A statistical mechanics approach to autopoietic immune networks

J. Stat. Mech. (2010) P07004

(http://iopscience.iop.org/1742-5468/2010/07/P07004)

Journal of Statistical Mechanics: Theory and Experiment
An IOP and SISSA journal

A statistical mechanics approach to autopoietic immune networks

Adriano Barra[1] **and Elena Agliari**[2,3,4]

[1] Dipartimento di Fisica, 'La Sapienza' Università di Roma, P le A Moro 2, 00185 Roma, Italy
[2] Dipartimento di Fisica, Università di Parma, Viale G P Usberti 7/a, 43100 Parma, Italy
[3] INFN, Gruppo Collegato di Parma, Viale G P Usberti 7/a, 43100 Parma, Italy
[4] Theoretische Polymerphysik, Albert-Ludwigs-Universität Freiburg, Hermann-Herder-Str. 3, 79104 Freiburg, Germany
E-mail: adriano.barra@roma1.infn.it and elena.agliari@fis.unipr.it

Received 26 February 2010
Accepted 10 June 2010
Published 1 July 2010

Online at stacks.iop.org/JSTAT/2010/P07004
doi:10.1088/1742-5468/2010/07/P07004

Abstract. In this work we aim to bridge theoretical immunology and disordered statistical mechanics. We introduce a model for the behavior of B-cells which naturally merges the clonal selection theory and the autopoietic network theory as a whole. From the analysis of its features we recover several basic phenomena such as low-dose tolerance, dynamical memory of antigens and self/non-self discrimination.

Keywords: disordered systems (theory), molecular networks (theory)

J. Stat. Mech. (2010) P07004

1742-5468/10/P07004+24$30.00

Contents

1. Introduction 2
2. Fundamentals of theoretical immunology 3
3. Antibodies as vectors in the base of idiotopes 5
4. One-body and two-body Hamiltonian 6
5. Topology and affinity pattern of the emergent network 8
6. Self/non-self recognition 15
7. Low-dose tolerance 17
8. Multiple responses and spurious states 18
9. Dynamical memory 20
10. Summary 21
Acknowledgments 22
References 23

1. Introduction

The purpose of the immune system is to detect and neutralize the molecules, or cells, dangerous to the body (antigens, which could be foreign invaders—e.g. viruses or bacteria—or deranged—e.g. cancerous—cells of the host), without damaging healthy cells [1]. Despite the evident differences, to accomplish its function the immune system exhibits properties analogous to the nervous system [11]: it 'learns' not to attack healthy cells and it 'develops a memory' of the pathogens encountered as time goes by. In theoretical immunology there are two main strands to explain the functioning of the immune system that ultimately represent two approaches, *reductionist* and *systemic*, for the modeling of nature in general. In the first and most popular approach, lymphocytes basically operate independently or, better, the researcher focuses on the action of the single lymphocyte and on the details of its interactions (i.e. internal cascade signals, etc) rather than on the global behavior of all the lymphocytes interacting with each other. In the second approach, pioneered in immunology by Elrich [19] and Jerne [24], the immune system is thought of as a whole and designed as a network of cells stimulated to proliferate by the affinity interactions of their exchanging antibodies (a functional idiotypic network [27]). Interestingly, the two approaches are not incompatible but complementary. While the former deals primarily with the response to a stimulus, the latter allows us to explain the ability to learn and memorize of the immune system and the tolerance to low doses of antigen [14]. In the past, the immune network theory has been investigated, although not exhaustively, with disordered statistical mechanics tools [31]. However, recent and deep advances in the field of statistical mechanics of highly diluted networks [6, 22, 2, 21, 30, 36] now allow us to combine the two viewpoints described above

doi:10.1088/1742-5468/2010/07/P07004

J. Stat. Mech. (2010) P07004

and to develop a unified and quantitative theory. Indeed, reductionist and systemic approaches can be recovered as special cases of null and non-negligible connectivity, respectively. This unification should be extremely promising from a biological as well as a mathematical point of view.

We start by introducing the *one-body* theory and noticing that in immunology it corresponds to what we call a 'Burnet-like behavior'. Then we extend our model including the *two-body* theory and show that it recovers what we call a 'Jerne-like behavior'; as a natural consequence, we will show how this leads to the approach of Counthino–Varela for the systemic self/non-self distinction [39, 40]. After these results, we show how hysteresis, with its remanent magnetization, can play the role of the generator of memory cells from plasma cells, according to the Clonal Selection Theory [7, 8]. Finally, we show how low- and high-dose tolerances, as well as the bell-shaped response, appear as emergent features in our model, while in theoretical immunology analysis they are often postulated *a priori*.

Even though not exhaustive, our model may act as an alternative starting backbone for this field of research.

2. Fundamentals of theoretical immunology

The main constituents of an (adaptive) immune system are B-lymphocytes (B-cells), together with T-lymphocytes, and free antibodies produced by B-cells. B-cells and T-cells have specific protein molecules on their surfaces, called receptors. The receptors of B-cells are antibodies (Immunoglobulin, Ig), which can recognize and connect to antigens in order to neutralize them. Finally, the purpose of killer T-cells is to attack and kill infected or deranged cells. The receptors of B- and T-cells have specific three-dimensional structures, called 'idiotypes'. A family of B-cells generated by a proliferating B-cell are called 'clones'; a clone and the antibodies which it produces have the same idiotypes.

In a healthy human body at rest, it is estimated that the total number of 'sentinel' clones generated from a single B-cell (the amount of identical lymphocytes) is about 10^2 to 10^4, the total number of clones amounts to some 10^{12}–10^{14}, such that diverse clones are around 10^{10}–10^{12}, and the number of antibodies is about 10^{18}. Remarkably, the amount of epitopes/idiotopes belonging to a given antibody are present in a smaller number, i.e. of the order of 10^2.

When antigens enter the body, those clones which recognize it will bind to it. Aided by helper T-cells, B-cells of an activated clone will proliferate, becoming antibody-producing cells. The latter will secrete large numbers of free antibodies, which attach to the antigen, neutralize it and trigger killer cells into action.

Each clone of B-cells always produces the same antibody (hyper-mutations apart, which will not be discussed here—see, for instance, [23, 38]).

The above is (a part of) 'clonal selection theory' and has been confirmed experimentally.

We notice that this approach, pioneered by Burnet [7], takes into account an enormous amount of different data, and absolutely does not rely on interactions among different lymphocytes, as it deals with the external antigen interaction with the immune system, where the network works at a completely hidden level.

The idea of an internal network appeared early in immunology [19], and its concretization happened when Jerne, in the 1970s, suggested that each antibody must

J. Stat. Mech. (2010) P07004

have several idiotopes which are detected by other antibodies. Via this mechanism, an effective network of interacting antibodies is formed, in which antibodies not only detect antigens, but also function as individual internal images of certain antigens and are themselves being detected and acted upon. These mutual interactions provide a 'dynamical memory' of the immune system, by keeping the concentrations of antibodies (especially those representing encountered antigens) at appropriate levels. This can be understood as follows: at a given time a virus is introduced into the body and starts replication. As a consequence, at high enough concentrations, it is found by the proper B-lymphocyte counterpart: let us consider, for simplicity, a virus as a string of information (i.e. 1001001). The complementary B-cell producing the antibody Ig1, which can be thought of as the string 0110110 (the dichotomy of a binary alphabet in strings mirrors the one of the electromagnetic field governing chemical bonds) which then will start a clonal expansion and will release high levels of Ig1. As a consequence, after a while, another B-cell will meet 0110110 and, as this string never (macroscopically) existed before, attacks it by releasing the complementary string 1001001, which, actually, is a 'copy' (internal image) of the original virus but with no DNA or RNA charge inside. The interplay among these keeps a memory of the past infection. However, in the 1990s, the network theory was considered to be marginal: it did not appear as a part of a whole although it gave an appealing mechanism for the implementation of memory in the immune system.

Beyond memory storage, another feature of the immune system is very impressive: it is able to attack antigens but not host molecules or cells. Immunologists name this ability as the distinction among *self* and *non-self*: self/non-self discrimination is of fundamental importance as several diseases may appear if it is non-properly working (this is the case of auto-immune pathologies [34]).

In a nutshell, following the classical vision and according to the (sometimes called reductionist [17]) antigen-driven view of the immune system, newborn lymphocytes learn from the beginning the difference among self and non-self (it is assumed the existence of an *a priori* learning in specific regions of the body—i.e. thymus—where all the lymphocytes are made to interact with self and all the responding ones are killed). As a consequence, the presence of auto-immunity in the system is due to a non-proper elimination of those B-cells which, at their early stage, failed to learn such a difference. This defines the allopoietic viewpoint.

It must be stressed that, without a two-body interaction, which makes possible the existence of a network, this property cannot be spread on the whole system and, indeed, we must assume that each lymphocyte stores the whole required information by itself, namely the reductionist viewpoint.

However, within the idiotypic network theory started by Jerne, the emergence of an interaction network allows the following speculations on autopoiesis due to Varela, Counthino and co-workers [39, 40]: the mutual interaction among lymphocytes rules out the need for an *a priori* learning for these cells, as tolerance to self may turn out to be an emerging property of the immune network thought of as a whole. In fact, it is the modulation and the mutual influence among interacting immunoglobulins (and their corresponding clones indirectly) that makes clones to be either in a quiescent or responsive state as a consequence of a given stimulus, which may be due to 'self' or 'non-self' agents. In other words, antibodies are randomly produced and, as a consequence, may react against anything (their idiotopes form somehow a 'base' in a proper space): however,

J. Stat. Mech. (2010) P07004

clones producing self-reacting antibodies are always taken to be quiescent, in such a way that they can produce only low—but not zero—concentrations of Igs [34] by the interaction with the network of all the others. Indeed, we stress that experimentally low doses of self antibodies are commonly found in healthy bodies [35, 33].

