Typical kernel size and number of sparse random matrices over Galois fields: A statistical physics approach

Alamino, R.C., Saad D.
Aston University, Neural Computing Research Group, Birmingham B4 7ET, UK

Using methods of statistical physics, we study the average number and kernel size of general sparse random matrices over $GF(q)$, with a given connectivity profile, in the thermodynamical limit of large matrices. We introduce a mapping of $GF(q)$ matrices onto spin systems using the representation of the cyclic group of order q as the q-th complex roots of unity. This representation facilitates the derivation of the average kernel size of random matrices using the replica approach, under the replica symmetric ansatz, resulting in saddle point equations for general connectivity distributions. Numerical solutions are then obtained for particular cases by population dynamics. Similar techniques also allow us to obtain an expression for the exact and average number of random matrices for any general connectivity profile. We present numerical results for particular distributions.

PACS numbers: 02.10.Yn, 02.70.-c,05.10.-a
Keywords: random matrices, Galois fields, statistical mechanics, replica theory

I. INTRODUCTION

Random matrices over $GF(q)$ are highly important in a number of application areas ranging from biology to computer science and telecommunication. One of the areas where they play a particularly important role is coding theory [1]. In particular, linear codes are defined by the kernel of a parity-check matrix, where each kernel vector is termed a codeword and is associated with an original uncoded message vector by a linear operation defined by a generator matrix. Well known examples include the Hadamard codes, where properties of the kernel and rank play an important role [2], and low-density parity-check codes (LDPC) which provide the best performance to date in many noise regimes. Although the most studied and applied case of LDPC codes is of binary codes over $GF(2)$ there is a significant body of work, of both practical and theoretical nature [3], on codes over more general finite fields showing an improvement in performance with respect to the binary version. In particular, statistical physics based analysis of LDPC codes over $GF(q)$ has been reported in [4].

Low-density parity-check codes are based on random sparse matrices, where the fraction of non-zero elements goes to zero as the size of the matrix increases. In most studies of LDPC codes, it is assumed that a parity-check matrix with M rows (parity-checks) and N columns defines a code of rate $R = 1 - M/N$, *exactly*, which is equivalent to the assertion that the number of vectors in the kernel (and therefore the number of codewords) is exactly q^{NR}.

In addition to being an interesting applied problem, the properties of these matrices are also of great interest from the pure mathematical point of view and a number of papers has already tried to answer related questions in different instances with a mathematical rigorous approach [5–7].

Random matrices are a well studied topic in the physics community where they are important in a range of applications from classical physics to quantum chaos. Recently there has been a lot of activity in the area boosted by the application of techniques originated in the statistical mechanics of disordered systems [8–13]. These techniques have been used to analyse ensemble properties and the replica method has proved to be a valuable tool in several of these approaches. Differently from this paper, however, most of other works concentrate in the spectral properties of the matrices. Also, in most cases, the studied matrices are real, while the restriction to $GF(q)$ matrices considered here makes the solution of the problem more involved.

In this contribution, we address two key properties of sparse random matrices over $GF(q)$, namely the average dimension of their kernel and the number of matrices for a given connectivity profile, in the case of *large matrices*. When the matrices are large, keeping $N \to \infty$ with M/N constant, the problem can be mapped into a system of interacting "spins" and the powerful machinery developed for the study of disordered spin lattices in condensed matter physics can then be used, under some assumptions, to obtain the required properties.

In order to keep this paper as self-contained as possible and make it accessible to a broad readership, we provide in section II a brief introduction to $GF(q)$ matrices and their properties, and to the basic statistical physics methodology on which we have based our analysis. In section III, the usual statistical physics approach to the analysis of LDPC codes over the binary field $GF(2)$ is generalized in such a way that it can be efficiently applied to any $GF(q)$ for a general connectivity distribution of non-zero elements and then used to calculate the average kernel dimension of sparse random matrices (SRM) in section IV. Making use of techniques developed in section IV, the number of matrices for a given distribution of non-zero elements is then obtained for various connectivity profiles, in section V. Finally, we present a discussion of the obtained results in section VI.

II. KEY CONCEPTS

A. $GF(q)$-Matrices

A Galois field $GF(q)$ is a finite field with q elements, i.e., a set of q elements $\{0, ..., q-1\}$, which we symbolize by integers for convenience, which is a commutative group under addition $\oplus : GF(q) \to GF(q)$, defined as integer addition *mod* q, and with a monoid structure with respect to a commutative multiplication operation $\otimes : GF(q) \to GF(q)$. The field also includes the zero element '0', mapping every other element to itself, and the identity '1'; an additional requirement is that the multiplication and addition have the algebraic distributive property. This last requirement restricts the number of elements to be $q = p^n$, where p is a prime number and n an integer.

Entries in matrices over $GF(q)$ take values of numbers in the field $GF(q)$, where the usual additions and multiplications involved in their algebra are defined by the corresponding operations over the Galois field. The kernel, or *null space*, of an $M \times N$ matrix A is defined as the set of vectors $\mathbf{v} \in GF(q)^N$ such that $A\mathbf{v} = 0$, with all operations in the field $GF(q)$. The kernel is a linear vector space and therefore will have $q^{d(A)}$ vectors, where $d(A)$ is the kernel dimension. The rank $r(A)$ of the matrix is obtained by the rank-nullity theorem as $r(A) = N - d(A)$.

B. Disordered Systems

An interacting spin problem has two main elements: an interaction defined between a number of spin units, collectively represented by the vector $\boldsymbol{\sigma} = (\sigma_1, ..., \sigma_N)$, in a lattice and a local field which acts in each variable σ_i separately. Disordered spin systems are systems where one or both of these elements (interaction and field) is a random variable. Usually, we are interested in the properties of very large systems, where the number N of spins becomes infinite, the so-called thermodynamic limit.

The main properties of the system in the thermodynamic limit can be derived from a key quantity, the free-energy f, which in probabilistic terms corresponds to the cummulant generating function. For disordered systems, in the cases where the free-energy is self-averaging with respect to the disorder, we can calculate this quantity as

$$f = -\lim_{N\to\infty} \frac{1}{\beta N} \langle \ln Z \rangle, \tag{1}$$

where $\langle \cdot \rangle$ indicates the disorder average, $Z = \sum_{\boldsymbol{\sigma}} e^{-\beta \mathcal{H}(\boldsymbol{\sigma})}$ is the partition function and $\mathcal{H}(\boldsymbol{\sigma})$ is the Hamiltonian of the system. Although the self-averaging property should be rigorously investigated for each system, we will assume it holds here.

In order to obtain the free-energy, a powerful technique is to make use of the replica method, based on the identity

$$\left[\frac{\partial}{\partial n} \ln \langle Z^n \rangle \right]_{n=0} = \langle \ln Z \rangle \ . \tag{2}$$

Average quantities can then be calculated for integer n and then analytically continued to zero. The replica theory is commonly used in the area of disordered systems and is known to provide exact results in many regimes, which include both physical and non-physical systems [14, 15].

Many problems in computing and communication theory can be mapped to spin systems. For instance, error-correcting codes, in particular LDPC codes [16] and hard computational problems such as K-SAT [17] and graph-coloring [18, 19], can be mapped to diluted spin systems with random p-spin interactions and local fields. In the coding example, interactions are defined by the parity-check constraints, while the local fields are induced by the codeword and received message. In the statistical physics treatment, for mathematical convenience, the message bits $\{0, 1\}$ and '$\oplus$' operation are mapped onto spin values $\{+1, -1\}$ and multiplication using the mapping $x \to (-1)^x$. Variables over a general finite field $GF(q)$, $q \neq 2$ are typically first mapped onto a binary string and then, using the spin values representation, transformed into a spin system [4].

III. MAPPING $GF(q)$ MATRICES INTO SPIN SYSTEMS

The transformation

$$\sigma(v) = (-1)^v, \tag{3}$$

where $\sigma \in \{+1, -1\}$ and $v \in \{0, 1\}$, is usually employed to map the $GF(2)$ variables onto the binary representation. This mapping can be generalized to any $GF(q)$ without an intermediate use of the binary field.

