Von Neumann's expanding model on random graphs

A De Martino‡, C Martelli¶, R Monasson§, and I Pérez Castillo†‡

‡CNR-INFM (ISC) and Dipartimento di Fisica, Università di Roma "La Sapienza", p.le Aldo Moro 2, 00185 Roma, Italy

¶Dipartimento di Scienze Biochimiche, Università di Roma "La Sapienza", p.le Aldo Moro 2, 00185 Roma, Italy

§Laboratoire de Physique Théorique de l'ENS, 24 rue Lhomond - 75231 Paris Cedex 05, France

†Department of Mathematics, King's College London, Strand, London WC2R 2LS, United Kingdom

Abstract. Within the framework of Von Neumann's expanding model, we study the maximum growth rate $\rho^\star$ achievable by an autocatalytic reaction network in which reactions involve a finite (fixed or fluctuating) number D of reagents. $\rho^\star$ is calculated numerically using a variant of the Minover algorithm, and analytically via the cavity method for disordered systems. As the ratio between the number of reactions and that of reagents increases the system passes from a contracting ($\rho^\star < 1$) to an expanding regime ($\rho^\star > 1$). These results extend the scenario derived in the fully connected model ($D \to \infty$), with the important difference that, generically, larger growth rates are achievable in the expanding phase for finite D and in more diluted networks. Moreover, the range of attainable values of $\rho^\star$ shrinks as the connectivity increases.

1. Introduction

Von Neumann's expansion problem was initially formulated to describe growth in production economies as an autocatalytic process and has played a key role in the development of the theory of economic growth [1–3]. In a nutshell, it concerns the calculation of the maximum growth rate achievable by an autocatalytic system of M reactants (labeled by Greek indices like $\mu, \nu, \ldots$) interconnected by N chemical reactions (labeled by Roman indices like $i, j, \ldots$) specified by a given stoichiometric matrix. The basic ingredients are however sufficiently simple to be applicable in different contexts ranging from economics to systems biology.

In order to state the optimization problem in mathematical terms (see however [4] for a more detailed description), it is convenient to separate the matrix of input stoichiometric coefficients $\boldsymbol{A} = \{a_i^\mu \geq 0\}$ (input matrix for brevity) from that of outputs $\boldsymbol{B} = \{b_i^\mu \geq 0\}$. The relevant microscopic variables are the reaction fluxes, denoted by

s_i, which are assumed to be non-negative. The problem amounts to finding a vector $\boldsymbol{s} = \{s_i \geq 0\}$ of positive fluxes and a number $\rho > 0$ such that ρ is maximum subject to

$$\boldsymbol{B}\boldsymbol{s} \geq \rho \boldsymbol{A}\boldsymbol{s} \tag{1}$$

where the inequality is to be understood component-wise *i.e.* valid for each reactant μ. The trivial null solution $\boldsymbol{s} = \boldsymbol{0}$, *viz.* $\{s_i = 0\ \forall i\}$, is not accepted: at least one component of $\boldsymbol{s}$ must be strictly positive.

Condition (1) requires that the total output of every reactant is at least ρ times the total input. So the maximum feasible ρ, $\rho^\star$, measures the largest uniform growth rate achievable by the given set of reactions. If $\rho^\star < 1$, the optimal state of the system is a contracting one. Otherwise for $\rho^\star > 1$ it is expanding. The case $\rho^\star = 1$ describes instead a system in which at optimality reaction fluxes are arranged so as to guarantee mass balance.

Von Neumann's problem has been studied recently in a statistical mechanics perspective under the assumptions that M scales linearly with N (with $N \to \infty$) and that the stoichiometric matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ have quenched random independent and identically distributed elements [4]. With a fully connected network of reactions (that is, one in which every reaction uses each reactant both as an input and as an output), one finds a transition from a contracting to an expanding regime when the parameter $n = N/M$ exceeds the critical value $n_c = 1$.

In this work we extend the analysis of [4] to the finitely connected case where reactions use a finite number of inputs to produce a finite number of outputs. By analogy with integer programming problems (see e.g. [5]), we represent our autocatalytic network as a bipartite (factor) graph (see Fig. 1). We denote reagents as squares and reactions as circles. Each circle has a number of incoming (resp. outgoing) connections to squares, representing the inputs (resp. outputs) of the reaction, and each of these links carries the input (resp. output) stoichiometric coefficient. Similarly, each square has a number of incoming (resp. outgoing) connections to circles, denoting the reactions where that reagent enters as an output (resp. input). The notation $i \in \mu$ identifies a reaction i in which a reactant μ is involved either as an input or as an output (and vice-versa for $\mu \in i$). To each reaction node a variable $s_i \geq 0$ is attached, representing the flux of reaction i. Each reactant node μ carries instead a function c^μ given by

$$c^\mu = \sum_{i \in \mu} s_i (b_i^\mu - \rho a_i^\mu) \tag{2}$$

We are interested in finding non-null flux configurations satisfying (1), that is such that $c^\mu \geq 0$ for each μ and specifically in finding the largest value of ρ for which a configuration of this type exists. Our approach is developed along two main lines. On the one hand (Section 2), we employ a suitably modified Minover algorithm to compute optimal growth rates numerically on any graph. Results thus derived will be theoretically validated in Section 3, where we use the cavity method to describe solutions analytically for locally tree-like instances. This approximation turns out to be in good

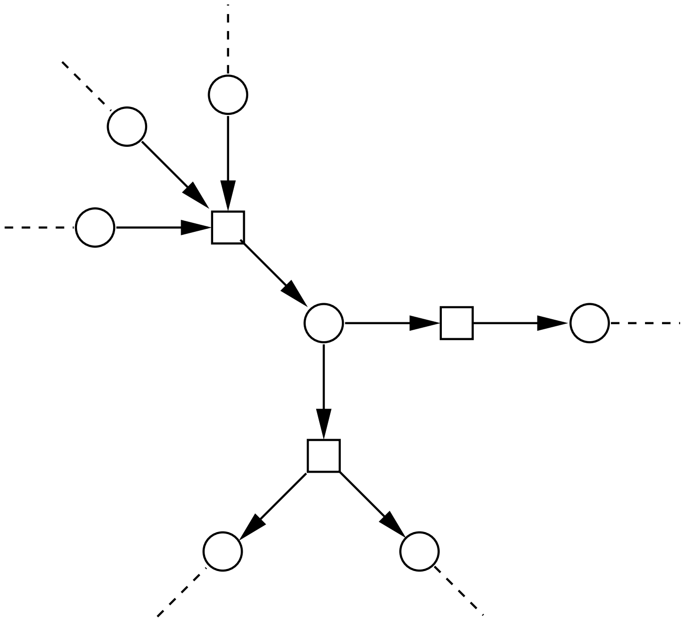

Figure 1. Factor graph representation of a network of reactions and reagents. Reactions are represented by circles, reagents by squares. Stoichiometric coefficients label the directed links.

qualitative agreement with numerical data. The connection of the latter approach with the fully connected theory developed in [4] is presented in two detailed Appendices.

2. Numerical analysis

2.1. Algorithm

Formula (1) is reminiscent of pattern storage conditions in the perceptron, a formal neural network model [6]. This analogy allowed us to devise an exact algorithm to solve Von Neumann's problem on any graph in a time growing only polynomially with the number N of nodes. The algorithm is an extension of the so-called Minover algorithm [7] enforcing the constraint of positivity for the fluxes, and hereafter referred to as Minover$^+$.

The core part of the algorithm is a subroutine, Minover$^+(\rho)$, we now describe. We assume that the matrices $\boldsymbol{A}, \boldsymbol{B}$ are given, and ρ denotes a real, non-negative number.

(i) Step $\ell = 0$: all fluxes are null, $s_i = 0$.

(ii) Step $\ell + 1$: Calculate the M functions $c^\mu(\ell)$ defined in (2) from the values $s_i(\ell)$ of the fluxes at step ℓ, and look for the reactant with least value,

$$\mu_0(\ell) = \arg\ \min_\mu c^\mu(\ell) \ ; \tag{3}$$

* if $c^{\mu_0(\ell)} \geq 0$ then OUTPUT $\boldsymbol{s}(\ell)$ and HALT (the null solution is not accepted);
* otherwise, if $c^{\mu_0(\ell)} < 0$, update the fluxes according to the rule

$$s_i(\ell + 1) = \max\{0, s_i(\ell) + b_i^{\mu_0(\ell)} - \rho\, a_i^{\mu_0(\ell)}\} \ , \tag{4}$$

increase ℓ by one, and GO TO (ii).

In case of ties, one may select μ_0 at random uniformly among those with the same (and lowest) values of c^μ.