The last point we highlight is the so-called 'tolerance phenomenon': in a nutshell the immune system is not reactive when stimulated by a 'too low dose' of antigen, whose particular value strongly depends on the particular antigen. Furthermore, even too high a dose of antigen may yield tolerance as well.

3. Antibodies as vectors in the base of idiotopes

We model the B-core of the immune system as an imitative, eliciting system: stimulation is expressed by a firing lymphocyte towards its nearest neighbors while suppression is expressed by a quiescent one. Hence, in our scheme, it is not the sign of the coupling to establish the kind of interaction (either imitative or anti-imitative), which here is always positive or zero, but the state of the single lymphocyte itself.

We want to formalize this scenario within a statistical mechanics context where interacting antibodies ultimately reflect the interaction among lymphocytes due to the one-to-one postulate previously introduced: from a 'field theory language' [28], the antibodies are the 'fields' that the lymphocytes produce for interacting: these can interact both among themselves and directly with the lymphocytes. As the ratio among the amount of antibodies versus lymphocytes is much greater than 1 we focus primarily on the antibody–antibody interaction: how to extend this to the former case is straightforward as lymphocytes display antibodies on their external surface.

In order to get a network of Igs links, we relax the earlier simplifying assumption of 'a perfect mirror of a mirror' for the interacting Igs. In fact, we are going to consider interactions among antibodies formed by idiotopes such that the better the matches among idiotopes, the stronger the stimulus received by the respective clones via their immunoglobulins.

We consider the most generic antibody as a chain made up of the possible expression of L idiotopes. The assumption that each antibody can be thought of as a string of the same length is based on two observations: the molecular weight for each Igs is very accurately close to 15×10^4 and each idiotope on average is as large as each other [16].

Therefore, the elementary L idiotopes are

$$\begin{aligned} &\xi^1 = (1, 0, 0, \ldots, 0) \\ &\xi^2 = (0, 1, 0, \ldots, 0) \\ &\ldots \\ &\xi^L = (0, 0, 0, \ldots, 1), \end{aligned} \tag{1}$$

and they form an orthogonal base in the L-dimensional space of the antibodies Υ. A generic antibody ξ_i can then be decomposed as a linear combination of these eigenvectors $\xi_i = \lambda_i^1 \xi^1 + \lambda_i^2 \xi^2 + \cdots + \lambda_i^L \xi^L$, with $\lambda_i^\mu \in (0, 1)$ accounting for the expression (1) of a particular μth idiotope on the ith antibody or its lacking (0). For example, both the strings $(1001000), (1001001)$ are reactive with (0110110), but the second is better as it matches all the entries. As a counterpart the strings with several differences

J. Stat. Mech. (2010) P07004

in idiotope/epitope linking (i.e. 0111110 in the same example) do not match and the corresponding lymphocytes are disconnected in the network they belong to (it is straightforward to understand that there are no links inside the lymphocytes of the same clone, namely they act paramagnetically among each other). The fact that the interaction of two Igs is stronger when their relative strings are more complementary responds to the kind of interaction among their proteic structures: protein–protein interactions are dominated by weak, short-range non-covalent forces which arise when the geometry of the two proteins is complementary.

This naturally enlarges the idea of 'a mirror of a mirror' into an effective affinity matrix $J_{ij} \geq 0$, which, although described throughout in the next section, we use now as the starting point of the following speculation.

4. One-body and two-body Hamiltonian

The Hamiltonian H encodes the interactions among lymphocytes as well as the interactions among lymphocytes and the external antigens, providing a measure for the 'energy' of the system.

First of all, let us formalize the interactions taking place within the system. We consider an ensemble of M identical lymphocytes σ_i^α, $\alpha = 1, \ldots, M$, all belonging to the ith clone and N all different clones $i = 1, \ldots, N$. In principle M, the size of available lymphocytes exhibiting the same idiotypicity (in an healthy human body at rest), can depend on the clone itself. Anyhow, here, for the sake of simplicity, we are going to use the same M for all the clones.

If the match among antibodies had to be perfect for recognizing each other, then in order to reproduce all possible antibodies obtained by the L epitopes, the immune system would need $N \sim \mathcal{O}(2^L)$ lymphocytes. Conversely, if we relax the hypothesis of the perfect match, only a fraction of such a quantity is retained to manage the repertoire, such that we can define the following scaling among lymphocytes and antibodies:

$$N = f(L)\exp(\gamma L), \tag{2}$$

where $\gamma \in [0, 1]$ encodes for the ratio of the involved lymphocytes (the order of magnitude) and $f(L)$ is a generic rational monomial in L for the fine tuning (as often introduced in complex systems [37], we will see that $f(L) \sim \sqrt{L}$).

Interestingly, a far-from-complete system is consistent with the fact that binding between antigens and antibodies can occur even when the match is not perfect: experimental measurements showed that the affinity among antibody and anti-antibody is of the order of 65/70% or more (but strictly less than 100%) [29, 7, 9, 32]. Furthermore, the experimental existence of more than one antibody responding to a given stimulus (multiple attachment [10]) confirms this statement.

We can think of each lymphocyte as a binary variable $\sigma_i^\alpha = \pm 1$ (where i stands for the ith clone in some ordering and α for the generic element in the i subset) such that, when it assumes the value -1, it is quiescent (low level of antibody secretion) and when it is $+1$ it is firing (high level of antibody secretion).

The ability of newborn lymphocytes to spontaneously secrete a low dose of its antibody (corresponding to its genotype) even when not stimulated is fundamental in order to retain the network equilibrium and can be deepened in [25]. We stress once

J. Stat. Mech. (2010) P07004

again that within our approach the upper bound of the available firing lymphocytes is conserved $M \neq M(t)$ and the exponential growth of a clone i exposed to the external antigen is translated here in the evolution from a scenario where almost all its M are in the state $\sigma_i = -1$ to a scenario with all $\sigma_i = +1$, due to the effect of a positive local field acting on the ith subset.

To check immune responses we need to introduce the N order parameters m_i as local magnetizations:

$$m_i = \frac{1}{M} \sum_{\alpha=1}^{M} \sigma_i^{\alpha}, \tag{3}$$

where i labels the clone and α the lymphocyte inside the clone's family; the global magnetization is given by the average of all the m_i as $\langle m \rangle = N^{-1} \sum_i^N m_i$.

It is important to stress that the magnetizations, which play the role of the principal order parameters, account for the averaged concentration of firing lymphocytes into the immune network, such that, as $m_i \in [-1, 1]$, we can define the concentrations of the firing ith lymphocytes as

$$c_i \equiv \exp\left[\tau \frac{(m_i + 1)}{2}\right], \qquad \tau = \log M. \tag{4}$$

Note that, in general, the state of lymphocytes is not constant and, as a result, m_i and c_i are time-dependent. Moreover, the concentration is not normalized and ranges over several orders of magnitude, from $\mathcal{O}(10^0)$ when no firing lymphocyte is present up to $\mathcal{O}(10^{12}) \sim M$ when all the lymphocytes of the ith clone are firing. Strictly speaking, the quiescence of a given clone is a collective state where $\sim 10^2/10^3$ clones are present; this can be understood, within a thermodynamical framework, relaxing the idea that the system works at 'zero temperature' (that is not really physical), in fact, a small amount of noise would change the quiescent concentration from strictly 1 to a slightly higher value.

Now, let us turn to the external field and start with the ideal case of perfect coupling among a given antigen and its lymphocyte counterpart: let us associate the field $\mathbf{h^k}$ with the antigen displaying a sharp match with the kth antibody, hence described by the string $\bar{\xi}_k$ with entries $\bar{\xi}_k^{\mu} = 1 - \xi_k^{\mu}$. In general, the coupling with an arbitrary antibody i is h_i^k.

Following classical statistical mechanics [5, 18], the interaction among the two can be described as

$$H_1 = -\sum_{i=1}^{N} h_i^k m_i. \tag{5}$$

To take into account a 'network' of clones we should include their interaction term H_2, that is

$$H_2 = -\frac{1}{N} \sum_{i<j}^{N,N} J_{ij} m_i m_j, \tag{6}$$

where the prefactor $1/N$ accounts for the fact that the summation is performed on two variables and it correctly yields an extensive average of the energy.

Hence, with $J_{ij} > 0$, two generic clones i and j in mutual interaction, tend to imitate one another: if i is quiescent, it tries to make j quiescent as well—suppression—while if

J. Stat. Mech. (2010) P07004

the former is firing it tries to make the latter fire—stimulation—and symmetrically j acts on i.

The matrix J_{ij} encodes how the generic i and j elements are coupled together such that its high positive value stands for an high affinity among the two; conversely, a null value accounts for the missing interaction.

The complete Hamiltonian, taking into account both antigenic stimulus and interactions among lymphocytes, is $H = H_1 + H_2$.

If the affinity matrix is symmetric, so that detailed balance holds, the stochastic evolution of our immune model approaches the Maxwell–Boltzmann distribution, which determines the thermodynamic equilibria.

Finally, we notice that, when mapping from physics, we found implicitly paved the bridge with immunology; in fact, as suggested in [12], the two important 'thermodynamic observables' of the immune system are its *economy* and its *specificity*. Still following [12], if we assume that the immune system tries to maximize its specificity (entropy in our parallel) and to minimize its cost (energy in the same parallel) the way to statistical mechanics is naturally merged.