Under the operation $\oplus$, $GF(q)$ is homeomorphic to the cyclic group of order q and therefore has a representation as the complex q-th roots of unity with the group homeomorphism $\sigma : GF(q) \to \mathbb{C}$ given by

$$\sigma(v) = \exp\left(\frac{2\pi i}{q} v\right), \tag{4}$$

such that for every $v_1, v_2 \in GF(q)$

$$\begin{aligned} \sigma(v_1 \oplus v_2) &= \exp\left[\frac{2\pi i}{q}(v_1 \oplus v_2)\right] \\ &= \exp\left[\frac{2\pi i}{q}(v_1 + v_2)\right] \\ &= \exp\left(\frac{2\pi i}{q} v_1\right) \exp\left(\frac{2\pi i}{q} v_2\right) \\ &= \sigma(v_1)\sigma(v_2). \end{aligned} \tag{5}$$

This mapping has a clear geometric interpretation: $2\pi v/q$ is an angle in the unit circle, such that each element of the Galois field is being mapped onto a spin variable "pointing" in one of q possible angles. Using this mapping allows one to write the null-space constraint for a general vector $\mathbf{v} = \left(v^1, ..., v^N\right) \in GF(q)^N$ as

$$\delta(A\mathbf{v}, 0) = \prod_{i=1}^{M} \delta\left[\bigoplus_{j=1}^{N} \left(A_{ij} \otimes v^j\right), 0\right], \tag{6}$$

with

$$\delta\left[\bigoplus_{j=1}^{N} \left(A_{ij} \otimes v^j\right), 0\right] = \frac{1}{\Delta(q)} \prod_{m=1}^{q-1} \left\{1 - \exp\left(-\frac{2\pi i}{q} m\right) \prod_{j=1}^{N} \exp\left[\frac{2\pi i}{q}\left(A_{ij} \otimes v^j\right)\right]\right\}, \tag{7}$$

and

$$\Delta(q) = \prod_{m=1}^{q-1} \left[1 - \exp\left(-\frac{2\pi i}{q} m\right)\right]. \tag{8}$$

Using the properties of the complex roots of unity, the above quantity $\Delta(q)$ can be shown (see appendix A) to be real and equal to the order q of the field.

Based on this representation, we can now define the "magnetization" of the original system in analogy with the spin system as

$$m = \frac{1}{N} \sum_{j=1}^{N} \sigma^j, \tag{9}$$

and the overlap between two configurations $\boldsymbol{\sigma}$ and $\boldsymbol{\sigma}'$ as

$$\rho = \frac{1}{N} \sum_{j=1}^{N} \sigma^j \sigma'^j, \tag{10}$$

where we are now working with the spin variables already mapped to the the complex field $\mathbb{C}$ and therefore the operations of multiplication and addition correspond to the usual ones in $\mathbb{C}$.

It turns out that this kind of representation allows a factorization of the terms simplifying the equations and making the replica calculations simpler, as we will see in the following.

IV. AVERAGE PROPERTIES OF THE KERNEL

The dimension of the kernel of an $M \times N$ matrix A over $GF(q)$ can be written as $d(A) = \log_q \Omega$ where

$$\Omega = \sum_{\mathbf{v}} \delta(A\mathbf{v}, 0), \tag{11}$$

is the number of vectors in the kernel, δ is the Kroenecker delta and $\mathbf{v} \in GF(q)^N$. Direct calculation of Ω from equation (11) by straightforwardly substituting the Kroenecker delta by its integral representation trivially reproduces the rank-nullity theorem. This calculation is not presented here.

The quantity we are interested in here is the average kernel dimension, more specifically, its density in the limit of large matrices, defined as Ts where

$$s \equiv \frac{1}{T} \lim_{N\to\infty} \frac{\langle d(A)\rangle_A}{N} = \lim_{N\to\infty} \frac{1}{N} \langle \ln \Omega \rangle_A, \tag{12}$$

where $1/T = \ln q$ and $M/N \equiv \lambda$, with λ a finite positive constant. Using the replica identity (2), we can write

$$s = \lim_{N\to\infty} \left[\frac{\partial}{\partial n} \ln \langle \Omega^n \rangle_A \right]_{n=0}. \tag{13}$$

The randomly chosen sparse matrices A have exactly K_i non-zero elements in the i-th row with probability $\mathcal{P}(\mathbf{K})$, $\mathbf{K} \equiv (K_1, ..., K_M)$, and C_j elements in the j-th column with probability $\mathcal{P}(\mathbf{C})$, $\mathbf{C} \equiv (C_1, ..., C_N)$, obeying the constraint $\Lambda \equiv \sum_i K_i = \sum_j C_j$, where Λ is the total number of non-zero elements of the matrix. The elements of A are sampled from the finite field $GF(q)$ with independent equal probabilities $\mathcal{P}(A_{ij})$.

Let us define, for brevity of notation, $\mathcal{Z}_n \equiv \langle \Omega^n \rangle_A$. Although the calculations, presented in appendix B, are similar to related calculations in [20, 21], we will use a different approach which is conceptually clearer and has the advantage of allowing later generalizations. In this approach, we sum directly over all entries of the matrix instead of defining a connectivity tensor as used elsewhere [20, 21],

$$\begin{aligned} \mathcal{Z}_n = \Bigg\langle \frac{1}{\mathcal{N}} \sum_{\{A_{ij}\}} \left[\prod_{i,j} \mathcal{P}(A_{ij}) \right] \left[\prod_{i=1}^{M} \delta\left(\sum_{j=1}^{N} \chi(A_{ij}), K_i \right) \right] \left[\prod_{j=1}^{N} \delta\left(\sum_{i=1}^{M} \chi(A_{ij}), C_j \right) \right] \\ \times \prod_{a=1}^{n} \left[\sum_{\mathbf{v}_a} \delta(A\mathbf{v}_a, 0) \right] \Bigg\rangle_{\mathbf{K},\mathbf{C},\Lambda}, \end{aligned} \tag{14}$$

where the average is over the probability distribution $\mathcal{P}(\mathbf{K}, \mathbf{C}, \Lambda)$ with $\chi(A_{ij}) = 0$ if $A_{ij} = 0$ and 1 otherwise, and the normalization $\mathcal{N}$ gives the number of matrices which obey the constraints averaged over the distributions of the entries. In this way, any type of constraint on the matrix can be readily included in the calculation, which could be rather cumbersome in other approaches, based on the introduction of a connectivity tensor as the corresponding constraints have to be written in terms of the tensor elements, which can be extremely complicated.

We refer the reader to appendix B for details of the calculations. Using the replica symmetric ansatz, which is shown to be exact for this problem (see appendix D) we arrive at the following self-consistent saddle point equations

$$\hat{\pi}(\hat{x}) = \frac{1}{\alpha\epsilon(\alpha)} \sum_{i=1}^{M} \left\langle \frac{\alpha^\Lambda}{\Lambda!} K_i \delta\left(\hat{x} - \prod_{l=1}^{K_i - 1} x_l \right) \right\rangle_{\mathbf{x},\mathbf{K},\mathbf{C},\Lambda}, \tag{15}$$

$$\pi(x) = \frac{1}{\alpha\epsilon(\alpha)} \sum_{j=1}^{N} \left\langle \frac{\alpha^\Lambda}{\Lambda!} C_j \delta\left(x - \frac{\prod_{l=1}^{C_j-1} [1+(q-1)\hat{x}_l] - \prod_{l=1}^{C_j-1} (1-\hat{x}_l)}{\prod_{l=1}^{C_j-1} [1+(q-1)\hat{x}_l] + (q-1)\prod_{l=1}^{C_j-1} (1-\hat{x}_l)} \right) \right\rangle_{\hat{\mathbf{x}},\mathbf{K},\mathbf{C},\Lambda}, \tag{16}$$

$$0 = \left\langle \frac{\alpha^\Lambda}{\Lambda!} \left(1 - \frac{\Lambda}{\alpha} \right) \right\rangle_{\mathbf{K},\mathbf{C},\Lambda}, \tag{17}$$

with

$$\epsilon(\alpha) = \left\langle \frac{\alpha^\Lambda}{\Lambda!} \right\rangle_{\mathbf{K},\mathbf{C},\Lambda}, \tag{18}$$

and to the corresponding expression for s

$$\begin{aligned} s = &-\lambda \ln q - \frac{\alpha}{N} \langle \ln [1 + (q-1)x\hat{x}] \rangle_{x,\hat{x}} \\ &+ \frac{1}{N\epsilon(\alpha)} \sum_i \left\langle \frac{\alpha^\Lambda}{\Lambda!} \left\langle \ln \left[1 + (q-1) \prod_{l=1}^{K_i} x_l \right] \right\rangle_{\mathbf{x}} \right\rangle_{\mathbf{K},\mathbf{C},\Lambda} \\ &+ \frac{1}{N\epsilon(\alpha)} \sum_j \left\langle \frac{\alpha^\Lambda}{\Lambda!} \left\langle \ln \left\{ \prod_{l=1}^{C_j} [1 + (q-1)\hat{x}_l] + (q-1) \prod_{l=1}^{C_j} (1 - \hat{x}_l) \right\} \right\rangle_{\hat{\mathbf{x}}} \right\rangle_{\mathbf{K},\mathbf{C},\Lambda}. \end{aligned} \tag{19}$$

The probability distributions $\pi(x)$ and $\hat{\pi}(\hat{x})$ in equations (15) and (16) are of the auxiliary effective field x and its conjugate $\hat{x}$, respectively. They represent a generating function for the order parameters defined as site averages of a given number of system replicas. They can also be interpreted as message probability distributions within the microscopic message passing framework, also known as belief propagation[22]; a more detailed explanation is provided in [23, 24]. Their exact definition is given in equations (B17) and (B18) of appendix B.

