INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND GENERAL

J. Phys. A: Math. Gen. **35** (2002) 2761–2809 PII: S0305-4470(02)28758-2

Hierarchical self-programming in recurrent neural networks

T Uezu[1] **and A C C Coolen**

Department of Mathematics, King's College London, The Strand, London WC2R 2LS, UK

E-mail: uezu@ki-rin.phys.nara-wu.ac.jp and tcoolen@mth.kcl.ac.uk

Received 6 September 2001, in final form 10 December 2001
Published 15 March 2002
Online at stacks.iop.org/JPhysA/35/2761

Abstract

We study self-programming in recurrent neural networks where both neurons (the 'processors') and synaptic interactions ('the programme') evolve in time simultaneously, according to specific coupled stochastic equations. The interactions are divided into a hierarchy of L groups with adiabatically separated and monotonically increasing time-scales, representing sub-routines of the system programme of decreasing volatility. We solve this model in equilibrium, assuming ergodicity at every level, and find as our replica-symmetric solution a formalism with a structure similar but not identical to Parisi's L-step replica symmetry breaking scheme. Apart from differences in details of the equations (due to the fact that here interactions, rather than spins, are grouped into clusters with different time-scales), in the present model the block sizes m_i of the emerging ultrametric solution are not restricted to the interval [0, 1], but are independent control parameters, defined in terms of the noise strengths of the various levels in the hierarchy, which can take any value in $[0, \infty\rangle$. This is shown to lead to extremely rich phase diagrams, with an abundance of first-order transitions especially when the level of stochasticity in the interaction dynamics is chosen to be low.

PACS numbers: 87.10.+e, 05.20.−y, 84.35.+i

1. Introduction

In this paper we study recurrent networks of binary neuronal state variables, represented as Ising spins, with symmetric couplings (or synaptic interactions) J_{ij}, taken to be of infinite range. In contrast to most standard neural network models, not only the neuron states but also the interactions are allowed to evolve in time (simultaneously), driven by correlations in

[1] On leave from: Graduate School of Human Culture, Nara Women's University, Kitauoyanishimachi, Nara City, 630-8506, Japan.

the states of the neurons (albeit slowly compared to the dynamics of the latter), reflecting the effect of 'learning' or 'long-term potentiation' in real nervous tissue. Since the interactions represent the 'programme' of the system, and since the slow interaction dynamics are driven by the states of the neurons (the 'processors'), such models can be regarded as describing self-programming information-processing systems, which can be expected to exhibit highly complex dynamical behaviour.

The first papers in which self-programming recurrent neural networks were studied appear to be [1, 2]. In the language of self-programming systems one could say that these authors were mostly concerned with the stability properties of embedded 'programmes' (usually taken to be those implementing content-addressable or associative memories). In both [1, 2], the programme dynamics, i.e. that of the $\{J_{ij}\}$, was defined to be adiabatically slow compared to the neuronal dynamics, and fully deterministic. However, the authors already made the important observation that the natural type of (deterministic) programme dynamics (from a biological point of view), so-called Hebbian learning, could be written as a gradient descent of the interactions $\{J_{ij}\}$ on the free energy surface of a symmetric recurrent neural network equipped with these interactions.

In order to study more generally the potential of such self-programming systems, several authors (simultaneously and independently) took the natural next step [3–8]: they generalized the interaction dynamics by adding Gaussian white noise to the deterministic laws, converting the process into one described by conservative Langevin equations, and were thus able to set up an equilibrium statistical mechanics of the self-programming process. This was (surprisingly) found to take the form of a replica theory with finite replica dimension, whose value was given by the ratio of the noise levels in the neuronal dynamics and the interaction dynamics, respectively. Furthermore, adding explicit quenched disorder to the problem in the form of additional random (but frozen) forces in the interaction dynamics led to theories with two nested levels of replicas, one representing the disorder (with zero replica dimension) and the other representing the adiabatically slow dynamics of the interactions [6, 9, 10] (with non-zero replica dimension). The structure of these latter theories was found to be more or less identical to those of ordinary disordered spin systems such as the SK model [11], with fixed interactions but quenched disorder, when described by replica theories with one step replica symmetry breaking (RSB) [12]. The only (yet crucial) difference was that in ordinary disordered spin systems the size m of the level-1 block in the Parisi solution is determined by extremization of the free energy, which forces m to lie in the interval [0, 1] (see also [13]), whereas in the self-programming neural networks of [6, 9, 10] m was an independent control parameters, given by the ratio of two temperatures, which can take any non-zero value. As a consequence one can observe in the latter systems, for sufficiently large values of such dimensions, much more complicated scenarios of (generally discontinuous) phase transitions.

Two further classes of neural network theory studies deserve mentioning at this stage (apart from those where a coupled dynamics of fast neurons and slow synapses is initially defined, but where in working out equations the authors sever the two-way link between the two after all, such as [14, 15]). The first class consists of studies aimed at modelling networks with increased biological realism, including the effects of, e.g., synaptic exhaustion during operation; here one is forced either to resort fully to simulations [16] or to restrict one's analysis to the behaviour of special solutions [17, 18]. The second class, closer in spirit to [3–8] and the present study, aims at solving the coupled dynamics of neurons and synapses in extremely diluted networks, exploiting the simplifications resulting from having Gaussian distributed local fields [19, 20].

In contrast to the previously studied models involving coupled dynamics of fast neurons and slow interactions, in this paper we study systems in which the interactions do not evolve

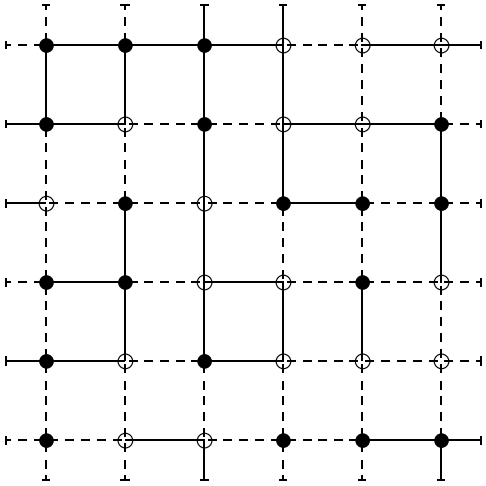

Figure 1. Schematic illustration of the structure and ingredients of our recurrent self-programming model for $L = 2$. The binary spin variables (spin up: $\circ$, spin down: $\bullet$) evolve on time-scales of order 1. The level-1 interactions (solid line segments) evolve on time-scales $\tau_1 \gg 1$. The level-2 interactions (dashed line segments) evolve on time-scales $\tau_2 \gg \tau_1$. The interactions are randomly allocated to levels. Our present model differs from this simple picture in two ways: firstly, it is fully connected (i.e. infinite dimensional), and secondly, we allow for an arbitrary number L of interaction types.

on a single time-scale, but are divided into a hierarchy of L different groups, each with their own characteristic time-scale τ_ℓ and noise level T_ℓ ($\ell = 1, \ldots, L$), describing a hierarchy of increasingly non-volatile programming levels. This appears to be a much more realistic representation of self-programming systems; conventional programmes generally take the form of hierarchies of routines, sub-routines and so on, and it would appear appropriate to allow low-level sub-routines to be more easily modifiable than high-level ones. In order to retain analytical solvability we choose the different groups of interactions randomly (prescribing only their sizes), see figure 1. We solve the model in equilibrium, and find, upon making the replica-symmetric (i.e. ergodic) ansatz within each level of our hierarchy, a theory which resembles, but is not identical to, Parisi's L-level replica symmetry breaking solution for spin systems with frozen disorder. Although Parisi's solution can also be traced back to the existence of a hierarchy of adiabatically separated time-scales [13], in the present model the *interactions* are grouped into clusters with different time-scales, rather than the spins. Apart from quantitative differences in the details of the order parameter equations, a major consequence of this difference is that in the present model the block sizes m_i of the emerging ultrametric solution are not restricted to the interval $[0, 1]$, but are independent control parameters, defined in terms of the noise strengths of the various levels in the hierarchy. They can consequently take any value in the interval $[0, \infty\rangle$. We show that this leads to extremely rich phase diagrams, with an abundance of first-order transitions especially when the level of stochasticity in the interaction dynamics is chosen to be low, i.e. when the dimensions $\{m_i\}$ become large. We study our model in full detail for the choices $L = 2$ and $L = 3$, including phase diagrams, and we study the asymptotic properties of our model in the limits $m_1 \to \infty$ for fixed T (deterministic dynamics of the level-1 interactions) and $m_1 \to 0$ for fixed T_1 (deterministic dynamics of the neuronal processors), for arbitrary L.

2. Model definitions

In this paper we will refer to our binary neurons as spins and to the synaptic interactions as couplings. We will write the N-spin state vector as $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_N) \in \{-1, 1\}^N$, and the matrix of interactions as $\boldsymbol{J} = \{J_{ij}\}$. The spins are taken to have a stochastic Glauber-type dynamics such that for *stationary* choices of the couplings the microscopic spin probability density would evolve towards a Boltzmann distribution

$$p_\infty(\boldsymbol{\sigma}) = \frac{\mathrm{e}^{-\beta H(\boldsymbol{\sigma})}}{Z} \qquad Z = \sum_{\boldsymbol{\sigma}} \mathrm{e}^{-\beta H(\boldsymbol{\sigma})} \tag{1}$$

with the conventional Hamiltonian

$$H(\boldsymbol{\sigma}) = -\sum_{i<j} J_{ij}\sigma_i\sigma_j \tag{2}$$

where $i, j \in \{1, \ldots, N\}$, and with the inverse temperature $\beta = T^{-1}$.

The couplings J_{ij} also evolve in a stochastic manner, in response to the states of the spins, but adiabatically slowly compared to the spins, such that on the time-scales of the couplings the spins are always in an equilibrium state described by (1). For the coupling dynamics the following Langevin equations are proposed:

$$\tau_{ij}\frac{\mathrm{d}}{\mathrm{d}t}J_{ij} = \frac{1}{N}\langle\sigma_i\sigma_j\rangle_{\mathrm{sp}} - \mu_{ij}J_{ij} + \eta_{ij}(t)\sqrt{\frac{\tau_{ij}}{N}} \qquad i < j = 1, \ldots, N \tag{3}$$

with $\tau_{ij} \gg 1$. In the adiabatic limit $\tau_{ij} \to \infty$, the term $\langle\sigma_i\sigma_j\rangle_{\mathrm{sp}}$, representing spin correlations associated with the coupling J_{ij}, becomes an average over the Boltzmann distribution (1) of the spins, given the instantaneous couplings $\boldsymbol{J}$. $\eta_{ij}(t)$ represent Gaussian white noise contributions, of zero mean and covariance $\langle\eta_{ij}(t)\eta_{kl}(t')\rangle = 2T_{ij}\delta_{ik}\delta_{jl}\delta(t - t')$, with associated temperature $T_{ij} = \beta_{ij}^{-1}$. Appropriate factors of N have been introduced in order to ensure non-trivial behaviour in the limit $N \to \infty$. We classify the spin pairs (i, j) according to the characteristic time-scale τ_{ij} and the control parameters (T_{ij}, μ_{ij}) associated with their interactions J_{ij}. In contrast to papers such as [3, 4, 6, 9, 10], where $\tau_{ij} = \tau$, $T_{ij} = \tilde{T}$ and $\mu_{ij} = \mu$ for all (i, j), here the various time-scales, temperatures and decay rates are no longer assumed to be identical, but to come in L distinct adiabatically separated groups I_ℓ (always with $i < j$):

$$I_\ell = \{(i, j)|\tau_{ij} = \tau_\ell, T_{ij} = T_\ell, \mu_{ij} = \mu_\ell\} \qquad l = 1, \ldots, L \tag{4}$$

with $1 \ll \tau_1 \ll \tau_2 \ll \cdots \ll \tau_L$. Thus $\{(i, j)\} = \bigcup_{\ell\leqslant L} I_\ell$. We will write the set of spin-interactions with time-scale τ_ℓ as $\boldsymbol{J}^\ell = \{J_{ij}|(i, j) \in I_\ell\}$. The interactions in group I_2 are adiabatically slow compared to those in group I_1, and so on. The rationale of this set-up is that, in information processing terms, this would represent a stochastic self-programming neural information processing system equipped with a programme which consists of a hierarchy of increasingly less volatile and less easily modifiable sub-routines.

Finally we have to define the detailed partitioning of the $\frac{1}{2}N(N-1)$ interactions into the L volatility groups. We introduce $\epsilon_{ij}(\ell) \in \{0, 1\}$ such that $\epsilon_{ij}(\ell) = 1$ if and only if $(i, j) \in I_\ell$, so $\sum_{\ell=1}^{L}\epsilon_{ij}(\ell) = 1$ for all (i, j). In order to arrive at a solvable mean-field problem, with full equivalence of the N sites, we will choose the $\epsilon_{ij}(\ell)$ independently at random for each pair (i, j) with $i < j$, with probabilities

$$\mathrm{Prob}[\epsilon_{ij}(\ell) = 1] = \epsilon_\ell \qquad \sum_{\ell=1}^{L}\epsilon_\ell = 1. \tag{5}$$

These time-scale and temperature allocation variables $\{\epsilon_{ij}(\ell)\}$ thus introduce quenched disorder into our problem. Averaging over this disorder will be denoted by $\overline{\cdots}$, as usual.

3. Replica analysis of the stationary state

3.1. Equilibrium statistical mechanics of the couplings

We denote averages over the probability distribution of the couplings at level ℓ in the hierarchy as $\langle\cdots\rangle_\ell$. At every level ℓ, the stochastic equation (3) for those couplings which evolve on that particular time-scale τ_ℓ has now become conservative:

$$\tau_\ell \frac{\mathrm{d}}{\mathrm{d}t} J_{ij} = \frac{1}{N}\langle\cdots\langle\langle\sigma_i\sigma_j\rangle_{\mathrm{sp}}\rangle_1\cdots\rangle_{\ell-1} - \mu_\ell J_{ij} + \eta_{ij}(t)\sqrt{\frac{\tau_\ell}{N}}$$
$$= -\frac{1}{N}\frac{\partial}{\partial J_{ij}} H_\ell(\boldsymbol{J}^\ell,\ldots,\boldsymbol{J}^L) + \eta_{ij}(t)\sqrt{\frac{\tau_\ell}{N}} \tag{6}$$

with the following effective Hamiltonian for the couplings at level ℓ:

$$H_1(\boldsymbol{J}^1,\ldots,\boldsymbol{J}^L) = -\beta^{-1}\log Z[\boldsymbol{J}^1,\ldots,\boldsymbol{J}^L] \tag{7}$$
$$H_{\ell+1}(\boldsymbol{J}^{\ell+1},\ldots,\boldsymbol{J}^L) = -\beta_\ell^{-1}\log Z_\ell[\boldsymbol{J}^{\ell+1},\ldots,\boldsymbol{J}^L] \qquad (1\leqslant\ell<L) \tag{8}$$

and with the partition functions

$$Z[\boldsymbol{J}^1,\ldots,\boldsymbol{J}^L] = \sum_{\boldsymbol{\sigma}} \mathrm{e}^{-\beta H(\boldsymbol{\sigma},\boldsymbol{J})} \tag{9}$$
$$Z_\ell[\boldsymbol{J}^{\ell+1},\ldots,\boldsymbol{J}^L] = \int \mathrm{d}\boldsymbol{J}^\ell\ \mathrm{e}^{-\beta_\ell H_\ell(\boldsymbol{J}^l,\ldots,\boldsymbol{J}^L)} \qquad (1\leqslant\ell<L) \tag{10}$$
$$Z_L = \int \mathrm{d}\boldsymbol{J}^L\ \mathrm{e}^{-\beta_L H_L(\boldsymbol{J}^L)} \tag{11}$$

in which

$$H(\boldsymbol{\sigma},\boldsymbol{J}) = -\sum_{i<j} J_{ij}\sigma_i\sigma_j + \frac{1}{2}N\sum_{i<j}\mu_{ij}J_{ij}^2. \tag{12}$$

This describes a hierarchy of nested equilibrations. At each time-scale τ_ℓ the interactions $\boldsymbol{J}^\ell$ equilibrate to a Boltzmann distribution, with effective Hamiltonian H_ℓ which is the free energy of the previous level $\ell+1$, starting from the overall Hamiltonian (12) (for spins and couplings) at the fastest (spin) level. As a result of having different effective temperatures T_ℓ associated with each level, the partition functions are found to generate replica theories with replica dimensions $m_\ell \geqslant 0$ which represent the ratios of the effective temperatures of the two levels involved. This follows from substitution of (8) into (10):

$$Z_\ell[\boldsymbol{J}^{l+1},\ldots,\boldsymbol{J}^L] = \int \mathrm{d}\boldsymbol{J}^\ell \left\{Z_{\ell-1}[\boldsymbol{J}^l,\ldots,\boldsymbol{J}^L]\right\}^{m_\ell} \tag{13}$$
$$m_\ell = \beta_\ell/\beta_{\ell-1} \qquad m_1 = \beta_1/\beta. \tag{14}$$

The statics of the system, including the effect of the quenched disorder, are governed by the disorder-averaged free energy $\mathcal{F}$ associated with the partition function Z_L in (11), where the slowest variables have finally been integrated out:

$$\mathcal{F} = -\frac{1}{\beta_L}\overline{\log Z_L} = -\lim_{m_{L+1}\to 0}\frac{1}{m_{L+1}\beta_L}\log\overline{Z_L^{m_{L+1}}}. \tag{15}$$

This function is found to act as the general generator of equilibrium values for observables at any level in the hierarchy, since upon adding suitable generating terms to the Hamiltonian (12), i.e. $H(\boldsymbol{\sigma},\boldsymbol{J}) \to H(\boldsymbol{\sigma},\boldsymbol{J}) + \lambda\Phi(\boldsymbol{\sigma},\boldsymbol{J})$, one finds

$$\overline{\langle\cdots\langle\langle\Phi(\boldsymbol{\sigma},\boldsymbol{J})\rangle_{\mathrm{sp}}\rangle_1\cdots\rangle_L} = \lim_{\lambda\to 0}\frac{\partial}{\partial\lambda}\mathcal{F}. \tag{16}$$

We can now combine our previous results and write down an explicit expression for $\mathcal{F}$, involving multiple replications due to (13) and (15). We find $\mathcal{F} = \lim_{m_{L+1}\to 0} F[m_1,\ldots,m_{L+1}]$, with $F[m_1,\ldots,m_{L+1}] = -(m_{L+1}\beta_L)^{-1}\log\overline{Z_1^{m_{L+1}}}$, and where $Z_L^{m_{L+1}}$ is written as

$$Z_L^{m_{L+1}} = \int\left[\prod_{\alpha_{L+1}} \mathrm{d}\boldsymbol{J}^{L,\alpha_{L+1}}\right]\prod_{\alpha_{L+1}}\{Z_{L-1}[\boldsymbol{J}^{L,\alpha_{L+1}}]\}^{m_L}$$
$$= \sum_{\{\boldsymbol{\sigma}^{\alpha_1,\ldots,\alpha_{L+1}}\}} \mathrm{e}^{\frac{\beta}{2N}\sum_{\ell\leqslant L}\frac{1}{m_1\cdots m_\ell\mu_\ell}\sum_{(i<j)\in I_\ell}\sum_{\alpha_{\ell+1},\ldots,\alpha_{L+1}}\left[\sum_{\alpha_1,\ldots,\alpha_\ell}\sigma_i^{\alpha_1,\ldots,\alpha_{L+1}}\sigma_j^{\alpha_1,\ldots,\alpha_{L+1}}\right]^2} \tag{17}$$

(modulo irrelevant constants), in which always $\alpha_\ell = 1,\ldots,m_\ell$.