5. Topology and affinity pattern of the emergent network

We consider a system made of N idiotypically different clones, each denoted by an italic letter i and associated with a binary string ξ_i of length L encoding the specificity of the antibody produced. Each entry μ of the ith string is extracted randomly according to the discrete uniform distribution in such a way that $\xi_i^\mu = 1$ ($\xi_i^\mu = 0$) with probability $1/2$; this choice corresponds to a minimal assumption which can be possibly modified, yet preserving the structure of our model.

Now, given a couple of clones, say i and j, the μh entries of the corresponding strings are said to be complementary, iff $\xi_i^\mu \neq \xi_j^\mu$. Therefore, the number of complementary entries $\chi_{ij} \in [0, L]$ can be written as

$$\chi_{ij} = \sum_{\mu=1}^{L} [\xi_i^\mu (1 - \xi_j^\mu) + \xi_j^\mu (1 - \xi_i^\mu)] = \sum_{\mu=1}^{L} [\xi_i^\mu + \xi_j^\mu - 2\xi_i^\mu \xi_j^\mu]. \tag{7}$$

The affinity between two antibodies and, more generally, among two entities described by a vector in the idiotype basis, is expected to depend on how complementary their structures are (see also [4, 20]). In fact, the non-covalent forces acting among antibodies depend on the geometry, on the charge distribution and on hydrophilic–hydrophobic effects which give rise to an attractive (repulsive) interaction for any complementary (non-complementary) match. Consequently, in our model we assume that each complementary/non-complementary entry yields an attractive/repulsive contribution. In general, attractive and repulsive contributions can have different intensities and we quantify their ratio by a parameter $\alpha \in \mathbb{R}^+$. Hence, we introduce the functional $f_{\alpha,L} : \Upsilon \times \Upsilon \to \mathbb{R}$ as

$$f_{\alpha,L}(\xi_i, \xi_j) \equiv [\alpha \chi_{ij} - (L - \chi_{ij})], \tag{8}$$

which provides a measure of how 'affine' ξ_i and ξ_j are. In principle, $f_{\alpha,L}(\xi_i, \xi_j)$ can range from $-L$ (when $\xi_i = \xi_j$) to αL (when all entries are complementary, i.e. $\xi_i = \bar{\xi}_j$). Now,

when the repulsive contribution prevails, that is $f_{\alpha,L} < 0$, the two antibodies do not see each other and the coupling among the corresponding lymphocytes $J_{ij}(\alpha, L)$ is set equal to zero; conversely, we take $J_{ij}(\alpha, L) = \exp[f_{\alpha,L}(\xi_i, \xi_j)]/\langle \tilde{J} \rangle_{\alpha,L}$, $\langle \tilde{J} \rangle_{\alpha,L}$ being a proper normalizing factor.

Otherwise stated, nodes can interact pairwise according to a coupling $J_{ij}(\alpha, L)$, which is defined as

$$J_{ij}(\alpha, L) \equiv \Theta(f_{\alpha,L}(\xi_i, \xi_j)) \frac{\exp[f_{\alpha,L}(\xi_i, \xi_j)]}{\langle \tilde{J} \rangle_{\alpha,L}}, \tag{9}$$

where $\Theta(x)$ is the discrete Heaviside function returning x if $x > 0$ and 0 if $x \leq 0$.

Some remarks are in order here. The choice of an exponential law connecting the affinity $f_{\alpha,L}(\xi_i, \xi_j)$ between two strings and their relevant coupling J_{ij} follows empirical arguments. In fact, we expect the latter to depend sensitively on how complementary the two strings are, possibly spanning several orders of magnitude. Notice that this choice is also consistent with general experimental findings [41] and Parisi's intuition [31]. Moreover, the prefactor $1/\langle \tilde{J} \rangle_{\alpha,L}$ is taken in such a way that J_{ij} has[5] a finite (unitary) average for any value of α and L. More precisely, $\langle \tilde{J} \rangle_{\alpha,L}$ is just the average of $\tilde{J}_{ij}(\alpha, L) \equiv \Theta(f_{\alpha,L}(\xi_i, \xi_j)) \exp[f_{\alpha,L}(\xi_i, \xi_j)]$, calculated over all possible matchings between ξ_i and ξ_j.

This system can be envisaged by means of a graph $\mathcal{G}$, whose nodes represent lymphocytes and a link between them is drawn whenever the pertaining coupling is positive (see figure 1 for a sketch of the graph). The number of nearest neighbors of the generic site i, referred to as the coordination number or degree, can be recovered as a sum of adjacency matrix elements: $k_i = \sum_{j \in V} A_{ij}$.

In our model the graph describing the interaction among lymphocytes is a random graph where links are drawn with probability $p_{\alpha,L}$ which, in general, depend on the way strings ξ's are extracted and on the way affinity $f_{\alpha,L}$ is defined.

Here, due to the uniform distribution underlying the extraction of ξ's, we have that the probability that ξ_i^μ and ξ_j^μ are complementary equals 1/2 independently of i, j and μ. Therefore, the probability that they display χ_{ij} (hereafter simply χ) complementary entries follows a binomial distribution which is

$$\mathcal{P}(\chi) = \left(\frac{1}{2}\right)^L \binom{L}{\chi}. \tag{10}$$

Correspondingly, we have that lymphocytes i and j are connected together, namely that $f_{\alpha,L}(\xi_i, \xi_j) > 0$, when $\chi_{ij}(\alpha + 1) - L$ is positive (see equation (8)) and this occurs with probability

$$p_{\alpha,L} = \sum_{\chi = \lfloor L/(\alpha+1) \rfloor + 1}^{L} \mathcal{P}(\chi), \tag{11}$$

where $\lfloor x \rfloor = \max\{n \in N | n \leq x\}$.

[5] Henceforth we will drop the dependence on α and L, if not ambiguous.

J. Stat. Mech. (2010) P07004

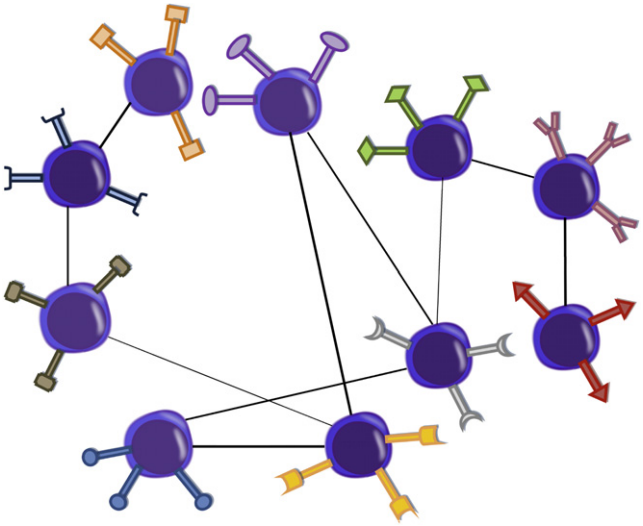

Figure 1. Representation of the idiotypic network. Each clone is represented by only one of its lymphocytes; the thickness of links denotes the strength of the corresponding coupling.

Now the emergent graph, denoted as $\mathcal{G}(N, p_{\alpha,L})$, is a weighted, random graph exhibiting non-trivial correlations among links. In fact, it is easy to see that, if i is connected to both j and k, it is rather unlikely that the latter are also connected together, that is, triangles correspond to frustrated configurations. Nonetheless, when α is large the requirements for two strings to be connected is weaker and the complementarity they need to display can be rather small. Under such conditions the link probability $p_{\alpha,L}$ (see equation (11)) can be assumed as independent of the chosen couple, so that the degree distribution for $\mathcal{G}(N, p_{\alpha,L})$ can be written as

$$P(k) = \binom{N}{k} p_{\alpha,L}^{k}(1 - p_{\alpha,L})^{N-k}, \tag{12}$$

representing the probability that a generic node has k nearest neighbors; the average degree follows as $\langle k \rangle = p_{\alpha,L}(N-1)$ or, more simply, for N large, we use $\langle k \rangle = p_{\alpha,L}N$. Notice that, consistent with the above-mentioned assumptions, a binomial degree distribution is typical of Erdös–Renyi random graphs [26]. In figure 2 we show some numerical results which corroborate the validity of equation (12) when α is close to 1.

As is well known, by increasing the link probability from 0 upwards, the (infinite) Erdös–Renyi random graph undergoes a percolation transition; namely there exists a critical link probability p_c such that, when the link probability starts to get larger than $p_c = 1/N$, a so-called 'giant component', displaying a size $\mathcal{O}(N)$, i.e. infinite in the thermodynamic limit, suddenly appears [26].

Analogously, for $\mathcal{G}(N, p)$ large values of $p_{\alpha,L}$ will give rise to a complete graph of N vertices, while small values of $p_{\alpha,L}$ will give rise to a graph disconnected in a set of components. Therefore, it is important to analyze in more detail the behavior of $p_{\alpha,L}$ as a function of α and L. For $\alpha = 1$ it is straightforward to see that $p_{1,L} = 1/2$, due to the symmetry of the distribution $\mathcal{P}(\chi)$ with respect to $\chi = L/2$.

J. Stat. Mech. (2010) P07004

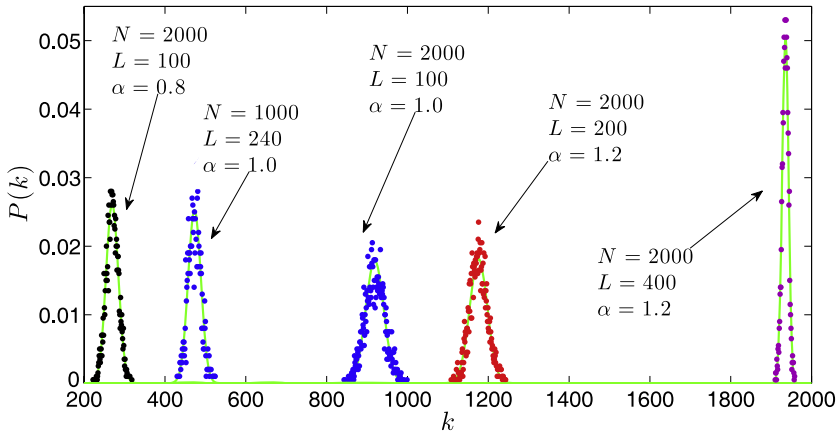

Figure 2. Degree distribution $P(k)$ for different values of N, L and α; data from numerical (•) and analytical estimates (green lines), see equation (12), are compared: the good agreement confirms the analytical derivation based on the lack of correlation among links.