We claim that, if $\rho < \rho^\star$, the largest growth rate associated to matrices $\boldsymbol{A}$ and $\boldsymbol{B}$, Minover$^+(\rho)$ will halt after a finite number of steps, and output a set of non-negative fluxes guaranteeing a growth rate larger than ρ. The proof is simple and goes as follows. Assume $\rho < \rho^\star$. Then there exists a strictly positive number Δ, hereafter called stability, and a vector $\boldsymbol{s}^\star \neq \boldsymbol{0}$ of fluxes such that‡

$$\boldsymbol{s}^\star \cdot (\boldsymbol{b}^\mu - \rho \boldsymbol{a}^\mu) \geq \Delta \, |\boldsymbol{s}^\star| \quad \text{for each } \mu \; . \tag{5}$$

Let

$$W = \max_\mu \sum_i (b_i^\mu - \rho \, a_i^\mu)^2 > 0 \; . \tag{6}$$

and consider the functions $\mathcal{N}(\ell) = \boldsymbol{s}^\star \cdot \boldsymbol{s}(\ell)$ and $\mathcal{D}(\ell) = |\boldsymbol{s}(\ell)|^2$. We have, from (4),

$$\begin{aligned} \mathcal{N}(\ell+1) &= \sum_i s_i^\star \; \max\{0, s_i(\ell) + b_i^{\mu_0(\ell)} - \rho \, a_i^{\mu_0(\ell)}\} \\ &\geq \sum_i s_i^\star \; (s_i(\ell) + b_i^{\mu_0(\ell)} - \rho \, a_i^{\mu_0(\ell)}) \\ &\geq \mathcal{N}(\ell) + \Delta \, |\boldsymbol{s}^\star| \end{aligned} \tag{7}$$

from the non-negativity of $s_i^\star$ and (5). We thus obtain $\mathcal{N}(\ell) \geq \ell \, \Delta \, |\boldsymbol{s}^\star|$ for any step ℓ. Turning to function $\mathcal{D}(\ell)$ we get

$$\begin{aligned} \mathcal{D}(\ell+1) &= \sum_i \max\{0, s_i(\ell) + b_i^{\mu_0(\ell)} - \rho \, a_i^{\mu_0(\ell)}\}^2 \\ &\leq \sum_i (s_i(\ell) + b_i^{\mu_0(\ell)} - \rho \, a_i^{\mu_0(\ell)})^2 \\ &= \mathcal{D}(\ell) + 2 \, c^{\mu_0(\ell)} + \sum_i (b_i^{\mu_0(\ell)} - \rho \, a_i^{\mu_0(\ell)})^2 \\ &\leq \mathcal{D}(\ell) + W \end{aligned} \tag{8}$$

from definition (6) and the assumption that the algorithm has not halted at step ℓ *i.e.* $c^{\mu_0(\ell)} < 0$. Hence $\mathcal{D}(\ell) \leq \ell \, W$. Now let

$$f(\ell) = \frac{\boldsymbol{s}(\ell) \cdot \boldsymbol{s}^\star}{|\boldsymbol{s}(\ell)| \, |\boldsymbol{s}^\star|} \tag{9}$$

By Cauchy-Schwarz inequality $f(\ell) \leq 1$. But according to the above calculations

$$f(\ell) = \frac{\mathcal{N}(\ell)}{|\boldsymbol{s}^\star| \sqrt{\mathcal{D}(\ell)}} \geq \frac{\Delta}{\sqrt{W}} \, \sqrt{\ell} \; . \tag{10}$$

Therefore the algorithm halts after

$$\ell_0 = \frac{W}{\Delta^2} \tag{11}$$

‡ The set $\mathcal{S}(\rho)$ of fluxes fulfilling constraints (1) is, by definition of $\rho^\star$, non empty. Any vector $\boldsymbol{s}^\star$ in the interior of $\mathcal{S}(\rho)$ satisfies (5) for some positive Δ. A natural choice is the optimal vector of fluxes, *i.e.* that associated to $\rho^\star$.

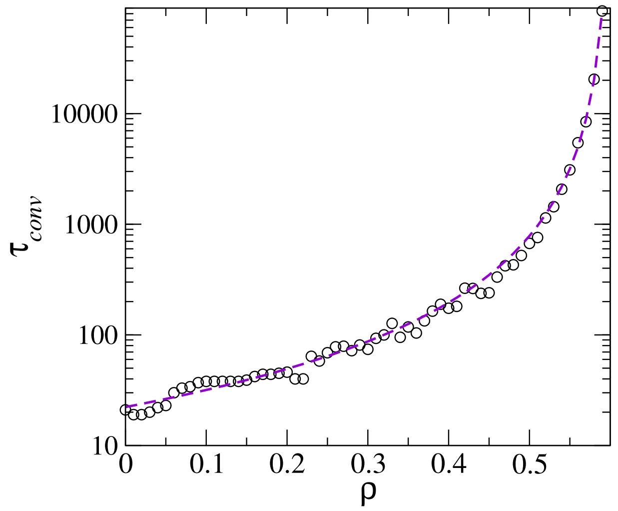

Figure 2. Convergence time (in individual steps) versus ρ for a *single* network with $M = 100$ reagents and $n = 0.66$. The number of reagents per reaction is fixed at $D = 5$, the number of reactions per reagent is a Poisson random variable with mean $Dn = 3.3$. The dashed line represents the best fit with $\tau_{conv} \sim (\rho^\star - \rho)^{-2}$. For this network, $\rho^\star = 0.5994$.

steps at most. Now call ℓ_s the step at which the algorithm stops. The halting condition ensures that $c^{\mu_0(\ell_s)} \geq 0$, *i.e.* condition (1) is satisfied, with a non-null vector of fluxes (by construction the fluxes are non-negative at any step). Hence $\boldsymbol{s}(\ell_s)$ is a solution to Von Neumann's problem with growth rate ρ.

Our complete algorithm, Minover$^+$, is an iteration of Minover$^+(\rho)$ for increasing values of ρ. Starting from $\rho = \epsilon$, a small positive value for which (1) surely admits a solution §, one can measure the convergence time (number of steps prior to halting) ℓ_s for increasing values of ρ. The convergence time increases too. Now we expect Δ to vanish as $\rho^\star - \rho$ when ρ gets closer and closer to its optimal value, thus $\ell_s \sim (\rho^\star - \rho)^{-2}$ from (11) (see Fig. 2). One may then estimate the maximal growth rate by extrapolating from the log-log plot of the convergence time versus ρ.

2.2. Survey of results

In what follows we restrict ourselves to purely autocatalytic systems, that is we assume that every reagent is produced and consumed by at least one reaction. Indeed a reagent that only serves as input gives rise to a constraint of the form

$$c^\mu = -\rho \sum_{i\in\mu} s_i a_i^\mu \geq 0 \tag{12}$$

which is satisfied by taking $\rho = 0$. Such a situation would immediately force the result $\rho^\star = 0$‖. On the other hand, "sink" reagents (namely those that are not inputs of any

§ We can always choose $\epsilon = \min_{i,\mu} b_i^\mu / a_i^\mu > 0$.

‖ The a priori probability to generate a factor graph without such pathologic function nodes in the Poissonian model discussed later is given by $(1-e^{-Dn}(1-e^{-Dn}))^M$, where M is the number of reactants.

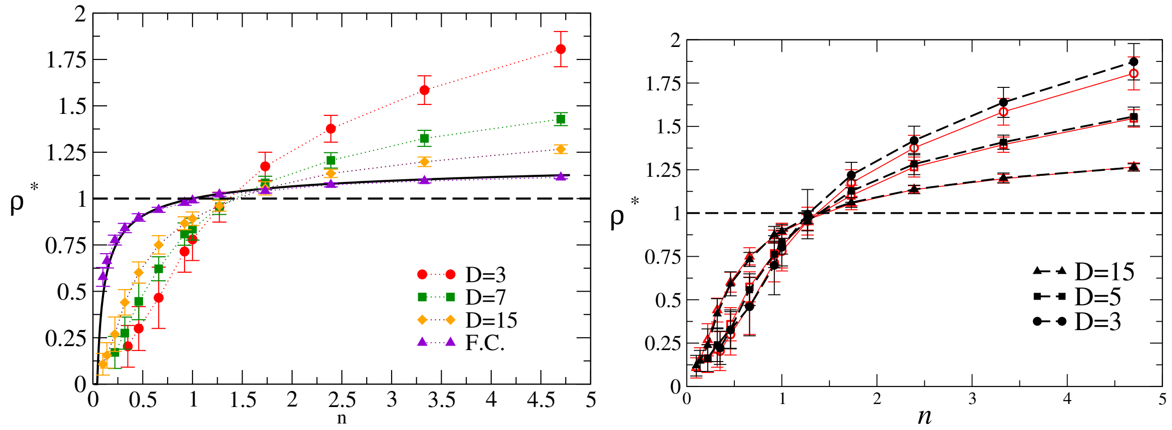

Figure 3. Left panel: $\rho^\star$ vs n for Regular(N)/Poisson(M) networks of 100 reagents for various D. The continuous line is the fully-connected limit for a system of the same size [4] while F.C. labels numerical results for a fully connected system. Right panel: $\rho^\star$ vs n for Regular(N)/Poisson(M) (open markers) and Regular(N)/SF(M) (closed markers) networks for different values of D (see text for details). Averages over 50 disorder samples.

reaction) provide constraints that are ρ-independent and trivially satisfied. Furthermore, for the sake of simplicity, we fix the number M of reagents and assume that the quenched random stoichiometric matrix has Gaussian elements with mean 1 and variance 1/2.

For a start, we consider Von Neumann's problem on factor graphs in which every reaction has D inputs and D outputs, while for reagents we assume that the in(out)-degree distribution is Poissonian, $P(k) = \lambda^k e^{-\lambda}/k!$ with mean degree $\lambda = Dn$. For short, we write Regular(N)/Poisson(M) to denote this type of situation. Our analysis focuses on values of $n = N/M > 1/D$, which ensure that the resulting factor graph is connected. Results for $\rho^\star$ in this case are displayed in the left panel of Fig. 3. One sees that the overall qualitative behaviour reproduces the results obtained for the fully connected model. The system passes from an expanding ($\rho^\star > 1$) to a contracting ($\rho^\star < 1$) phase when n is lowered below a critical value. Interestingly, the maximum growth rate achievable in the diluted system is larger than in the fully connected model in the expanding phase and higher growth rates can be achieved in more diluted systems. This conclusion turns out to be valid generically for all types of networks we studied.

In the right panel of Fig. 3 we compare these results with those obtained for the case in which the number of reactions per reagent is power-law distributed (SF or scale-free for short) rather than Poisson, which enables the coexistence of widely used and rarely used reagents in the reaction network (reactions still have a δ-distributed connectivity). More precisely, the number of reactions per reagent is distributed as $P(k) \sim k^{-\gamma}$ with $2 < \gamma < 3$ (as in e.g. metabolic networks [9]). The particular value of the exponent does

This implies that such nodes are always present in the thermodynamic limit. In this work we discard them completely. However in practical applications (e.g. metabolic networks [8]) it is important to take these nodes into account as prescribed "sources" (uptakes).