3.2. Disorder averaging

In order to average (17) over the disorder, we note that the spin summation in the exponent can also be written in terms of the allocation variables $\epsilon_{ij}(\ell)$:

$$\sum_{\ell\leqslant L}\frac{1}{m_1\cdots m_\ell\mu_\ell}\sum_{(i<j)\in I_\ell}\sum_{\alpha_{l+1},\ldots,\alpha_{L+1}}\cdots$$
$$= \sum_{i<j}\sum_{\ell\leqslant L}\frac{\beta}{\beta_\ell}\frac{\epsilon_{ij}(\ell)}{\mu_\ell}\sum_{\alpha_{\ell+1},\ldots,\alpha_{L+1}}\left[\sum_{\alpha_1,\ldots,\alpha_\ell}\sigma_i^{\alpha_1,\ldots,\alpha_{L+1}}\sigma_j^{\alpha_1,\ldots,\alpha_{L+1}}\right]^2$$

where we used $m_1\cdots m_\ell = \beta_\ell/\beta$. This allows us to carry out the average in (17):

$$F[\cdots] = -\frac{1}{m_{L+1}\beta_L}\log\sum_{\{\boldsymbol{\sigma}^{\alpha_1,\ldots,\alpha_{L+1}}\}}\prod_{i<j}\overline{\left[\mathrm{e}^{\frac{\beta^2}{2N}\sum_{\ell\leqslant L}\frac{\epsilon_{ij}(\ell)}{\beta_\ell\mu_\ell}\sum_{\alpha_{\ell+1},\ldots,\alpha_{L+1}}\left[\sum_{\alpha_1,\ldots,\alpha_\ell}\sigma_i^{\alpha_1,\ldots,\alpha_{L+1}}\sigma_j^{\alpha_1,\ldots,\alpha_{L+1}}\right]^2}\right]}$$
$$= -\frac{1}{m_{L+1}\beta_L}\log\sum_{\{\boldsymbol{\sigma}^{\alpha_1,\ldots,\alpha_{L+1}}\}}\mathrm{e}^{\mathcal{O}(N^0)}$$
$$\times\exp\left\{\frac{N\beta^2}{4}\sum_{\alpha_1,\ldots,\alpha_{L+1}}\sum_{\beta_1,\ldots,\beta_{L+1}}\sum_{\ell\leqslant L}\frac{\epsilon_\ell}{\beta_\ell\mu_\ell}\delta_{(\alpha_{\ell+1},\ldots,\alpha_{L+1}),(\beta_{\ell+1},\ldots,\beta_{L+1})}\right.$$
$$\left.\times\left[\frac{1}{N}\sum_i\sigma_i^{\alpha_1,\ldots,\alpha_{L+1}}\sigma_i^{\beta_1,\ldots,\beta_{L+1}}\right]^2\right\}$$

(again modulo irrelevant constants). We abbreviate $a = (\alpha_1,\ldots,\alpha_{L+1})$, and introduce the spin-glass replica order parameters $q_{ab} = \frac{1}{N}\sum_i\sigma_i^{\alpha_1,\ldots,\alpha_{L+1}}\sigma_i^{\beta_1,\ldots,\beta_{L+1}}$ by inserting appropriate integrals over δ-distributions, and arrive at

$$F[\cdots] = -\frac{1}{m_{L+1}\beta_L}\log\int\{\mathrm{d}q_{ab}\,\mathrm{d}\hat{q}_{ab}\}\,\mathrm{e}^{NG[\{q_{ab},\hat{q}_{ab}\}]+\mathcal{O}(N^0)} \tag{18}$$

$$G[\{q_{ab},\hat{q}_{ab}\}] = \mathrm{i}\sum_{ab}\hat{q}_{ab}\,q_{ab} + \frac{\beta^2}{4}\sum_{\ell\leqslant L}\frac{\epsilon_\ell}{\beta_\ell\mu_\ell}\sum_{ab}\delta_{(\alpha_{\ell+1},\ldots,\alpha_{L+1}),(\beta_{\ell+1},\ldots,\beta_{L+1})}q_{ab}^2$$
$$+\log\sum_{\{\sigma_{\alpha_1,\ldots,\alpha_{L+1}}\}}\mathrm{e}^{-\mathrm{i}\sum_{ab}\hat{q}_{ab}\sigma_{\alpha_1,\ldots,\alpha_{L+1}}\sigma_{\beta_1,\ldots,\beta_{L+1}}}. \tag{19}$$

For $N\to\infty$ the above integral can be evaluated by steepest descent, and upon elimination of the conjugate order parameters $\{\hat{q}_{ab}\}$ by variation of $\{q_{ab}\}$, the disorder-averaged free energy

per spin $f = \lim_{N\to\infty} \mathcal{F}/N$ is found to be

$$f = \lim_{m_{L+1}\to 0} \frac{1}{m_{L+1}\beta_L} \left\{ \frac{\beta^2}{4} \sum_{\ell\leqslant L} \frac{\epsilon_\ell}{\beta_\ell\mu_\ell} \sum_{ab} \delta_{(\alpha_{\ell+1},\ldots,\alpha_{L+1}),(\beta_{\ell+1},\ldots,\beta_{L+1})} q_{ab}^2 \right.$$
$$\left. - \log \sum_{\{\sigma_a\}} \mathrm{e}^{\frac{\beta^2}{2}\sum_{\ell\leqslant L}\frac{\epsilon_\ell}{\beta_\ell\mu_\ell}\sum_{ab}\delta_{(\alpha_{\ell+1},\ldots,\alpha_{L+1}),(\beta_{\ell+1},\ldots,\beta_{L+1})} q_{ab}\sigma_a\sigma_b} \right\}. \tag{20}$$

The saddle-point equations used to solve the $\{q_{ab}\}$ are given by

$$q_{cd} = \frac{\sum_{\{\sigma_a\}} \sigma_c\sigma_d\, \mathrm{e}^{\frac{\beta^2}{2}\sum_{\ell\leqslant L}\epsilon_\ell/(\beta_\ell\mu_\ell)\sum_{ab}\delta_{(\alpha_{\ell+1},\ldots,\alpha_{L+1}),(\beta_{\ell+1},\ldots,\beta_{L+1})} q_{ab}\sigma_a\sigma_b}}{\sum_{\{\sigma_a\}} \mathrm{e}^{\frac{\beta^2}{2}\sum_{\ell\leqslant L}\epsilon_\ell/(\beta_\ell\mu_\ell)\sum_{ab}\delta_{(\alpha_{\ell+1},\ldots,\alpha_{L+1}),(\beta_{\ell+1},\ldots,\beta_{L+1})} q_{ab}\sigma_a\sigma_b}} \tag{21}$$

with $a = (\alpha_1,\ldots,\alpha_{L+1})$ and $\alpha_\ell \in \{1,\ldots,m_\ell\}$ for all ℓ, where the dimensions m_ℓ are given by the ratios of the temperatures (14) of subsequent programming levels in the hierarchy.

4. Replica-symmetric solutions

4.1. Evaluation of the replica-symmetric free energy

Our full order parameters are $q_{ab} = \frac{1}{N}\sum_i \sigma_i^{\alpha_1,\ldots,\alpha_{L+1}}\sigma_i^{\beta_1,\ldots,\beta_{L+1}}$. Note that $\alpha_{L+1} = \beta_{L+1}$ and that $q_{aa} = 1$. Spin variables with $\alpha_L = \beta_L$ have identical level-L bonds. Those with $\alpha_L = \beta_L$ and $\alpha_{L-1} = \beta_{L-1}$ have identical level L bonds *and* identical level $L-1$ bonds, and so on. Hence, for the present model the replica-symmetric (RS) ansatz (describing ergodicity at each level of the hierarchy of time-scales) takes the following form:

$$\alpha_L \neq \beta_L: \qquad q_{ab} = q_L$$
$$(\alpha_{\ell+1},\ldots,\alpha_{L+1}) = (\beta_{\ell+1},\ldots,\beta_{L+1}), \qquad \alpha_\ell \neq \beta_\ell: \quad q_{ab} = q_\ell$$

or

$$q_{ab} = \sum_{\ell=1}^{L} q_\ell\, \delta_{(\alpha_{\ell+1},\ldots,\alpha_{L+1})(\beta_{\ell+1},\ldots,\beta_{L+1})}\bar{\delta}_{\alpha_\ell\beta_\ell} + \delta_{ab} \tag{22}$$

where $\bar{\delta}_{ij} = 1 - \delta_{ij}$, and where $0 \leqslant q_L \leqslant \cdots \leqslant q_1 \leqslant 1$. With a modest amount of foresight we introduce the abbreviation

$$\pi_\ell = \sum_{\ell'=\ell}^{L} \frac{\beta\epsilon_{\ell'}}{\beta_{\ell'}\mu_{\ell'}} = \sum_{\ell'=\ell}^{L} \frac{1}{m_1\cdots m_{\ell'}}\frac{\epsilon_{\ell'}}{\mu_{\ell'}}. \tag{23}$$

Insertion of the ansatz (22) into (20) gives (using the relation $\beta_L = \beta\prod_{\ell=1}^{L} m_\ell$, and with the notational convention $q_{L+1} = 0$ to simplify summations):

$$f = -\frac{1}{4}\pi_L - \frac{1}{\beta}\log 2 + \frac{1}{4}\sum_{\ell=2}^{L} q_\ell^2\pi_\ell m_1\cdots m_{\ell-1}(m_\ell - 1) + \frac{1}{4}q_1^2\pi_1(m_1-1) + \frac{1}{2}q_1\pi_1$$
$$-\frac{1}{\beta m_1\cdots m_L}\log\int Dz_L \left\{ \ldots \right.$$
$$\left. \times \left\{ \int Dz_1 \cosh^{m_1}\left[\sum_{\ell=1}^{L} z_\ell\sqrt{\beta(q_\ell\pi_\ell - q_{\ell+1}\pi_{\ell+1})} \right] \right\}^{m_2} \ldots \right\}^{m_L}$$

For $L = 1$ this reduces to (modulo the irrelevant constant):

$$f = \frac{1}{2}q_1\pi_1 + \frac{1}{4}q_1^2\pi_1(m_1-1) - \frac{1}{\beta}\log 2 - \frac{1}{\beta m_1}\log\int \mathrm{D}z_1 \cosh^{m_1}[z_1\sqrt{\beta q_1\pi_1}] \tag{24}$$

whereas for $L > 1$ we can simplify our result to (modulo the irrelevant constant):

$$\begin{aligned} f = {} & \frac{1}{2}\pi_1 q_1 - \frac{1}{\beta}\log 2 + \frac{1}{4}\sum_{\ell=2}^{L} q_\ell^2\pi_\ell\left[\prod_{k=1}^{\ell-1} m_k\right](m_\ell - 1) + \frac{1}{4}q_1^2\pi_1(m_1-1) \\ & - \frac{1}{\beta\prod_{\ell=1}^{L} m_\ell}\log\int \mathrm{D}z_L\left\{\int \mathrm{D}z_{L-1}\left\{\cdots\left\{\int \mathrm{D}z_1\right.\right.\right. \\ & \left.\left.\left.\times\left\{\cosh\left[z_L\sqrt{\beta q_L\pi_L} + \sum_{\ell=1}^{L-1} z_\ell\sqrt{\beta(q_\ell\pi_\ell - q_{\ell+1}\pi_{\ell+1})}\right]\right\}^{m_1}\right\}^{m_2}\cdots\right\}^{m_{L-1}}\right\}^{m_L}. \end{aligned} \tag{25}$$

For systems with a single coupling time-scale (24) we again observe the similarity with the free energy of Parisi's RSB-1 ansatz [12] (see also, e.g., [6, 13]). For systems with multiple coupling time-scales (25) this similarity is lost: Parisi's RSB-L solution [12, 13] emerges only when $\pi_\ell = \pi_L$ for all ℓ, i.e. when we return to a single coupling time-scale.

4.2. The single-level benchmark

A simple and convenient benchmark test of the above equations is obtained upon putting $\epsilon_\ell \to \delta_{r\ell}$ for some $r \in \{1, \ldots, L\}$. Here we retain only a single bond time-scale, and our theory should effectively reduce to that of [3]. For $L = 1$ this is indeed true, according to equation (24); here we will address the generic case $L > 1$ and $1 < r < L$. The ϵ_ℓ occur only in the quantities π_ℓ of (23), which now simplify to

$$\pi_{\ell\leqslant r} = \frac{\epsilon_r}{\mu_r m_1\cdots m_r} \qquad \pi_{\ell>r} = 0. \tag{26}$$

Insertion into the replica-symmetric disorder-averaged free energy per spin (25) gives

$$\begin{aligned} f = {} & \frac{1}{2}\pi_r q_1 - \frac{1}{\beta}\log 2 + \frac{1}{4}\pi_r\sum_{\ell=2}^{r} q_\ell^2\left[\prod_{k=1}^{\ell-1} m_k\right](m_\ell - 1) + \frac{1}{4}q_1^2\pi_r(m_1-1) \\ & - \frac{1}{\beta\prod_{\ell=1}^{r} m_\ell}\log\int \mathrm{D}z_r\left\{\int \mathrm{D}z_{r-1}\left\{\cdots\left\{\int \mathrm{D}z_1\right.\right.\right. \\ & \left.\left.\left.\times\left\{\cosh\left[z_r\sqrt{\beta q_r\pi_r} + \sum_{\ell=1}^{r-1} z_\ell\sqrt{\beta(q_\ell - q_{\ell+1})\pi_r}\right]\right\}^{m_1}\right\}^{m_2}\cdots\right\}^{m_{r-1}}\right\}^{m_r}. \end{aligned}$$

Since q_L measures the overlap between spin states with different level L bonds, q_{L-1} measures the overlap between spin states with identical level L bonds but different level $(L-1)$ bonds, and so on, the relevant state for $\epsilon_\ell = \delta_{r\ell}$ must be one where $q_\ell = q_r$ for all $\ell \leqslant r$. Insertion into the above expression gives

$$\begin{aligned} f = {} & \frac{1}{2}\pi_r q_r - \frac{1}{\beta}\log 2 + \frac{1}{4}\pi_r q_r^2\left(\prod_{k=1}^{r} m_k - 1\right) \\ & - \frac{1}{\beta\prod_{\ell=1}^{r} m_\ell}\log\int \mathrm{D}z_r\{\cosh[z_r\sqrt{\beta q_r\pi_r}]\}^{m_1 m_2\cdots m_r}. \end{aligned}$$

This expression is indeed equivalent to (24), from which it can be obtained by making the replacements $\pi_1 \to \pi_r$ and $m_1 \to m_1 \cdots m_r$.