More generally, for large L we can adopt a continuous description and write $p_{\alpha,L}$ (see equation (11)) as

$$p_{\alpha,L} \approx \int_{L/(\alpha+1)}^{L} \mathcal{P}(\chi)\,\mathrm{d}\chi \approx \int_{L/(\alpha+1)}^{L} \sqrt{\frac{2}{\pi L}}\,\mathrm{e}^{-(\chi-L/2)^2/(L/2)}\,\mathrm{d}\chi \tag{13}$$

$$= \frac{1}{2}\left[\mathrm{Erf}\left(\sqrt{\frac{L}{2}}\right) - \mathrm{Erf}\left(\frac{(1-\alpha)}{(1+\alpha)}\sqrt{\frac{L}{2}}\right)\right], \tag{14}$$

where we replaced the distribution $\mathcal{P}(\chi)$ with the normal distribution, having mean $L/2$ and variance $L/4$; in fact, for L large enough, the skew of the distribution $\mathcal{P}(\chi)$ is not too great and we can approximate the binomial distribution by the normal distribution. From equation (13) we can calculate the derivative of $p_{\alpha,L}$ with respect to α, which is

$$\frac{\partial p_{\alpha,L}}{\partial \alpha} \approx \sqrt{\frac{2L}{\pi}}\frac{1}{(1+\alpha)^2}\mathrm{e}^{-(L/2)\tilde{\alpha}^2}, \tag{15}$$

where we called $\tilde{\alpha} \equiv (1-\alpha)/(1+\alpha)$.

We now turn to the coupling strength introduced in equation (9) and we notice that we can write $\tilde{J}_{ij} = \exp[\chi_{ij}(\alpha+1) - L]$ whenever $\chi_{ij} > L/(\alpha+1)$, otherwise $\tilde{J}_{ij} = 0$. Hence, its mean value, averaged over all possible matchings between two binary strings, can be written as (see equations (9) and (11))

$$\langle \tilde{J} \rangle_{\alpha,L} \approx \int_{L/(\alpha+1)}^{L} \mathrm{e}^{\chi(\alpha+1)-L}\sqrt{\frac{2}{\pi L}}\,\mathrm{e}^{-(\chi-L/2)^2/(L/2)}\,\mathrm{d}\chi$$

$$= \frac{1}{2}\mathrm{e}^{(L/8)(\alpha^2+6\alpha-3)}\left\{\mathrm{Erf}\left[\frac{\alpha^2+4\alpha-1}{2(1+\alpha)}\sqrt{\frac{L}{2}}\right] + \mathrm{Erf}\left[\frac{1-\alpha}{2}\sqrt{\frac{L}{2}}\right]\right\}. \tag{16}$$

J. Stat. Mech. (2010) P07004

Now, we focus on the regime $L \gg 1$ and, according to the value of the (finite) parameter α, we distinguish among the following cases:

- $\alpha = 1$
  The expressions in equations (14)–(16) can be evaluated exactly obtaining, respectively:

$$p_{1,L} \approx \frac{1}{2}\mathrm{Erf}\left(\sqrt{\frac{L}{2}}\right) = \frac{1}{2}\left[1 - \mathcal{O}\left(\frac{\mathrm{e}^{-L/2}}{\sqrt{L}}\right)\right], \tag{17}$$

$$\frac{\partial p_{\alpha,L}}{\partial \alpha}\Big|_{\alpha=1} \approx \frac{1}{2}\sqrt{\frac{L}{2\pi}}, \tag{18}$$

  and

$$\langle \tilde{J}\rangle_{1,L} \approx \frac{1}{2}\,\mathrm{e}^{L/2}\,\mathrm{Erf}\left(\sqrt{\frac{L}{2}}\right) = \frac{1}{2}\mathrm{e}^{L/2}\left[1 - \mathcal{O}\left(\frac{\mathrm{e}^{-L/2}}{\sqrt{L}}\right)\right] \approx \mathrm{e}^{L/2} p_{\alpha,L}. \tag{19}$$

- $\alpha < 1$:

$$p_{\alpha,L} \approx \sqrt{\frac{1}{2\pi L}}\,\frac{1}{\tilde{\alpha}}\,\mathrm{e}^{-L/2\tilde{\alpha}^2}\left[1 + \mathcal{O}\left(\frac{1}{L}\right)\right], \tag{20}$$

  hence $p_{\alpha,L} \to 0$ as $L \to \infty$. Moreover, for $\alpha > -1 + \sqrt{2} \approx 0.41$:

$$\langle \tilde{J}\rangle_{\alpha,L} \approx \mathrm{e}^{(L/8)(\alpha^2+6\alpha-3)}\left[1 - \mathcal{O}\left(\frac{\mathrm{e}^{-L(1-\alpha)^2/8}}{\sqrt{L}}\right)\right], \tag{21}$$

  which is diverging for $\alpha > -3 + 2\sqrt{3} \approx 0.46$.
- $\alpha > 1$:

$$p_{\alpha,L} \approx 1 - \mathcal{O}\left(\frac{\mathrm{e}^{-(L/2)\tilde{\alpha}^2}}{\sqrt{L}}\right), \tag{22}$$

  hence $p_{\alpha,L} \to 1$ as $L \to \infty$. Moreover, analogously to the previous case:

$$\langle \tilde{J}\rangle_{\alpha,L} \approx \mathrm{e}^{(L/8)(\alpha^2+6\alpha-3)}\left[1 - \mathcal{O}\left(\frac{\mathrm{e}^{-L(1-\alpha)^2/8}}{\sqrt{L}}\right)\right]. \tag{23}$$

The asymptotic expressions above are all consistent with numerical results which indeed confirm the validity of the Gaussian approximation already for $L \sim 10^2$. Moreover, we notice that, in the limit $L \to \infty$, the link probability $p_{1,L}$ is a step function with diverging derivative in $\alpha = 1$ and $p_{\alpha,L} = 1$ for $\alpha > 1$, while $p_{\alpha,L} = 0$ for $0 \leq \alpha < 1$. In figure 3 we show the behavior of $p_{\alpha,L}$ as L and α are varied.

In order to characterize the dilution of the graph under study, a proper parameter is the average coordination number $\langle k\rangle = p_{\alpha,L}N$, which must be finite in order to have a well-defined thermodynamic limit for $N \to \infty$ [2, 15, 21]; all other cases would be either trivial ($\langle k\rangle \to 0$) or unphysical ($\langle k\rangle \to \infty$).

doi:10.1088/1742-5468/2010/07/P07004

J. Stat. Mech. (2010) P07004

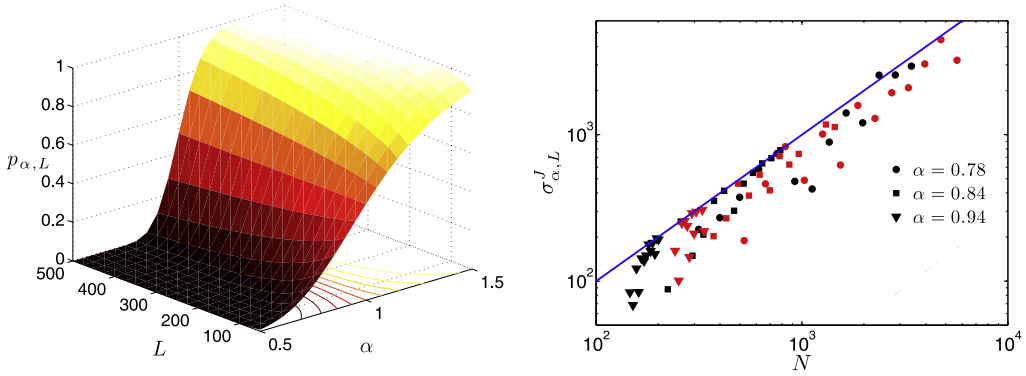

Figure 3. Left: link probability $p_{\alpha,L}$ for a system of $N = 4000$ lymphocytes as a function of α and L. Right: standard deviation $\sigma^J_{\alpha,L}$ as a function of the system size N; the parameter L is properly rescaled in order to keep $\langle k \rangle$ fixed and equal to 60 (red) or 100 (black); different symbols represent different choices for α, as shown by the legend. The straight line represents the first bisector.

Then, the number of nodes N follows as

$$N = \frac{\langle k \rangle}{p_{\alpha,L}} \approx \frac{2\langle k \rangle}{\mathrm{Erf}\left(\sqrt{\frac{L}{2}}\right) - \mathrm{Erf}\left(\tilde{\alpha}\sqrt{\frac{L}{2}}\right)}, \tag{24}$$

where we used equation (14). In particular, for $\alpha < 1$, one can write (see equation (20))

$$N \approx \sqrt{2\pi L}\, \langle k \rangle \tilde{\alpha}\ \exp\left(\frac{L}{2}\tilde{\alpha}^2\right), \tag{25}$$

which should be compared with equation (2), to get $f(L) \sim \sqrt{L}$.