It must be noted that the above equations are only meaningful if $\Lambda \propto N$. A striking property of the above equations is that they are *completely independent* of the specific distribution of the individual elements of the matrix, depending only on the distribution of $\mathbf{K}$ and $\mathbf{C}$ (and, obviously, of Λ).

There exists two straightforward analytical solutions of the above equations, namely, the paramagnetic one given by

$$\hat{\pi}(\hat{x}) = \delta(\hat{x}),\ \pi(x) = \delta(x), \tag{20}$$

and the ferromagnetic solution

$$\hat{\pi}(\hat{x}) = \delta(\hat{x} - 1),\ \pi(x) = \delta(x - 1). \tag{21}$$

When substituted in the above equations, the paramagnetic solution gives the average kernel density as $Ts = 1 - \lambda = 1 - M/N$ independently of the order q of the finite field used. In the case of LDPC codes defined by such matrices, this corresponds to random parity-check matrices that defines a code of rate $R = 1 - \lambda$. The average rank density in this case is λ. The ferromagnetic solution gives $Ts = 0$ and the matrix is full rank; which incidentally means that such matrices cannot be used to define a parity-check code due to the lack of redundancy.

These quantities can be associated to analogous quantities in the statistical mechanics framework. We start by associating the average rank density with the free-energy f and writing

$$f \equiv \frac{\langle r(A) \rangle_A}{N} = 1 - Ts, \tag{22}$$

which allows one to associate s with the entropy and the internal energy density being constrained to be $u = 1$. Defining $\beta = 1/T$, equation (22) becomes

$$\begin{aligned} \beta f &= 1 - \frac{1}{N} \left\langle \ln \sum_{\mathbf{v}} \delta(A\mathbf{v}, 0) \right\rangle_A \\ &= \frac{1}{N} \left\langle N \ln e^{\beta} - \ln \sum_{\mathbf{v}} \delta(A\mathbf{v}, 0) \right\rangle_A \\ &= -\frac{1}{N} \left\langle \ln \sum_{\mathbf{v}} \delta(A\mathbf{v}, 0)\, e^{-\beta N} \right\rangle_A \\ &= -\frac{1}{N} \left\langle \ln \sum_{\mathbf{v}} e^{-\beta \mathcal{H}(\mathbf{v})} \right\rangle_A, \end{aligned} \tag{23}$$

where the Hamiltonian of the corresponding statistical mechanical system is formally

$$\mathcal{H}(\mathbf{v}) \equiv N - \ln \delta(A\mathbf{v}, 0). \tag{24}$$

We solved the saddle point equations by means of population dynamics for three different cases, in all of which we keep K fixed

1. Regular matrices - C and K fixed;
2. Fixed K and $\mathbf{C}$ drawn from a multinomial uniform probability

$$\mathcal{P}(\mathbf{C}) = \frac{(MK)!}{\prod_j C_j!} \frac{1}{N^{MK}}; \tag{25}$$

3. Fixed K while $\mathbf{C}$ values are drawn from a Poisson integer distribution of mean $\Lambda/N = \lambda K$, for each column separately, until the limit of MK non-zero elements is reached.

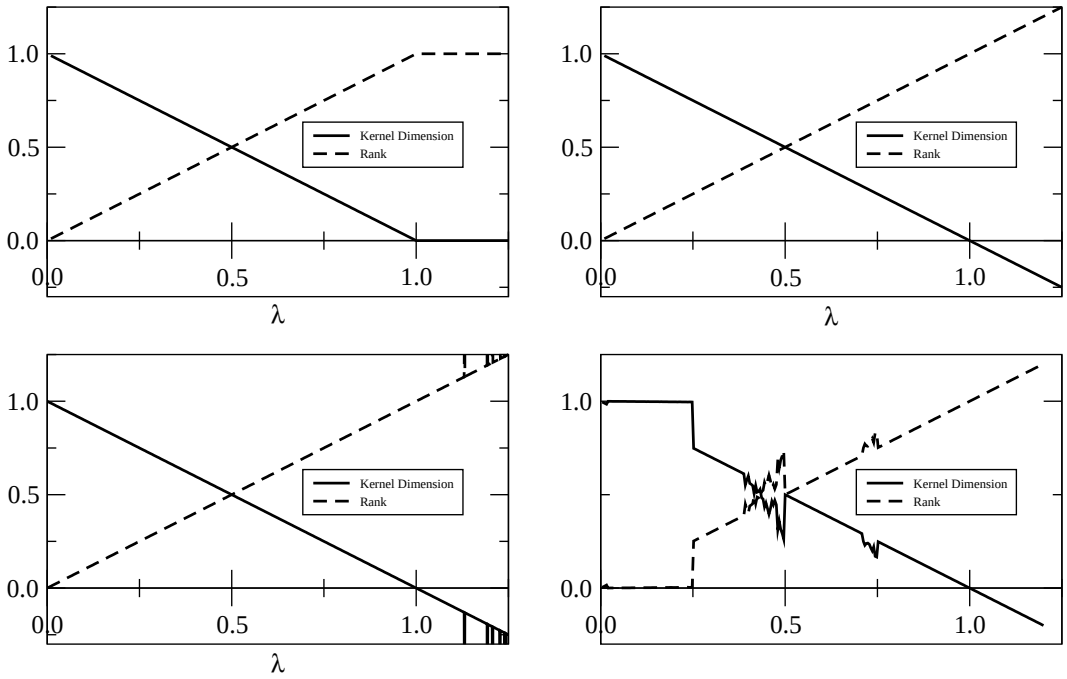

FIG. 1: Average kernel dimension density (continuous lines) and average rank density (dashed lines) calculated as solutions to the replica symmetric saddle point equations. The top left plot shows the thermodynamically favored solution (paramagnetic for $0 \leq \lambda \leq 1$ and ferromagnetic for $\lambda > 1$). The top right shows the regular case (i) for fixed K and C. Cases (ii) and (iii) are presented at the bottom left and right, respectively. Note that numerical instabilities occur for specific λ values.

Population dynamics [25] is an iterative method to obtain the probability distributions of the auxiliary fields. Firstly, a large population of fields x and $\hat{x}$ is generated and each individual is initialised to a random value in the interval $[-1, +1]$. Then, the whole population is sampled in a random order and updated according to the relations defined by the saddle-point equations. This process is repeated a large number of times and averaged over all realisations. The number of iterations in the algorithm is fixed due to the fact that close to the critical points it exhibits a critical slowing down and convergence of the population is extremely slow.

Results for the various cases are presented in Fig. 1. The top left plot shows the theoretical thermodynamically dominant solutions (paramagnetic in the range $0 \leq \lambda \leq 1$ and ferromagnetic for $\lambda > 1$) having the lower free energy.

The top right plot shows the results for the regular case (i). Solutions were obtained numerically by iterating equations (15) and (16) for the case of $q = 4$ and $K = 200$; C was varied from 2 to 250. Repeating the calculations for different values of q and K have produced similar results. We see that the stable solution is always paramagnetic, but becomes unphysical at $\lambda = 1$ once the entropy, and consequently the dimension of the kernel, becomes negative.

In the case of parity-check codes, this result means that the typical parity-check matrix defines a code of rate exactly $(N - M)/N$. This is assumed for any parity-check matrix in most calculations in the literature and is confirmed by our results to be true on average; however, it is important to point out that the result is true in the limit of large matrices and is likely to have finite size corrections which may affect practical applications.

Cases (ii) and (iii) are presented, respectively, at the bottom left and right of Fig. 1. Although these cases do not rigorously obey the constraint that each C_j must be at most M, for large matrices and small values of K (which is what happens in practice) C_j is unlikely to exceed this value. However, instabilities can and indeed occur for specific λ values, presumably due to instances where C_j takes higher values.

The bottom left plot shows results for the case (ii), with $q = 3$, $K = 4$, $N = 1000$ and $1 \leq M \leq 1250$. Also in this case, the stable dominant solution is paramagnetic. Numerical instabilities, which disappear slowly with the increase in the number of fields and steps in the population dynamics, emerge in the unphysical region and are shown in the figure.