4.3. Nature of the physical saddle-point

Due to our previous elimination of (imaginary) conjugate order parameters and the possible curvature sign changes occurring in replica theories, we can no longer be sure that the relevant saddle-point (24) and (25) gives the minimum of f. In order to allow us to select the physical saddle-point in situations where multiple saddle-points exist, we determine the nature of the physical saddle-point by inspection of the high-temperature state. Since there are multiple temperatures in this problem (one for the spins, and L for the various bond levels), there are also different ways to send all temperatures to infinity, with potentially different outcomes. Here we consider the limit which appears most natural, where $\beta \to 0$ for fixed $\{m_1, \ldots, m_L\}$. We then obtain from expression (25), with $\tilde{f} = f + \frac{1}{\beta}\log 2$:

$$\begin{aligned}\lim_{\beta\to 0}\tilde{f} &= \frac{1}{2}\pi_1 q_1 + \frac{1}{4}\sum_{\ell=2}^{L} q_\ell^2\pi_\ell\left[\prod_{k=1}^{l-1} m_k\right](m_\ell - 1) + \frac{1}{4}q_1^2\pi_1(m_1-1)\\ &\quad - \lim_{\beta\to 0}\frac{1}{\beta\prod_{\ell=1}^{L} m_\ell}\log\int \mathrm{D}z_L\left\{\int \mathrm{D}z_{L-1}\left\{\cdots\left\{1+\frac{1}{2}\beta m_1\int \mathrm{D}z_1\right.\right.\right.\\ &\quad \left.\left.\left.\times\left[z_L\sqrt{q_L\pi_L} + \sum_{\ell=1}^{L-1} z_\ell\sqrt{q_\ell\pi_\ell - q_{\ell+1}\pi_{\ell+1}}\right]^2 + \mathcal{O}(\beta^2)\right\}^{m_2}\cdots\right\}^{m_{L-1}}\right\}^{m_L}\\ &= \frac{1}{4}\sum_{\ell=2}^{L} q_\ell^2\pi_\ell\left[\prod_{k=1}^{\ell-1} m_k\right](m_\ell - 1) + \frac{1}{4}q_1^2\pi_1(m_1-1) + \mathcal{O}(\beta^2). \end{aligned} \tag{27}$$

This shows explicitly that the nature of the physical state, for $\beta \to 0$ given by $q_\ell = 0$ for all ℓ, depends on the specific choice made for the various coupling temperatures, which determine the replica dimensions $\{m_\ell\}$ via (14). This is in sharp contrast with the standard Parisi ansatz [12] where one always has $m_\ell \leqslant 1$. Here we find

$$\begin{aligned} m_\ell > 1: &\qquad \text{minimization of } f \text{ with respect to } q_\ell\\ m_\ell < 1: &\qquad \text{maximization of } f \text{ with respect to } q_\ell. \end{aligned} \tag{28}$$

4.4. The full saddle-point equations

In this section we derive the full saddle point equations of (25) for the replica-symmetric order parameters $0 \leqslant q_L \leqslant q_{L-1} \leqslant \cdots q_1 \leqslant 1$, in the general L-level hierarchy. This is most efficiently done by first writing the free energy (25) as

$$\begin{aligned} f &= \frac{1}{2}\pi_1 q_1 - \frac{1}{\beta}\log 2 + \frac{1}{4}\sum_{\ell=2}^{L} q_\ell^2\pi_\ell\left[\prod_{k=1}^{\ell-1} m_k\right](m_\ell - 1)\\ &\quad + \frac{1}{4}q_1^2\pi_1(m_1-1) - \left[\beta\prod_{\ell=1}^{L} m_\ell\right]^{-1}\log K_L \end{aligned} \tag{29}$$

$$K_L = \int \mathrm{D}z_L\left\{\int \mathrm{D}z_{L-1}\left\{\cdots\left\{\int \mathrm{D}z_1\{\cosh\Xi\}^{m_1}\right\}^{m_2}\cdots\right\}^{m_{L-1}}\right\}^{m_L} \tag{30}$$

$$\Xi = z_L\sqrt{\beta q_L \pi_L} + \sum_{\ell=1}^{L-1} z_\ell \sqrt{\beta(q_\ell \pi_\ell - q_{\ell+1}\pi_{\ell+1})} = \sum_{\ell=1}^{L} a_l z_l \tag{31}$$

$$a_{\ell(<L)} = \sqrt{\beta(q_\ell \pi_\ell - q_{\ell+1}\pi_{\ell+1})} \qquad a_L = \sqrt{\beta q_L \pi_L}. \tag{32}$$

From this it follows, upon taking derivatives of f with respect to q_1 and $q_{\ell>1}$, respectively, that the saddle-point equations can be written as

$$0 = \frac{1}{2}\pi_1 + \frac{1}{2}q_1\pi_1(m_1 - 1) - \frac{1}{\beta m_1 \cdots m_L}\frac{1}{K_L}\frac{\partial K}{\partial q_1} \tag{33}$$

$$0 = \frac{1}{2}q_\ell \pi_\ell \left[\prod_{k=1}^{\ell-1} m_k\right](m_\ell - 1) - \frac{1}{\beta m_1 \cdots m_L}\frac{1}{K_L}\frac{\partial K}{\partial q_\ell} \qquad \text{for} \quad \ell > 1. \tag{34}$$

The remaining problem is to calculate derivatives of K_L, which is complicated by the nesting of integrals. Note that K_L can be defined iteratively,

$$K_0 = \cosh \Xi \qquad K_L = \int \mathrm{D}z_L K_{L-1}^{m_L}. \tag{35}$$

To suppress notation we also define

$$M_0 = 1 \qquad M_{\ell>0} = \prod_{\ell'=1}^{\ell} m_{\ell'} \qquad \langle f \rangle_\ell = K_\ell^{-1} \int \mathrm{D}z_\ell \, K_{\ell-1}^{m_\ell} f(z_\ell). \tag{36}$$

In appendix B we show that, with this shorthand, the relevant derivatives of K are given by the following expressions:

$$\frac{\partial K_L}{\partial q_1} = \frac{1}{2}\beta\pi_1 M_L K \left\{1 + (m_1 - 1)\left\langle \cdots \langle \tanh^2 \Xi \rangle_1 \cdots \right\rangle_L\right\} \tag{37}$$

$$\frac{\partial K_L}{\partial q_{\ell>1}} = \frac{1}{2}\beta\pi_\ell M_L M_{\ell-1}(m_\ell - 1) K \left\langle \cdots \left\langle [\langle \cdots \langle \tanh \Xi \rangle_1 \cdots \rangle_{\ell-1}]^2 \right\rangle_\ell \cdots \right\rangle_L. \tag{38}$$

Hence the saddle-point equations (33) and (34) reduce to the, in view of the definition (22), appealing and transparent relations

$$q_1 = \left\langle \cdots \langle \tanh^2 \Xi \rangle_1 \cdots \right\rangle_L \tag{39}$$

$$q_{\ell>1} = \left\langle \cdots \left\langle [\langle \cdots \langle \tanh \Xi \rangle_1 \cdots \rangle_{\ell-1}]^2 \right\rangle_\ell \cdots \right\rangle_L. \tag{40}$$

5. The two-level hierarchy: $L = 2$

5.1. General properties

We next apply our results to the case where we have just two time-scales in the bond dynamics, and calculate the phase diagram. We found in studying the single level limit that our theory will simply self-consistently 'forget' about non-existent intermediate levels (as it should), and substitute the right temperature definitions in the expressions that would have been obtained if we had considered subsequent levels. Hence we may choose $L = 2$, without loss of generality, and simply put $L = 2$ in expression (25):

$$f_{L=2}(q_1, q_2) = \frac{1}{2}\pi_1 q_1 - \frac{1}{\beta}\log 2 + \frac{1}{4}q_2^2 \pi_2 m_1(m_2 - 1) + \frac{1}{4}q_1^2\pi_1(m_1 - 1) - \frac{1}{\beta m_1 m_2}\log$$
$$\times \int \mathrm{D}z_2 \left\{\int \mathrm{D}z_1 \left\{\cosh\left[z_2\sqrt{\beta q_2 \pi_2} + z_1\sqrt{\beta(q_1\pi_1 - q_2\pi_2)}\right]\right\}^{m_1}\right\}^{m_2}. \tag{41}$$

For $L = 2$ the saddle-point equations (39) and (40) reduce in explicit form to

$$q_1 = \frac{\int \mathrm{D}z_2 \left[\int \mathrm{D}z_1 \cosh^{m_1} \Xi\right]^{m_2} \left[\frac{\int \mathrm{D}z_1 \cosh^{m_1} \Xi \tanh^2 \Xi}{\int \mathrm{D}z_1 \cosh^{m_1} \Xi}\right]}{\int \mathrm{D}z_2 \left[\int \mathrm{D}z_1 \cosh^{m_1} \Xi\right]^{m_2}} \tag{42}$$

$$q_2 = \frac{\int \mathrm{D}z_2 \left[\int \mathrm{D}z_1 \cosh^{m_1} \Xi\right]^{m_2} \left[\frac{\int \mathrm{D}z_1 \cosh^{m_1} \Xi \tanh \Xi}{\int \mathrm{D}z_1 \cosh^{m_1} \Xi}\right]^2}{\int \mathrm{D}z_2 \left[\int \mathrm{D}z_1 \cosh^{m_1} \Xi\right]^{m_2}} \tag{43}$$

(note: $q_1 \geqslant q_2$), with

$$\Xi = z_2\sqrt{\beta q_2 \pi_2} + z_1\sqrt{\beta(q_1\pi_1 - q_2\pi_2)}.$$

We can distinguish between three phases. The first is a paramagnetic phase (P), corresponding to the solution $q_1 = q_2 = 0$ of (42), (43); according to (27) it is the sole saddle-point of (41) in the high-temperature regime, as it should be. The second is a spin-glass phase (SG_1), corresponding to a solution of the form $q_1 > 0, q_2 = 0$, describing a (spin-glass type) state with freezing of spins but not of the couplings. Insertion of $q_2 = 0$ into (42) and (43) shows that in this SG_1 phase q_1 is the solution of

$$q_1 = \frac{\int \mathrm{D}z \cosh^{m_1}(z\sqrt{\beta q_1 \pi_1}) \tanh^2(z\sqrt{\beta q_1 \pi_1})}{\int \mathrm{D}z \cosh^{m_1}(z\sqrt{\beta q_1 \pi_1})}. \tag{44}$$

In the third phase (SG_2) one has $q_1 \neq 0$ and $q_2 \neq 0$, here both spins and level-1 couplings 'freeze' into a state determined by the level-2 couplings, but the latter slowly but continually evolve, given sufficient time. Now one has to solve (42) and (43) in full. We indicate the various potential transition temperatures separating these phases as follows:

$$\begin{array}{lll} T_{p,1}^{\mathrm{2nd}}: & P \to SG_1 & \text{2nd order} \\ T_{1,2}^{\mathrm{2nd}}: & SG_1 \to SG_2 & \text{2nd order} \\ T_{p,1}^{\mathrm{1st}}: & P \to SG_1 & \text{1st order} \\ T_{p,2}^{\mathrm{1st}}: & P \to SG_2 & \text{1st order.} \end{array} \tag{45}$$

We will find that this list contains all transitions occurring.

5.2. $P \to SG_1$ transitions

The locations and properties of the two types of $P \to SG_1$ transitions follow upon expanding $f(q_1, 0)$ in powers of q_1:

$$-\beta m_1 f(q_1, 0) = F_0 + a(T)q_1^2 + b(T)q_1^3 + c(T)q_1^4 + \cdots.$$

The coefficients are found to be

$$\begin{aligned} a(T) &= \tfrac{1}{4}m_1(m_1 - 1)\beta\pi_1(\beta\pi_1 - 1) \\ b(T) &= \tfrac{1}{6}m_1(m_1 - 1)(m_1 - 2)(\beta\pi_1)^3 \\ c(T) &= \tfrac{1}{96}m_1(\beta\pi_1)^4\big(3m_1^3 - 72m_1^2 + 128m_1 - 68\big). \end{aligned} \tag{46}$$

Note that $c(T) < 0$ for $0 < m_1 \leqslant m_+$, where $m_+ \approx 15$. A continuous bifurcation occurs when $a(T) = 0$, i.e. when $T = \pi_1$, so

$$T_{p,1}^{\mathrm{2nd}} = \epsilon_1/m_1\mu_1 + \epsilon_2/m_1 m_2 \mu_2. \tag{47}$$

This bifurcation, dependent on the values of $\{m_1, m_2\}$, can, however, be preceded by a discontinuous one. Upon varying m_1 one finds the following scenario:

$$\begin{aligned}
0 < m_1 < 1: \quad & b(T) > 0 \qquad && f(q_1, 0) \text{ maximal at } q_1 = 0 \text{ for } T > T_{p,1}^{2\text{nd}} \\
& && f(q_1, 0) \text{ maximal at } q_1 > 0 \text{ for } T < T_{p,1}^{2\text{nd}} \\
1 < m_1 < 2: \quad & b(T) < 0 && f(q_1, 0) \text{ minimal at } q_1 = 0 \text{ for } T > T_{p,1}^{2\text{nd}} \\
& && f(q_1, 0) \text{ minimal at } q_1 > 0 \text{ for } T < T_{p,1}^{2\text{nd}} \\
2 < m_1: \quad & b(T) > 0 && f(q_1, 0) \text{ minimal at } q_1 = 0 \text{ for } T > T_{p,1}^{2\text{nd}} \\
& && f(q_1, 0) \text{ minimal at } q_1 > 0 \text{ for } T < T_{p,1}^{1\text{st}}
\end{aligned}$$

where $T_{p,1}^{1\text{st}}$ signals a discontinuous (i.e. first-order) transition, with $T_{p,1}^{1\text{st}} > T_{p,1}^{2\text{nd}}$. The condition for this latter transition follows from putting the derivative of the right-hand side of (44) with respect to q_1 equal to unity, giving

$$1 = \frac{1}{2}\beta\pi_1 \left[m_1(1-m_1)q_1^2 + 4(m_1-2)q_1 + 2 + (m_1-2)(m_1-3)\frac{\int \mathrm{D}z \cosh^{m_1}(z\sqrt{\beta q_1 \pi_1}) \tanh^4(z\sqrt{\beta q_1 \pi_1})}{\int \mathrm{D}z \cosh^{m_1}(z\sqrt{\beta q_1 \pi_1})} \right]. \tag{48}$$

Together with (28) we may now conclude that the $P \to SG_1$ transition is of second order for $m_1 < 2$, in which case it is given by (47), and first order for $m_1 > 2$, in which case it is given by the solution of (48). In a subsequent section we will show that for $m_1 \to \infty$ the first-order transition temperature $T_{p,1}^{1\text{st}}$ tends to a finite and non-zero value.

5.3. Transitions to SG_2

Two such transition temperatures: $T_{p,2}^{1\text{st}}$ (where one goes from a paramagnetic state to one with $q_1 \geqslant q_2 > 0$) and $T_{1,2}^{2\text{nd}}$ (where one goes from $q_1 > q_2 = 0$ to $q_1 \geqslant q_2 > 0$) are found. The transition condition defining $T_{1,2}^{2\text{nd}}$ follows from putting the derivative of the right-hand side of (43) with respect to q_2 equal to unity, followed by putting $q_2 = 0$, which gives

$$1 = \sqrt{\beta\pi_2}[1 + (m_1 - 1)q_1]. \tag{49}$$

We will show in appendix A that $T_{1,2}^{2\text{nd}} \propto \sqrt{T/m_1} = \sqrt{T_1}$ as $m_1 \to \infty$.

To determine $T_{p,2}^{1\text{st}}$, in contrast, one must demand an instability of (42) and (43) with $q_{1,2} > 0$. This implies meeting the more general condition

$$\left(\frac{\partial\varphi_1}{\partial q_1} - 1\right)\left(\frac{\partial\varphi_2}{\partial q_2} - 1\right) = \frac{\partial\varphi_1}{\partial q_2}\frac{\partial\varphi_2}{\partial q_1} \tag{50}$$

where $\varphi_1(q_1, q_2)$ and $\varphi_2(q_1, q_2)$ denote the right-hand sides of (42) and (43), respectively. For $m_1 \to \infty$ we will show, as with $T_{p,1}^{1\text{st}}$, that $T_{p,2}^{1\text{st}}$ also tends to a finite and non-zero value, and that $T_{p,2}^{(1\text{st})} > T_{p,1}^{(1\text{st})}$. Hence, $T_{p,2}^{1\text{st}}$ will ultimately become the true physical transition as $m_1 \to \infty$ (or, equivalently, $T_1 \to 0$ for fixed T).

5.4. The $L = 2$ phase diagram

For $L = 2$ there are still five independent control parameters in our model, namely T, m_1, m_2, π_1 and π_2 (where $\pi_2 = \epsilon_2/m_1m_2\mu_2$ and $\pi_1 = \pi_2 + \epsilon_1/m_1\mu_1$) or any combination of these. Hence we can only present representative cross-sections of the full phase diagram.

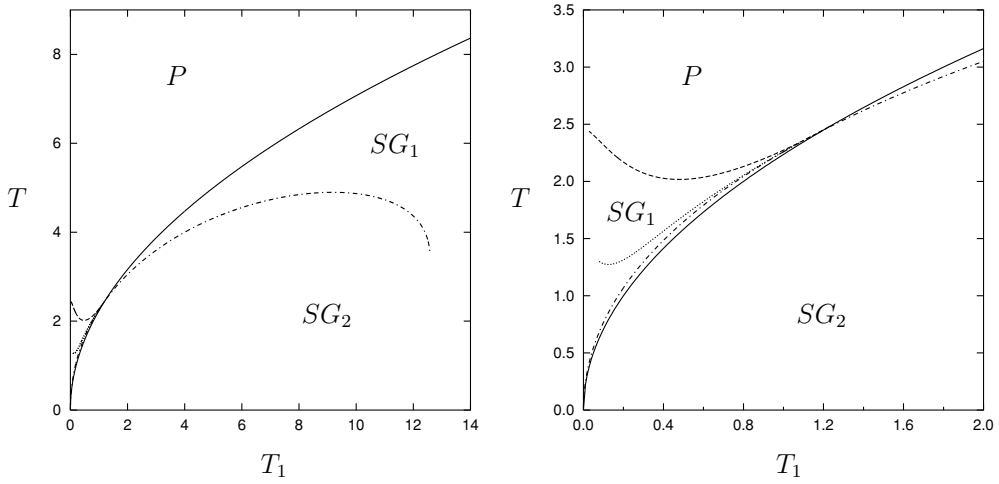

Figure 2. Phase diagram in the (T_1, T) plane for $L = 2$, $\epsilon_1/\mu_1 = 1$, $\epsilon_2/\mu_2 = 2$ and $m_2 = 0.5$, obtained by solving numerically the equations defining the various bifurcation lines (note: $T_1 = T/m_1$). Left panel: the large-scale structure. Right panel: enlargement of the low T_1 (or high m_1) region, where the first-order transitions are found. The relevant phases are P (paramagnetic phase, $q_1 = q_2 = 0$), SG_1 (first spin-glass phase, $q_1 > 0, q_2 = 0$), and SG_2 (second spin-glass phase, $q_1 > 0, q_2 > 0$). The transition lines shown, as defined in (45), are $T_{p,1}^{\rm 2nd}$ (solid curve), $T_{1,2}^{\rm 2nd}$ (dot-dash curve), $T_{p,1}^{\rm 1st}$ (dashed curve), and $T_{p,2}^{\rm 1st}$ (dotted curve). Note: the curve $T_{1,2}^{\rm 2nd}$ jumps to zero discontinuously at $T_1 = T_c^{(0)} \approx 12.575$ (see main text).