The role of topology in the idiotypic network has been stressed by Varela *et al* [39], who showed that the larger the number of antibodies recognized by a given antibody ξ_i (i.e. the larger k_i), the lower the reactivity of the antibody itself, i.e. the greater their degree of tolerance. As a result, self/non-self discrimination turns out to be an emergent property of the immune network which is therefore able to organize the mature repertoire. In our model the adjacency matrix is actually weighted since links are endowed with a weight J_{ij}, so that we introduce an effective, or weighted, degree w_i as

$$w_i(\alpha, L, \langle k \rangle) \equiv \sum_{j=1}^{N} J_{ij}(\alpha, L). \tag{26}$$

Notice that the local quantity w_i provides finer information with respect to k_i, being directly connected with the 'internal' stimulus felt by lymphocyte i: recalling the Hamiltonian of equation (6), and assuming, for the sake of simplicity, the zero-noise limit so that all lymphocytes are quiescent ($m_j = -1, \forall j$), the local field acting on i is just $\varphi_i = -\sum_{j=1}^{N} J_{ij} m_j = w_i$.

For a given realization of the system $(\alpha, L, \langle k \rangle)$ the average weighted degree can be calculated as

$$\bar{w} = \frac{\sum_{i=1}^{N} w_i}{N} = \frac{\sum_{i=1}^{N}\sum_{j=1}^{N} J_{ij}}{N} = (N-1)\bar{J} \approx (N-1)\langle J \rangle_{\alpha,L}, \tag{27}$$

J. Stat. Mech. (2010) P07004

where the last approximation holds for large N and L; analogously, we can write $\langle w\rangle_{\alpha,L,\langle k\rangle} \approx \bar{w}$. Hence, being that $\langle J\rangle_{\alpha,L}$ equals 1 by definition, we have that $\langle w\rangle_{\alpha,L,\langle k\rangle}$ scales linearly with N.

The variance for J_{ij} can be estimated via equations (9) and (10):

$$\begin{aligned}\langle J^2\rangle_{\alpha,L} &= \frac{1}{\langle \tilde{J}\rangle^2_{\alpha,L}} \int_{L/(\alpha+1)}^{L} \exp\left[2\chi(\alpha+1)-2L\right]\sqrt{\frac{2}{\pi L}}\mathrm{e}^{-(\chi-L/2)^2/L/2}\,\mathrm{d}\chi \\ &= \frac{\mathrm{e}^{(L/2)(\alpha^2+4\alpha-1)}}{2\langle \tilde{J}\rangle^2_{\alpha,L}}\left[\mathrm{Erf}\left(\frac{\alpha(3+\alpha)}{1+\alpha}\sqrt{\frac{L}{2}}\right)-\mathrm{Erf}\left(\alpha\sqrt{\frac{L}{2}}\right)\right].\end{aligned}$$

Now, with some algebra and recalling equation (16), we get the following estimate:

$$\langle J^2\rangle_{\alpha,L} \approx 2\mathrm{e}^{(L/4)(\alpha+1)^2}\frac{\left[\mathrm{Erf}\left(\frac{\alpha(3+\alpha)}{1+\alpha}\sqrt{\frac{L}{2}}\right)-\mathrm{Erf}\left(\alpha\sqrt{\frac{L}{2}}\right)\right]}{\left[\mathrm{Erf}\left(\frac{\alpha^2+4\alpha-1}{2(1+\alpha)}\sqrt{\frac{L}{2}}\right)+\mathrm{Erf}\left(\frac{1-\alpha}{2}\sqrt{\frac{L}{2}}\right)\right]^2} \tag{28}$$

$$= \frac{1}{2}\mathrm{e}^{(L/4)(-\alpha^2+2\alpha+1)}\frac{1}{\alpha\sqrt{2\pi L}}\left[1-\mathcal{O}\left(\frac{1}{L}\right)\right]. \tag{29}$$

After noticing that, for $1-\sqrt{2} < \alpha < 1+\sqrt{2}$ the exponent is positive, yielding $\langle J^2\rangle_{\alpha,L} \gg \langle J\rangle^2_{\alpha,L} = 1$, we can write the standard deviation for the coupling strength as

$$\sigma^J_{\alpha,L} \approx \sqrt{\langle J^2\rangle_{\alpha,L}} \sim \frac{1}{\sqrt[4]{L}}\mathrm{e}^{(L/8)(-\alpha^2+2\alpha+1)}. \tag{30}$$

Moreover, due to the uncorrelatedness among J_{ij}'s, one can use Bienayme's theorem and write

$$\sigma^w_{\alpha,L} \approx \sqrt{N\langle J^2\rangle_{\alpha,L}} \approx \mathrm{e}^{(L/8)(-\alpha^4+6\alpha^2+3)/(1+\alpha)^2}, \tag{31}$$

where in the last expression we used equation (25).

In figure 3 we show, as a function of N and for several choices of α, the standard deviation $\sigma^J_{\alpha,L}$; notice that, while N is varied, L is properly scaled in order to keep $\langle k\rangle$ fixed. The log–log scale plot highlights a regime, for large enough N, where a power law growth for $\sigma^J_{\alpha,L}$ holds.

We conclude this section with a comment concerning the properties of small-length loops. Due to the lack of a perfect match among antibodies, a given lymphocyte, say σ_1, undergoing clonal expansion, may elicit one (or more) of the best Jerne counterparts (even spurious states may respond), say σ_2. The latter undergoing clonal expansion too may elicit another lymphocyte among the best Jerne spurious states, say σ_3, and so on in a cascade fashion. Now, since σ_1 and σ_3 both have large affinity, i.e. complementarity, with σ_2, they are expected to be similar. As a result, possible loops of length $l = 3$ are expected to be 'weak', corresponding to a frustrated configuration where at least one weak edge is present, while for $l = 4$ this kind of frustration can be avoided. In order to check the intrinsic robustness of a given circuit $\ell = \{i_1, i_2, \ldots, i_l\}$ of length l, we can introduce the overall strength as $J_\ell = l^{-1}\sum_{k=1}^{l} J_{i_k,i_{k+1}}$, with $i_{l+1} \equiv i_1$. The measurements of J_ℓ performed over several different realizations confirmed the previous remark: the average

overall strength for $l = 4$ turns out to be more than twice the one pertaining to $l = 3$; the former is larger than 1, while the latter is smaller than 1 (we recall that 1 is the expected coupling strength).

Moreover, while small loops are rather likely to occur, extensive (of the order of N) circuits (i.e. a so-called Jerne cascade) do not emerge; the reason is, at least, twofold: firstly, extensive loops would yield the activation of self lymphocytes which are particularly inhibited due to their large weighted connectivity, and secondly, each link implies a loss of information (due to the lack of a perfect match among immunoglobulins) in such a way that, after some iterations, no other (indirect) responses are possible. Interestingly, these points are intrinsic in our model and in agreement with experiments [9].

6. Self/non-self recognition

Let us consider again the whole idiotypic network made of N different clones, each characterized by a specific string of L idiotopes; once α is fixed, the affinity between two different nodes is specified by equation (8) from which the coupling in equation (9) follows.

Before turning to the analysis of the distribution $P(w)$ and showing how it naturally allows us to distinguish between self-and non-self-addressed antibodies, it is worth recalling the famous experiment led by Stewart, Varela and Coutinho [39, 40]: they measured the affinity of a collection of antibodies and analyzed the related affinity matrices, finding that these matrices are organized in blocks. More precisely, they distinguished a high affinity block, two blocks of groups which are mirrors of each other, and a low affinity remnant; then they showed that various groups play different roles: the mirror groups provide their model with various oscillation periods, while highly connected nodes maintain a 'basic network background level' and may be looked on as self-addressed antibodies. Indeed, their large connectivity prevents them from readily react to a stimulus. This point of view is extremely interesting as it outlines a natural interpretation of the topological properties of the immune network. Moreover, from an autopoietic point of view it also sheds light on the auto-immunity diseases: their origin would therefore lie on the 'inadequate connection' of self-reactive clones.

We considered different systems $(\alpha, L, \langle k \rangle)$ and by numerical analysis we derived the distribution $P(w)$, which, on a semilogarithmic scale, can be fitted by a Gaussian distribution; the relevant best fits are represented by the green curves in figure 4. Such distributions naturally outline three main groups characterized by high (right-hand-side tail), intermediate (central region) and low (left-hand-side tail) weighted degree, respectively. Hence, recalling that a large weight implies a low reactiveness, lymphocytes displaying low and high weighted degree can be labeled as non-self-and self-addressed lymphocytes, respectively. It is important to notice that, since the distribution $P(w)$ covers several orders of magnitude, the former will easily react even by low dose of affine agents, implicitly defining the low-dose tolerance; on the other hand, for larger w, the ability to react decreases progressively, up to prohibitive values of antigenic concentrations.