The behavior for case (iii) is a little more complex due to the nature of the distribution chosen. Using the average value λK for the variables C_j implies that, as λ varies, their average value also changes. The plot shown was obtained for $q = 2$, $K = 4$, $N = 250$ and $1 \leq M \leq 300$. There are clearly several special points in this plot, which distinguish it from the previous cases. The first point separates λ values which give rise to average connectivity values lower/higher than 1 (left and right, respectively). Up to this point, the matrix has too many zero columns, pushing the kernel size to cover the full space of vectors; this transition is of a different nature to the ones described previously. The other

two points, where numerical instabilities emerge, are related to the percolation transition. Further calculations with different K values indicate that these points appear around the extremes of the interval $2/K \leq \lambda \leq 3/K$. Inside this interval, the average value of the C_j's equal to 2 (once we take it to be an integer). This value marks the percolation transition for binary matrices; the numerical instabilities in this case are associated with a critical slowing down of the algorithm close to the transition point as the algorithm stops after a pre-determined number of iterations. Apart from these differences, the resulting curve seems to coincide with those obtained for the previous cases.

The solution of kernel size problem is mathematically equivalent to the solution of LDPC in channels with infinite noise. As the solution in the latter is paramagnetic, we are led to speculate that it is the dominant solution also here up to the point where the quantity s, analogous to the entropy, becomes negative. From this point and on the solution becomes ferromagnetic. The numerical results seem to support this conjecture, although more careful calculations, varying all the parameters involved must be carried out to confirm this hypothesis more generally.

V. NUMBER OF MATRICES

The number of $GF(q)$ matrices given a connectivity profile is of significant interest within the discrete mathematics community. Exact results have been obtained for the case of *finite* binary matrices [26] in the form of a formula that facilitates the calculation of their precise number. In this paper we will analyze the case of large $GF(q)$ matrices and provide an expression for both their exact and average number. Given the precise number of non-zero elements per row $\mathbf{K} = (K_1, ..., K_M)$ and per column $\mathbf{C} = (C_1, ..., C_N)$, one can write the number of matrices as

$$N_A = \sum_{\{A_{ij}\}} \left[\prod_{i=1}^{M} \delta\left(\sum_{j=1}^{N} \chi(A_{ij}), K_i \right) \right] \left[\prod_{j=1}^{N} \delta\left(\sum_{i=1}^{M} \chi(A_{ij}), C_j \right) \right]. \tag{26}$$

Note that we are using the summation directly over the entries of the matrix instead of the introduction of a connectivity tensor. In this way, the calculations are similar to the ones for obtaining the kernel dimension with the details given in C. The final result is

$$N_A = (q-1)^{\Lambda} \frac{\Lambda!}{\prod_i K_i! \prod_j C_j!}. \tag{27}$$

Note that the component on the right represents the number of binary matrices with the given non-zero elements profile. The factor $(q-1)^{\Lambda}$ is the multiplicity of the non-zero entries which can have any non-zero value in the Galois field.

If we consider a distribution $\mathcal{P}(\mathbf{K}, \mathbf{C}, \Lambda)$, we can look at the *average* number of matrices

$$\bar{N}_A = \left\langle (q-1)^{\Lambda} \frac{\Lambda!}{\prod_i K_i! \prod_j C_j!} \right\rangle_{\mathbf{K},\mathbf{C},\Lambda}. \tag{28}$$

Note that we can write the joint probability distribution as

$$\mathcal{P}(\mathbf{K}, \mathbf{C}, \Lambda) = \mathcal{P}(\mathbf{K}|\Lambda, \mathbf{C}) \mathcal{P}(\Lambda|\mathbf{C}) \mathcal{P}(\mathbf{C}), \tag{29}$$

and that $\mathcal{P}(\Lambda|\mathbf{C}) = \delta\Big(\Lambda, \sum_j C_j\Big)$. Therefore, we have obtained for the average number of matrices

$$\bar{N}_A = \sum_{\mathbf{K}} \sum_{\mathbf{C}} \mathcal{P}(\mathbf{K}|\mathbf{C}) \mathcal{P}(\mathbf{C}) (q-1)^{\sum_j C_j} \frac{\Big(\sum_j C_j\Big)!}{\prod_i K_i! \prod_j C_j!}, \tag{30}$$

where the distribution $\mathcal{P}(\mathbf{K}|\mathbf{C})$ includes the constraint $\delta\Big(\sum_i K_i, \sum_j C_j\Big)$.

A simple calculation shows that for the regular case, where all C_j's and K_i's are fixed (to C and K, respectively), and $q = 2$, the number of matrices scales as N^{CN}. Therefore, a more appropriate quantity to calculate instead of the average number of matrices would be the quenched entropy

$$\Xi \equiv \left\langle \frac{1}{N} \ln N_A \right\rangle = \frac{1}{N} \sum_{\mathbf{K}} \sum_{\mathbf{C}} \mathcal{P}(\mathbf{K}|\mathbf{C}) \mathcal{P}(\mathbf{C}) \ln \left[(q-1)^{\sum_j C_j} \frac{\Big(\sum_j C_j\Big)!}{\prod_i K_i! \prod_j C_j!} \right], \tag{31}$$

TABLE I: Asymptotic values of Ξ^* for large λ

$\bar{C}$	As. Value
5	29.66
10	60.73
20	123.20

which scales as $\ln N$.

We analyze the behavior of this quantity for three different cases. We choose each C_j to be i.i.d. and K to be chosen from a multinomial distribution

$$\mathcal{P}(\mathbf{K}) = \frac{(\sum_i K_i)!}{\prod_i K_i!} \frac{1}{N^{\sum_i K_i}} \delta\left(\sum_i K_i, \sum_j C_j\right), \tag{32}$$

for each realization of $\mathbf{C}$. The three probability distributions for the variables C_j to be analyzed are

1. uniform in the interval $[0, 2\bar{C}]$

$$\mathcal{P}(C_j) = 1/(2\bar{C}+1); \tag{33}$$

2. binomial in the interval $[0, M]$

$$\mathcal{P}(C_j) = \binom{M}{C_j} \left(\frac{\bar{C}}{M}\right)^{C_j} \left(1 - \frac{\bar{C}}{M}\right)^{M-C_j}; \tag{34}$$

3. Zipf distribution for $C_j = 1, ..., M$

$$\mathcal{P}(C_j) = \frac{C_j^{-s}}{\sum_{n=1}^{K} n^{-s}}, \tag{35}$$

where $\bar{C}$ is the mean of the distributions. The motivation for choosing these connectivity profiles is that they appear to be the most commonly analyzed and feature (especially the latter) in recent analysis and modelling of networks.

Results for the binomial (dashed line) and uniform (dotted line) distributions with means $\bar{C} = 5.0, 10.0, 20.0$, $q = 2$ and $N = 300$ are plotted in Fig. 2, together with the value of Ξ with constant $C_j = \bar{C}$ and $K_j = \bar{C}/\lambda$ values for all i and j. This function is explicitly given by

$$\Xi^* = \bar{C}\ln(q-1) - \ln\bar{C}! + \frac{1}{N}\ln(NC)! - \lambda\ln\left(\bar{C}/\lambda\right)!, \tag{36}$$

and we can obtain its asymptotic behavior for small and large λ as

$$\lambda \ll 1 \Rightarrow \Xi^* = \bar{C}\ln(q-1) - \ln\bar{C}! + \bar{C}\ln\lambda N, \tag{37}$$

$$\lambda \gg 1 \Rightarrow \Xi^* = \bar{C}\ln(q-1) - \ln\bar{C}! + \bar{C}\ln\bar{C}N + (\gamma-1)\bar{C}, \tag{38}$$

where $\gamma \approx 0.577216$ is the Euler-Mascheroni constant. Asymptotic limits for large λ are given in table I.

For large λ values the result for constant $\mathbf{C}$ and $\mathbf{K}$ upper-bounds the other two distributions. Additional calculations seem to indicate that it is always the case for any distribution, although a proof for this conjecture is still sought. This implies that if we keep the number of columns constant and increase the ratio λ by adding rows, whenever the number of rows is much larger than the number of columns, the average number of matrices becomes independent of both the ratio and number of rows. The plots also suggest that the average number of matrices in these cases are basically defined by the average value of the $\mathbf{C}$ distributions.

For small values of λ, the uniform distribution continues to be upper-bounded by the constant distribution. The binomial distribution, however, is higher for a small interval around zero. This behavior is shown in the inset where lower $\mathbf{C}$ values give rise to higher Ξ as λ becomes smaller.

Figure 3 shows the results for the Zipf distribution with different values for the power s compared with a uniform distribution in the range $[0, M]$. In this case, the mean of the distributions vary with λ. We see that, although the average value of the Zipf distributions increasingly differs from the uniform value $M/2$ as s increases, the average number of matrices actually becomes highly similar.