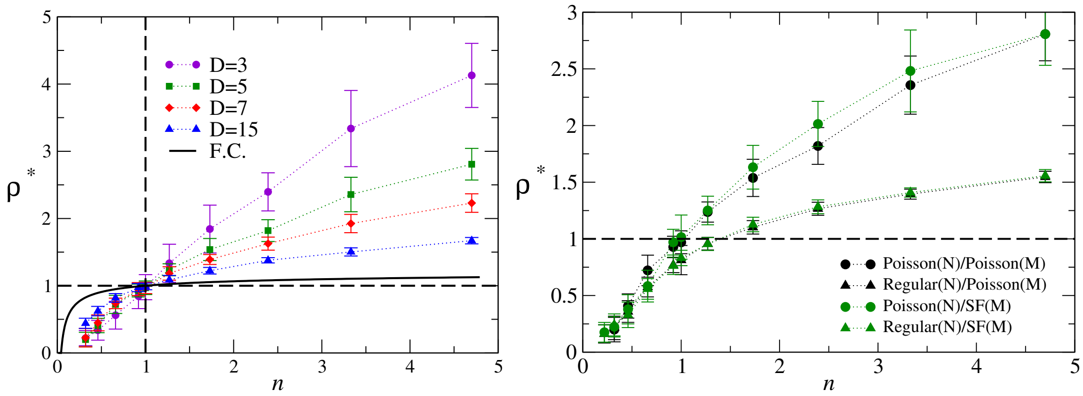

Figure 4. Left panel: $\rho^\star$ vs n for Poisson(N)/Poisson(M) networks. Right panel: $\rho^\star$ vs n for different types of factor graphs with $D = 5$. Averages over 50 disorder samples; system of 100 reagents.

not affect results in a marked way. Note that a significant difference (which however lies inside the error bar) is observed only for the smallest values of D.

Next we compare networks with degree regular reactions, where, as said above, the in/out-degree of reactions is fixed equal to D, with Poisson(N)/Poisson(M) networks where also reactions have fluctuating degree with mean D. From Fig. 4 one sees that optimal growth rates for all values of n are always larger in Poissonian networks than in regular ones, independently of the degree distribution of reagents (right panel). Moreover a fluctuating connectivity for reactions makes the expanding regime achievable with fewer reactions (left panel), and in particular the critical point where $\rho^\star$ becomes larger than 1 is $n_c \simeq 1$, similar to what is found in the fully connected case. In few words, we can say that topological regularity has a strong influence on the maximum growth rate achievable and that an asymmetric choice of degree distributions for reactions and reagents moves n_c from its fully connected value. In the case above, it takes a larger repertoire of reactions to sustain an expanding regime.

It is possible to find a simple expression that allows to re-scale the data obtained for different D's onto a single curve (which appears to be topology-dependent), at least for large n. It turns out that

$$\rho^\star(n, D) \simeq \rho^\star_{\mathrm{F.C.}}(n) + \frac{1}{D} f(n) \tag{13}$$

where $\rho^\star_{\mathrm{F.C.}}(n)$ is the optimal growth rate of the fully connected system (see Fig. 5). This scaling form, which gives $\rho^\star$'s for diluted systems as $1/D$ corrections to the fully connected limit, implies that the range of achievable growth rates shrinks as D increases.

In Fig. 6, we show the values of $\rho^\star$ for systems of different sizes and topologies (similar results hold for other types of networks). One sees that already for moderate system sizes finite size effect are negligible.

To conclude, we present (see Figures 7 and 8) the flux distribution at optimality for Regular(N)/Poisson(M) and Poisson(N)/Poisson(M) networks for values of n below and above the critical point. In the former case, distributions are well fitted by

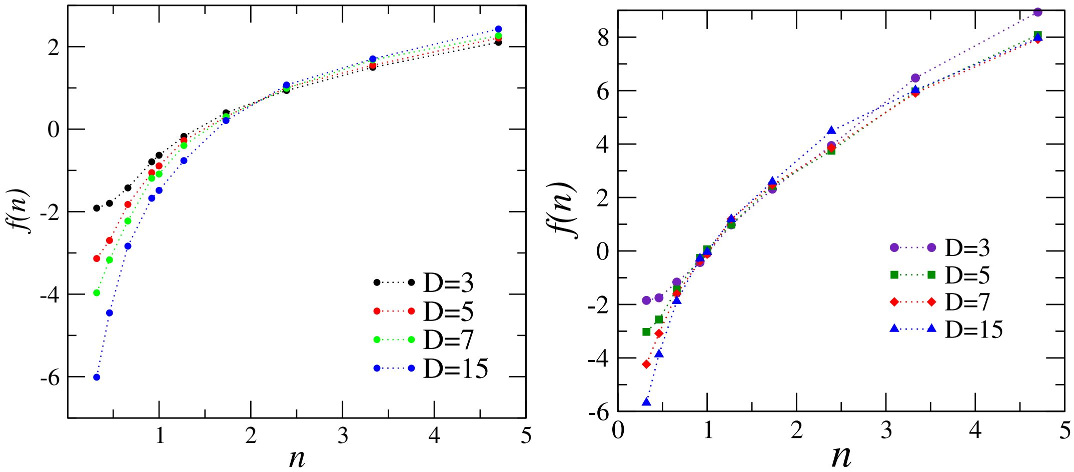

Figure 5. Scaling functions $f(n)$ for Regular(N)/Poisson(M) networks (left) and Poisson(N)/Poisson(M) networks (right). Systems of 100 reagents; error bars not reported.

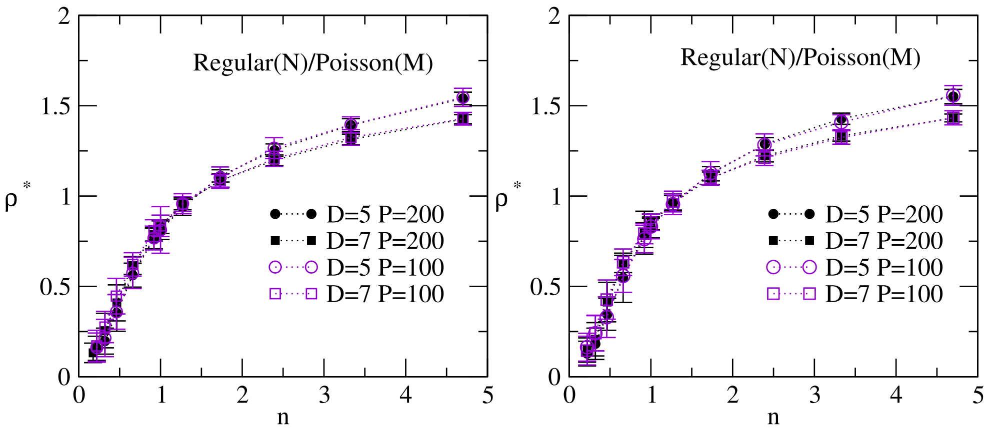

Figure 6. $\rho^{\star}$ vs n for Regular(N)/Poisson(M) and Regular(N)/SF(M) networks of 100 and 200 reagents for different D. Averages over 50 disorder samples.

an exponential form both for small and large n. In the latter case, an algebraic law provides the best fit in the expanding regime. Similar results are obtained for different connectivity distributions for reaction nodes. The particular value of the exponent appears to be topology dependent. However, it is interesting that such distributions arise in this context, specifically for their relevance in connection to metabolic networks [8,9]. Note finally that as in the fully-connected model the weight of the mass at $s = 0$ increases with n, though less drastically.

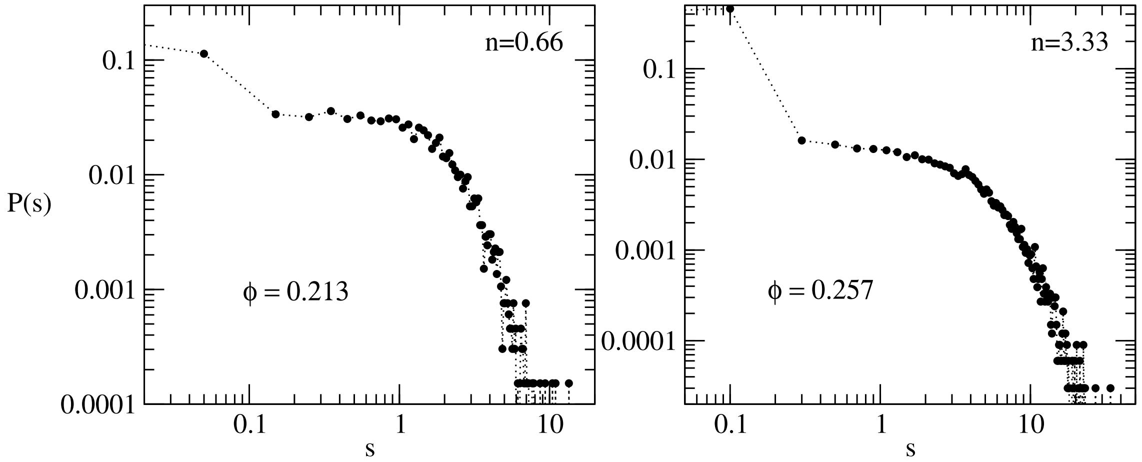

Figure 7. Flux distribution for a Regular(N)/Poisson(M) network of reagents with $D = 5$ inputs and outputs per reaction and different values of n (average over 50 disorder samples). System of 200 reagents. The δ-peak at $s = 0$ (carrying an intensive weight ϕ reported inside the panels) is not shown.