Now, starting from $P(w)$ we want to focus on couples of Ig and anti-Ig and figure out possible correlations in their weighted degrees. We first select non-self-addressed lymphocytes, namely nodes in the network corresponding to the left-hand side of the weighted degree distribution, and we look for their most tightly connected neighbors

J. Stat. Mech. (2010) P07004

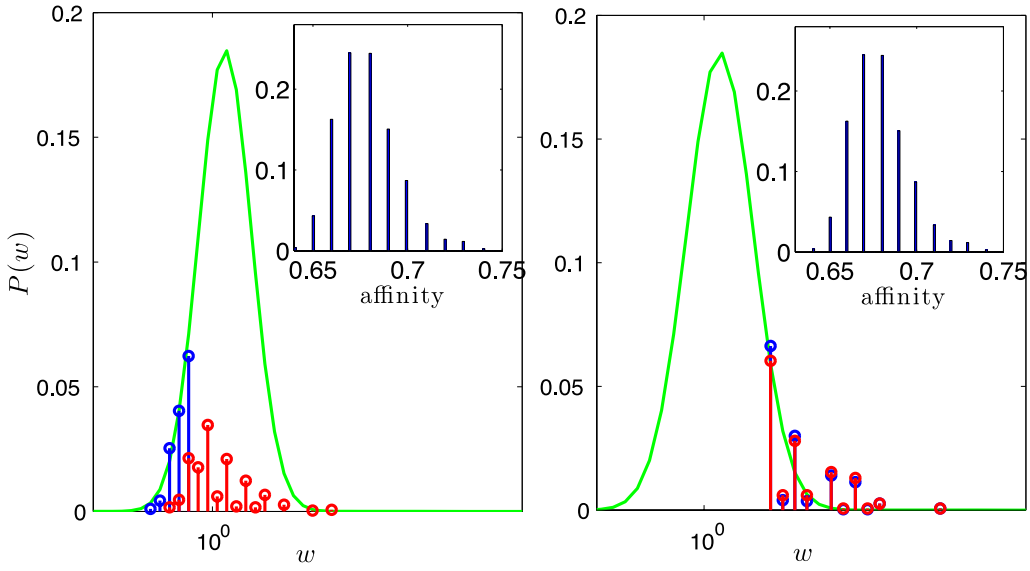

Figure 4. Weighted connectivity distributions (green line) for systems made up of $N = 4000$ lymphocytes characterized by idiotypic strings of length $L = 100$ and $\alpha = 0.78$; the average connectivity $\langle k \rangle = 290$. In blue we show the distribution of low-connected clones (left panel) or high-connected clones (right panel) chosen and in red the distribution of the pertaining 'anti-clones'. Notice the semilogarithmic scale plot. The insets show the relative matching χ_{ij}/L among all the couples Ig and anti-Ig detected in the whole system.

among the remaining $N - 1$. This way, we distinguish couples $(i, \bar{i})$, where i should be meant as a lymphocyte producing Ig directed against non-self agents and $\bar{i}$ as a lymphocyte producing so-called anti-Ig, able to respond to a significant growth in Ig concentration, according to Jerne's idea of an idiotypic network.

As shown in figure 4, when i belongs to the low weighted degree region, the corresponding $\bar{i}$ typically falls in the intermediate region of the distribution, hence fitting the 'mirror block'; this holds for several choices of α, L and $\langle k \rangle$ (see the left panel).

Let us now turn to the right-hand side of the weighted degree distribution and, analogously, we distinguish couples $(j, \bar{j})$, where j represents a lymphocyte producing Ig directed against self agents and $\bar{j}$ is the relevant anti-Ig (see the right panel). In this case anti-Ig still belongs to the highly connected group, that is they should as well be meant as directed to self agents. This result is easy to see: since by definition j exhibits a large weight w_j, it follows that, typically, $J_{j\bar{j}} \sim w_j$ as the main contribution to w_j comes from $J_{j\bar{j}}$ and, analogously, $J_{j\bar{j}} \sim w_{\bar{j}}$. Otherwise stated, when a highly connected node is selected, there exists a correlation between w_j and $w_{\bar{j}}$; on the other hand, when a lowly connected node is considered, the contribution of $J_{i\bar{i}}$ to w_i and $w_{\bar{i}}$, respectively, is small enough not to bias $w_{\bar{i}}$. The very origin of such a different behavior of self and non-self anti-Ig lies in the wide range spanned by w.

As considered further in the following, anti-Ig's play a crucial role in the establishment of memory effects, so that the 'mirror block' here acquires the fundamental function of memory storage. Interestingly, such a memory storage here turns out to be effectively

J. Stat. Mech. (2010) P07004

managed since it is restricted to non-self-directed Ig only. Conversely, self-directed Ig and relevant anti-Ig both set up the highly connected group.

7. Low-dose tolerance

In this section we want to investigate the effects elicited by a concentration c of a given antigen. Let us consider the antigen with specificity $\bar{\xi}_k$, namely displaying a perfect match with antibody ξ_k. Therefore we can rewrite the Hamiltonian (2.7) as

$$H(\sigma, \mathbf{J}, \mathbf{h}) = -\frac{1}{N}\sum_{i<j}^{N,N} J_{ij} m_i m_j - \sum_{i=1}^{N} c_k h_i^k m_i, \tag{32}$$

where c_k can be possibly tuned to mimic variations in the antigen concentration. This way, an arbitrary lymphocyte i is subject to two stimuli, one deriving from the presence of the antigen and the other from the presence of the remaining lymphocytes. This can be formalized by saying that the field acting on the ith node is

$$\varphi_i = -\frac{1}{N}\sum_{j=1}^{N} J_{ij} m_j + c_k h_i^k. \tag{33}$$

It is worth underlining that the coupling between an antigen, say $\bar{\xi}_k$, and a clone, say ξ_i, follows the same rule as for the coupling between immunoglobulins (see equation (9)), that is h_i^k is proportional to $\exp(f_{\alpha,L}(\xi_i, \bar{\xi}_k))$ whenever the affinity is positive. As a result, the presence of an antigen is 'felt' differently by clones, according to their reciprocal affinities.

In the absence of any antigen it is reasonable to consider all lymphocytes in a quiescent state, i.e. $m_j = -1$ (under the assumption of negligible noise) for any j; this provides the initial state assumed to be stationary when no antigen is at work. Hence, as the field $\mathbf{h^k}$ is switched on, we have

$$\varphi_i = \frac{1}{N} w_i - c h_i^k, \tag{34}$$

where we used equation (26). Now, if φ_i is negative, the state for the ith lymphocyte which minimizes the energy is the firing one, namely $m_i = +1$. Hence, assuming that all lymphocytes are quiescent, the condition for lymphocyte i to fire is

$$c_k h_i^k > \frac{w_i}{N}. \tag{35}$$

This means that the minimal concentration necessary in order to elicit an immune response by i is directly proportional to its degree w_i and inversely proportional to its coupling h_i^k. This also suggests that, in the presence of the antigen $\bar{\xi}_k$, the most reactive idiotype is not necessarily ξ_i, but rather it may be a spurious one which exhibits the lowest ratio w_i/h_i^k. Interestingly, the threshold mechanism determined by equation (35) consistently mimics the low-dose tolerance phenomenon: the immune system attacks antigens or, more generally, proteins, if their concentration is larger than a minimum value, which depends on the particular protein.

Implicitly this mechanism suggests a possible interpretation even of the high-dose tolerance: as what elicits a given lymphocyte is the product of the weighted connectivity

J. Stat. Mech. (2010) P07004

with another agent (antigen or internal molecules) times its concentration (properly expressed via a magnetization function), it cannot distinguish among self or antigen in an high dose. Responding to a high dose of antigen should, in principle, allow a response even to the self, whose defense turns out to be a primary goal.

These analytical estimates have been checked by means of numerical simulations: for a given system $(\alpha, L, \langle k \rangle)$ we run several experiments, each for a different applied field $\bar{\xi}_i$, where the concentration of the antigen is tuned from 0 up to the minimal value $\tilde{c}_i$ necessary to elicit an immune response of i; data are reported in figure 6. On the x axis we set the weighted degree of the first reactive lymphocyte (typically the ith one) and on the y axis we set the minimal concentration $\tilde{c}_i$, multiplied by $h_i^i = \exp(\alpha L)$; a linear dependence between w and $\tilde{c}$ is evidenced by the fit, in agreement with equation (35). We therefore recover the important result from Varela *et al* [39, 40] that the reactivity of an antibody is closely related to its degree, where, here, the degree is more specifically meant as the weighted degree.

The linear scale between $\tilde{c}$ and w has some significant consequences: the tolerated concentration of antigen directly reflects the (weighted) inhomogeneity of the graph. Otherwise stated, if the weighted degree spans a range, say $\mathcal{O}(10^k)$, the tolerated concentration relevant to all lymphocytes making up the system is expected to span an analogously wide range. Hence, recalling that non-self-addressed Ig's belong to the left tail of the weighted degree distribution, while self-addressed Ig's lie on the right tail, we have that the doses typically tolerated by the former are of k orders of magnitude less that those tolerated by the latter. Now, as previously shown, the weighted degree distribution $P(w)$ exhibits a standard deviation scaling exponentially with L or, analogously, algebraically with N, being $1/2$ a lower bound for the exponent. As a consequence, we expect that the region spanned by w grows not slower than $\sqrt{N}$; this means that for real systems the difference between doses tolerated by self and non-self is at least $\mathcal{O}(10^7)$.

We finally stress that the low-dose tolerance emerges as a genuine collective effect directly related to the properties of the idiotypic network and, in particular, on the distribution of the weighted degree.

8. Multiple responses and spurious states

As explained in section 7, the introduction of a concentration c_i of a given antigen described by the external field $\mathbf{h^i}$ is able to increase the magnetization (concentration) of the node (lymphocyte) i, provided that c is sufficiently high. In general, if the concentration is large enough, several clones, different from i, may prompt a response: some of them, say $j_1, \ldots, j_p$ (hereafter called spurious, once again in order to stress similarities with neural networks), respond because of a non-null, though small, coupling with the external field. Some others, say $j'_1, \ldots, j'_{p'}$ (hereafter called Jerne states for consistency), respond because they display a strong interaction $J_{jj'}$ with the former or with the specific Ig i. Spurious states can be very numerous according to the particular antigen considered and to its concentration; under proper conditions the response of spurious states can be even more intensive than the specific response from i. In fact, the reactivity of a given node j is determined not only by the relevant antigenic stimulus h_j^i, but also by its local environment, namely by the concentration of firing lymphocytes to which j is connected.