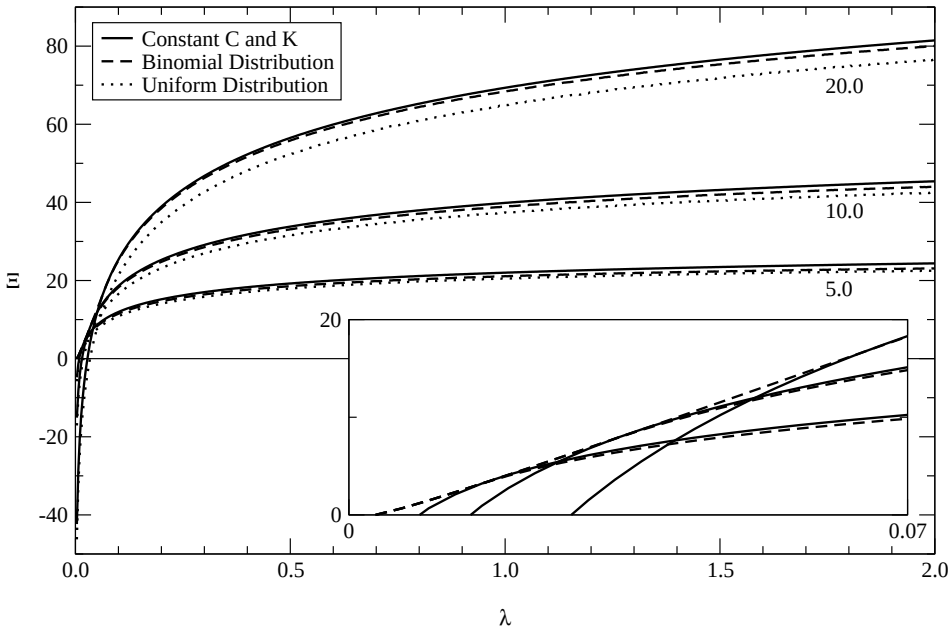

FIG. 2: Values of the quenched entropy Ξ versus λ for the different distributions and various $\mathbf{C}$ values ($\mathbf{C} = 5, 10, 20$), with multinomial $\mathbf{K}$: constant (continuous line), binomial (dashed line) and uniform (dotted line). The inset shows in detail the small λ regime, where just the binomial and constant distributions are represented. The higher lines on the right correspond to the higher $\mathbf{C}$ values.

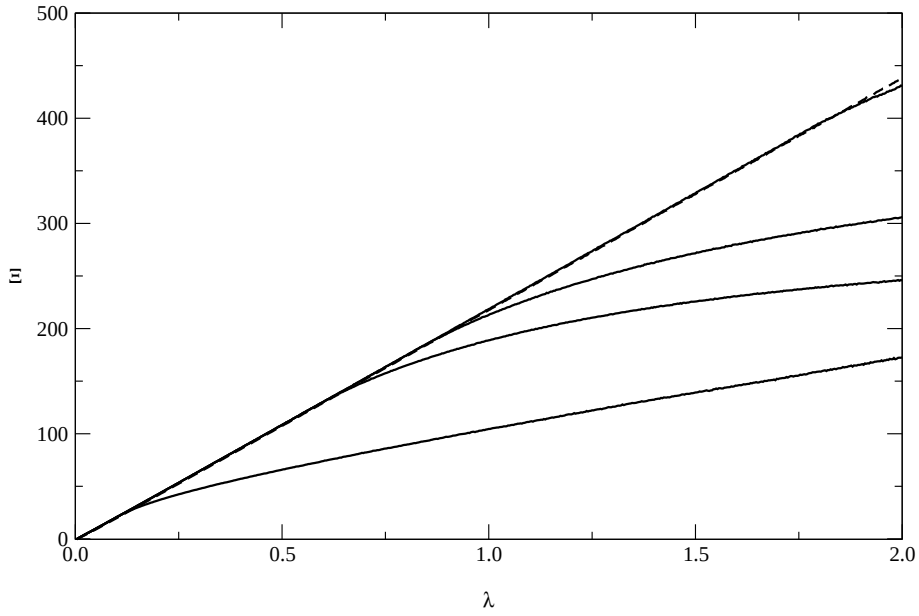

FIG. 3: Values of Ξ versus λ for the uniform distribution (dashed line) and the Zipf distribution (continuous lines) for $s = 1, 3, 4, 10$, respectively, from bottom to top.

VI. CONCLUSIONS

We have introduced a new mapping of Galois matrices to spin systems based on the group homeomorphism between $GF(q)$ under addition mod q (denoted by $\oplus$) and the complex q-th roots of unity. In addition, we have introduced a different way for summing over random matrices that can be generalized to include any kind of connectivity constraint and is conceptually cleaner and simpler than the existing approaches. The new mapping and alternative summation over random matrices allows for a factorization of the constraints, which simplifies calculations of the kernel and the

number of matrices under various connectivity profiles.

Using the replica approach and these new introduced techniques, we calculated the average dimension of the kernel for a general distribution of non-zero entries and solved the resulting equations numerically, finding that the average kernel density is $1 - M/N$ in all cases studied. We conjecture that this result is always valid. Based on the analogy with thermodynamical quantities corresponding to free energy, internal energy and Hamiltonian, we showed that the replica symmetric ansatz in this case must be exact. With the same techniques, we were also able to find the total number of large matrices for fixed $\mathbf{K}$ and $\mathbf{C}$ and their average number, which was then computed for different distributions of theoretical and practical relevance.

The results presented have practical relevance in a number of areas, including coding network modelling and some biological models. With respect to LDPC codes, the average kernels density result implies that randomly generated LDPC codes typically define codes of rate exactly $1 - M/N$, an assumption which is generally made but lacks rigorous derivations. Also, as the parity-parity check matrix can represent the connectivities in graphs (see [27]), the results obtained for the average number of matrices provide a principled approach to determine the average number of possible graphs with a given connectivity distributions of a more general nature than the connectivity profiles examined in this paper.

Acknowledgements

We would like to thank the invaluable comments and suggestions of the anonymous referees. Support from EPSRC grant EP/E049516/1 is gratefully acknowledged. R.C.A. would also like to thank Dr. Juan P. Neirotti for useful discussions.

APPENDIX A: PROOF OF $\Delta(q) = q$

In this appendix we prove the statement made in section IV that $\Delta(q) = q$ where [29]

$$\Delta(q) = \prod_{m=1}^{q-1} \left[1 - \exp\left(-\frac{2\pi i}{q} m\right)\right]. \tag{A1}$$

From the above equation, we have

$$\begin{aligned} \Delta(q) &= \prod_{m=1}^{q-1} \exp\left(-\frac{\pi i}{q} m\right) \left[\exp\left(\frac{\pi i}{q} m\right) - \exp\left(-\frac{\pi i}{q} m\right)\right] \\ &= (2i)^{q-1} \exp\left(-\frac{\pi i}{q} \sum_{m=1}^{q-1} m\right) \prod_{m=1}^{q-1} \sin\left(m \frac{\pi}{q}\right). \end{aligned} \tag{A2}$$

The identity

$$\exp\left(-\frac{\pi i}{q} \sum_{m=1}^{q-1} m\right) = i^{1-q}, \tag{A3}$$

implies the equation

$$\Delta(q) = 2^{q-1} \prod_{m=1}^{q-1} \sin\left(m \frac{\pi}{q}\right). \tag{A4}$$

Using the known identity [28]

$$\sin(qx) = 2^{q-1} \prod_{m=0}^{q-1} \sin\left(x + m \frac{\pi}{q}\right), \tag{A5}$$

divided by $\sin x$ and taking $x \to 0$, one obtains

$$\prod_{m=1}^{q-1} \sin\left(m \frac{\pi}{q}\right) = \frac{q}{2^{q-1}}, \tag{A6}$$

which by substituting into equation (A4) gives the desired result.