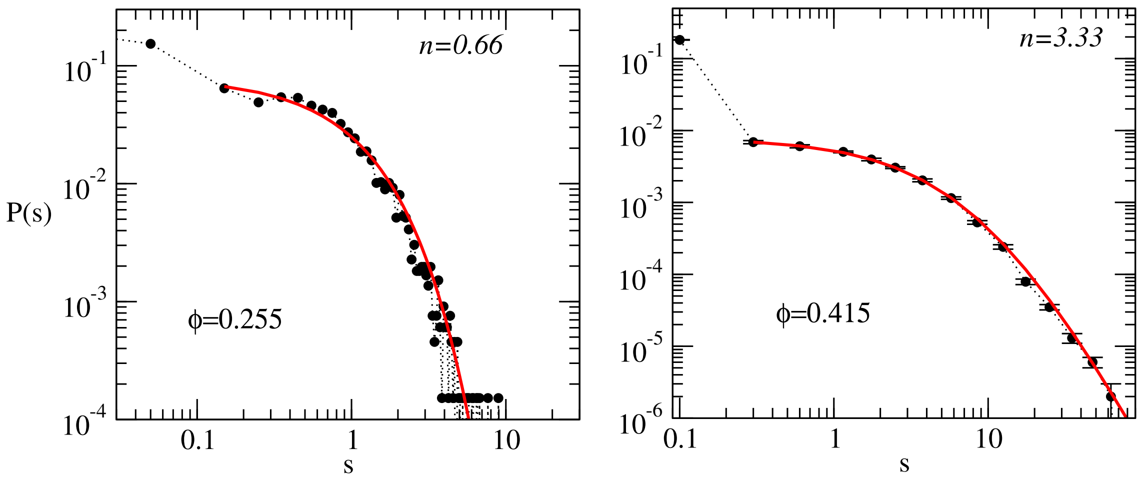

Figure 8. Flux distribution for a Poisson(N)/Poisson(M) network of reagents with $D = 5$ inputs and outputs per reaction and different values of n (average over 50 disorder samples). System of 200 reagents. The δ-peak at $s = 0$ (carrying an intensive weight ϕ reported inside the panels) is not shown. The red lines represent best fits with an exponential distribution of the form $p(x) = Ae^{-Bx}$ (left panel) and with an algebraic distribution of the form $p(x) = A(B + x)^{-C}$ (right panel).

3. Cavity theory

Von Neumann's problem possesses the natural cost function

$$E = \sum_{\mu=1}^{M} \Theta(-c^{\mu}), \tag{14}$$

that, given ρ, counts the number of reagents for which the condition (1) is violated. It is clear that $\rho^{\star}$ corresponds to the largest ρ for which one can arrange fluxes in such a

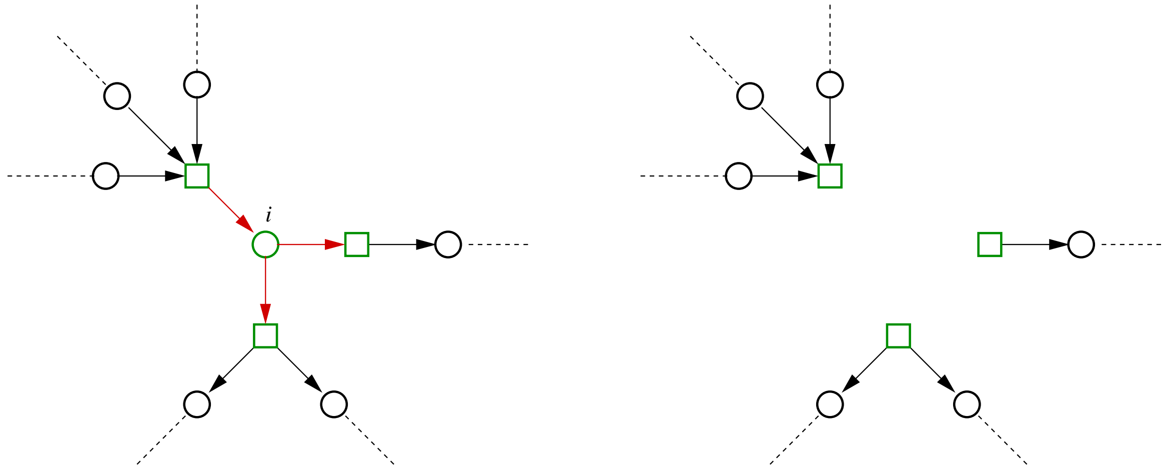

Figure 9. Left: reagents participating in reaction i are correlated. Right: If reaction i is removed those reagents become uncorrelated on a tree-like graph.

way that $E = 0$. More generally, we can look for configurations that minimize E, so it is convenient to introduce the "partition function"

$$Z = \int \gamma(\boldsymbol{s}) e^{-\beta E} d\boldsymbol{s} \equiv \mathrm{Tr}_{\boldsymbol{s}}\, e^{-\beta E} \tag{15}$$

with $\gamma(\boldsymbol{s})$ some constraint over the configuration space (e.g. a linear constraint). Minima of the cost function are selected in the limit $\beta \to \infty$.

Let us assume that the factor graph has a locally tree-like structure and consider, as in Figure 9, a reaction i and the reagents involved in it. It is clear that the joint distribution of consumptions $\boldsymbol{c}^{\mu\in i} = \{c^\mu\}_{\mu\in i}$ does not factorize into single-reagent terms, *i.e.*

$$q(\boldsymbol{c}^{\mu\in i}) \neq \prod_{\mu\in i} q^\mu(c^\mu) \tag{16}$$

because the μ's share the common vertex i. However, if we remove the reaction i (here and in what follows indexes enclosed in parentheses like (x) denote quantities calculated removing the node x, while sums and products where x is not considered carry the index $\backslash x$) the resulting joint distribution of consumptions $q_{(i)}(\boldsymbol{c}^{\mu\in i})$ does factorize:

$$q_{(i)}(\boldsymbol{c}^{\mu\in i}) = \prod_{\mu\in i} q^\mu_{(i)}(c^\mu) \tag{17}$$

Similarly, with $\boldsymbol{s}_{i\in\mu} = \{s_i\}_{i\in\mu}$, we have

$$p(\boldsymbol{s}_{i\in\mu}) \neq \prod_{i\in\mu} p_i(s_i) \tag{18}$$

but, removing reagent μ,

$$p^{(\mu)}(\boldsymbol{s}_{i\in\mu}) = \prod_{i\in\mu} p_i^{(\mu)}(s_i) \tag{19}$$

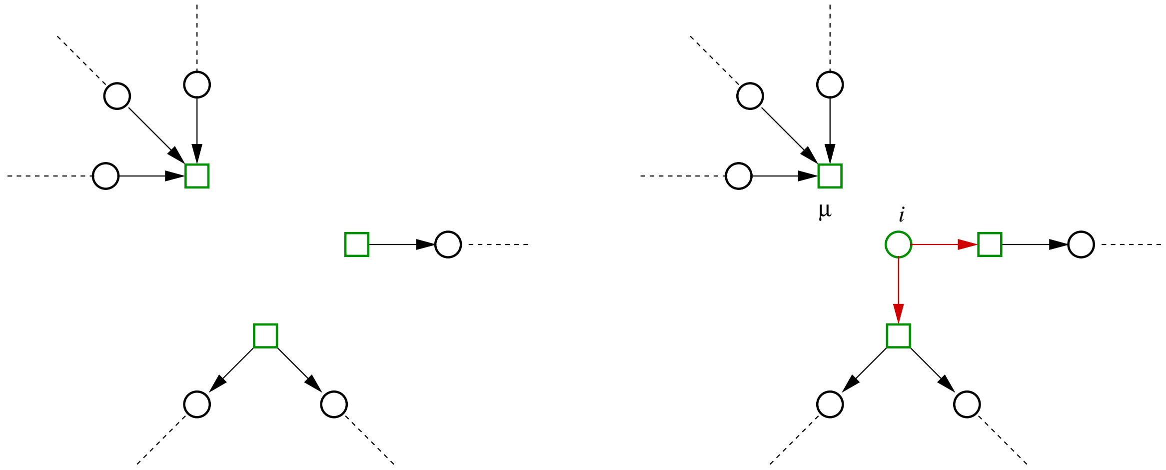

Figure 10. From the graph where reaction i has been removed (left), one calculates easily the distribution $p_i^{(\mu)}(s_i)$ in the absence of the reagent μ by merging back all reagents participating in reaction i but reagent μ.