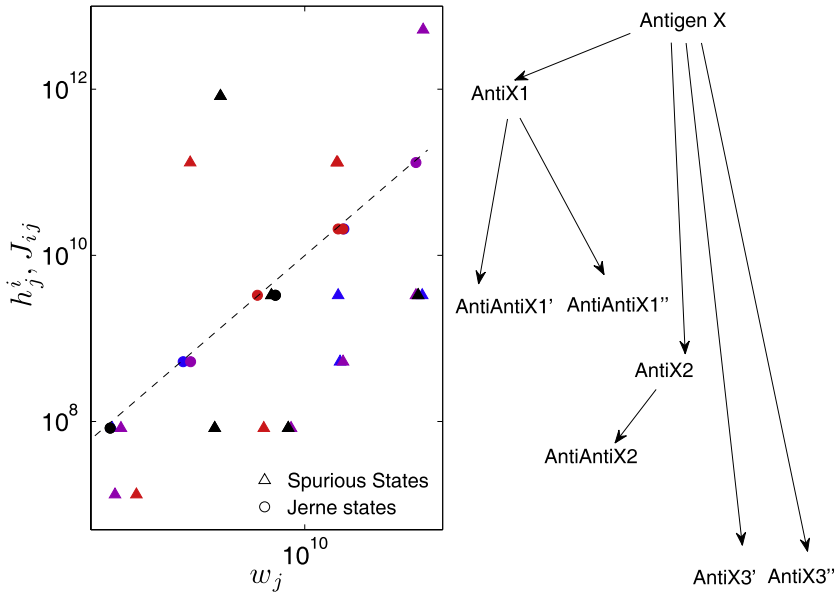

Figure 5. Analysis of spurious states for different realizations of a system made up of $N = 1000$ families and $M = 10$ clones per family; we assumed $\alpha = 0.84$ and $L = 140$. Left panel: field coupling h_j^i (for reactive spurious states (triangles)) and interaction strength J_{ij} (for reactive anti-antibodies (circles)) as a function of their weighted connectivity; each realization is depicted in a different color. Right panel: schematic representation of connections between reactive antibodies.

While in a neural network framework spurious states correspond to 'errors' during the retrieval (once a stimulus is presented) and should be avoided, in an immune network spurious states are fundamental to an effective functioning of the whole machinery. In fact, when an antigen is introduced into the body, all the set of responders (proper lymphocyte and spurious states) do contribute to attack the enemy and neutralize it. Moreover, the reaction of spurious states can, in turn, have important consequences on the generation of memory cells and on the effectiveness of the secondary response. Accordingly, one can investigate whether it is possible to figure out proper strategies which can limit or increase the number of such spurious states.

Here, we just want to analyze the overall response of the system, outlining which kind of clone does react, that is, we distinguish between spurious states and Jerne states. Results for different realizations of a system $(\alpha, L, \langle k \rangle)$, where the antigenic concentration is set as $c \sim 10^2 \tilde{c}$, are shown in figure 5; different symbols are used for spurious states (triangles) and for Jerne states (circles). For such concentrations the number of reactive spurious states is approximately twice the number of reactive Jerne states and a clear correlation between their weighted connectivity $J_{j'i}/J_{jj'}$ is also shown. Interestingly, Jerne states require a larger stimulus to react and this is due to the fact that the reaction is not directed, but rather mediated by the specific Ig i or by a spurious state j.

Furthermore it is also important to stress that apparently, without the introduction of spurious states, the amount of antibodies is greater than the amount of lymphocytes and this would be in conflict with the first postulate of immunology: consistency is obtained thanks to the lack of a perfect match among antibodies (or antibody and antigen), which

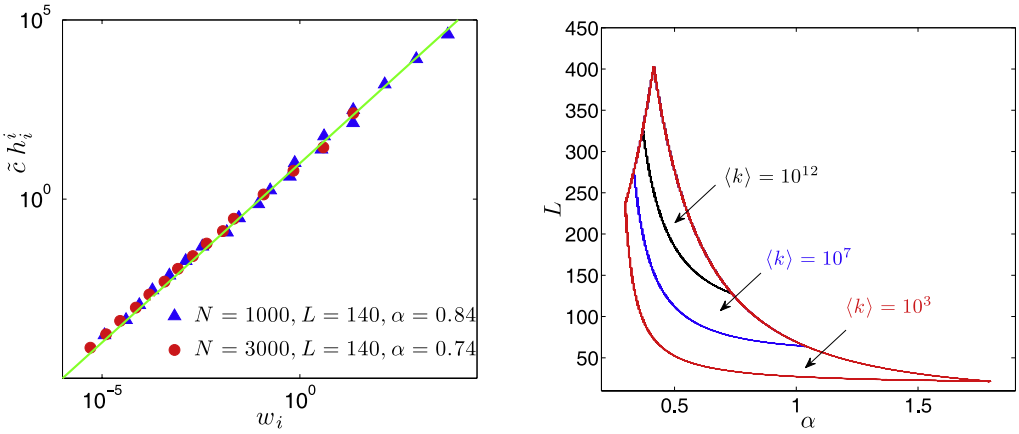

Figure 6. Left: threshold concentration $\tilde{c}$ of antigens ξ_{h^i} as a function of the weighted connectivity w_i of the first reacting family, in systems with parameters $N = 1000, \alpha = 0.84, \langle k \rangle \sim 150$ (triangles) and $N = 3000, \alpha = 0.75, \langle k \rangle \sim 40$ (circles), respectively. For both cases we used $M = 100, L = 140$ and a level of noise much below the critical value. Right: numerical solution of equation (39): different values of $\langle k \rangle$ correspond to different closed curves, whose internal region provides the values of α and L which satisfy the inequality of equation (39), namely the region of retrieval.

allows multiple attachments, ultimately accounting for a large over-counting of different responses.

9. Dynamical memory

As is well known, the immune system is able to develop memory effects; for this to happen the mutual interaction among lymphocytes is crucial. The activation of a lymphocyte i must therefore be followed by the activation of the relevant anti-antibody $\bar{i}$, which, in turn, may elicit the anti-anti-antibody $\bar{\bar{i}}$ and so on in a cascade fashion. It is just the modulation and mutual influence among such interacting antibodies that keeps the concentration of antibodies themselves at appropriate levels, which provides memory storage within the system. In this section we want to analyze under which conditions, if any, a full firing state of lymphocyte i, i.e. $m_i = 1$, can determine a non-null concentration for the anti-antibody $\bar{i}$ to react.

Let us consider a system characterized by parameters α, L and $\langle k \rangle$ in such a way that N is determined by equation (24). Assuming that $m_i = 1$ and $m_k = -1$ for any $k \neq i$, we have that, all in all, node $\bar{i}$ is subjected to a field $\varphi_{\bar{i}}$ given by the presence of the other $N - 1$ families:

$$\varphi_{\bar{i}} = -\frac{1}{N}\left(\sum_{j \neq \bar{i}} J_{ij} m_j + J_{i\bar{i}} m_i\right) = \frac{1}{N}(w_{\bar{i}} - 2J_{i\bar{i}}), \tag{36}$$

where we used $\sum_i J_{ij} = w_i$. We therefore derive that $\bar{i}$ is also firing if

$$w_{\bar{i}} < 2J_{i\bar{i}}. \tag{37}$$

J. Stat. Mech. (2010) P07004

As shown in section 6, anti-antibodies corresponding to non-self-addressed Ig typically belong to the so-called mirror block of the affinity matrix and they display an average connectivity $w_{\bar{i}} \approx \langle J \rangle_{\alpha,L}(N-1)$. Moreover, the affinity between i and $\bar{i}$ can be estimated as $J_{i\bar{i}} \approx \langle J \rangle_{\alpha,L} + 2\sigma^J_{\alpha,L}$, since the coupling between Ig and anti-Ig lies on the right tail of the coupling distribution. Therefore, recalling $\langle J \rangle_{\alpha,L} = 1$, we can rewrite equation (37) as

$$N - 1 < 2(1 + 2\sigma^J_{\alpha,L}). \tag{38}$$

Now, we can use equations (24) and (28) to write the previous expression as a function of L, α and $\langle k \rangle$:

$$\frac{2\langle k \rangle}{\mathrm{Erf}\left(\sqrt{\frac{L}{2}}\right) - \mathrm{Erf}\left(\tilde{\alpha}\sqrt{\frac{L}{2}}\right)} < 3 + 4\sqrt{2\mathrm{e}^{(L/4)(\alpha+1)^2}} \times \frac{\sqrt{\mathrm{Erf}\left(\frac{\alpha(3+\alpha)}{1+\alpha}\sqrt{\frac{L}{2}}\right) - \mathrm{Erf}\left(\alpha\sqrt{\frac{L}{2}}\right)}}{\mathrm{Erf}\left(\frac{\alpha^2+4\alpha-1}{2(1+\alpha)}\sqrt{\frac{L}{2}}\right) + \mathrm{Erf}\left(\frac{1-\alpha}{2}\sqrt{\frac{L}{2}}\right)}. \tag{39}$$

A better insight into the previous expression can be achieved from figure 6 (right panel) which shows its numerical solution for different values of $\langle k \rangle$ (each depicted in a different color): for a given average degree $\langle k \rangle$, the region of the plan (α, L) contained within the pertaining curve satisfies equation (39). For instance, let us assume $\langle k \rangle = 10^{12}$ and $L = 140$, then α must not be larger than approximately 0.75 if we want that a response from the anti-anti-Ig follows the reaction of a specific Ig. By analyzing figure 6, we notice that the less diluted the network, the smaller the region of 'retrieval'. Indeed, if we fix a given point on the (α, L) plane, increasing $\langle k \rangle$ implies a larger N and this contrasts with the satisfiability of equation (39). On the other hand, this result is rather intuitive as, when the coordination is large, the anti-antibody is less reactive with respect to the stimulus. As for the region corresponding to large L and relatively large α, this never overlaps with the retrieval region. In fact, for those values $\sigma^J_{\alpha,L}$ is relatively small.