APPENDIX B: REPLICA SYMMETRIC SADDLE POINT EQUATIONS

Using integral representations for the first two sets of Kroenecker delta functions, we can write the averaged replicated kernel size defined in equation (14) as

$$\begin{aligned}\mathcal{Z}_n = \Bigg\langle \frac{1}{\mathcal{N}} \sum_{\{\mathbf{v}_a\}} \oint DWDZ \sum_{\{A_{ij}\}} \left[\prod_{i,j} \mathcal{P}(A_{ij}) (W_i Z_j)^{\chi(A_{ij})} \right] \\ \times \prod_{i=1}^{M} \prod_{a} \delta \left[\bigoplus_{j=1}^{N} (A_{ij} \otimes v_a^j), 0 \right] \Bigg\rangle_{\mathbf{K},\mathbf{C},\Lambda},\end{aligned} \tag{B1}$$

where $\otimes$ and $\oplus$ indicate multiplication and summation on $GF(q)$, respectively, and

$$DWDZ = \left[\prod_{i=1}^{M} \frac{dW_i}{W_i^{K_i+1}} \right] \left[\prod_{j=1}^{N} \frac{dZ_j}{Z_j^{C_j+1}} \right]. \tag{B2}$$

Using the representation of the parity-check constraint given in equation (6), the product over replica indices of the delta function can be written as

$$\begin{aligned}\prod_a \delta \left[\bigoplus_{j=1}^{N} (A_{ij} \otimes v_a^j), 0 \right] &= \prod_a \frac{1}{q} \prod_{m=1}^{q-1} \left\{ 1 - \exp\left(-\frac{2\pi i}{q} m \right) \prod_{j=1}^{N} \exp\left[\frac{2\pi i}{q} (A_{ij} \otimes v_a^j) \right] \right\} \\ &= \frac{1}{q^n} \prod_a \left[1 + \sum_{s=1}^{q-1} F_i(s,a) G(s) \right] \\ &= \frac{1}{q^n} \sum_{r=0}^{n} \sum_{\langle a_1 \cdots a_r \rangle} \sum_{s_1,\ldots,s_r} G(s_1) \cdots G(s_r) F_i(s_1,a_1) \cdots F_i(s_r,a_r),\end{aligned} \tag{B3}$$

with

$$G(s) \equiv \sum_{\langle m_1 \cdots m_s \rangle} (-1)^s \exp\left(-\frac{2\pi i}{q} m_1 \right) \cdots \exp\left(-\frac{2\pi i}{q} m_s \right), \tag{B4}$$

and

$$\begin{aligned}F_i(s,a) &\equiv \exp\left[\frac{2\pi i}{q} (A_{i1} \otimes v_a^1) \right] \cdots \exp\left[\frac{2\pi i}{q} (A_{iN} \otimes v_a^N) \right] \\ &= \prod_{j=1}^{N} \gamma_j(s,a,A_{ij}),\end{aligned} \tag{B5}$$

where we defined, for simplicity,

$$\gamma_j(s,a,A_{ij}) \equiv \exp\left[\frac{2\pi i}{q} s (A_{ij} \otimes v_a^j) \right]. \tag{B6}$$

We can now write the partition function as

$$\mathcal{Z}_n = \left\langle \frac{1}{\mathcal{N}} \sum_{\{\mathbf{v}_a\}} \oint DZ \prod_{i=1}^{M} \frac{1}{q^n} \sum_{r=0}^{n} \sum_{\langle a_1 \cdots a_r \rangle} \sum_{s_1,\ldots,s_r} G(s_1) \cdots G(s_r) \oint \frac{dW_i}{2\pi i} \frac{1}{W_i^{K_i+1}} \Gamma_i \right\rangle_{\mathbf{K},\mathbf{C},\Lambda}, \tag{B7}$$

where

$$\begin{aligned}\Gamma_i &= \sum_{A_{i1},\ldots,A_{iN}} \left[\prod_j \mathcal{P}(A_{ij})(W_i Z_j)^{\chi(A_{ij})}\right] \prod_j \gamma_j(s_1,a_1,A_{ij})\cdots\gamma_j(s_r,a_r,A_{ij}) \\ &= \prod_j \sum_{A_{ij}} \mathcal{P}(A_{ij})(W_i Z_j)^{\chi(A_{ij})}\gamma_j(s_1,a_1,A_{ij})\cdots\gamma_j(s_r,a_r,A_{ij}) \\ &= p^N \prod_j \left[1+\frac{1}{p}\sum_{h=1}^{q-1}\mathcal{P}(A_{ij}=h)W_i Z_j \gamma_j(s_1,a_1,h)\cdots\gamma_j(s_r,a_r,h)\right],\end{aligned} \tag{B8}$$

where we define, for convenience, $p \equiv \mathcal{P}(A_{ij}=0)$. Let us define a probability distribution over the values of h as

$$\mathcal{P}(h) = \frac{\mathcal{P}(A_{ij}=h)}{1-p}, \tag{B9}$$

in such a way that h varies from 1 to $q-1$ and the probability over this range is correctly normalized. Then

$$\begin{aligned}\Gamma_i &= p^N \prod_j \left[1+\left(\frac{1-p}{p}\right) W_i Z_j \langle \gamma_j(s_1,a_1,h)\cdots\gamma_j(s_r,a_r,h)\rangle_h\right] \\ &= p^N \sum_{l=0}^{N} \sum_{\langle j_1\cdots j_l\rangle} \left(\frac{1-p}{p}\right)^l W_i^l Z_{j_1}\cdots Z_{j_l} \\ &\quad \times \langle\gamma_{j_1}(s_1,a_1,h)\cdots\gamma_{j_1}(s_r,a_r,h)\rangle_h \cdots \langle\gamma_{j_l}(s_1,a_1,h)\cdots\gamma_{j_l}(s_r,a_r,h)\rangle_h.\end{aligned} \tag{B10}$$

The integrals over the W_i's, acting on the Γ_i's, select the power of W_i to be K_i and we therefore obtain

$$\begin{aligned}\mathcal{Z}_n &= \left\langle \kappa \sum_{\{\mathbf{v}_a\}} \oint DZ \prod_{i=1}^{M} \left\{ \sum_{r=0}^{n} \sum_{\langle a_1\cdots a_r\rangle} \sum_{s_1,\ldots,s_r} G(s_1)\cdots G(s_r) \sum_{\langle j_1\cdots j_{K_i}\rangle} Z_{j_1}\cdots Z_{j_{K_i}} \right.\right. \\ &\quad \left.\left. \times \langle\gamma_{j_1}(s_1,a_1,h)\cdots\gamma_{j_1}(s_r,a_r,h)\rangle_h \cdots \left\langle\gamma_{j_{K_i}}(s_1,a_1,h)\cdots\gamma_{j_{K_i}}(s_r,a_r,h)\right\rangle_h \right\}\right\rangle_{\mathbf{K},\mathbf{C},\Lambda} \\ &\approx \left\langle \kappa \sum_{\{\mathbf{v}_a\}} \oint DZ \prod_{i=1}^{M} \left\{ \sum_{r=0}^{n} \sum_{\langle a_1\cdots a_r\rangle} \sum_{s_1,\ldots,s_r} G(s_1)\cdots G(s_r) \right.\right. \\ &\quad \left.\left. \times \frac{N^{K_i}}{K_i!}\left[\frac{1}{N}\sum_{j=1}^{N} Z_j \langle\gamma_j(s_1,a_1,h)\cdots\gamma_j(s_r,a_r,h)\rangle_h\right]^{K_i} \right\}\right\rangle_{\mathbf{K},\mathbf{C},\Lambda}\end{aligned} \tag{B11}$$

where

$$\kappa = p^{NM}\left(\frac{1-p}{p}\right)^{\sum_i K_i} \mathcal{N}^{-1} q^{-nM}. \tag{B12}$$

The calculation of $\mathcal{N}$ is similar to the calculation of the number of matrices shown in appendix C and we end up with

$$\kappa = \frac{1}{q^{nM} N_A^{(2)}}, \tag{B13}$$

where $N_A^{(2)}$ is exactly the number of binary matrices ($q=2$) as calculated in appendix C. Introducing the replica overlaps

$$Q^{s_1,\ldots,s_r}_{\langle a_1\cdots a_r\rangle} \equiv \frac{1}{N}\sum_{j=1}^{N} Z_j \langle\gamma_j(s_1,a_1,h)\cdots\gamma_j(s_r,a_r,h)\rangle_h, \tag{B14}$$

and the corresponding auxiliary variables $\hat{Q}^{s_1,\ldots,s_r}_{\langle a_1\cdots a_r\rangle}$ by means of Dirac delta functions, we can express the partition function as

$$\begin{aligned}
\mathcal{Z}_n &= \int DQD\hat{Q}\exp\Big(-N\sum Q^{s_1,\ldots,s_r}_{\langle a_1\cdots a_r\rangle}\hat{Q}^{s_1,\ldots,s_r}_{\langle a_1\cdots a_r\rangle}\Big)\\
&\quad\times\left\langle \kappa\frac{N^{\sum_i K_i}}{\prod_i K_i!}\prod_i\left[\sum G(s_1)\cdots G(s_r)\Big(Q^{s_1,\ldots,s_r}_{\langle a_1\cdots a_r\rangle}\Big)^{K_i}\right]\right.\\
&\quad\left.\times\prod_j\left\{\sum_{\{v_a^j\}}\oint DZ_j\exp\Big[Z_j\sum\hat{Q}^{s_1,\ldots,s_r}_{\langle a_1\cdots a_r\rangle}\langle\gamma_j(s_1,a_1,h)\cdots\gamma_j(s_r,a_r,h)\rangle_h\Big]\right\}\right\rangle_{\mathbf{K},\mathbf{C},\Lambda}\\
&= \int DQD\hat{Q}\exp\Big(-N\sum Q^{s_1,\ldots,s_r}_{\langle a_1\cdots a_r\rangle}\hat{Q}^{s_1,\ldots,s_r}_{\langle a_1\cdots a_r\rangle}\Big)\\
&\quad\times\left\langle q^{-nM}\frac{N^{\sum_i K_i}}{(\sum_i K_i)!}\prod_i\left[\sum G(s_1)\cdots G(s_r)\Big(Q^{s_1,\ldots,s_r}_{\langle a_1\cdots a_r\rangle}\Big)^{K_i}\right]\right.\\
&\quad\left.\times\prod_j\left\{\sum_{\{v_a^j\}}\Big[\sum\hat{Q}^{s_1,\ldots,s_r}_{\langle a_1\cdots a_r\rangle}\langle\gamma_j(s_1,a_1,h)\cdots\gamma_j(s_r,a_r,h)\rangle_h\Big]^{C_j}\right\}\right\rangle_{\mathbf{K},\mathbf{C},\Lambda}
\end{aligned}\tag{B15}$$

where

$$DQD\hat{Q}\equiv\left(\prod\frac{dQ\,d\hat{Q}}{2\pi i/N}\right),\tag{B16}$$

and the summations run over all the allowed values of r, $\langle a_1\cdots a_r\rangle$ and $s_1,\ldots s_r$.