In this section we will restrict ourselves to showing the transition lines in the (T_1, T) plane, solved numerically from the various appropriate conditions as derived above, for the choice $\epsilon_1/\mu_1 = 1$, $\epsilon_2/\mu_2 = 2$ with $m_2 \in \{0.5, 1.5\}$ (which was found to cover more or less the generic scenarios). According to (47), for this choice of parameters the second-order $P \to SG_1$ transition temperature $T_{p,1}^{\rm 2nd}$ will be at

$$T_{p,1}^{\rm 2nd} = \frac{1}{m_1}\left(1 + \frac{2}{m_2}\right) \tag{51}$$

and it is replaced by a first-order $P \to SG_1$ transition at $m_1 = 2$, which, together with the relation $m_1 = T/T_1$, tells us that the first- and second-order $P \to SG_1$ lines meet at

$$T_{p,1}^{\rm 1st} = T_{p,2}^{\rm 2nd}: \qquad T_1 = \frac{1}{4}\left(1 + \frac{2}{m_2}\right). \tag{52}$$

In section 7 we will prove that, for any L, the discontinuous transition temperatures $T_{p,1}^{\rm 1st}$ and $T_{p,2}^{\rm 1st}$ always tend to a finite non-zero value as $m_1 \to \infty$, whereas $\lim_{m_1\to\infty} T_{p,1}^{\rm 2nd} = \lim_{m_1\to\infty} T_{1,2}^{\rm 2nd} = 0$. Thus, for sufficiently small T_1 (i.e. large m_1) the first-order transitions are always the ones which will actually take place. Secondly, we will find that there is always a critical value $T_c^{(0)}$ for T_1 such that $T_{1,2}^{\rm 2nd} = 0$ at $T_1 = T_c^{(0)}$, with $T_{1,2}^{\rm 2nd}$ simply absent for $T_1 > T_c^{(0)}$; as a consequence, the phase SG_2 no longer exists for $T_1 > T_c^{(0)}$. These two properties will be shown to be specific cases of more general results concerning spin-glass phases of arbitrary order.

Let us first turn to $m_2 = 0.5$, i.e. relatively high noise levels in the level-2 couplings compared to that of the spins. In figure 2 we show the phase diagram and the various bifurcation lines of the $L = 2$ system for $m_2 = 0.5$, as obtained by

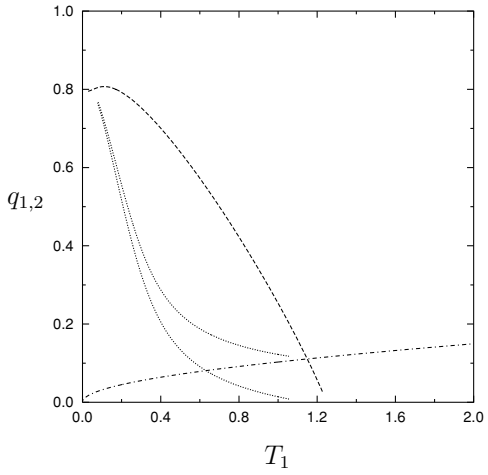

Figure 3. Values of the order parameters q_1 and q_2 as functions of T_1, along some of the transition lines of the previous figure, for $L = 2$, $\epsilon_1/\mu_1 = 1$, $\epsilon_2/\mu_2 = 2$ and $m_2 = 0.5$. The values shown are q_1 along $T_{1,2}^{\rm 2nd}$ (dot-dash curve), q_1 along $T_{p,1}^{\rm 1st}$ (dashed curve), and both q_1 and q_2 for $T_{p,2}^{\rm 1st}$ (dotted curves). Note: the order parameters which are not shown in this figure are zero by definition.

numerical solution of the relevant equations. The actual physical transitions are the ones which occur first as the temperature T is lowered from within the paramagnetic region. It will be clear from the figure that upon lowering T one can find different types of transition sequences, dependent on the value of T_1, which prompts us to define the following critical values for T_1 (these depend on the values chosen for the control parameters):

$$\begin{array}{ll} T_c^{(0)} \approx 12.575: & \text{creation point of phase } SG_2 \\ T_c^{(1)} = 1.25: & \text{creation point of } T_{p,1}^{\rm 1st} \quad T_{p,1}^{\rm 2nd} = T_{p,1}^{\rm 1st} \\ T_c^{(2)} \approx 1.15: & T_{p,1}^{\rm 1st} \text{ and } T_{1,2}^{\rm 2nd} \text{ touch tangentially} \\ T_c^{(3)} \approx 1.06: & \text{creation point of } T_{p,2}^{\rm 1st} \quad T_{1,2}^{\rm 2nd} = T_{p,2}^{\rm 1st} \end{array} .$$

Note: $T_c^{(1)}$ is the point given by expression (52). In terms of these critical values we can classify the different transition scenarios encountered upon reducing T down to zero, starting in the paramagnetic region, as follows:

$$\begin{array}{lll} T_c^{(0)} < T_1: & P \to SG_1(\text{2nd order}) & \\ T_c^{(1)} < T_1 < T_c^{(0)}: & P \to SG_1(\text{2nd order}) & SG_1 \to SG_2(\text{2nd order}) \\ T_c^{(2)} < T_1 < T_c^{(1)}: & P \to SG_1(\text{1st order}) & SG_1 \to SG_2(\text{2nd order}) \\ T_c^{(3)} < T_1 < T_c^{(2)}: & P \to SG_1(\text{1st order}) & SG_1 \to SG_2(\text{2nd order}) \\ T_1 < T_c^{(3)}: & P \to SG_1(\text{1st order}) & SG_1 \to SG_2(\text{1st order}). \end{array}$$

In figure 3 we show the corresponding values of the order parameters along the various transition lines, as functions of T_1 (with line types identical to those used in figure 2 for the same transitions), except for those order parameters which are zero by definition (such as the values of q_2 when bifurcating continuously from zero). Curves in figures 2 and 3 which terminate for small but non-zero values of T_1 do so due to computational limitations; they can be shown to extend down to $T_1 = 0$. We will calculate the values of the order parameters in the $m_1 \to \infty$ (i.e. $T_1 \to 0$) limit analytically in a subsequent section.

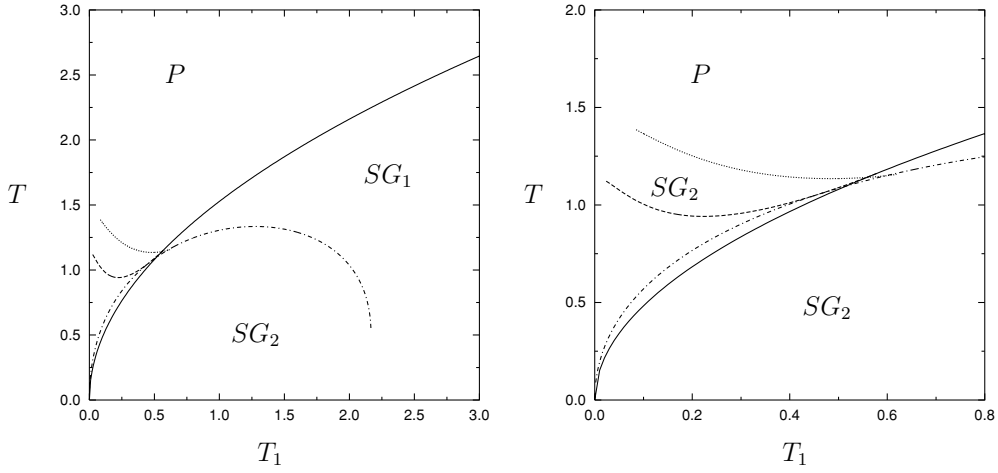

Figure 4. Phase diagram in the (T_1, T) plane for $L = 2$, $\epsilon_1/\mu_1, \epsilon_2/\mu_2 = 2$ and $m_2 = 1.5$, obtained by solving numerically the equations defining the various bifurcation lines (note: $T_1 = T/m_1$). Left panel: the large-scale structure. Right panel: enlargement of the low T_1 (or high m_1) region, where the first-order transitions are found. The relevant phases are P (paramagnetic phase, $q_1 = q_2 = 0$), SG_1 (first spin-glass phase, $q_1 > 0, q_2 = 0$), and SG_2 (second spin-glass phase, $q_1 > 0, q_2 > 0$). The transition lines shown, as defined in (45), are $T_{p,1}^{\rm 2nd}$ (solid curve), $T_{1,2}^{\rm 2nd}$ (dot-dash curve), $T_{p,1}^{\rm 1st}$ (dashed curve), and $T_{p,2}^{\rm 1st}$ (dotted curve). Note: the curve $T_{1,2}^{\rm 2nd}$ jumps to zero discontinuously at $T_1 = T_c^{(0)} \approx 2.17$ (see main text).

The general tendency for the system in the phase diagram, as expected, is to have more phases, increasingly discontinuous transitions separating them, and hence the most interesting physics, for small values of T_1 (i.e. large m_1). We should emphasize once more in this context that in our model we have the freedom to choose $m_1 > 1$ (or $T_1 < T$), in contrast to the standard Parisi solution for ordinary complex systems [12] (where one always has $m_1 \leqslant 1$).

Let us next turn to $m_2 = 1.5$, i.e. relatively low noise levels in the level-2 couplings compared to that of the spins. In figure 4 we show the phase diagram and the various bifurcation lines of the $L = 2$ system for $m_2 = 1.5$, as obtained by numerical solution of the relevant equations. It is again clear from the figure that upon lowering T one can find different types of transition sequences, dependent on the value of T_1. Now the relevant critical values for T_1 are as follows:

$T_c^{(0)} \approx 2.17$:	creation point of phase SG_2
$T_c^{(1)} \approx 0.597$:	creation point of $T_{p,2}^{\rm 1st}$ $\quad T_{1,2}^{\rm 2nd} = T_{p,2}^{\rm 1st}$
$T_c^{(2)} = 7/12 \approx 0.583$:	$T_{p,1}^{\rm 2nd} = T_{p,1}^{\rm 1st}$
$T_c^{(3)} \approx 0.5572$:	$T_{p,1}^{\rm 1st} = T_{p,2}^{\rm 1st}$.

Note: $T_c^{(2)}$ is the point given by expression (52). In terms of these critical values we can classify the different transition scenarios encountered upon reducing T, as before, as follows:

$T_c^{(0)} < T_1$:	$P \to SG_1$(2nd order)	
$T_c^{(1)} < T_1 < T_c^{(0)}$:	$P \to SG_1$(2nd order)	$SG_1 \to SG_2$(2nd order)
$T_c^{(2)} < T_1 < T_c^{(1)}$:	$P \to SG_1$(2nd order)	$SG_1 \to SG_2$(1st order)
$T_c^{(3)} < T_1 < T_c^{(2)}$:	$P \to SG_1$(1st order)	$SG_1 \to SG_2$(1st order)
$T_1 < T_c^{(3)}$:	$P \to SG_2$(1st order).	

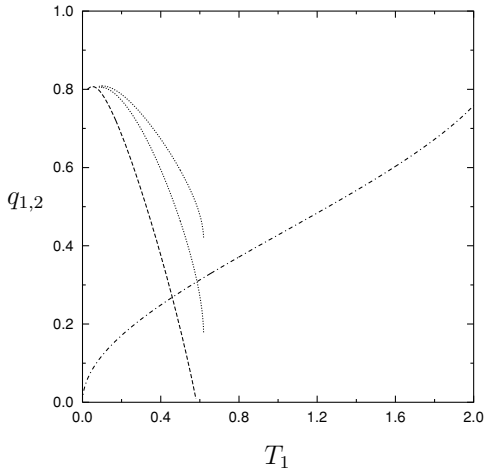

Figure 5. Values of the order parameters q_1 and q_2 as functions of T_1, along some of the transition lines of the previous figure, for $L = 2$, $\epsilon_1/\mu_1 = 1$, $\epsilon_2/\mu_2 = 2$ and $m_2 = 1.5$. The values shown are q_1 along $T_{1,2}^{\rm 2nd}$ (dot-dash curve), q_1 along $T_{p,1}^{\rm 1st}$ (dashed curve), and both q_1 and q_2 for $T_{p,2}^{\rm 1st}$ (dotted curves). Note: the order parameters which are not shown in this figure are zero by definition.

In figure 5 we show the corresponding values of the order parameters along the various transition lines, as functions of T_1 (with line types identical to those used in figure 4 for the same transitions), except for those which are zero by definition. Again, curves in figures 4 and 5 which terminate for small but non-zero values of T_1 do so due to computational limitations. The values of the order parameters in the $m_1 \to \infty$ (i.e. $T_1 \to 0$) limit will be calculated analytically in a subsequent section.

The main difference between $m_2 = 1.5$ and $m_2 = 0.5$ is that in the former case, for sufficiently low T_1, when lowering T the system goes straight from the paramagnetic state $q_1 = q_2 = 0$ into the SG_2 state $q_{1,2} \neq 0$, without an intermediate SG_1 phase. This is in agreement with the physical picture sketched so far: high values of the replica dimension m_2, corresponding to low coupling temperatures (here T_2), are found to induce more pronounced discontinuities, similar to the effect of choosing large values for m_1.

5.5. Comparison to the $L = 1$ system

It might be helpful at this stage to compare the phenomenology of the $L = 2$ model, as derived above, with that of the $L = 1$ model studied originally in [3, 4] (where there is just one interaction time-scale)[2]. For $L = 1$ the free energy per spin reduces to (24); there is only the order parameter q_1, and the possible phases are P and SG_1. One easily convinces oneself that all $L_1 = 1$ formulae can be obtained by simply putting $q_2 = 0$ in those derived for $L = 2$. In particular, q_1 is solved from (44), and the two remaining transition temperatures $T_{p,1}^{\rm 2nd}$ and $T_{p,1}^{\rm 1st}$ (which are derived from (44)) remain exactly as calculated earlier for $L = 2$, including the cross-over point at $m_1 = 2$. Hence the $L = 1$ phase diagram is obtained from the $L = 2$ phase diagram simply by elimination of the phase SG_2 and its associated transition lines. The result is shown in figure 6.

[2] Note that in [3, 4] and related papers the phases were shown in the (m_1, T) plane, whereas in this paper we have consistently opted for (T_1, T) phase diagrams, in view of the richest phenomenology occurring for small T_1.

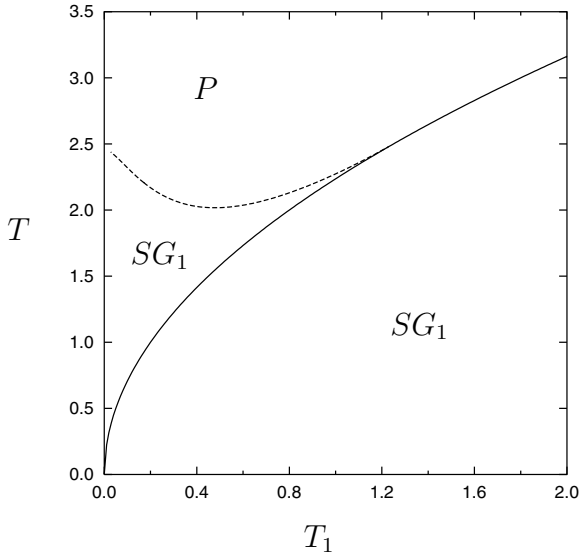

Figure 6. Phase diagram in the (T_1, T) plane for $(L = 1)$ and $\mu_1 = 1$, obtained by solving numerically the equations defining the bifurcation lines, to be compared with that of $L = 2$. The phases are P (paramagnetic phase, $q_1 = 0$) and SG_1 (spin-glass phase, $q_1 > 0$). The transition lines shown are $T_{p,1}^{\rm 2nd}$ (solid curve) and $T_{p,1}^{\rm 1st}$ (dashed curve).

Increasing the number of interaction time-scales is thus found to enable the creation of new phases, with new order parameters which measure the degree of freezing at even larger time-scales. However, although these newly created phases replace the previous phases in those regions of the phase diagram where the associated new order parameter is non-zero, they cannot alter the locations of the previous phases in those regions where the new order parameter remains zero (provided we keep the parameters $\{\phi_\ell\}$ unchanged). This is found to be true not only when comparing the $L = 1$ with the $L = 2$ case (where the novel phase is SG_2), but also when inspecting larger values of L; see e.g. the results for $L = 3$ in the next section. The explanation is reasonably clear. The phase SG_ℓ describes freezing at time-scale $\tau_{\ell-1}$; at this time-scale the system simply cannot feel the difference between truly frozen higher level bonds, and bonds which undergo (pseudo-) random motion at time-scales $\tau_\ell \gg \tau_{\ell-1}$.