The quantities $q^\mu_{(i)}(c^\mu)$ and $p_i^{(\mu)}(s_i)$ are known as *cavity distributions*. Now for all $i = 1, \ldots, N$ and $\mu \in i$ it is possible to find equations for the cavity distributions by simply merging back all disconnected nodes but one (see Fig. 10). It is easily seen in particular that

$$p_i^{(\mu)}(s_i) = \frac{1}{Z_i^{(\mu)}} \int e^{-\beta \sum_{\nu\in i\setminus\mu} \Theta\left[-c^\nu_{(i)} - s_i(b_i^\nu - \rho a_i^\nu)\right]} \prod_{\nu\in i\setminus\mu} q^\nu_{(i)}(c^\nu) dc^\nu$$

$$\equiv \frac{1}{Z_i^{(\mu)}} \mathrm{Tr}_{\boldsymbol{c}^{\nu\in i\setminus\mu}} \left[\prod_{\nu\in i\setminus\mu} q^\nu_{(i)}(c^\nu)\right] e^{-\beta \sum_{\nu\in i\setminus\mu} \Theta\left[-c^\nu_{(i)} - s_i(b_i^\nu - \rho a_i^\nu)\right]} \tag{20}$$

(where $Z_i^{(\mu)}$ is a normalization factor ensuring that $\int p_i^{(\mu)}(s_i) ds_i = 1$) while for each $\mu = 1, \ldots, M$ and $i \in \mu$ we have

$$q^\mu_{(i)}(c^\mu) = \mathrm{Tr}_{\boldsymbol{s}_{j\in\mu\setminus i}} \left[\prod_{j\in\mu\setminus i} p_j^{(\mu)}(s_j)\right] \delta \left[c^\mu - \sum_{j\in\mu\setminus i} s_j(b_j^\mu - \rho a_j^\mu)\right] \tag{21}$$

Note that these in turn imply that

$$p_i(s_i) = \frac{1}{Z_i} \mathrm{Tr}_{\boldsymbol{c}^{\nu\in i}}\, q_{(i)}(\boldsymbol{c}^{\nu\in i}) e^{-\beta \sum_{\nu\in i} \Theta\left[-c^\nu_{(i)} - s_i(b_i^\nu - \rho a_i^\nu)\right]} \tag{22}$$

$$q^\mu(c^\mu) = \mathrm{Tr}_{\boldsymbol{s}_{j\in\mu}}\, p^{(\mu)}(\boldsymbol{s}_{j\in\mu}) \delta \left[c^\mu - \sum_{j\in\mu} s_j(b_j^\mu - \rho a_j^\mu)\right] \tag{23}$$

so that the whole probability distribution can be reconstructed from the set of cavity distributions. The statistical average cost function is finally given by

$$\left\langle \frac{E}{N} \right\rangle = \frac{1}{N} \sum_{\mu=1}^{M} \mathrm{Tr}_c \Theta(-c) \frac{q^\mu(c)\left[e^{-\beta} + (1-e^{-\beta})\Theta(c)\right]}{\mathrm{Tr}_c q^\mu(c)\left[e^{-\beta} + (1-e^{-\beta})\Theta(c)\right]} \tag{24}$$

The cavity theory developed here recovers the replica theory of [4] in the fully-connected limit. This is shown in detail in Appendices A, where a slightly revised replica theory is discussed, and B, where the fully connected limit is constructed explicitly.

Because the configurational variables s_i and c^μ are continuous, solving equations (20) and (21), which in the zero temperature ($\beta \to \infty$) limit take the form

$$p_i^{(\mu)}(s_i) = \frac{1}{Z_i^{(\mu)}} \mathrm{Tr}_{\boldsymbol{c}^{\nu\in i\setminus\mu}} \left[\prod_{\nu\in i\setminus\mu} q^\nu_{(i)}(c^\nu)\right] \prod_{\nu\in i\setminus\mu} \Theta\left[c^\nu_{(i)} + s_i(b_i^\nu - \rho a_i^\nu)\right] \tag{25}$$

$$q^\mu_{(i)}(c^\mu) = \mathrm{Tr}_{\boldsymbol{s}_{j\in\mu\setminus i}} \left[\prod_{j\in\mu\setminus i} p_j^{(\mu)}(s_j)\right] \delta\left[c^\mu - \sum_{j\in\mu\setminus i} s_j(b_j^\mu - \rho a_j^\mu)\right], \tag{26}$$

by the standard method of population dynamics requires studying a population of populations, as *e.g.* in [10] (in other words, it is equivalent to a one-step RSB calculation already at the replica-symmetric level). We therefore decided to focus our analysis on the calculation of the optimal growth rate, for which it is possible to obtain good results at a modest computational cost.

We are thus interested in finding the critical line $\rho^\star(n)$. One may proceed as follows. Assume there is no linear constraint (the value of $\rho^\star$ is unaffected by the presence of a linear constraint). Then the equations always admit the trivial solution ($\boldsymbol{s} = \mathbf{0}$) for all values of ρ. In fact, for $\rho > \rho^\star$ this is the only zero-energy solution while below $\rho^\star$ other non-null solutions exist. Thus, starting from random initial conditions for the fluxes, we would expect the average flux to vanish (resp. remain different from zero) under the iteration of the cavity equations for values of ρ above (resp. below) $\rho^\star$. This obviously assumes that the iteration of the cavity equations for $\rho < \rho^\star$ keeps the fluxes away from their trivial value, which turns out to be the case.

To check the validity of this assumption, we have first considered the cavity equations (25) and (26) in their fully connected limit (see Appendices for details), *viz.*

$$x^\mu_{(i)} = M_0\left(h^\mu_{(i)}, \phi^\mu_{(i)}\right) \tag{27}$$

$$y^\mu_{(i)} = M_1\left(h^\mu_{(i)}, \phi^\mu_{(i)}\right) \tag{28}$$

$$m_\ell^{(\mu)} = f_1\left(A_\ell^{(\mu)}, B_\ell^{(\mu)}\right) \tag{29}$$

$$L_\ell^{(\mu)} = f_2\left(A_\ell^{(\mu)}, B_\ell^{(\mu)}\right) \tag{30}$$

with

$$h^\mu_{(i)} = \frac{1}{\sqrt{K}} \sum_{\ell\in\mu\setminus i} \xi_\ell^\mu(\rho) m_\ell^{(\mu)} \tag{31}$$

$$\phi^\mu_{(i)} = \frac{1}{K} \sum_{\ell\in\mu\setminus i} [\xi_\ell^\mu(\rho)]^2 [L_\ell^{(\mu)} - [m_\ell^{(\mu)}]^2] \tag{32}$$

$$A_i^{(\mu)} = \frac{1}{\sqrt{K}} \sum_{\nu\in i\setminus\mu} \xi_i^\nu(\rho) x^\mu_{(i)} \tag{33}$$

$$B_i^{(\mu)} = \frac{1}{2K}\sum_{\nu\in i\backslash\mu}[\xi_i^\nu(\rho)]^2\left(y^\nu_{(i)}+[x^\nu_{(i)}]^2\right) \tag{34}$$

$$\xi_i^\mu(\rho) = b_i^\mu - \rho a_i^\mu \tag{35}$$

and where $M_s(a,b)$ and $f_k(a,b)$ are defined as follows ($H(x) = \mathrm{erfc}(x/\sqrt{2})/2$)

$$M_s(a,b) = \frac{\left(1-e^{-\beta}\right)\left(\frac{\partial}{\partial c}\right)^s \mathcal{G}_c(a,b)\Big|_{c=0}}{e^{-\beta}+(1-e^{-\beta})H\left(\frac{-a}{\sqrt{b}}\right)} \quad , \qquad (\beta\to\infty) \tag{36}$$

$$f_k(a,b) = \frac{\mathrm{Tr}_s e^{as-bs^2}s^k}{\mathrm{Tr}_s e^{as-bs^2}} \tag{37}$$

with $\mathcal{G}_c(a,b)$ a Gaussian distribution for the variable c with mean a and variance b. For computational reasons, it is more convenient to iterate the corresponding TAP equations [11], namely

$$x^\mu = M_0\left(\frac{1}{\sqrt{K}}\sum_{\ell\in\mu}\xi_\ell^\mu(\rho)m_\ell - x^\mu\phi^\mu, \phi^\mu\right) \tag{38}$$

$$y^\mu = M_1\left(\frac{1}{\sqrt{K}}\sum_{\ell\in\mu}\xi_\ell^\mu(\rho)m_\ell - x^\mu\phi^\mu, \phi^\mu\right) \tag{39}$$

$$m_i = f_1\left(\frac{1}{\sqrt{K}}\sum_{\nu\in i}\xi_i^\nu(\rho)x^\mu + m_iB_i, B_i\right) \tag{40}$$

$$L_i = f_2\left(\frac{1}{\sqrt{K}}\sum_{\nu\in i}\xi_i^\nu(\rho)x^\mu + m_iB_i, B_i\right) \tag{41}$$

$$B_i = \frac{1}{2K}\sum_{\nu\in i}[\xi_i^\nu(\rho)]^2\left(y^\nu+(x^\nu)^2\right) \tag{42}$$

$$\phi^\mu = \frac{1}{K}\sum_{i\in\mu}[\xi_i^\mu(\rho)]^2\left(L_i - m_i^2\right) \tag{43}$$

Given a matrix $\xi_i^\mu(\rho)$, we solve the preceding set of coupled equations by iteration: starting from a small value of ρ we monitor the average flux and determine $\rho^\star$ when, under iteration of the equations, the average flux goes to zero. In Fig. 11 we compare the results for a system of $M = 100$ reagents (average over 10 disorder samples) with the theoretical prediction. The agreement is fairly good.