Under the assumption of replica symmetry in the form

$$Q^{s_1,\ldots,s_r}_{\langle a_1\cdots a_r\rangle} = Q_0\langle x^r\rangle_x,\tag{B17}$$

$$\hat{Q}^{s_1,\ldots,s_r}_{\langle a_1\cdots a_r\rangle} = \hat{Q}_0\langle\hat{x}^r\rangle_{\hat{x}},\tag{B18}$$

where the averages over x and $\hat{x}$ are taken with respect to the field distributions $\pi(x)$ and $\hat{\pi}(\hat{x})$ respectively, we can show by straightforward algebraic manipulations that

$$\sum Q^{s_1,\ldots,s_r}_{\langle a_1\cdots a_r\rangle}\hat{Q}^{s_1,\ldots,s_r}_{\langle a_1\cdots a_r\rangle} = Q_0\hat{Q}_0\langle[1+(q-1)x\hat{x}]^n\rangle_{x,\hat{x}},\tag{B19}$$

$$\sum G(s_1)\cdots G(s_r)\Big(Q^{s_1,\ldots,s_r}_{\langle a_1\cdots a_r\rangle}\Big)^{K_i} = Q_0^{K_i}\left\langle\left\{1+\left[\sum_s G(s)\right]\prod_{l=1}^{K_i}x_l\right\}^n\right\rangle_{\mathbf{x}},\tag{B20}$$

where it is easy to see that

$$\sum_s G(s) = \Delta(q)-1 = q-1,\tag{B21}$$

and

$$\begin{aligned}
&\sum_{\{v_a^j\}}\Big[\sum\hat{Q}^{s_1,\ldots,s_r}_{\langle a_1\cdots a_r\rangle}\langle\gamma_j(s_1,a_1,h)\cdots\gamma_j(s_r,a_r,h)\rangle_h\Big]^{C_j} =\\
&\qquad\hat{Q}_0^{C_j}\left\langle\left\{\sum_{v=0}^{q-1}\prod_{l=1}^{C_j}[1+\omega(v,h_l)\hat{x}_l]\right\}^n\right\rangle_{\hat{\mathbf{x}},\mathbf{h}},
\end{aligned}\tag{B22}$$

with

$$\omega(v,h_l)\equiv\sum_{s=1}^{q-1}\exp\left[i\frac{2\pi s}{q}(h_l\otimes v)\right] = \begin{cases} q-1, & \text{if } h_l\otimes v=0,\\ -1, & \text{otherwise.}\end{cases}\tag{B23}$$

We can simplify the last equation by noting that

$$\sum_{v=0}^{q-1}\prod_{l=1}^{C_j}[1+\omega(v,h_l)\hat{x}_l]=\prod_{l=1}^{C_j}[1+(q-1)\hat{x}_l]+(q-1)\prod_{l=1}^{C_j}(1-\hat{x}_l). \tag{B24}$$

Let us write

$$\mathcal{Z}_n=\int DQD\hat{Q}\,e^{N\tilde{s}}, \tag{B25}$$

with

$$\tilde{s}=-\frac{1}{N}\ln N_A^{(2)}-n\lambda\ln q-Q_0\hat{Q}_0\langle[1+(q-1)x\hat{x}]^n\rangle_{x,\hat{x}}+\frac{1}{N}\ln\Phi, \tag{B26}$$

where

$$\begin{aligned}\Phi=\left\langle\frac{N^\Lambda}{\Lambda!}Q_0^\Lambda\hat{Q}_0^\Lambda\prod_i\left\langle\left[1+(q-1)\prod_{l=1}^{K_i}x_l\right]^n\right\rangle_{\mathbf{x}}\right.\\ \left.\times\prod_j\left\langle\left\{\prod_{l=1}^{C_j}[1+(q-1)\hat{x}_l]+(q-1)\prod_{l=1}^{C_j}(1-\hat{x}_l)\right\}^n\right\rangle_{\hat{\mathbf{x}}}\right\rangle_{\mathbf{K},\mathbf{C},\Lambda}\end{aligned} \tag{B27}$$

Let us define $\alpha\equiv NQ_0\hat{Q}_0$. For $n\ll 1$, we can consider only the leading contributions in the number of replicas, which gives

$$\begin{aligned}\ln\Phi=\ln\epsilon(\alpha)+\frac{n}{\epsilon(\alpha)}\sum_i\left\langle\frac{\alpha^\Lambda}{\Lambda!}\left\langle\ln\left[1+(q-1)\prod_{l=1}^{K_i}x_l\right]\right\rangle_{\mathbf{x}}\right\rangle_{\mathbf{K},\mathbf{C},\Lambda}\\ \frac{n}{\epsilon(\alpha)}\sum_j\left\langle\frac{\alpha^\Lambda}{\Lambda!}\left\langle\ln\left\{\prod_{l=1}^{C_j}[1+(q-1)\hat{x}_l]+(q-1)\prod_{l=1}^{C_j}(1-\hat{x}_l)\right\}\right\rangle_{\hat{\mathbf{x}}}\right\rangle_{\mathbf{K},\mathbf{C},\Lambda},\end{aligned} \tag{B28}$$

with

$$\epsilon(\alpha)=\left\langle\frac{\alpha^\Lambda}{\Lambda!}\right\rangle_{\mathbf{K},\mathbf{C},\Lambda}. \tag{B29}$$

Substituting the above formulas in $\tilde{s}$ for $n\to 0$, the extremization with respect to Q_0, $\hat{Q}_0$, $\pi(x)$ and $\hat{\pi}(\hat{x})$ leads to the saddle point equations (15), (16) and (17).

APPENDIX C: NUMBER OF MATRICES

Here we give the detailed calculation of the average number of $GF(q)$ $(M)\times N$ matrices for large N and N. Repeating the formula given in section V, we have

$$N_A=\sum_{\{A_{ij}\}}\left[\prod_{i=1}^{M}\delta\left(\sum_{j=1}^{N}\chi(A_{ij}),K_i\right)\right]\left[\prod_{j=1}^{N}\delta\left(\sum_{i=1}^{M}\chi(A_{ij}),C_j\right)\right]. \tag{C1}$$

with $\chi(A_{ij}) = 0$ if $A_{ij} = 0$ and 1 otherwise. Following a similar procedure as in B, we use the integral representations of the Kroenecker delta functions to write it as

$$\begin{aligned} N_A &= \oint DWDZ \prod_{i,j} \sum_{A_{ij}} (W_i Z_j)^{\chi(A_{ij})} \\ &= \oint DWDZ \prod_{i,j} [1 + (q-1) W_i Z_j] \\ &= \oint DWDZ \prod_{i} \left[1 + \sum_{r=1}^{N} (q-1)^r W_i^r \sum_{\langle j_1 \cdots j_r \rangle} Z_{j_1} \cdots Z_{j_r} \right] \\ &= \oint DWDZ \left[1 + \sum_{s=1}^{M} \sum_{\langle i_1 \cdots i_s \rangle} \sum_{r_1,\ldots,r_s} (q-1)^{r_1+\cdots+r_s} W_{i_1}^{r_1} \cdots W_{i_s}^{r_s} F(r_1, Z) \cdots F(r_s, Z) \right], \end{aligned} \tag{C2}$$

where

$$F(r, Z) \equiv \sum_{\langle j_1 \cdots j_r \rangle} Z_{j_1} \cdots Z_{j_r}. \tag{C3}$$