6. The three-level hierarchy: $L = 3$

6.1. General properties

We now apply our results to the case $L = 3$, where we have three time-scales in the bond dynamics. Putting $L = 3$ in expression (25) for the free energy per spin gives:

$$f_{L=3}(q_1,q_2,q_3) = \frac{1}{2}\pi_1 q_1 - \frac{1}{\beta}\log 2 + \frac{1}{4}q_1^2\pi_1(m_1-1) + \frac{1}{4}q_2^2\pi_2 m_1(m_2-1) + \frac{1}{4}q_3^2\pi_3 m_1 m_2(m_3-1) - \frac{1}{\beta m_1 m_2 m_3}\log\int {\rm D}z_3\left\{\int {\rm D}z_2\left\{\int {\rm D}z_1\{\cosh[\Xi]\}^{m_1}\right\}^{m_2}\right\}^{m_3} \tag{53}$$

with

$$\Xi = z_3\sqrt{\beta q_3\pi_3} + z_2\sqrt{\beta(q_2\pi_2 - q_3\pi_3)} + z_1\sqrt{\beta(q_1\pi_1 - q_2\pi_2)}.$$

Due to the increased complexity of formulae induced by the nesting of integrals, it is, for $L = 3$, no longer helpful to write the saddle-point equations for the three order parameters $q_1 \geqslant q_2 \geqslant q_3$ in explicit form. However, in order to suppress notation in deriving the conditions for the various phase transitions, we will introduce the following abbreviations for the right-hand sides of the saddle-point equations (39) and (40):

$$q_1 = \left\langle\left\langle\langle\tanh^2\Xi\rangle_1\right\rangle_2\right\rangle_3 = \varphi_1(q_1, q_2, q_3)$$
$$q_2 = \left\langle\left\langle\{\langle\tanh\Xi\rangle_1\}^2\right\rangle_2\right\rangle_3 = \varphi_2(q_1, q_2, q_3)$$
$$q_3 = \left\langle\{\langle\langle\tanh\Xi\rangle_1\rangle_2\}^2\right\rangle_3 = \varphi_3(q_1, q_1, q_3).$$

For $L = 3$ we can distinguish between four phases. The first is a paramagnetic phase P, where $q_1 = q_2 = q_3 = 0$; this always solves (39) and (40), and according to (27), is the sole saddle-point of (53) in the high-temperature regime. In addition, we now have three spin-glass phases: SG_1, corresponding to a solution of the form $q_1 > q_2 = q_3 = 0$ (i.e. freezing of spins only), SG_2, corresponding to a solution $q_1 \geqslant q_2 > q_3 = 0$ (i.e. freezing of spins and couplings on time-scale τ_1 but not on the time-scale $\tau_2 \gg \tau_1$), and SG_3, corresponding to a solution $q_1 \geqslant q_2 \geqslant q_3 > 0$ (i.e. freezing of spins and couplings on time-scale τ_2). It immediately follows from the simple relation $f_{L=3}(q_1, q_2, 0) = f_{L=2}(q_1, q_2)$ that the phases SG_1 and SG_2 are fully identical to those of the $L = 2$ model, including the values of the order parameters and the locations of transition lines. However, there will now be new transition lines which are related to transitions into phase SG_3.

All partial derivatives of the type $\partial\varphi_\ell/\partial q_{\ell'}$ which occur in the transition conditions discussed below are calculated in detail in appendix D. In the various derivations we will save ink and paper by using the shorthand $\psi_\ell = \langle\cdots\langle\tanh\Xi\rangle_1\cdots\rangle_\ell$, e.g.

$$\psi_1 = \langle\tanh\Xi\rangle_1 \qquad \psi_2 = \langle\langle\tanh\Xi\rangle_1\rangle_2 \text{ etc.}$$

The result is

$$\frac{\partial\varphi_1}{\partial q_2} = \frac{\beta\pi_2}{2}m_1(m_2-1)\left\{-m_1m_2m_3\varphi_1\varphi_2 + m_1(m_2-2)\left\langle\left\langle\psi_1^2\langle\tanh^2\Xi\rangle_1\right\rangle_2\right\rangle_3 + m_1m_2(m_3-1)\left\langle\left\langle\langle\tanh^2\Xi\rangle_1\right\rangle_2\left\langle\psi_1^2\right\rangle_2\right\rangle_3 + 2(m_1-2)\left\langle\left\langle\psi_1\langle\tanh^3\Xi\rangle_1\right\rangle_2\right\rangle_3 + 4\varphi_2\right\} \tag{54}$$

$$\frac{\partial\varphi_2}{\partial q_1} = \frac{\beta\pi_1}{2}(m_1-1)\left\{-m_1m_2m_3\varphi_1\varphi_2 + m_1(m_2-2)\left\langle\left\langle\psi_1^2\langle\tanh^2\Xi\rangle_1\right\rangle_2\right\rangle_3 + m_1m_2(m_3-1) \times\left\langle\left\langle\langle\tanh^2\Xi\rangle_1\right\rangle_2\left\langle\psi_1^2\right\rangle_2\right\rangle_3 + 2(m_1-2)\left\langle\left\langle\psi_1\langle\tanh^3\Xi\rangle_1\right\rangle_2\right\rangle_3 + 4\varphi_2\right\} \tag{55}$$

$$\frac{\partial\varphi_1}{\partial q_1} = \frac{\beta\pi_1}{2}\left\{-(m_1-1)m_1m_2m_3(\varphi_1)^2 + m_1(m_1-1)(m_2-1)\left\langle\left\langle\langle\tanh^2\Xi\rangle_1^2\right\rangle_2\right\rangle_3 + (m_1-1)m_1m_2(m_3-1)\left\langle\left\langle\langle\tanh^2\Xi\rangle_1\right\rangle_2^2\right\rangle_3 + (m_1-2)(m_1-3)\left\langle\left\langle\langle\tanh^4\Xi\rangle_1\right\rangle_2\right\rangle_3 + 4(m_1-2)\varphi_1 + 2\right\} \tag{56}$$

$$\frac{\partial\varphi_2}{\partial q_2} = \frac{\beta\pi_2}{2}\Big\{-m_1^2m_2m_3(m_2-1)(\varphi_2)^2 + m_1^2(m_2-2)(m_2-3)\left\langle\left\langle\psi_1^4\right\rangle_2\right\rangle_3 + m_1^2m_2(m_2-1)(m_3-1)\left\langle\left\langle\psi_1^2\right\rangle_2^2\right\rangle_3 + 4m_1(m_2-2)\varphi_2 + 4m_1(m_1-1)(m_2-2) \times \left\langle\left\langle\psi_1^2\langle\tanh^2\Xi\rangle_1\right\rangle_2\right\rangle_3 + 2 + 4(m_1-1)\varphi_1 + 2(m_1-1)^2\left\langle\left\langle\langle\tanh^2\Xi\rangle_1^2\right\rangle_2\right\rangle_3\Big\} \quad (57)$$

$$\frac{\partial\varphi_3}{\partial q_1} = \frac{\beta\pi_1}{2}(m_1-1)\Big\{2m_1(m_2-1)\left\langle\psi_2\left\langle\langle\tanh^2\Xi\rangle_1\psi_1\right\rangle_2\right\rangle_3 + (m_3-2)m_1m_2\left\langle\left\langle\langle\tanh^2\Xi\rangle_1\right\rangle_2\psi_2^2\right\rangle_3 + 2(m_1-2)\left\langle\psi_2\left\langle\langle\tanh^3\Xi\rangle_1\right\rangle_2\right\rangle_3 + 4\varphi_3 - m_1m_2m_3\varphi_1\varphi_3\Big\} \quad (58)$$

$$\frac{\partial\varphi_3}{\partial q_2} = \frac{\beta\pi_2}{2}(m_2-1)\Big\{m_1^2m_2(m_3-2)\left\langle\left\langle\psi_1^2\right\rangle_2\psi_2^2\right\rangle_3 + 4m_1(m_1-1)\left\langle\psi_2\left\langle\psi_1\langle\tanh^2\Xi\rangle_1\right\rangle_2\right\rangle_3 + 2m_1^2(m_2-2)\left\langle\psi_2\left\langle\psi_1^3\right\rangle_2\right\rangle_3 + 4m_1\varphi_3 - m_1^2m_2m_3\varphi_2\varphi_3\Big\} \quad (59)$$

$$\frac{\partial\varphi_3}{\partial q_3} = \frac{\beta\pi_3}{2}\Big\{2 - (m_3-1)m_1^2m_2^2m_3\varphi_3^2 + m_1^2m_2^2(m_3-2)(m_3-3)\left\langle\psi_2^4\right\rangle_3 + 4(m_3-2)m_1m_2\varphi_3 + 4(m_1-1)\varphi_1 + 4m_1(m_2-1)\varphi_2 + 4(m_1-1)(m_3-2)m_1m_2\left\langle\left\langle\langle\tanh^2\Xi\rangle_1\right\rangle_2\psi_2^2\right\rangle_3 + 4m_1^2m_2(m_2-1)(m_3-2)\left\langle\left\langle\psi_1^2\right\rangle_2\psi_2^2\right\rangle_3 + 4m_1(m_1-1)(m_2-1)\left\langle\left\langle\langle\tanh^2\Xi\rangle_1\right\rangle_2\left\langle\psi_1^2\right\rangle_2\right\rangle_3 + 2(m_1-1)^2\left\langle\left\langle\langle\tanh^2\Xi\rangle_1\right\rangle_2^2\right\rangle_3 + 2m_1^2(m_2-1)^2\left\langle\left\langle\psi_1^2\right\rangle_2^2\right\rangle_3\Big\} \quad (60)$$

$$\frac{\partial\varphi_1}{\partial q_3} = \frac{\beta\pi_3}{2}m_1m_2(m_3-1)\Big\{4\varphi_3 - m_1m_2m_3\varphi_1\varphi_3 + 2m_1(m_2-1)\left\langle\psi_2\left\langle\langle\tanh^2\Xi\rangle_1\psi_1\right\rangle_2\right\rangle_3 + 2(m_1-2)\left\langle\psi_2\left\langle\langle\tanh^3\Xi\rangle_1\right\rangle_2\right\rangle_3 + (m_3-2)m_1m_2\left\langle\left\langle\langle\tanh^2\Xi\rangle_1\right\rangle_2\psi_2^2\right\rangle_3\Big\} \quad (61)$$

$$\frac{\partial\varphi_2}{\partial q_3} = \frac{\beta\pi_3}{2}m_2(m_3-1)\Big\{4m_1\varphi_3 - m_1^2m_2m_3\varphi_2\varphi_3 + 2m_1^2(m_2-2)\left\langle\psi_2\left\langle\psi_1^3\right\rangle_2\right\rangle_3 + 4m_1(m_1-1)\left\langle\psi_2\left\langle\psi_1\langle\tanh^2\Xi\rangle_1\right\rangle_2\right\rangle_3 + m_1^2(m_3-2)m_2\left\langle\left\langle\psi_1^2\right\rangle_2\psi_2^2\right\rangle_3\Big\}. \quad (62)$$

6.2. Transitions identical to those of $L = 2$

All transitions relating only to the three phases P, SG_1 and SG_2, as well as the second-order transition $SG_1 \to SG_2$, where $q_3 = 0$, must be identical to those derived in the previous section for $L = 2$. The general condition for such transitions is

$$0 = \left(\frac{\partial\varphi_1}{\partial q_1} - 1\right)\left(\frac{\partial\varphi_2}{\partial q_2} - 1\right) - \frac{\partial\varphi_1}{\partial q_2}\frac{\partial\varphi_2}{\partial q_1}. \quad (63)$$

This involves only the partial derivatives (54)–(57). Due to $q_3 = 0$ the averages $\langle\cdots\rangle_3$ drop out, and together with the identities

$$\left\langle\langle\tanh^2\Xi\rangle_1\right\rangle_2 = q_1 \qquad \left\langle\langle\tanh\Xi\rangle_1^2\right\rangle_2 = q_2 \qquad \left\langle\psi_1^2\right\rangle_2 = q_2 \qquad \psi_2 = 0 \quad (64)$$

the four relevant partial derivatives simplify considerably to

$$\frac{\partial\varphi_1}{\partial q_2} = \frac{\beta\pi_2}{2}m_1(m_2-1)\big\{(4 - m_1m_2q_1)q_2 + m_1(m_2-2)\left\langle\psi_1^2\langle\tanh^2\Xi\rangle_1\right\rangle_2 + 2(m_1-2)\left\langle\psi_1\langle\tanh^3\Xi\rangle_1\right\rangle_2\big\} \quad (65)$$

$$\frac{\partial\varphi_2}{\partial q_1} = \frac{\beta\pi_1}{2}(m_1-1)\big\{(4-m_1m_2q_1)q_2 + m_1(m_2-2)\left\langle\psi_1^2\langle\tanh^2\Xi\rangle_1\right\rangle_2 + 2(m_1-2)\left\langle\psi_1\langle\tanh^3\Xi\rangle_1\right\rangle_2\big\} \tag{66}$$

$$\frac{\partial\varphi_1}{\partial q_1} = \frac{\beta\pi_1}{2}\big\{2-(m_1-1)m_1m_2q_1^2 + m_1(m_1-1)(m_2-1)\left\langle\langle\tanh^2\Xi\rangle_1^2\right\rangle_2 + (m_1-2)(m_1-3)\left\langle\langle\tanh^4\Xi\rangle_1\right\rangle_2 + 4(m_1-2)q_1\big\} \tag{67}$$

$$\frac{\partial\varphi_2}{\partial q_2} = \frac{\beta\pi_2}{2}\Big\{-m_1^2m_2m_3(m_2-1)q_2^2 + m_1^2(m_2-2)(m_2-3)\left\langle\psi_1^4\right\rangle_2 + m_1^2m_2(m_2-1)(m_3-1)\left\langle\psi_1^2\right\rangle_2^2 + 4m_1(m_2-2)q_2 + 4m_1(m_1-1)(m_2-2) \times\left\langle\psi_1^2\langle\tanh^2\Xi\rangle_1\right\rangle_2 + 2 + 4(m_1-1)q_1 + 2(m_1-1)^2\left\langle\langle\tanh^2\Xi\rangle_1^2\right\rangle_2\Big\}. \tag{68}$$

For the $P \to SG_1$ transition and the continuous $SG_1 \to SG_2$ transition we have to put $q_2 = 0$, in (65)–(68). We then arrive, together with the dropping out of $\langle\cdots\rangle_2$ and the relations $\psi_1 = \psi_2 = 0$, $\langle\tanh^2\Xi\rangle_1 = q_1$, at the simple expressions $\partial\varphi_1/\partial q_2 = \partial\varphi_2/\partial q_1 = 0$ and

$$\frac{\partial\varphi_1}{\partial q_1} = \frac{\beta\pi_1}{2}\left\{2-m_1(m_1-1)q_1^2 + (m_1-2)(m_1-3)\langle\tanh^4\Xi\rangle_1 + 4(m_1-2)q_1\right\}$$

$$\frac{\partial\varphi_2}{\partial q_2} = \beta\pi_2[1+(m_1-1)q_1]^2.$$

Thus (63) gives us the transition conditions

$$\frac{\beta\pi_1}{2}\left\{2-m_1(m_1-1)q_1^2 + (m_1-2)(m_1-3)\langle\tanh^4\Xi\rangle_1 + 4(m_1-2)q_1\right\} = 1 \tag{69}$$

$$\beta\pi_2[1+(m_1-1)q_1]^2 = 1. \tag{70}$$

For $q_1 = 0$ (where also $\Xi = 0$) we recover from (69) the condition (47) for the line $T_{p,1}^{2\text{nd}}$, for $q_1 > 0$ equation (69) is seen to be identical to condition $T = \pi_1$ (which for $L = 2$ led to (48)) for the line $T_{p,1}^{1\text{st}}$, and (70) is identical to condition (49) for the line $T_{1,2}^{2\text{nd}}$. Finally, the condition for the first-order $SG_1 \to SG_2$ transition line $T_{1,2}^{1\text{st}}$ of $L = 2$ is recovered by combining the full expressions (65)–(68) with the bifurcation condition (63), as it should. Thus, from our $L = 3$ saddle-point equations we do indeed extract fully those transitions encountered earlier for $L = 2$, which do not involve the SG_3 phase.

6.3. Transitions to SG_3

The novel transitions induced by going from $L = 2$ to $L = 3$ are those where the new phase SG_3 is concerned. The condition for second-order $SG_2 \to SG_3$ transitions is simply given by $\partial\varphi_3/\partial q_3|_{q_3=0} = 1$. Inserting $q_3 = 0$ into (60), followed by usage of the simplifying relations (64) which follow from $q_3 = 0$, leads to

$$\frac{\partial\varphi_3}{\partial q_3}\Big|_{q_3=0} = \beta\pi_3[1+(m_1-1)q_1 + m_1(m_2-1)q_2]^2.$$

And the condition defining the critical temperature $T_{2,3}^{2\text{nd}}$ for the second-order transition $SG_2 \to SG_3$ is seen to be simply

$$\beta\pi_3[1+(m_1-1)q_1 + m_1(m_2-1)q_2]^2 = 1. \tag{71}$$

Finally, the most general bifurcation condition defining first-order $SG_2 \to SG_3$ transitions is given by

$$\begin{vmatrix} \partial\varphi_1/\partial q_1 - 1 & \partial\varphi_1/\partial q_2 & \partial\varphi_1/\partial q_3 \\ \partial\varphi_2/\partial q_1 & \partial\varphi_2/\partial q_2 - 1 & \partial\varphi_2/\partial q_3 \\ \partial\varphi_3/\partial q_1 & \partial\varphi_3/\partial q_2 & \partial\varphi_3/\partial q_3 - 1 \end{vmatrix} = 0 \tag{72}$$

where $q_1 \geqslant q_2 \geqslant q_3 > 0$. Here there are no further simplifying properties to be exploited, and hence (72) must be solved numerically with the full expressions (55)–(62) for the nine partial derivatives.

6.4. The $L = 3$ phase diagram

For $L = 3$ we have seven independent control parameters, namely T, m_1, m_2, m_3, π_1, π_2 and π_3 (where $\pi_3 = \epsilon_3/m_1m_2m_3\mu_3$, $\pi_2 = \pi_3 + \epsilon_2/m_1m_2\mu_2$ and $\pi_1 = \pi_2 + \epsilon_1/m_1\mu_1$), so we again have to resort to phase diagram cross-sections. As in the previous case $L = 2$ we will show the transition lines in the (T_1, T) plane, solved numerically from the various bifurcation conditions, now for the choice $\epsilon_1/\mu_1 = \epsilon_2/\mu_2 = 1$, $\epsilon_3/\mu_3 = 2$ with $m_2 \in \{0.5, 1.5\}$ and with $m_3 = 2$ (i.e. relatively low noise levels in the level-3 bonds). According to (47), for this choice of parameters the second-order $P \to SG_1$ transition condition $T_{p,1}^{\rm 2nd} = \pi_1$ will again reduce to

$$T_{p,1}^{\rm 2nd} = \frac{1}{m_1}\left(1 + \frac{2}{m_2}\right) \tag{73}$$

and is replaced by a first-order $P \to SG_1$ transition at $m_1 = 2$, or

$$T_{p,1}^{\rm 1st} = T_{p,2}^{\rm 2nd}: \quad T_1 = \frac{1}{4}\left(1 + \frac{2}{m_2}\right). \tag{74}$$

As in the case $L = 2$, we will generally find (see section 7) that for any L and for any $\ell \geqslant 1$ there are always non-zero critical values for T_1 above which $T_{\ell,\ell+1}^{\rm 2nd} = 0$. Hence all spin-glass phases $SG_{\ell>1}$ will at some finite point cease to exist as T_1 is increased. The values of the order parameters in the $m_1 \to \infty$ (i.e. $T_1 \to 0$) limit will be also be calculated analytically in section 7.