We have then considered the same approach to calculate $\rho^\star$ for a Regular(N)/Poisson(M) network with connectivity $D = 7$ (see Fig. 11). To do so, we have considered the ensemble version of equations (25) and (26), which take the form (we suppress cavity indexes (x) for simplicity)

$$p_i(s) = \frac{1}{Z_i}\prod_{\nu=1}^{2D-1}\left[\mathrm{Tr}_{c^\nu}\, q^\nu(c^\nu)\Theta\left[c^\nu + s(b_i^\nu - \rho a_i^\nu)\right]\right] \tag{44}$$

$$q^\mu(c) = \left[\prod_{j=1}^{k_\mu-1}\mathrm{Tr}_{s_j}\, p_j(s_j)\right]\delta\left[c - \sum_{j=1}^{k_\mu-1}s_j(b_j^\mu - \rho a_j^\mu)\right] \tag{45}$$

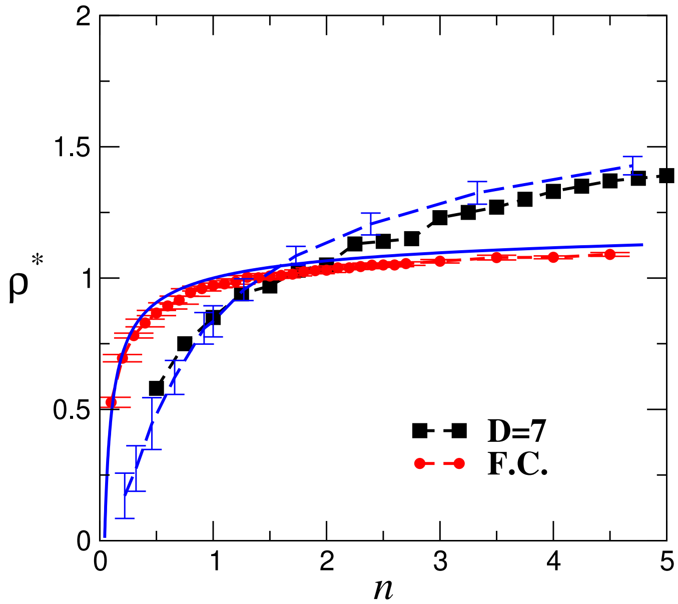

Figure 11. Results from cavity theory. The continuous (blue) line is the theoretical (replica) prediction in the fully connected limit. The dashed line represents Minover results for Regular(N)/Poisson (M) graphs, with error bars. Markers correspond to iteration of the fully connected (circles, average over 10 disorder samples, number of reagents $M = 100$) and Regular(N)/Poisson (M) cavity equations (squares).

with k_μ a random Poisson number with average nD. In the ensemble, which is obtained after the limit $N \to \infty$ has been performed and where n plays the role of an external parameter, the indexes i and μ now take values $i, \mu = 1, 2, 3, \ldots$ and describe an infinite population of probabilities $\{q^\mu(c)\}_{\mu=1,2,\ldots}$ and $\{p_i(s)\}_{i=1,2,\ldots}$. We have studied a finite population of probabilities, *i.e.* $i, \mu = 1, \ldots, \mathcal{N}_{pop}$, which are initialised randomly. In order to find $\rho^\star$, we start from a small initial value for ρ. The population is iterated via equations (44) and (45). If after a certain (sufficiently large) number of iteration steps the average flux is different from zero, we increase the value of ρ by a small value and restart the procedure. Eventually we can locate a value of ρ at which the average flux vanishes. This defines $\rho^\star$. In Fig. 11 we display the resulting curve $\rho^\star(n)$ for a population of $\mathcal{N}_{pop} = 50$ distributions and $D = 7$, which qualitatively agrees with the results found by the Minover$^+$ algorithm. (We remind that cavity equations are obtained for a tree-like topology and the network we considered in the numerical solution is relatively small.)

4. Summary

Despite evident quantitative differences, the overall picture obtained for the fully-connected Von Neumann's model is roughly preserved in the case of finite connectivity, discussed here. The major difference appears to be that much larger growth rates can be achieved in diluted networks (reactions with few inputs/outputs are evidently more efficient) and, generically, when reactions have stochastic connectivities. The framework of Von Neumann's problem, including its global optimization aspect, is sufficiently simple to ensure that the methods developed in this work can be applied in several different real-world contexts and provides an alternative approach with respect to the typical assumption that the optimal state of such systems is the one maximizing a particular set of fluxes. Metabolic networks represent an important test ground for these ideas. In fact in this context Von Neumann's global optimization principle as a tool to predict e.g. reaction flux distributions appears to be alternative to the schemes based on local mass-balance conditions coupled with the maximization of a subset of fluxes (e.g. flux-balance analysis). Their analysis however requires the inclusion of some additional ingredients, most notably that of an input layer of metabolites (source nodes), which provide the environment surrounding the cell and which have been neglected in the present study. Work along these lines is in progress.

Acknowledgments

It is our pleasure to thank G Bianconi, S Cocco, E Marinari, M Marsili, A Pagnani, G Parisi, F Ricci Tersenghi, T Rizzo and MA Virasoro for valuable discussions and suggestions.

References

[1] Von Neumann J 1937 Ergebn. eines Math. Kolloq. **8**. English translation: Von Neumann J 1945 Rev. Econ. Studies **13** 1
[2] Gale D 1956 In Kuhn HW and Tucker AW (Eds), *Linear inequalities and related systems* (Ann. Math. Studies **38**, Princeton, NJ)
[3] McKenzie LW 1986 "Optimal Economic Growth, Turnpike Theorems and Comparative Dynamics", in Arrow KJ and Intriligator MD (eds), *Handbook of Mathematical Economics, Vol. III* (North-Holland, Amsterdam)
[4] De Martino A and Marsili M 2005 JSTAT L09003
[5] Mézard M and Zecchina R 2002 Phys. Rev. E **66** 056126
[6] Hertz J, Krogh A and Palmer RG 1991 *Introduction to the Theory of Neural Computation* (Addison-Wesley, Redwood City, CA)
[7] Krauth W and Mézard M 1987 J. Phys. A: Math. Gen. **20** L745
[8] De Martino A, Marinari E, Marsili M, Martelli C and Perez Castillo I 2006 in preparation
[9] Jeong H, Tombor B, Albert R, Oltvai ZN, Barabasi AL 2000 Nature **407** 651
[10] Skantzos N, Perez Castillo I and Hatchett J 2005 Phys. Rev. E **72** 066127
[11] Shamir M and Sompolinsky H 2000 Phys. Rev. E **61** 1839

Appendix A. The fully-connected model revisited

We shall briefly re-consider here the replica approach of [4] for the fully connected model in the Derrida-Gardner framework (as opposed to the Gardner one employed in [4]). This allows us to derive equations for the order parameters that are easily compared with those which shall be derived in the following section as the fully connected limit of the cavity theory.

Our starting point is the partition function, which we write in this case

$$Z = \mathrm{Tr}_{\boldsymbol{c}} \prod_{\mu=1}^{M} \left[e^{-\beta} + (1-e^{-\beta})\Theta(c^\mu)\right] \mathrm{Tr}_{\boldsymbol{s}} \prod_{\mu=1}^{M} \delta\left(c^\mu - \frac{1}{\sqrt{N}}\sum_{i=1}^{N} s_i \xi_i^\mu(\rho)\right) \tag{A.1}$$

where

$$\xi_i^\mu(\rho) = b_i^\mu - \rho a_i^\mu \tag{A.2}$$

As in [4], we write $b_i^\mu = \overline{b}(1+\beta_i^\mu)$ and $a_i^\mu = \overline{a}(1+\alpha_i^\mu)$ with Gaussian α_i^μ and β_i^μ and further set $\overline{a} = \overline{b} = 1$. Moreover, we re-cast ρ as $\rho = 1 + g/\sqrt{N}$ so that

$$\xi_i^\mu(\rho) = \beta_i^\mu - \alpha_i^\mu - \frac{g}{\sqrt{N}}(1+\alpha_i^\mu) \tag{A.3}$$

The replicated and disorder-averaged partition function reads

$$\begin{aligned}\overline{Z^r} = \mathrm{Tr}_{\{\boldsymbol{c}_1,\ldots,\boldsymbol{c}_r\}} &\prod_{\ell=1}^{r}\prod_{\mu=1}^{M} \left[e^{-\beta} + (1-e^{-\beta})\Theta(c_\ell^\mu)\right] \\ \times \mathrm{Tr}_{\{\boldsymbol{s}_1,\ldots,\boldsymbol{s}_r\}} &\overline{\prod_{\ell=1}^{r}\prod_{\mu=1}^{M} \delta\left(c_\ell^\mu - \frac{1}{\sqrt{N}}\sum_{i=1}^{N} s_{i\ell}\xi_i^\mu(\rho)\right)}\end{aligned} \tag{A.4}$$

If the δ-distributions are written in their Fourier representation (with multipliers $\widehat{c}_\ell^\mu$) the disorder average can be carried out easily, generating the standard order parameters

$$q_{\ell\ell'} = \frac{1}{N}\sum_i s_{i\ell}s_{i\ell'} \qquad m_\ell = \frac{1}{N}\sum_i s_{i\ell} \tag{A.5}$$

which can be forced into the partition function with proper Lagrange multipliers $\widehat{q}_{\ell\ell'}$ and $\widehat{m}_\ell$. One finally finds

$$\overline{Z^r} = \int D\boldsymbol{q}D\widehat{\boldsymbol{q}}D\boldsymbol{m}D\widehat{\boldsymbol{m}}\; e^{N[\sum_{\ell\le\ell'}\widehat{q}_{\ell\ell'}q_{\ell\ell'} + \sum_\ell \widehat{m}_\ell m_\ell + F(\boldsymbol{q},\boldsymbol{m}) + S(\widehat{\boldsymbol{q}},\widehat{\boldsymbol{m}})]} \tag{A.6}$$

where ($\alpha = M/N, \kappa \equiv \overline{(\beta_i^\mu - \alpha_i^\mu)^2}$)

$$\begin{aligned}F(\boldsymbol{q},\boldsymbol{m}) &= \alpha\log\int\left[\prod_{\ell=1}^{r}\frac{dc_\ell d\widehat{c}_\ell}{2\pi}\right] e^{i\sum_{\ell=1}^{r}\widehat{c}_\ell c_\ell + \mathrm{i}g\sum_{\ell=1}^{r}\widehat{c}_\ell m_\ell - \frac{\kappa}{2}\sum_{\ell,\ell'}\widehat{c}_\ell\widehat{c}_{\ell'}q_{\ell\ell'}} \\ &\qquad\times\prod_{\ell=1}^{r}\left[e^{-\beta} + (1-e^{-\beta})\Theta(c_\ell)\right] \\ S(\widehat{\boldsymbol{q}},\widehat{\boldsymbol{m}}) &= \frac{1}{N}\log\mathrm{Tr}_{\{\boldsymbol{s}_\ell\}}\; e^{-\sum_{\ell\le\ell'}\widehat{q}_{\ell\ell'}\sum_{i=1}^{N}s_{i\ell}s_{i\ell'} - \sum_\ell\widehat{m}_\ell\sum_{i=1}^{N}s_{i\ell}}\end{aligned}$$