The integrals over the W's can pass through the summations and will factorize to give the corresponding Kroenecker delta functions resulting in

$$\begin{aligned} N_A &= (q-1)^{\sum_i K_i} \oint DZ F(K_1, Z) \cdots F(K_M, Z) \\ &= (q-1)^{\Lambda} \oint DZ F(K_1, Z) \cdots F(K_M, Z) \\ &= (q-1)^{\Lambda} \oint DZ \prod_i \sum_{\langle j_1 \cdots j_{K_i} \rangle} Z_{j_1} \cdots Z_{j_{K_i}} \\ &\approx (q-1)^{\Lambda} \oint DZ \prod_i \frac{1}{K_i!} \left(\sum_{j=1}^{N} Z_j \right)^{K_i} \\ &= (q-1)^{\Lambda} \oint DZ \frac{1}{\prod_i K_i!} \left(\sum_{j=1}^{N} Z_j \right)^{\sum_i K_i} \\ &= \frac{(q-1)^{\Lambda}}{\prod_i K_i!} \oint DZ \sum_{j_1,\ldots,j_\Lambda} Z_{j_1} \cdots Z_{j_\Lambda} \\ &= \frac{(q-1)^{\Lambda}}{\prod_i K_i!} \binom{\Lambda}{C_1} \binom{\Lambda - C_1}{C_2} \cdots \binom{\Lambda - C_1 - \cdots - C_{N-1}}{C_N}, \end{aligned} \tag{C4}$$

which gives the final result

$$N_A = \frac{(q-1)^{\Lambda} \Lambda!}{\prod_i K_i! \prod_j C_j!}. \tag{C5}$$

APPENDIX D: PROOF OF REPLICA SYMMETRY

Using the fact that the random matrices can be seen as statistical physics systems with Hamiltonian $\mathcal{H}(\mathbf{v}) \equiv N - \ln \delta(A\mathbf{v}, 0)$ we now prove that this implies that the replica symmetric solution is the exact one. In fact, the form of the Hamiltonian implies that

$$\mathcal{P}(\mathbf{v}) = \left[\sum_{\mathbf{v}} \delta(A\mathbf{v}, 0) \right]^{-1} = q^{-d(A)}. \tag{D1}$$

The distribution of the overlaps of the spins is given by

$$\begin{aligned}\mathcal{P}(\rho) &= \left\langle \delta\left(\rho - \frac{1}{N}\sum_{j=1}^{N}\sigma^j\sigma'^j\right)\right\rangle_{\boldsymbol{\sigma},\boldsymbol{\sigma}'} \\ &= q^{-2d(A)}\sum_{\mathbf{v},\mathbf{v}'}\delta(A\mathbf{v},0)\delta(A\mathbf{v}',0)\delta\left[\rho - \frac{1}{N}\sum_{j=1}^{N}\exp\left(\frac{2\pi i}{q}\left(v^j + v'^j\right)\right)\right].\end{aligned} \tag{D2}$$

Let us call

$$g(\mathbf{v},\mathbf{v}') \equiv \delta\left[\rho - \frac{1}{N}\sum_{j=1}^{N}\exp\left(\frac{2\pi i}{q}\left(v^j + v'^j\right)\right)\right], \tag{D3}$$

and note that $g(\mathbf{v},\mathbf{v}') = g(0,\mathbf{v}\oplus\mathbf{v}')$. Therefore we can write

$$\begin{aligned}\mathcal{P}(\rho) &= q^{-2d(A)}\sum_{\mathbf{v},\mathbf{v}'}\delta(A\mathbf{v},0)\delta(A\mathbf{v}',0)g(0,\mathbf{v}\oplus\mathbf{v}') \\ &= q^{-2d(A)}\sum_{\mathbf{v},\mathbf{v}'}\delta(A\mathbf{v},0)\delta(A\mathbf{v}',0)\sum_{\mathbf{u}}\delta(\mathbf{u},\mathbf{v}\oplus\mathbf{v}')g(0,\mathbf{u}) \\ &= q^{-2d(A)}\sum_{\mathbf{u}}g(0,\mathbf{u})\left[\sum_{\mathbf{v}}\delta(A\mathbf{v},0)\sum_{\mathbf{v}'}\delta(A\mathbf{v}',0)\delta(\mathbf{u},\mathbf{v}\oplus\mathbf{v}')\right] \\ &= q^{-2d(A)}\sum_{\mathbf{u}}g(0,\mathbf{u})\left[\sum_{\mathbf{v}}\delta(A\mathbf{v},0)\delta(A(\mathbf{u}\oplus(-\mathbf{v})),0)\right] \\ &= q^{-d(A)}\sum_{\mathbf{u}}\delta(A\mathbf{u},0)g(0,\mathbf{u}) \\ &= \left\langle\delta\left(\rho - \frac{1}{N}\sum_{j=1}^{N}\sigma^j\right)\right\rangle_{\boldsymbol{\sigma}}.\end{aligned} \tag{D4}$$

Therefore, the distribution of the overlaps is the same as the distribution of the magnetization in the spin systems. This implies that there is no spin glass phase in the system and, therefore, no replica symmetry breaking [15]. The above calculation can also be viewed as a consequence of the *gauge invariance* of the Hamiltonian with respect to the transformation $\mathbf{v} \to \mathbf{v}\oplus\mathbf{v}'$, where $A\mathbf{v}' = 0$, which leads basically to the same calculation above.

[1] R. McEliece, *Theory of Information & Coding* (Cambridge University Press, Cambridge, MA, 2002 2nd edition).
[2] K. T. Phelps, J. Rifà, and M. Villanueva, IEEE Trans. Inf. Theory **51**, 3931 (2005).
[3] M. Davey and D. MacKay, IEEE Communications Letters **2**, 165 (1998).
[4] K. Nakamura, Y. Kabashima, and D. Saad, Eurphys. Lett. **56**, 610 (2001).
[5] C. Cooper, Random Structures and Algorithms **16**, 209 (2000).
[6] J. Blömer, R. Karp, and E. Weiz, Random Structures and Algorithms **10**, 407 (1998).
[7] X. Feng and Z. Zhang, Applied Mathematics and Computation **185**, 689 (2007).
[8] M. Mezard, G. Parisi, and A. Zee, Nuclear Physics B **559**, 689 (1999).
[9] E. Kanzieper, Nuclear Physics B **596**, 548 (2001).
[10] T. Nagao and T. Tanaka, Journal of Physics A: Mathematical and Theoretical **40**, 4973 (2007).
[11] G. Biroli, J.-P. Bouchaud, and M. Potters, Europhysics Letters **78**, 10001 (2007).
[12] G. Biroli, J.-P. Bouchaud, and M. Potters, Journal of Statistical Mechanics: Theory and Experiment P07019 (2007).
[13] O. Bohigas, J. X. de Carvalho, and M. P. Pato, Physical Review E **77**, 011122 (2008).
[14] M. Mézard, G. Parisi, and M. Virasoro, *Spin Glass Theory and Beyond* (World Scientific Publishing Co., Singapore, 1987).
[15] H. Nishimori, *Statistical Physics of Spin Glasses and Information Processing* (Oxford University Press, Oxford, UK, 2001).
[16] Y. Kabashima and D. Saad, J. Phys. A. **37**, R1 (2004).
[17] R. Monasson and R. Zecchina, Phys. Rev. Lett. **76**, 3881 (1996).
[18] J. van Mourik and D. Saad, Phys. Rev. E **66**, 056120 (2002).

[19] R. Mulet, A. Pagnani, M. Weigt, and R. Zecchina, Phys. Rev. Lett. **89**, 268701 (2002).
[20] R. C. Alamino and D. Saad, J. Phys A: Math. Theor. **40**, 12259 (2007).
[21] T. Tanaka and D. Saad, Technical report (unpublished).
[22] J. Pearl, *Probabilistic Reasoning in Intelligent Systems* (Morgan Kaufmann Publishers, Inc., San Francisco, CA, 1988).
[23] R. Vicente, D. Saad, and Y. Kabashima, in *Advances in Imaging and Electron Physics*, edited by P. Hawkes (Academic Press, USA, 2002), Vol. 125, pp. 232–353.
[24] R. Vicente, D. Saad, and Y. Kabashima, J. Phys. A **33**, 6527 (2000).
[25] M. Mézard and G. Parisi, Eur. Phys. J. B **20**, 217 (2001).
[26] B.-Y. Wang and F. Zhang, Discrete Mathematics **187**, 211 (1998).
[27] R. Vicente, D. Saad, and Y. Kabashima, Europhys. Lett. **51**, 698 (2000).
[28] I. S. Gradshteyn and I. M. Ryzhik, in *Table of Integrals, Series, and Products*, edited by A. Jeffrey and D. Zwillinger (Academic Press, USA, 1993).
[29] We are indebted to one of the referees who indicated a simplification in the original proof.