In figure 7 we show the phase diagram and the various bifurcation lines of the $L = 3$ system for $m_2 = 0.5$ (relatively high noise levels in the level-2 couplings), obtained by numerical solution of the relevant equations. The physical transitions are the ones which occur first as the temperature T is lowered from within the paramagnetic region. In the present case the different T_1-dependent types of transition sequences are separated by the following critical values for T_1:

$T_c^{(0)} \approx 12.575$:	creation point of phase SG_2
$T_c^{(1)} = 1.25$:	creation point of $T_{p,1}^{\rm 1st}$ $\quad T_{p,1}^{\rm 2nd} = T_{p,1}^{\rm 1st}$
$T_c^{(2)} \approx 1.15$:	$T_{p,1}^{\rm 1st}$ and $T_{1,2}^{\rm 2nd}$ touch tangentially
$T_c^{(3)} \approx 1.06$:	creation point of $T_{p,2}^{\rm 1st}$ $\quad T_{1,2}^{\rm 2nd} = T_{p,2}^{\rm 1st}$
$T_c^{(4)} \approx 0.692$:	creation point of phase SG_3
$T_c^{(5)} \approx 0.38$:	creation point of $T_{p,3}^{\rm 1st}$ $\quad T_{2,3}^{\rm 2nd} = T_{p,3}^{\rm 1st}$
$T_c^{(6)} \approx 0.325$:	$T_{p,2}^{\rm 1st} = T_{p,3}^{\rm 1st}$.

Note: $T_c^{(1)}$ is the point given by expression (74). In terms of these critical values we can classify the different transition scenarios encountered upon reducing T down to zero (and the orders of the transitions), starting in the paramagnetic region, as follows:

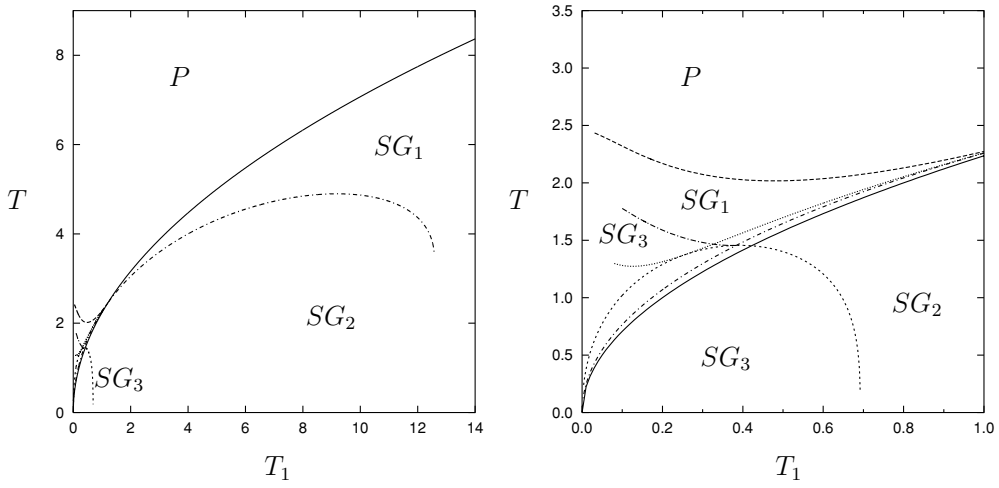

Figure 7. Phase diagram in the (T_1, T) plane for $L = 3$, $\epsilon_1/\mu_1 = \epsilon_2/\mu_2 = 1$, $\epsilon_3/\mu_3 = 2$, $m_3 = 2$ and $m_2 = 0.5$, obtained by solving numerically the equations defining the various bifurcation lines (note: $T_1 = T/m_1$). Left panel: the large-scale structure. Right panel: enlargement of the low T_1 (or high m_1) region, where the first-order transitions are found. The relevant phases are P (paramagnetic phase, $q_1 = q_2 = 0$), SG_1 (first spin-glass phase, $q_1 > q_2 = q_3 = 0$), SG_2 (second spin-glass phase, $q_1 \geqslant q_2 > q_3 = 0$), and SG_3 (third spin-glass phase, $q_1 \geqslant q_2 \geqslant q_3 > 0$). The transition lines shown are $T_{p,1}^{\rm 2nd}$ (solid curve), $T_{1,2}^{\rm 2nd}$ (dot-dash curve), $T_{p,1}^{\rm 1st}$ (dashed curve), $T_{p,2}^{\rm 1st}$ (dotted curve), $T_{2,3}^{\rm 2nd}$ (short dashes), and $T_{2,3}^{\rm 1st}$ (dot-long dash curve). Note: the curves $T_{1,2}^{\rm 2nd}$ and $T_{2,3}^{\rm 2nd}$ jump to zero discontinuously at $T_1 = T_c^{(0)} \approx 12.575$ and $T_1 = T_c^{(4)} \approx 0.692$, respectively (see main text).

$T_c^{(0)} < T_1$:	$P \to SG_1$(2nd)		
$T_c^{(1)} < T_1 < T_c^{(0)}$:	$P \to SG_1$(2nd)	$SG_1 \to SG_2$(2nd)	
$T_c^{(2)} < T_1 < T_c^{(1)}$:	$P \to SG_1$(1st)	$SG_1 \to SG_2$(2nd)	
$T_c^{(3)} < T_1 < T_c^{(2)}$:	$P \to SG_1$(1st)	$SG_1 \to SG_2$(2nd)	
$T_c^{(4)} < T_1 < T_c^{(3)}$:	$P \to SG_1$(1st)	$SG_1 \to SG_2$(1st)	
$T_c^{(5)} < T_1 < T_c^{(4)}$:	$P \to SG_1$(1st)	$SG_1 \to SG_2$(1st)	$SG_2 \to SG_3$(2nd)
$T_c^{(6)} < T_1 < T_c^{(5)}$:	$P \to SG_1$(1st)	$SG_1 \to SG_2$(1st)	$SG_2 \to SG_3$(1st)
$T_1 < T_c^{(6)}$:	$P \to SG_1$(1st)	$SG_1 \to SG_3$(1st).	

In figure 8 we show the corresponding values of the order parameters along the various transition lines, as functions of T_1 (with line types identical to those used in figure 7 for the same transitions), except for those order parameters which are zero by definition (such as the values at the transition of order parameters which bifurcate continuously from zero). Curves in figures 7 and 8 which terminate for small but non-zero values of T_1 do so due to computational limitations.

In figure 9 we show the phase diagram and the various bifurcation lines of the $L = 3$ system for $m_2 = 1.5$ (i.e. relatively low noise levels in the level-2 couplings), as obtained by numerical solution of the relevant equations. Now the relevant critical values for T_1 are as follows:

$T_c^{(0)} \approx 2.17$:	creation point of phase SG_2
$T_c^{(1)} \approx 1.3$:	creation point of phase SG_3, via $T_{2,3}^{\rm 1st}$

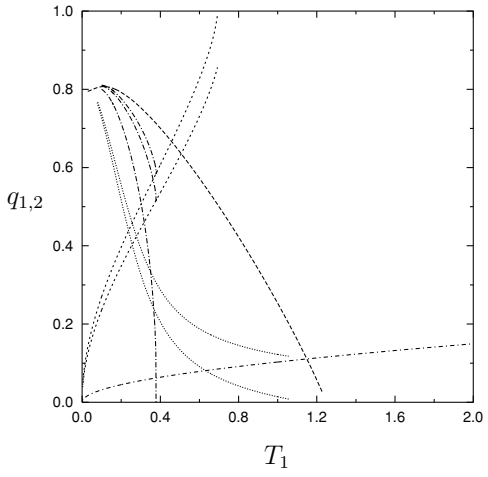

Figure 8. Values of the order parameters q_1 and q_2 as functions of T_1, along some of the transition lines of the previous figure, for $L = 3$, $\epsilon_1/\mu_1 = \epsilon_2/\mu_2 = 1$, $\epsilon_3/\mu_3 = 2$, $m_3 = 2$ and $m_2 = 0.5$. The values shown are q_1 along $T_{1,2}^{\rm 2nd}$ (dot-dash curve), q_1 along $T_{p,1}^{\rm 1st}$ (dashed curve), q_1 and q_2 along $T_{p,2}^{1st}$ (dotted curves), q_1 and q_2 along $T_{2,3}^{\rm 2nd}$ (short dashes), and q_1, q_2 and q_3 for $T_{2,3}^{\rm 1st}$ (dot-long dash curves) Note: the order parameters which are not shown in this figure are zero by definition.

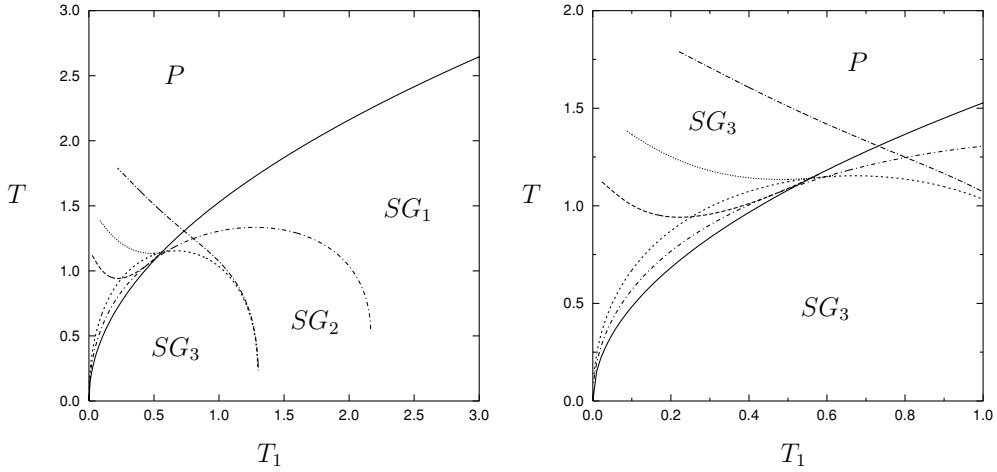

Figure 9. Phase diagram in the (T_1, T) plane for $L = 3$, $\epsilon_1/\mu_1 = \epsilon_2/\mu_2 = 1, \epsilon_3/\mu_3 = 2, m_3 = 2$ and $m_2 = 1.5$, obtained by solving numerically the equations defining the various bifurcation lines (note: $T_1 = T/m_1$). Left panel: the large-scale structure. Right panel: enlargement of the low T_1 (or high m_1) region, where the first-order transitions are found. The relevant phases are P (paramagnetic phase, $q_1 = q_2 = 0$), SG_1 (first spin-glass phase, $q_1 > q_2 = q_3 = 0$), SG_2 (second spin-glass phase, $q_1 \geqslant q_2 > q_3 = 0$), and SG_3 (third spin-glass phase, $q_1 \geqslant q_2 \geqslant q_3 > 0$). The transition lines shown are $T_{p,1}^{\rm 2nd}$ (solid curve), $T_{1,2}^{\rm 2nd}$ (dot-dash curve), $T_{p,1}^{\rm 1st}$ (dashed curve), $T_{p,2}^{\rm 1st}$ (dotted curve), $T_{2,3}^{\rm 2nd}$ (short dashes), and $T_{2,3}^{\rm 1st}$ (dot-long dash curve). Note: the curves $T_{1,2}^{\rm 2nd}$ and $T_{2,3}^{\rm 2nd}$ jump to zero discontinuously at $T_1 = T_c^{(0)} \approx 2.17$ and $T_1 = T_c^{(1)} \approx 1.3$, respectively (see main text).

$T_c^{(2)} \approx 1.3$:	creation point of $T_{2,3}^{\rm 2nd}$
$T_c^{(3)} \approx 0.8$:	$T_{1,2}^{\rm 2nd} = T_{2,3}^{\rm 1st}$
$T_c^{(4)} \approx 0.73$:	$T_{p,1}^{\rm 2nd} = T_{2,3}^{\rm 1st}$
$T_c^{(5)} \approx 0.597$:	creation point of $T_{p,2}^{\rm 1st}$ $\quad T_{1,2}^{\rm 2nd} = T_{p,2}^{\rm 1st}$
$T_c^{(6)} = 7/12 \approx 0.583$:	$T_{p,1}^{\rm 2nd} = T_{p,1}^{\rm 1st}$
$T_c^{(7)} \approx 0.5572$:	$T_{p,1}^{\rm 1st} = T_{p,2}^{\rm 1st}$.

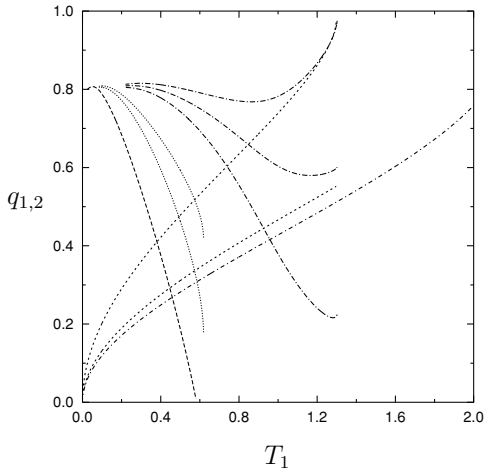

Figure 10. Values of the order parameters q_1 and q_2 as functions of T_1, along some of the transition lines of the previous figure, for $L = 3$, $\epsilon_1/\mu_1 = \epsilon_2/\mu_2 = 1$, $\epsilon_3/\mu_3 = 2$, $m_3 = 2$ and $m_2 = 1.5$. The values shown are q_1 along $T^{\rm 2nd}_{1,2}$ (dot-dash curve), q_1 along $T^{\rm 1st}_{p,1}$ (dashed curve), q_1 and q_2 along $T^{\rm 1st}_{p,2}$ (dotted curves), q_1 and q_2 along $T^{\rm 2nd}_{2,3}$ (short dashes), and q_1, q_2 and q_3 for $T^{\rm 1st}_{2,3}$ (dot-long dash curves) Note: the order parameters which are not shown in this figure are zero by definition.

Note: $T_c^{(2)} < T_c^{(1)}$. In terms of these critical values we can classify the different transition scenarios encountered upon reducing T, as before, as follows:

$T_c^{(0)} < T_1$:	$P \to SG_1$(2nd)		
$T_c^{(1)} < T_1 < T_c^{(0)}$:	$P \to SG_1$(2nd)	$SG_1 \to SG_2$(2nd)	
$T_c^{(2)} < T_1 < T_c^{(1)}$:	$P \to SG_1$(2nd)	$SG_1 \to SG_2$(2nd)	$SG_2 \to SG_3$(1st)
$T_c^{(3)} < T_1 < T_c^{(2)}$:	$P \to SG_1$(2nd)	$SG_1 \to SG_3$(1st)	
$T_1 < T_c^{(3)}$:	$P \to SG_3$(1st).		

In figure 10 we show the corresponding values of the order parameters along the various transition lines, as functions of T_1 (with line types identical to those used in figure 9 for the same transitions), except for those which are zero by definition. Again, curves in figures 9 and 10 which terminate for small but non-zero values of T_1 do so due to computational limitations.

The general picture for $L = 3$ is obviously more complicated but qualitatively similar to that which we arrived at for $L = 2$, with increasingly discontinuous and non-trivial phase behaviour of the system for decreasing relative noise levels of the couplings, whether measured by m_1, m_2 or m_3. For instance, for $L = 2$ we observed that for sufficiently large values of m_1 and m_2, upon lowering T the system goes straight from the paramagnetic state $q_1 = q_2 = 0$ into the SG_2 state $q_1 \geqslant q_2 > 0$, without an intermediate SG_1 phase. Here we see, similarly, that for sufficiently large values of m_1, m_2 and m_2, upon lowering T the system goes straight from the paramagnetic state $q_1 = q_2 = q_3 = 0$ into the SG_3 state $q_1 \geqslant q_2 \geqslant q_3 > 0$, without intermediate SG_1 or SG_2 phases. In the next section we will show that the jump of the $SG_\ell \to SG_{\ell+1}$ transitions, as observed in the $L = 2$ and $L = 3$ phase diagrams for small T, reflect non-physical re-entrance phenomena, which can be regarded as indicators that replica symmetry will be broken for values of $m_1 = T/T_1$ close to zero.

7. Asymptotic behaviour in L-level hierarchy

In this section we study the saddle-point equations of the general L-level hierarchy in the two limits $T_1 \to 0$ for fixed T (i.e. $m_1 \to \infty$, fully deterministic evolution of level-1 couplings) and $T \to 0$ for fixed T_1 (i.e. $m_1 \to 0$, fully deterministic evolution of spins). For the particular choices $L = 2$ and $L = 3$ this implies solving our model along the left and bottom boundaries in the phase diagrams of figures 2, 4, 7 and 9.