In the limit $N\to\infty$, the integral (A.6) is evaluated by a saddle-point method. Let us consider the saddle-point equation for the order parameter $\widehat{q}_{\ell\ell'}$ ($\ell\neq\ell'$):

$$\widehat{q}_{\ell\ell'} = \kappa\alpha\frac{\int\prod_\ell\frac{dc_\ell d\widehat{c}_\ell}{2\pi}\;\widehat{c}_\ell\widehat{c}_{\ell'}\; e^{i\sum_{\ell=1}^{r}\widehat{c}_\ell c_\ell + \mathrm{i}g\sum_\ell\widehat{c}_\ell m_\ell - \frac{\kappa}{2}\sum_{\ell,\ell'}\widehat{c}_\ell\widehat{c}_{\ell'}q_{\ell\ell'}}\prod_{\ell=1}^{r}\left[e^{-\beta} + (1-e^{-\beta})\Theta(c_\ell)\right]}{\int\prod_\ell\frac{dc_\ell d\widehat{c}_\ell}{2\pi}\; e^{i\sum_{\ell=1}^{r}\widehat{c}_\ell c_\ell + \mathrm{i}g\sum_\ell\widehat{c}_\ell m_\ell - \frac{\kappa}{2}\sum_{\ell,\ell'}\widehat{c}_\ell\widehat{c}_{\ell'}q_{\ell\ell'}}\prod_{\ell=1}^{r}\left[e^{-\beta} + (1-e^{-\beta})\Theta(c_\ell)\right]} \tag{A.7}$$

The argument of the exponential in the denominator can be linearized by a Hubbard-Stratonovich transformation. Once this is done, it is clear that the remaining integrals over c_ℓ and $\widehat{c}_\ell$ factorize and one easily sees that in the replica limit $r \to 0$ the denominator tends to 1 and we can neglect it. It then remains to evaluate the numerator with the replica-symmetric (RS) Ansatz

$$q_{\ell\ell'} = (Q-q)\delta_{\ell\ell'} + q \qquad m_\ell = m \tag{A.8}$$

$$\widehat{q}_{\ell\ell'} = (\widehat{Q}-\widehat{q})\delta_{\ell\ell'} + \widehat{q} \qquad \widehat{m}_\ell = \widehat{m} \tag{A.9}$$

in the replica limit $r \to 0$. With standard manipulations one finds

$$\widehat{q} = -\kappa\alpha \int dx\ \mathcal{G}_x(0,\kappa q) \left(\frac{(1-e^{-\beta})\frac{1}{\sqrt{2\pi\kappa(Q-q)}} e^{-\frac{(gm+x)^2}{2\kappa(Q-q)}}}{e^{-\beta} + (1-e^{-\beta}) H\left(\frac{gm+x}{\sqrt{\kappa(Q-q)}}\right)} \right)^2 \tag{A.10}$$

where $\mathcal{G}_x(a,b)$ denotes a Gaussian distribution for the variable x with mean a and variance b. while $H(x) = \frac{1}{2}\mathrm{erfc}(x/\sqrt{2})$.

Along similar lines one can derive the corresponding equations for the remaining order parameters:

$$\widehat{Q} = -\frac{\kappa\alpha}{2} \int dx\ \mathcal{G}_x(0,\kappa q) \frac{(1-e^{-\beta})\partial_x \mathcal{G}_x(-gm, \kappa(Q-q))}{e^{-\beta} + (1-e^{-\beta}) H\left(\frac{gm+x}{\sqrt{\kappa(Q-q)}}\right)} \tag{A.11}$$

$$\widehat{m} = g\alpha \int dx\ \mathcal{G}_x(0,\kappa q) \frac{(1-e^{-\beta})\frac{1}{\sqrt{2\pi\kappa(Q-q)}} e^{-\frac{(gm+x)^2}{2\kappa(Q-q)}}}{e^{-\beta} + (1-e^{-\beta}) H\left(\frac{gm+x}{\sqrt{\kappa(Q-q)}}\right)} \tag{A.12}$$

$$q = \int dx\, \mathcal{G}_x(0,1) \left(\frac{\int ds\, e^{-\left(\widehat{Q}-\frac{\widehat{q}}{2}\right)s^2 + (ix\sqrt{\widehat{q}} - \widehat{m})s} s}{\int ds\, e^{-\left(\widehat{Q}-\frac{\widehat{q}}{2}\right)s^2 + (ix\sqrt{\widehat{q}} - \widehat{m})s}} \right)^2 \tag{A.13}$$

$$Q = \int dx\, \mathcal{G}_x(0,1) \left(\frac{\int ds\, e^{-\left(\widehat{Q}-\frac{\widehat{q}}{2}\right)s^2 + (ix\sqrt{\widehat{q}} - \widehat{m})s} s^2}{\int ds\, e^{-\left(\widehat{Q}-\frac{\widehat{q}}{2}\right)s^2 + (ix\sqrt{\widehat{q}} - \widehat{m})s}} \right) \tag{A.14}$$

$$m = \int dx\, \mathcal{G}_x(0,1) \left(\frac{\int ds\, e^{-\left(\widehat{Q}-\frac{\widehat{q}}{2}\right)s^2 + (ix\sqrt{\widehat{q}} - \widehat{m})s} s}{\int ds\, e^{-\left(\widehat{Q}-\frac{\widehat{q}}{2}\right)s^2 + (ix\sqrt{\widehat{q}} - \widehat{m})s}} \right) \tag{A.15}$$

Appendix B. The fully-connected limit of the cavity theory for the diluted model

In this Appendix we study the large connectivity limit of the cavity theory and derive equations for the relevant order parameters in this limit, in order to show that these equations are equivalent to those obtained in the revisited replica theory presented above. The first problem is to study the cavity distributions (20) and (21) in the fully connected limit for a fixed disorder realization. Next, we shall perform the disorder average.

Let us begin from (21), which we re-cast as

$$q^{\mu}_{(i)}(c) = \mathrm{Tr}_{\boldsymbol{s}_{j\in\mu\setminus i}} \left[\prod_{j\in\mu\setminus i} p_j^{(\mu)}(s_j)\right] \delta\left[c - \frac{1}{\sqrt{K}}\sum_{j\in\mu\setminus i} s_j \xi_j^{\mu}(\rho)\right] \tag{B.1}$$

$$= \int \frac{d\widehat{c}}{2\pi} e^{\mathrm{i}c\widehat{c}} \prod_{j\in\mu\setminus i} \mathrm{Tr}_{s_j}\, p_j^{(\mu)}(s_j) e^{-\frac{\mathrm{i}}{\sqrt{K}}\widehat{c} s_j \xi_j^{\mu}(\rho)} \tag{B.2}$$

where we introduced a re-scaling factor $1/\sqrt{K}$ for the fields (K being the average connectivity). For $K \gg 1$ we can expand the exponential. Keeping terms up to the second order we find

$$\prod_{j\in\mu\setminus i} \mathrm{Tr}_{s_j}\, p_j^{(\mu)}(s_j) e^{-\frac{\mathrm{i}}{\sqrt{K}}\widehat{c} s_j \xi_j^{\mu}(\rho)} = \exp\left[-\frac{\mathrm{i}}{\sqrt{K}}\widehat{c}\sum_{j\in\mu\setminus i} \xi_j^{\mu}(\rho) m_j^{(\mu)} - \frac{\widehat{c}^2}{2K}\sum_{j\in\mu\setminus i} [\xi_j^{\mu}(\rho)]^2 \Delta_j^{(\mu)}\right] \tag{B.3}$$

where

$$m_i^{(\mu)} = \langle s_i \rangle_{(\mu)} \tag{B.4}$$

$$\Delta_i^{(\mu)} = \left\langle s_i^2 \right\rangle_{(\mu)} - \langle s_i \rangle_{(\mu)}^2 \tag{B.5}$$

$$\langle f(s_i) \rangle_{(\mu)} = \mathrm{Tr}_s\, p_i^{(\mu)}(s) f(s_i) \tag{B.6}$$

Inserting (B.3) into (B.2) and integrating over $\widehat{c}$ one sees that for $K \gg 1$

$$q^{\mu}_{(i)}(c) = \mathcal{G}_c\left[\frac{1}{\sqrt{K}}\sum_{j\in\mu\setminus i} \xi_j^{\mu}(\rho) m_j^{(\mu)}, \frac{1}{K}\sum_{j\in\mu\setminus i} [\xi_j^{\mu}(\rho)]^2 \Delta_j^{(\mu)}\right] \tag{B.7}$$

where as before $\mathcal{G}_x(a,b)$ denotes a Gaussian distribution for the variable x with mean a and variance b.