10. Summary

In this paper we pioneered an alternative way for theoretical immunology by plugging it into a well-defined disordered statistical mechanics framework: our work is not meant as an exhaustive picture of the (adaptive response of the) immune system, but rather as a starting point in modeling its universal features by means of this technique.

We stress that, in our model, once the amount of available epitopes and their distributions is given (namely the amount of ξ's together with their distribution), everything can be worked out. In particular, in the complex system framework we developed, the immune network naturally achieves/recovers, qualitatively and partially quantitatively, the following properties:

- The Burnet clonal expansion theory appears as the standard one-body response of the system.
- The multi-attachment among antibodies is a natural property of the system and gives rise to the Jerne network.

J. Stat. Mech. (2010) P07004

- The Jerne antibody network, which is obtained as a random graph, encodes dynamically the memory of the encountered antigens.
- The Varela–Counthino self/non-self distinction appears as an emerging property of such a network.
- The low-dose tolerance is the inertia of the network when subjected to a response to external fields.
- The existence of several antibodies acting against a given antigen play the role of dynamical spurious states, generalizing the neural network static counterpart.
- The high-dose tolerance appears as a mechanism avoiding the breaking of self recognition.
- Increasing the noise, both the quality and the quantity of the available retrievals decrease.

All these different aspects of the immune system appear as features of the very same unified theory, which relies on simple, minimal assumptions. Furthermore, a quantitative agreement with experimental data also holds: for instance, the average connectivity of the network as well as the reciprocal affinities of the cascade of complementary antibodies are in good agreement with experiments.

With purely physical eyes our model describes the equilibrium of the immune system as a (thermodynamically) symmetry broken random-bond diluted ferromagnet. However, its non-equilibrium states (when antigens are present) map the latter into a random-field random-bond diluted model, conferring to the system a glassy flavor.

Among the several outlooks, surely the out-of-equilibrium thermodynamics has to be investigated as the model is shown to display a very rich ensemble of timescales and aging is expected. Another important point is its learning, which would merge the approach of neural networks [13] and dynamical graph theory with information theory. The extension of the concept of Hopfield statistical memories into a dynamical counterpart should be deepened as well as the Gardner saturation bound [3], which may play a key role in the breaking of defenses in the body. The transition from a 'simple' system to a 'spin glass' due to the increase of pasted random fields also needs a deep analysis as it is concerned with the genesis of auto-immune responses. The interplay among B-cells and T helper cells should also be taken into account as T helpers play the role of a spin glass self-regulation, adding a considerable amount of complex self-regulation. In the end, as our results are qualitatively quite robust and the framework very stable under the change in the epitope distributions, other, mathematically challenging (i.e. due to correlations), choices for such distributions are surely biologically plausible and should be investigated. We plan to report soon on several of the outlined directions of research.

Acknowledgments

This work was supported by FIRB-MIUR project no. RBFR08EKEV, EDENET onlus. The authors are grateful to R Burioni, M Casartelli, P Contucci, A C C Coolen, S Franz, F Guerra, G Ruocco, N Shayeghi and R Silvestri for useful discussions.

References

[1] Abbas A K, Lichtman A H and Pober J S, 2007 *Cellular and Molecular Immunology* (Amsterdam: Elsevier)
[2] Agliari E, Barra A and Camboni F, *Criticality in diluted ferromagnets*, 2008 *J. Stat. Mech.* P10003
[3] Amit D J, 1992 *Modeling Brain Function: The World of Attractor Neural Network* (Cambridge: Cambridge University Press)
[4] Bagley R J, Farmer J D, Packard N H, Perelson A S and Stadnyk I M, *Modeling adaptive biological systems*, 1989 *Biosystems* **23** 113
[5] Barra A, *The mean field Ising model through interpolating techniques*, 2008 *J. Stat. Phys.* **132** 787–809
[6] Barra A, Camboni F and Contucci P, *Structural stability of dilution in mean field models*, 2009 *J. Stat. Mech.* P03028
[7] Burnet F M, 1959 *The Clonal Selection Theory of Acquired Immunity* (Nashville: Vanderbilt University Press)
[8] Burnet F M and Fenner F, 1949 *The Production of Antibodies* (Melbourne: MacMillan)
[9] Cazenave P A, *Idiotypic–anti-idiotypic regulation of antibody synthesis in rabbits*, 1977 *Proc. Nat. Acad. Sci.* **74** 5122
[10] Celada F, *In search of T-cell help for the internal image of the antigen*, 1988 *Theoretical Immunology, Part 2, SFI Studies in the Sciences of Complexity* ed A S Perelson (Redwood City, CA: Addison-Wesley) pp 233–45
[11] Chun J, *Selected comparison of immune and nervous system development*, 2001 *Adv. Immunol.* **77** 297
[12] Colman P M, *Structure of antibody–antigen complexes: implication for immune recognition*, 1988 *Adv. Immunol.* **43** 99–132
[13] Coolen A C C, Kuehn R and Sollich P, 2005 *Theory of Neural Information Processing Systems* (Oxford: Oxford University Press)
[14] Delves P, Martin S, Burton D and Roitt I, 2006 *Roitt's Essential Immunology* (New York: Wiley)
[15] DeSanctis L and Guerra F, *Mean field dilute ferromagnet: high temperature and zero temperature behaviour*, 2008 *J. Stat. Phys.* **132** 759
[16] Dreyer W J and Bennett J C, 1965 *Molecular Bases of Antibody Formation: A Paradox* (*Nobel Prize Lecture*)
[17] Edelman G M, *The problem of molecular recognition by a selective system*, 1974 *Studies in the Philosophy of Biology* ed F J Ayala and T Dobzhansky (London: Macmillan) pp 45–56
[18] Ellis R S, 1985 *Large Deviations and Statistical Mechanics* (New York: Springer)
[19] Elrich P, 1910 *Studies in Immunity* (New York: Wiley)
[20] Farmer J D, Lapedes A, Packard N H and Wendross B, 1987 *Evolution, Games and Learning* (Amsterdam: North-Holland)
[21] Guerra F and Toninelli F L, *The high temperature region of the Viana–Bray diluted spin glass model*, 2004 *J. Stat. Phys.* **115** 531–55
[22] Hatchett J P L, Pérez Castillo I, Coolen A C C and Skantzos N S, *Dynamical replica analysis of disordered Ising spin systems on finitely connected random graphs*, 2005 *Phys. Rev. Lett.* **95** 117204
[23] Jacobs H and Bross L, *Towards an understanding of somatic hypermutation*, 2001 *Curr. Opin. Immunol.* **13** 208
[24] Jerne N K, *Towards a network theory of the immune system*, 1974 *Collect. Ann. Inst. Pasteur* **125C** 373–89
[25] Ledemberg J, 1959 *Genes and Antibodies* (*Nobel Prize Lecture*)
[26] Lemke H and Lange H, *Generalization of single immunological experiences by idiotypically mediated clonal connections*, 2002 *Adv. Immunol.* **80** 203–41
[27] Lundkvist I, Coutinho A, Varela F and Holmberg D, *Evidence for a functional idiotypic network among natural antibodies in normal mice*, 1989 *Proc. Nat. Acad. Sci.* **86** 5074
[28] Mandl F and Shaw G, 1984 *Quantum Field Theory* (New York: Wiley)
[29] Nesterenko V G, *Symmetry and asymmetry in the immune network*, 1988 *Theoretical Immunology* ed A S Perelson (New York: Addison-Wesley) pp 321–44
[30] Newman M, Watts D and Barabasi A-L, 2006 *The Structure and Dynamics of Networks* (Princeton, NJ: Princeton University Press)
[31] Parisi G, *A simple model for the immune network*, 1990 *Proc. Nat. Acad. Sci.* **87** 29–433
[32] Paul W E and Bona C, *Regulatory idiotopes and immune networks: a hypothesis*, 1982 *Immunol. Today* **3** 9
[33] Pereira P, Forni L, Larsson E L, Cooper M, Heusser C and Coutinho A, *Autonomous activation of B and T cells in antigen free mice*, 1986 *Eur. J. Immunol.* **16** 685–8
[34] Rose N R and Mackay I R, 2006 *The Autoimmune Diseases* (Amsterdam: Elsevier)

J. Stat. Mech. (2010) P07004

[35] Seigneurin J M, Guilbert B, Bourgeat M J and Avrameas S, *Polyspecific natural antibodies and autoantibodies secreted by human lymphocytes immortalized with Epstein–Barr virus*, 1988 *Blood* **71** 581
[36] Semerjian G and Weigt M, *Approximation schemes for the dynamics of diluted spin models: the Ising ferromagnet on a Bethe lattice*, 2004 *J. Phys. A: Math. Gen.* **37** 5525
[37] Sollich P, *Soft glassy rheology*, 1997 *Phys. Rev. Lett.* **78** 2020
[38] Storb U, *The molecular basis of somatic hypermutation of immunoglobulin genes*, 1996 *Curr. Opin. Immunol.* **8** 206
[39] Stewart J, Varela F J and Coutinho A, *The relationship between connectivity and tolerance as revealed by computer simulation of the immune network: some lessons for an understanding of autoimmunity*, 1989 *J. Autoimmun.* **2** 15–23
[40] Varela F J and Countinho A, *Second generation immune networks*, 1991 *Immunol. Today* **12** 159
[41] von Hippel P H and Berg O G, *On the specificity of DNA–protein interactions*, 1986 *Proc. Nat. Acad. Sci.* **83** 1608

J. Stat. Mech. (2010) P07004