7.1. The limit $m_1 \to \infty$ in the L-level hierarchy

In this section we study the order parameter equations in the ℓth spin-glass phase SG_ℓ, characterized by $q_1 \geqslant \cdots \geqslant q_\ell > 0$ and $q_{\ell+1} = \cdots = q_L = 0$. We calculate the order parameters in the limit $m_1 \to \infty$ (deterministic evolution of level-1 couplings), for fixed β, and derive in this limit an expression for the first-order $P \to SG_\ell$ transition temperature $T_{p,\ell}^{1\mathrm{st}}$. Insertion of $q_{\ell+1} = q_{\ell+2} = \cdots = q_L = 0$ into expression (29) for the free energy per spin gives

$$f = \frac{1}{2}\pi_1 q_1 - \frac{1}{\beta}\log 2 + \frac{1}{4}\sum_{k=2}^{\ell} q_k^2 \pi_k m_1 m_2 \cdots m_{k-1}(m_k - 1) + \frac{1}{4}q_1^2\pi_1(m_1 - 1) - \left[\beta \prod_{k=1}^{\ell} m_k\right]^{-1} \log K_\ell \tag{75}$$

$$K_\ell = \int \mathrm{D}z_\ell \left\{ \int \mathrm{D}z_{\ell-1} \left\{ \cdots \int \mathrm{D}z_2 \left\{ \int \mathrm{D}z_1 \{\cosh \Xi\}^{m_1} \right\}^{m_2} \cdots \right\}^{m_{\ell-1}} \right\}^{m_\ell}$$

with

$$\Xi = \sum_{k=1}^{\ell} z_k a_k \qquad a_\ell = \sqrt{\beta q_\ell \pi_\ell} \qquad a_k = \sqrt{\beta(q_k\pi_k - q_{k+1}\pi_{k+1})} \qquad (k < \ell)$$

It is not *a priori* obvious how the various order parameters will scale with m_1 for $m_1 \to \infty$. Here we will make a scaling ansatz which we then show to lead self-consistently to solutions of our saddle-point equations, which satisfy all physical and mathematical requirements. We assume that $\Xi = \mathcal{O}(m_1^0)$ as $m_1 \to \infty$, and consequently put

$$a_k = \tilde{a}_k \zeta_k \qquad z_k = \tilde{z}_k \zeta_k^{-1} \qquad (k = 1, \ldots, \ell)$$

such that $\Xi = \sum_{k=1}^{\ell} \tilde{z}_k \tilde{a}_k$, with $\tilde{a}_k = \mathcal{O}(m_1^0)$ and $\tilde{z}_k = \mathcal{O}(m_1^0)$, and with scaling functions $\zeta_k = \zeta_k(m_1)$ which will be determined in due course. As a consequence, we can work out the term K_ℓ in (76) for large m_1 as follows:

$$K_\ell = \int \mathrm{D}z_\ell \left\{ \int \mathrm{D}z_{\ell-1} \left\{ \cdots \int \mathrm{D}z_2 \left\{ \int \frac{\mathrm{d}\tilde{z}_1}{\sqrt{2\pi}\zeta_1} \mathrm{e}^{(\zeta_1^2 m_1 \ln\cosh \Xi - \frac{1}{2}\tilde{z}_1^2)/\zeta_1^2} \right\}^{m_2} \cdots \right\}^{m_{\ell-1}} \right\}^{m_\ell}.$$

This expression immediately suggests that the appropriate $m_1 \to \infty$ scaling is obtained upon choosing $\zeta_1 = m_1^{-1/2}$. As a result of this choice, the $\tilde{z}_1$ integration can, for $m_1 \to \infty$ in leading order, be carried out by steepest descent, and we arrive at

$$K_\ell \simeq \int \mathrm{D}z_\ell \left\{ \int \mathrm{D}z_{\ell-1} \left\{ \cdots \left\{ \int \mathrm{D}z_2 \, \mathrm{e}^{m_1 m_2 k_1(\tilde{z}_1^0, \tilde{z}_2, \ldots, \tilde{z}_\ell)} \right\}^{m_3} \cdots \right\}^{m_{\ell-1}} \right\}^{m_\ell}$$

$$k_1(\tilde{z}_1, \tilde{z}_2, \ldots, \tilde{z}_\ell) = \ln\cosh \Xi - \tfrac{1}{2}\tilde{z}_1^2$$

where $\tilde{z}_1^0$ denotes the saddle-point of k_1 with respect to variation of $\tilde{z}_1$, i.e.

$$\frac{\partial k_1}{\partial \tilde{z}_1}\Big|_{(\tilde{z}_1^0,\tilde{z}_2,\ldots,\tilde{z}_\ell)} = 0: \quad \tilde{z}_1^0 = \tilde{a}_1 \tanh \Xi\left(\tilde{z}_1^0, \tilde{z}_2, \ldots, \tilde{z}_\ell\right).$$

Hence $\tilde{z}_1^0$ is a function of $\tilde{z}_2, \ldots, \tilde{z}_\ell$. We now see that one can repeat this procedure and carry out all the Gaussian integrations iteratively by steepest descent, upon setting $\zeta_k = (m_1 m_2 \cdots m_k)^{-1/2}$, which gives in leading order

$$K_\ell \simeq \mathrm{e}^{m_1 m_2 \cdots m_{\ell-1} m_\ell k_\ell(\tilde{z}_1^0, \tilde{z}_2^0, \ldots, \tilde{z}_{\ell-1}^0, \tilde{z}_\ell^0)} \tag{76}$$

in which the functions $k_j(\cdots)$ are defined recursively via

$$k_j\left(\tilde{z}_1^0, \ldots, \tilde{z}_{j-1}^0, \tilde{z}_j, \ldots, \tilde{z}_\ell\right) = k_{j-1}\left(\tilde{z}_1^0, \ldots, \tilde{z}_{j-1}^0, \tilde{z}_j, \ldots, \tilde{z}_\ell\right) - \tfrac{1}{2}\tilde{z}_j^2 \quad (j > 1)$$

$$k_1\left(\tilde{z}_1, \ldots, \tilde{z}_\ell\right) = \ln\cosh \Xi\left(\tilde{z}_1, \ldots, \tilde{z}_\ell\right) - \tfrac{1}{2}\tilde{z}_1^2$$

giving for $j > 1$:

$$k_j\left(\tilde{z}_1^0, \ldots, \tilde{z}_{j-1}^0, \tilde{z}_j, \ldots, \tilde{z}_\ell\right) = \ln\cosh \Xi\left(\tilde{z}_1^0, \ldots, \tilde{z}_{j-1}^0, \tilde{z}_j, \ldots, \tilde{z}_\ell\right) - \frac{1}{2}\tilde{z}_j^2 - \frac{1}{2}\sum_{k=1}^{j-1}\left(\tilde{z}_k^0\right)^2. \tag{77}$$

The identity

$$\frac{\partial}{\partial \tilde{z}_j} k_j\left(\tilde{z}_1^0, \ldots, \tilde{z}_{j-1}^0, \tilde{z}_j, \ldots, \tilde{z}_\ell\right) = \tilde{a}_j \tanh \Xi\left(\tilde{z}_1^0, \ldots, \tilde{z}_{j-1}^0, \tilde{z}_j, \ldots, \tilde{z}_\ell\right) - \tilde{z}_j$$

shows that the values of all $\left\{\tilde{z}_j^0\right\}$ are ultimately to be solved from the following ℓ coupled saddle-point equations:

$$\tilde{z}_j^0 = \tilde{a}_j \tanh \Xi\left(\tilde{z}_1^0, \ldots, \tilde{z}_\ell^0\right) \qquad (j = 1, \ldots, \ell). \tag{78}$$

We insert (76) and (77) into (75), and find that for $m_1 \to \infty$ the free energy per spin reduces in leading order simply to

$$f \simeq \frac{1}{2}\pi_1 q_1 - \frac{1}{\beta}\log 2 + \frac{1}{4}\sum_{k=2}^{\ell} q_k^2 \pi_k m_1 m_2 \cdots m_{k-1}(m_k - 1)$$
$$+ \frac{1}{4}q_1^2 \pi_1 (m_1 - 1) - \beta^{-1}\left[\ln\cosh \Xi\left(\tilde{z}_1^0, \ldots, \tilde{z}_\ell^0\right) - \frac{1}{2}\sum_{k=1}^{\ell}\left(\tilde{z}_k^0\right)^2\right] \tag{79}$$

with $\Xi\left(\tilde{z}_1^0, \ldots, \tilde{z}_\ell^0\right) = \sum_{k=1}^{\ell} \tilde{a}_k \tilde{z}_k^0$.

We can now derive the saddle-point equations for $m_1 \to \infty$ by variation of (79) with respect to the order parameters $\{q_k\}$. Let us first derive the equation for q_1 in leading order:

$$0 = \frac{1}{2}\pi_1 + \frac{1}{2}q_1 \pi_1 (m_1 - 1) - \frac{1}{\beta}\sum_{k=1}^{\ell}\frac{\partial \tilde{z}_k^0}{\partial q_1}\frac{\partial}{\partial \tilde{z}_k^0}\left[\ln\cosh \Xi - \frac{1}{2}\sum_{k=1}^{\ell}\left(\tilde{z}_k^0\right)^2\right]$$
$$- \frac{1}{\beta}\sum_{k=1}^{\ell}\tanh \Xi \frac{\partial \Xi}{\partial \tilde{a}_k}\frac{\partial \tilde{a}_k}{\partial q_1} + \cdots \qquad (m_1 \to \infty).$$

The term with the partial derivatives $\partial/\partial \tilde{z}_k^0$ is identical zero due to the saddle-point equations (78), and hence, upon working out the derivatives $\partial \tilde{a}_k/\partial q_1 = \zeta_k^{-1}\partial a_k/\partial q_1$ and using (78) to eliminate occurrences of $\tilde{z}_1^0$, we just retain in leading order

$$0 = \tfrac{1}{2}\pi_1 m_1\left[q_1 - \tanh^2 \Xi\right] + \cdots \qquad (m_1 \to \infty).$$

Hence

$$q_1 = \tanh^2 \Xi \qquad (m_1 \to \infty).$$

Similarly we can deal with the derivatives of f (79) with respect to the other q_j. With the simple identity

$$a_k \frac{\partial a_k}{\partial q_j} = \frac{1}{2}\beta \frac{\partial}{\partial q_j}[q_k \pi_k - q_{k+1}\pi_{k+1}] = \frac{1}{2}\beta \pi_j [\delta_{j,k} - \delta_{j,k+1}]$$

the saddle-point equation for q_j with $j > 1$ is seen to become in leading order

$$\begin{aligned} 0 = \frac{1}{2} q_j \pi_j m_1 \cdots m_{j-1}(m_j - 1) - \frac{1}{\beta} \sum_{k=1}^{\ell} \frac{\partial \tilde{z}_k^0}{\partial q_j} \frac{\partial}{\partial \tilde{z}_k^0} \left[\ln \cosh \Xi - \frac{1}{2} \sum_{k=1}^{\ell} \left(\tilde{z}_k^0\right)^2 \right] \\ - \frac{1}{\beta} \sum_{k=1}^{\ell} \tanh \Xi \frac{\partial \Xi}{\partial \tilde{a}_k} \frac{\partial \tilde{a}_k}{\partial q_j} + \cdots \qquad (m_1 \to \infty) \\ = \frac{1}{2} q_j \pi_j m_1 \cdots m_{j-1}(m_j - 1) - \frac{1}{\beta} \tanh^2 \Xi \sum_{k=1}^{\ell} m_1 \cdots m_k a_k \frac{\partial a_k}{\partial q_j} + \cdots \\ = \frac{1}{2} \pi_j m_1 m_2 \cdots m_{j-1}(m_j - 1) \left[q_j - \tanh^2 \Xi \right] + \cdots . \end{aligned}$$

Hence, for all $j \geqslant 1$ one simply has

$$q_j = q^\star = \tanh^2 \Xi \qquad (m_1 \to \infty). \tag{80}$$

The final stage of our argument is to determine the amplitude $q^\star$ in (80). We note that, by virtue of (78), we can write $\Xi = \sum_{k=1}^{\ell} \tilde{a}_k \tilde{z}_k^0$ in leading order as $\Xi \simeq Y \tanh \Xi$, where $Y = \sum_{k=1}^{\ell} \tilde{a}_k^2$. It now follows, upon inserting (80) into this relation, that $q^\star$ is the solution of

$$q^\star = \tanh^2[Y\sqrt{q^\star}] \qquad (m_1 \to \infty). \tag{81}$$

In working out the quantity Y in leading order for $m_1 \to \infty$ we have to take care not to forget about the m_1 dependence of the factors π_k, defined in (23). In particular,

$$\pi_j \zeta_j^{-2} = \frac{\epsilon_j}{\mu_j} + \sum_{k=j+1}^{L} \frac{1}{m_{j+1} \cdots m_k} \frac{\epsilon_k}{\mu_k}.$$

Using this relation we find

$$\begin{aligned} Y = \beta q^\star \left[\sum_{j=1}^{\ell-1} (\pi_j - \pi_{j+1}) \zeta_j^{-2} + \pi_\ell \zeta_\ell^{-2} \right] + \cdots \qquad (m_1 \to \infty) \\ = \beta q^\star \left[\sum_{j=1}^{\ell} \frac{\epsilon_j}{\mu_j} + \sum_{j=\ell+1}^{L} \frac{1}{m_{\ell+1} \cdots m_j} \frac{\epsilon_j}{\mu_j} \right] + \cdots . \end{aligned}$$

Our final result on the limit $m_1 \to \infty$ is the following: in the phase SG_ℓ one has $q_1 = \cdots = q_\ell = q^\star > 0$ and $q_{\ell+1} = \cdots = q_L = 0$, where $q^\star$ is the solution of

$$q^\star = \tanh^2 \left[\beta \bar{\omega}_\ell (q^*)^{3/2} \right] \tag{82}$$

$$\bar{\omega}_\ell = \sum_{j=1}^{\ell} \frac{\epsilon_j}{\mu_j} + \sum_{j=\ell+1}^{L} \frac{1}{m_{\ell+1} \cdots m_j} \frac{\epsilon_j}{\mu_j}. \tag{83}$$

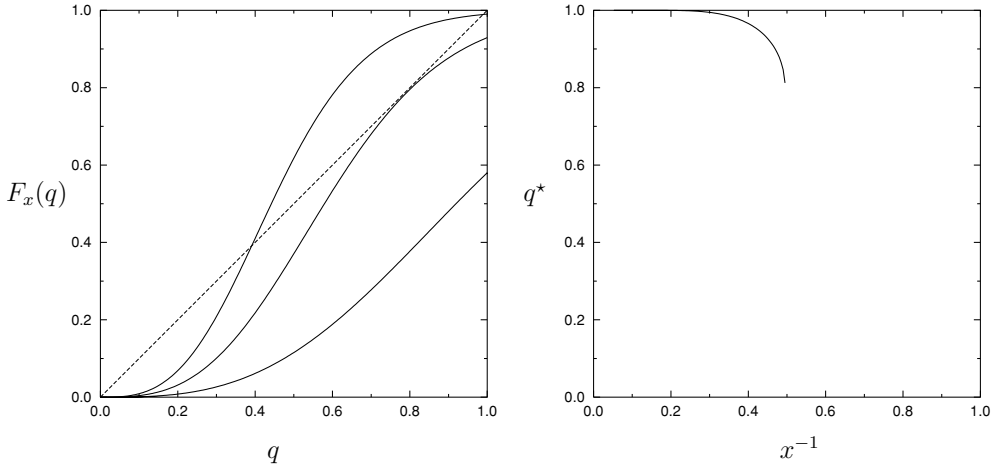

Figure 11. Properties of the equation $q = F_x(q)$, where $F_x(q) = \tanh^2(xq^{\frac{3}{2}})$, from which we solve the $m_1 \to \infty$ order parameter amplitude $q^\star$ (see main text). Left panel: shape of the function $F_x(q)$ for $x \in \{1, 2, 3\}$ (solid curves, from bottom to top), together with the diagonal (dashed) whose intersection with $F_x(q)$ defines $q^\star$. Right panel: largest solution $q^\star$ of the equation $q = F_x(q)$ as a function of x^{-1}. The jump occurs at $q^\star \approx 0.7911$ and $x^{-1} \approx 0.4958$.

We note that, for any $\bar{\omega}_\ell > 0$, equation (82) will have a non-zero solution (and hence phase SG_ℓ will indeed exist) when T is sufficiently low, and that $\lim_{T\to 0} q^\star = 1$. The behaviour of equation (82) as a function of the effective control parameter $x = \beta\bar{\omega}_\ell$ is illustrated in figure 11.

From (82) and (83) we can immediately extract the critical temperature $T^{1st}_{p,\ell}$, signalling the (first-order) transition $P \to SG_\ell$. The condition for non-zero solutions of (82) to bifurcate follows from solving (82) simultaneously with the equation $1 = \frac{\mathrm{d}}{\mathrm{d}q^\star}\tanh^2\left[\beta\bar{\omega}_\ell(q^*)^{3/2}\right]$, which leads us to the explicit and simple result

$$T^{1st}_{p,\ell} = 3\bar{\omega}_\ell q^\star(1-q^\star) \qquad q^\star = \tanh^2\left[\frac{\sqrt{q^\star}}{3(1-q^\star)}\right]. \tag{84}$$

Numerical solution of (84) shows that $q^\star \approx 0.7911$ and $T^{1st}_{p,\ell}/\bar{\omega}_\ell \approx 0.4958$. In order to assess the dependence on ℓ of the transition temperatures $T^{1st}_{p,\ell}$, we subtract

$$\begin{aligned} T^{1st}_{p,\ell+1} - T^{1st}_{p,\ell} &= 3q^\star(1-q^\star)[\bar{\omega}_{\ell+1} - \bar{\omega}_\ell] \\ &= 3q^\star(1-q^\star)\left[\frac{m_{\ell+1}-1}{m_{\ell+1}}\right]\left[\frac{\epsilon_{\ell+1}}{\mu_{\ell+1}} + \sum_{j=\ell+2}^{L}\frac{1}{m_{\ell+2}\cdots m_j}\frac{\epsilon_j}{\mu_j}\right]. \end{aligned}$$

Thus $T^{1st}_{p,\ell+1} > T_{p,\ell}$ for $m_{\ell+1} > 1$ (i.e. for relatively low noise levels in the level-$(\ell+1)$ couplings), whereas $T^{1st}_{p,\ell+1} < T_{p,\ell}$ for $m_{\ell+1} < 1$ (for relatively high noise levels in the level-$(\ell+1)$ couplings). Upon inserting the appropriate values of the control parameters, one can easily verify that the general results obtained in this section regarding the values of the order parameters in the phase SG_ℓ and of the $P \to SG_\ell$ transition temperatures $T^{1st}_{p,\ell}$ for arbitrary L are in perfect agreement with our solutions as calculated for the special choices $L = 2$ and $L = 3$ in previous sections.