The limit of (20) is a bit more complicated. For a start, we re-write it by introducing the usual re-scaling:

$$p_i^{(\mu)}(s) = \frac{1}{Z_s} \mathrm{Tr}_{\boldsymbol{c}^{\nu\in i\setminus\mu}} \left[\prod_{\nu\in i\setminus\mu} q^{\nu}_{(i)}(c^{\nu})\right] e^{-\beta\sum_{\nu\in i\setminus\mu}\Theta\left[-c^{\nu}_{(i)} - \frac{1}{\sqrt{K}} s \xi_i^{\nu}(\rho)\right]} \tag{B.8}$$

$$= \frac{1}{Z_s} \prod_{\nu\in i\setminus\mu} \mathrm{Tr}_c\, q^{\nu}_{(i)}(c) e^{-\beta\Theta\left[-c - \frac{1}{\sqrt{K}} s \xi_i^{\nu}(\rho)\right]} \tag{B.9}$$

We focus on

$$\mathrm{Tr}_c\, q^{\nu}_{(i)}(c) e^{-\beta\Theta\left[-c - \frac{1}{\sqrt{K}} s \xi_i^{\nu}(\rho)\right]} = \mathrm{Tr}_c\, q^{\nu}_{(i)}(c)\left[e^{-\beta} + \left(1 - e^{-\beta}\right)\Theta\left(c + \frac{1}{\sqrt{K}} s \xi_i^{\nu}(\rho)\right)\right] \tag{B.10}$$

Now for any K the Heaviside function can be disposed of via

$$\Theta\left(c + \frac{1}{\sqrt{K}} s \xi_i^{\nu}(\rho)\right) = \Theta(c) + \varepsilon^{\nu}(c,s) \tag{B.11}$$

$$\varepsilon^{\nu}(c,s) = \sum_{n\geq 1} \frac{1}{n!}\left(\frac{1}{\sqrt{K}} s\, \xi_i^{\nu}(\rho)\right)^n \delta^{(n-1)}(c) \tag{B.12}$$

so that

$$\mathrm{Tr}_c\, q^{\nu}_{(i)}(c) e^{-\beta\Theta\left[-c-\frac{1}{\sqrt{K}} s\xi^{\nu}_i(\rho)\right]} = \mathrm{Tr}_c\, q^{\nu}_{(i)}(c)\left[e^{-\beta}+\left(1-e^{-\beta}\right)\Theta(c)\right] + \left(1-e^{-\beta}\right)\mathrm{Tr}_c\, q^{\nu}_{(i)}(c)\varepsilon^{\nu}(c,s)$$

Factoring out a term $\mathrm{Tr}_c\, q^{\nu}_{(i)}(c)\left[e^{-\beta}+\left(1-e^{-\beta}\right)\Theta(c)\right]$ one sees that (B.9) becomes

$$p^{(\mu)}_i(s) = \frac{1}{\widetilde{Z}_s}\prod_{\nu\in i\setminus\mu}\left[1+\left(1-e^{-\beta}\right)\mathrm{Tr}_c\,\mathcal{Q}^{\nu}_{(i)}(c)\varepsilon^{\nu}(c,s)\right] \tag{B.13}$$

where $\widetilde{Z}_s$ is a new normalization factor and

$$\mathcal{Q}^{\nu}_{(i)}(c) = \frac{q^{\nu}_{(i)}(c)}{\mathrm{Tr}_c\, q^{\nu}_{(i)}(c)\left[e^{-\beta}+\left(1-e^{-\beta}\right)\Theta(c)\right]} \tag{B.14}$$

With the expression given in (B.12), we have

$$\mathrm{Tr}_c\,\mathcal{Q}^{\nu}_{(i)}(c)\varepsilon^{\nu}(c,s) = \sum_{n\geq 1}\frac{1}{n!}\left(\frac{1}{\sqrt{K}} s\,\xi^{\nu}_i(\rho)\right)^n (-1)^{n-1}\mathcal{Q}^{\nu,(n-1)}_{(i)}(0) \tag{B.15}$$

For $K \gg 1$ we can truncate the above expansion after the first two terms and approximate the resulting expression with an exponential. This gives us

$$p^{(\mu)}_i(s) = \frac{e^{A^{(\mu)}_i s - B^{(\mu)}_i s^2}}{\mathrm{Tr}_s\, e^{A^{(\mu)}_i s - B^{(\mu)}_i s^2}} \tag{B.16}$$

with

$$A^{(\mu)}_i = \left(1-e^{-\beta}\right)\frac{1}{\sqrt{K}}\sum_{\nu\in i\setminus\mu}\xi^{\nu}_i(\rho)\mathcal{Q}^{\nu}_{(i)}(0) \tag{B.17}$$

$$B^{(\mu)}_i = \left(1-e^{-\beta}\right)\frac{1}{2K}\sum_{\nu\in i\setminus\mu}\left[\xi^{\nu}_i(\rho)\right]^2\left(\mathcal{Q}^{\nu,(1)}_{(i)}(0)+\left(1-e^{-\beta}\right)\left[\mathcal{Q}^{\nu}_{(i)}(0)\right]^2\right) \tag{B.18}$$

(B.7) and (B.16) represent the $K \gg 1$ limits of the cavity distributions for a fixed disorder sample.

Recalling (A.2) and (A.3), we can now evaluate averages over the quenched disorder. To this aim, we set $\rho = 1 + g/\sqrt{K}$ so that

$$\xi^{\mu}_i(\rho) = \beta^{\mu}_i - \alpha^{\mu}_i - \frac{g}{\sqrt{K}}(1+\alpha^{\mu}_i) \tag{B.19}$$

Computing the statistics of the quantities $A^{(\mu)}_i$ and $B^{(\mu)}_i$ over disorder one obtains (with $\alpha = K/N$)

$$\overline{A^{(\mu)}_i} \equiv A = -g\alpha\left(1-e^{-\beta}\right)F \tag{B.20}$$

$$\overline{(A^{(\mu)}_i - A)^2} \equiv \sigma_A = \alpha\kappa\left(1-e^{-\beta}\right)^2 D \tag{B.21}$$

$$\overline{B^{(\mu)}_i} \equiv B = \alpha\kappa\left(1-e^{-\beta}\right)\left[C+\left(1-e^{-\beta}\right)D\right] \tag{B.22}$$

($B_i^{(\mu)}$ is sample-independent in the fully connected limit $K \to \infty$, so we neglect its fluctuations) where we used the following shorthands:

$$C = \overline{\mathcal{Q}_{(i)}^{\nu,(1)}(0)} \tag{B.23}$$

$$D = \overline{[\mathcal{Q}_{(i)}^{\nu}(0)]^2} \tag{B.24}$$

$$F = \overline{\mathcal{Q}_{(i)}^{\nu}(0)} \tag{B.25}$$

$$\kappa \equiv \overline{(\beta_i^\mu - \alpha_i^\mu)^2} \tag{B.26}$$

This implies that

$$\overline{[\langle f(s_i)\rangle_{(\mu)}]^k} = \int dH\ \mathcal{G}_H(A, \sigma_A) \left[\frac{\mathrm{Tr}_s\, f(s) e^{Hs - Bs^2}}{\mathrm{Tr}_s\, e^{Hs-Bs^2}}\right]^k \tag{B.27}$$

We now must evaluate C, D and F, recalling $\mathcal{Q}_{(i)}^{\nu}(c)$ as defined in Equations (B.14) and (B.7). The statistics of $h = \frac{1}{\sqrt{K}} \sum_{j\in\mu\setminus i} \xi_j^\mu(\rho) m_j^{(\mu)}$ is given by

$$\overline{h} = -gM, \qquad M = \overline{m_i^{(\mu)}} \tag{B.28}$$

$$\overline{h^2} = g^2 M^2 + \kappa q, \qquad q = \overline{\langle s_i\rangle_{(\mu)}^2} \tag{B.29}$$

$$\overline{(h - \overline{h})^2} = \kappa q \tag{B.30}$$

For the variance of $\mathcal{Q}_{(i)}^{\nu}(c)$ we obtain instead

$$\overline{\frac{1}{K}\sum_{j\in\mu\setminus i} [\xi_j^\mu(\rho)]^2 \Delta_j^{(\mu)}} = \kappa(Q - q), \qquad Q = \overline{\langle s_i^2\rangle_{(\mu)}} \tag{B.31}$$

We can therefore write

$$F = \int dh\, \mathcal{G}_h(-gM, \kappa q) \frac{\mathcal{G}_{c=0}\left[h, \kappa(Q-q)\right]}{\int dc\, \mathcal{G}_c\left[h, \kappa(Q-q)\right] \left[e^{-\beta} + (1 - e^{-\beta})\, \Theta\left(c\right)\right]} \tag{B.32}$$

$$D = \int dh\, \mathcal{G}_h(-gM, \kappa q) \left[\frac{\mathcal{G}_{c=0}\left[h, \kappa(Q-q)\right]}{\int dc\, \mathcal{G}_c\left[h, \kappa(Q-q)\right] \left[e^{-\beta} + (1 - e^{-\beta})\, \Theta\left(c\right)\right]}\right]^2 \tag{B.33}$$

$$C = \int dh\, \mathcal{G}_h(-gM, \kappa q) \frac{\mathcal{G}^{(1)}_{c=0}\left[h, \kappa(Q-q)\right]}{\int dc\, \mathcal{G}_c\left[h, \kappa(Q-q)\right] \left[e^{-\beta} + (1 - e^{-\beta})\, \Theta\left(c\right)\right]} \tag{B.34}$$

These form a closed system together with the equations for M, Q and q:

$$M = \overline{\langle s_i\rangle_{(\mu)}} = \int dH\ \mathcal{G}_H(A, \sigma_A) \left[\frac{\mathrm{Tr}_s\, s\ e^{Hs-Bs^2}}{\mathrm{Tr}_s\, e^{Hs-Bs^2}}\right] \tag{B.35}$$

$$Q = \overline{\langle s_i^2\rangle_{(\mu)}} = \int dH\ \mathcal{G}_H(A, \sigma_A) \left[\frac{\mathrm{Tr}_s\, s^2\ e^{Hs-Bs^2}}{\mathrm{Tr}_s\, e^{Hs-Bs^2}}\right] \tag{B.36}$$

$$q = \overline{\langle s_i\rangle_{(\mu)}^2} = \int dH\ \mathcal{G}_H(A, \sigma_A) \left[\frac{\mathrm{Tr}_s\, s\ e^{Hs-Bs^2}}{\mathrm{Tr}_s\, e^{Hs-Bs^2}}\right]^2 \tag{B.37}$$

With minor manipulations, these equations can be precisely identified with the equations obtained by the replica approach in RS approximation in the previous section.