea Publishing Series (American Mathematical Society, Rhode Island, Providence, 2009).

[2] J. Edmonds, Can. J. Math. **17**, 449 (1965).

[3] M. Chertkov, L. Kroc, F. Krzakala, M. Vergassola, and L. Zdeborová, Proc. Nat. Acad. Sci. U. S. A. **107**, 7663 (2010).

[4] A. Decelle, F. Krzakala, C. Moore, and L. Zdeborová, Phys. Rev. E **84**, 066106 (2011).

[5] M. Mézard, M. Tarzia, and C. Toninelli, J. Stat. Phys. **131**, 783 (2008).

[6] V. Bagaria, J. Ding, D. Tse, Y. Wu, and J. Xu, Oper. Res. **68**, 53 (2020).

[7] T. Richardson and R. Urbanke, *Modern Coding Theory* (Cambridge University Press, Cambridge, 2008).

[8] F. Ricci-Tersenghi, G. Semerjian, and L. Zdeborová, Phys. Rev. E **99**, 042109 (2019).

[9] E. Abbe, A. S. Bandeira, and G. Hall, IEEE Trans. Inform. Theory **62**, 471 (2015).

[10] M. Moharrami, C. Moore, and J. Xu, arXiv:1912.08880 [cs.DS] (2019).

[11] M. Mézard and A. Montanari, *Information, Physics, and Computation*, Oxford Graduate Texts (Oxford University Press, Oxford, 2009).

[12] J. M. Hammersley, Ann. Probab. **2**, 652 (1974).

[13] J. F. C. Kingman, Ann. Probab. **3**, 790 (1975).

[14] J. D. Biggins, J. Appl. Probab. **14**, 630 (1977).

[15] E. Brunet and B. Derrida, Phys. Rev. E **56**, 2597 (1997).

[16] S. N. Majumdar and P. L. Krapivsky, Phys. Rev. E **62**, 7735 (2000).

[17] U. Ebert and W. van Saarloos, Physica D **146**, 1 (2000).

[18] M. Bachmann, Adv. Appl. Probab. **32**, 159 (2000).

[19] D. J. Aldous and A. Bandyopadhyay, Ann. Appl. Probab. **15**, 1047 (2005).

[20] E. Aïdékon, Ann. Probab. **41**, 1362 (2013).

[21] Z. Shi, *Branching random walks, Ecole d'Eté de Probabilités de Saint-Flour XLII 2012* (Springer, Cham., 2015).

[22] M. Bramson, J. Ding, and O. Zeitouni, Ann. Inst. H. Poincaré Probab. Statist. **52**, 1897 (2016).

[23] M. Mézard and G. Parisi, J. Phys. Lett. (Paris) **46**, 771 (1985).

[24] M. Bayati, D. Shah, and M. Sharma, IEEE Trans. Inform. Theory **54**, 1241 (2008).

[25] J. Salez and D. Shah, Math. Oper. Res. **34**, 468 (2009).

[26] S. Sanghavi, D. Malioutov, and A. Willsky, IEEE Trans. Inform. Theory **57**, 2203 (2011).

[27] M. Bayati, C. Borgs, J. Chayes, and R. Zecchina, SIAM J. Discrete Math. **25**, 989 (2011).

[28] M. Mézard and G. Parisi, Eur. Phys. J. B **20**, 217 (2001).

[29] L. Zdeborová and M. Mézard, J. Stat. Mech. (2006) P05003.

[30] C. Bordenave, M. Lelarge, and J. Salez, Probab. Theory Relat. Fields **157**, 183 (2013).

[31] D. Gamarnik, T. Nowicki, and G. Swirszcz, Random Struct. Algor. **28**, 76 (2006).

[32] G. Parisi, G. Perrupato, and G. Sicuro, J. Stat. Mech. (2020) 033301.

[33] R. Karp and M. Sipser, in *Proceedings of the Twenty-Second Annual IEEE Symposium on Foundations of Computing*, Nashville Tennessee, (1981), pp. 364-375.

[34] M. Bauer and O. Golinelli, Eur. Phys. J. B **24**, 339 (2001).

[35] T. Lindvall, *Lectures on the Coupling Method*, Dover Books on Mathematics (Dover Publications, Mineola, NY, 2012).