7.2. The limit $m_1 \to 0$ in the L-level hierarchy

We next turn to the limit $m_1 \to 0$, for fixed T_1 (i.e. $T \to 0$). We observed in the phase diagrams of figures 2, 4, 7 and 9 for $L = 2$ and $L = 3$ that along the line $T = 0$ one finds critical values $T^{\rm 2nd}_{\ell-1,\ell}$ for T_1 which signal the creation of phase SG_ℓ (where $q_1 \geqslant \cdots \geqslant q_\ell > 0$ and $q_{\ell+1} = \cdots = q_L = 0$) from $SG_{\ell-1}$ (where $q_1 \geqslant \cdots \geqslant q_{\ell-1} > 0$ and $q_\ell = \cdots = q_L = 0$). In this section we calculate these critical values for the general L-level system. Our starting point is again expression (75) for the free energy per spin (obtained by putting $q_{\ell+1} = \cdots = q_L = 0$). In order to make all occurrences of m_1 explicit, we substitute $T = T_1 m_1$, and write the factors π_ℓ of (23) as

$$\pi_\ell = \frac{\tilde\pi_\ell}{m_1}: \quad \tilde\pi_1 = \frac{\epsilon_1}{\mu_1} + \sum_{\ell'=2}^{L} \frac{1}{m_2 \cdots m_{\ell'}} \frac{\epsilon_{\ell'}}{\mu_{\ell'}} \qquad \tilde\pi_{\ell>1} = \sum_{\ell'=\ell}^{L} \frac{1}{m_2 \cdots m_{\ell'}} \frac{\epsilon_{\ell'}}{\mu_{\ell'}}$$

so that $\tilde\pi_\ell = \mathcal{O}(m_1^0)$ for all ℓ as $m_1 \to 0$. This gives

$$m_1 f(q_1, \ldots, q_\ell) = \frac{1}{4}\tilde\pi_1 - \frac{1}{4}\tilde\pi_1 (q_1 - 1)^2 + \mathcal{O}\left(m_1^2\right) - \frac{T_1 m_1}{\prod_{k=2}^{\ell} m_k} \log K_\ell$$
$$+ \frac{1}{4} m_1 \left[q_1^2 \tilde\pi_1 + q_2^2 \tilde\pi_2 (m_2 - 1) + \sum_{k=3}^{\ell} q_k^2 \tilde\pi_k m_2 \cdots m_{k-1} (m_k - 1) \right] \tag{85}$$

with

$$K_\ell = \int {\rm D}z_\ell \left\{ \int {\rm D}z_{\ell-1} \left\{ \cdots \int {\rm D}z_2 \left\{ \int {\rm D}z_1 \{\cosh \Xi\}^{m_1} \right\}^{m_2} \cdots \right\}^{m_{\ell-1}} \right\}^{m_\ell}$$
$$\Xi = m_1^{-1} \tilde\Xi \qquad \tilde\Xi = \sum_{k=1}^{\ell} z_k \tilde a_k$$
$$\tilde a_\ell = \sqrt{T_1^{-1} q_\ell \tilde\pi_\ell} \qquad \tilde a_{k<\ell} = \sqrt{T_1^{-1}(q_k \tilde\pi_k - q_{k+1}\tilde\pi_{k+1})}.$$

Since $\tilde\Xi = \mathcal{O}(m_1^0)$ as $m_1 \to 0$, by virtue of our definitions, we may conclude that $K_\ell = \mathcal{O}(m_1^0)$:

$$\{\cosh \Xi\}^{m_1} = 2^{-m_1} [{\rm e}^{\tilde\Xi/m_1} + {\rm e}^{-\tilde\Xi/m_1}]^{m_1} = {\rm e}^{|\tilde\Xi|}(1 + \mathcal{O}(m_1)) \qquad (m_1 \to 0)$$

$$K_\ell = \int {\rm D}z_\ell \left\{ \int {\rm D}z_{\ell-1} \left\{ \cdots \int {\rm D}z_2 \left\{ \int {\rm D}z_1 \, {\rm e}^{|\tilde\Xi|} \right\}^{m_2} \cdots \right\}^{m_{\ell-1}} \right\}^{m_\ell} + \cdots. \tag{86}$$

Similarly one gets in leading order for $m_1 \to 0$:

$$\frac{1}{K_1} \int {\rm D}z_1 \{\cosh \Xi\}^{m_1} \tanh \Xi = \frac{\int {\rm D}z_1 {\rm e}^{|\tilde\Xi|} {\rm sgn}(\tilde\Xi)}{\int {\rm D}z_1 {\rm e}^{|\tilde\Xi|}} + \cdots. \tag{87}$$

According to (85) the saddle-point value of q_1 is of the form $q_1 = 1 - \mathcal{O}(\sqrt{m_1})$. If, however, we substitute $q_1 = 1 - \kappa\sqrt{m_1} + \mathcal{O}(m_1)$ into (85), we find that the saddle-point obeys $\kappa = 0$, hence the true scaling of q_1 is

$$q_1 = 1 - \kappa m_1 + \mathcal{O}\left(m_1^2\right). \tag{88}$$

Insertion of (88) into (85) gives

$$m_1 f(q_2, \ldots, q_\ell) = \frac{1}{4}(1 + m_1)\tilde\pi_1 + \mathcal{O}\left(m_1^2\right) - \frac{T_1 m_1}{\prod_{k=2}^{\ell} m_k} \log K_\ell|_{q_1=1}$$
$$+ \frac{1}{4} m_1 \left[q_2^2 \tilde\pi_2 (m_2 - 1) + \sum_{k=3}^{\ell} q_k^2 \tilde\pi_k m_2 \cdots m_{k-1} (m_k - 1) \right] \tag{89}$$

and hence the $m_1 \to 0$ saddle-point equations for the order parameters $q_{\ell'}$ with $1 < \ell' \leqslant \ell$ are given by

$$q_2 = \frac{2T_1}{\tilde{\pi}_2(m_2-1)\left(\prod_{k=2}^{\ell} m_k\right)} \frac{\partial}{\partial q_2} \log \tilde{K}_\ell \tag{90}$$

$$q_{\ell'} = \frac{2T_1}{\tilde{\pi}_{\ell'}(m_{\ell'}-1)\left(\prod_{k=2}^{\ell'-1} m_k\right)\left(\prod_{k=2}^{\ell} m_k\right)} \frac{\partial}{\partial q_{\ell'}} \log \tilde{K}_\ell \tag{91}$$

with $\tilde{K}_\ell(q_2,\ldots,q_\ell) = \lim_{q_1\to 1}\lim_{m_1\to 0} K_\ell$ (as given by (86)). At this stage it is advantageous to use our earlier results (39) and (40), which in the present context and in combination with (87) translate into

$$q_2 = \left\langle \cdots \left\langle \left[\frac{\int \mathrm{D}z_1 \mathrm{e}^{|\tilde{\Xi}|}\mathrm{sgn}(\tilde{\Xi})}{\int \mathrm{D}z_1 \mathrm{e}^{|\tilde{\Xi}|}} \right]^2 \right\rangle_2 \cdots \right\rangle_\ell \tag{92}$$

$$q_{k>2} = \left\langle \cdots \left\langle \left[\left\langle \cdots \left\langle \left[\frac{\int \mathrm{D}z_1 \mathrm{e}^{|\tilde{\Xi}|}\mathrm{sgn}(\tilde{\Xi})}{\int \mathrm{D}z_1 \mathrm{e}^{|\tilde{\Xi}|}} \right] \right\rangle_2 \cdots \right\rangle_{k-1} \right]^2 \right\rangle_k \cdots \right\rangle_\ell \tag{93}$$

with

$$\tilde{K}_1 = \int \mathrm{D}z_1\, \mathrm{e}^{|\tilde{\Xi}|} \qquad \tilde{K}_{\ell'} = \int \mathrm{D}z_{\ell'} \tilde{K}_{\ell'-1}^{m_{\ell'}}$$

$$\langle f\rangle_{\ell'} = \tilde{K}_{\ell'}^{-1} \int \mathrm{D}z_{\ell'} \tilde{K}_{\ell'-1}^{m_{\ell'}} f(z_{\ell'}) \qquad \tilde{\Xi} = \sum_{k=1}^{\ell} z_k \tilde{a}_k|_{q_1=1} = z_1\tilde{a}_1 + \tilde{\Xi}_1.$$

Working out the z_1 integrations in (92) and (93) leads to

$$\frac{\int \mathrm{D}z_1\, \mathrm{e}^{|\tilde{\Xi}|}\mathrm{sgn}(\tilde{\Xi})}{\int \mathrm{D}z_1\, \mathrm{e}^{|\tilde{\Xi}|}} = \frac{\sinh(\tilde{\Xi}_1) + \frac{1}{2}\mathrm{e}^{\tilde{\Xi}_1}\mathrm{Erf}\left[\frac{\tilde{a}_1^2+\tilde{\Xi}_1}{\tilde{a}_1\sqrt{2}}\right] - \frac{1}{2}\mathrm{e}^{-\tilde{\Xi}_1}\mathrm{Erf}\left[\frac{\tilde{a}_1^2-\tilde{\Xi}_1}{\tilde{a}_1\sqrt{2}}\right]}{\cosh(\tilde{\Xi}_1) + \frac{1}{2}\mathrm{e}^{\tilde{\Xi}_1}\mathrm{Erf}\left[\frac{\tilde{a}_1^2+\tilde{\Xi}_1}{\tilde{a}_1\sqrt{2}}\right] + \frac{1}{2}\mathrm{e}^{-\tilde{\Xi}_1}\mathrm{Erf}\left[\frac{\tilde{a}_1^2-\tilde{\Xi}_1}{\tilde{a}_1\sqrt{2}}\right]}. \tag{94}$$

The transition value $T_{\ell-1,\ell}^{\mathrm{2nd}}$ for T_1 is the one which gives a second-order bifurcation of q_ℓ away from zero in (93).

Let us work out these results first for $\ell = 2$. Here one simply has $\tilde{a}_1 = (\sqrt{\tilde{\pi}_1 - q_2\tilde{\pi}_2})/\sqrt{T_1}$ and $\tilde{\Xi}_1 = z_2\sqrt{q_2\tilde{\pi}_2}/\sqrt{T_1}$, with q_2 to be solved from

$$q_2 = \left\langle \left[\frac{\sinh(\tilde{\Xi}_1) + \frac{1}{2}\mathrm{e}^{\tilde{\Xi}_1}\mathrm{Erf}\left[\frac{\tilde{a}_1^2+\tilde{\Xi}_1}{\tilde{a}_1\sqrt{2}}\right] - \frac{1}{2}\mathrm{e}^{-\tilde{\Xi}_1}\mathrm{Erf}\left[\frac{\tilde{a}_1^2-\tilde{\Xi}_1}{\tilde{a}_1\sqrt{2}}\right]}{\cosh(\tilde{\Xi}_1) + \frac{1}{2}\mathrm{e}^{\tilde{\Xi}_1}\mathrm{Erf}\left[\frac{\tilde{a}_1^2+\tilde{\Xi}_1}{\tilde{a}_1\sqrt{2}}\right] + \frac{1}{2}\mathrm{e}^{-\tilde{\Xi}_1}\mathrm{Erf}\left[\frac{\tilde{a}_1^2-\tilde{\Xi}_1}{\tilde{a}_1\sqrt{2}}\right]} \right]^2 \right\rangle_2$$

$$= \left\langle \tilde{\Xi}_1^2 \left[\frac{1 + \mathrm{Erf}\left[\frac{\tilde{a}_1}{\sqrt{2}}\right] + \frac{\sqrt{2}}{\tilde{a}_1\sqrt{\pi}}\mathrm{e}^{-\frac{1}{2}\tilde{a}_1^2} + \mathcal{O}(\tilde{\Xi}_1)}{1 + \mathrm{Erf}\left[\frac{\tilde{a}_1}{\sqrt{2}}\right]} \right]^2 \right\rangle_2 \tag{95}$$

$$= \frac{q_2\tilde{\pi}_2}{T_1}\left[1 + \frac{\frac{\sqrt{2T_1}}{\sqrt{\tilde{\pi}_1}\sqrt{\pi}}\mathrm{e}^{-\frac{1}{2}\tilde{\pi}_1/T_1}}{1 + \mathrm{Erf}\left[\frac{\sqrt{\tilde{\pi}_1}}{\sqrt{2T_1}}\right]} \right]^2 + \mathcal{O}\left(q_2^{\frac{3}{2}}\right). \tag{96}$$

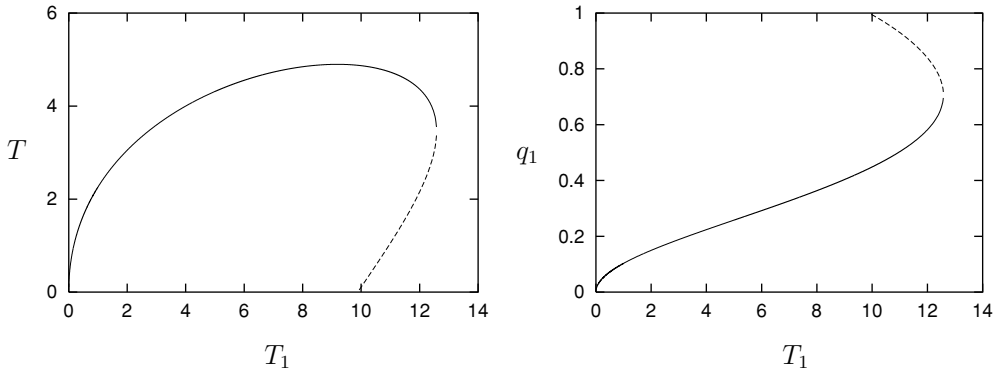

Figure 12. Re-entrance phenomena in the second-order $SG_1 \to SG_2$ transition for $L = 3$ and $m_2 = 0.5$. Left panel: phase diagram in the (T_1, T) plane. Right panel: value of the order parameter q_1 at the transition line as a function of T_1. In both cases the solid part of the curve indicates the physical part of the transition line; the dashed part indicates the non-physical (re-entrant) part.

Hence the condition for the second-order $SG_1 \to SG_2$ transition, $T_1 = T_{1,2}^{\rm 2nd}$, is to be solved from

$$1 = \sqrt{\frac{\tilde\pi_2}{T_1}}\left[1 + \frac{\left(\frac{2T_1}{\tilde\pi_1\pi}\right)^{\frac12}\,{\rm e}^{-\frac12\tilde\pi_1/T_1}}{1+{\rm Erf}\left[\left(\frac{\tilde\pi_1}{2T_1}\right)^{\frac12}\right]}\right]. \tag{97}$$

Equivalently, upon substituting $x^2 = \tilde\pi_1/2T_1$:

$$T_{1,2}^{\rm 2nd} = \frac{\tilde\pi_1}{2x^2} \qquad \sqrt{\frac{\tilde\pi_1}{2\tilde\pi_2}} = x + \frac{1}{\sqrt{\pi}}\frac{{\rm e}^{-x^2}}{1+{\rm Erf}[x]}. \tag{98}$$

Insertion of the parameter values appropriate for the $L = 2$ and $L = 3$ examples, as studied in detail in a previous section, sheds interesting new light on the nature of the jumps observed in the $SG_1 \to SG_2$ transition lines of the $L = 2$ and $L = 3$ phase diagrams for small T:

$$\begin{aligned} &m_2 = 0.5: \qquad T_1^{\rm jump} \approx 12.58 \qquad T_{1,2}^{\rm 2nd} \approx 9.884\\ &m_2 = 1.5: \qquad T_1^{\rm jump} \approx 2.17 \qquad\ T_{1,2}^{\rm 2nd} \approx 2.095. \end{aligned}$$

These results indicate that for low T the $SG_1 \to SG_2$ transition exhibits re-entrance, the extent of which decreases with increasing m_2. In fact, we found another second-order transition line from SG_1 to SG_2 which shows re-entrance for each m_2 (see figures 11 and 12). On reflection, this is not entirely surprising, since for $m_1 \to 0$ one should expect replica symmetry to break. It has been observed frequently in many disordered systems (see, e.g., [12]) that replica-symmetric solutions show unphysical re-entrance phenomena in RSB regions (which are removed by the proper RSB solution).

For general values of ℓ we have to work out the following continuous bifurcation condition of $q_\ell \neq 0$, derived from equation (93):

$$1 = \lim_{q_\ell\to 0}\frac{\partial}{\partial q_\ell} \times\left\langle\cdots\left\langle\left[\left\langle\cdots\left\langle\left[\frac{\sinh(\tilde\Xi_1)+\frac12\,{\rm e}^{\tilde\Xi_1}{\rm Erf}\left[\frac{\tilde a_1^2+\tilde\Xi_1}{\tilde a_1\sqrt{2}}\right]-\frac12\,{\rm e}^{-\tilde\Xi_1}{\rm Erf}\left[\frac{\tilde a_1^2-\tilde\Xi_1}{\tilde a_1\sqrt{2}}\right]}{\cosh(\tilde\Xi_1)+\frac12\,{\rm e}^{\tilde\Xi_1}{\rm Erf}\left[\frac{\tilde a_1^2+\tilde\Xi_1}{\tilde a_1\sqrt{2}}\right]+\frac12\,{\rm e}^{-\tilde\Xi_1}{\rm Erf}\left[\frac{\tilde a_1^2-\tilde\Xi_1}{\tilde a_1\sqrt{2}}\right]}\right]\right\rangle_2\cdots\right\rangle_{k-1}\right]^2\right\rangle_k\cdots\right\rangle_\ell .$$

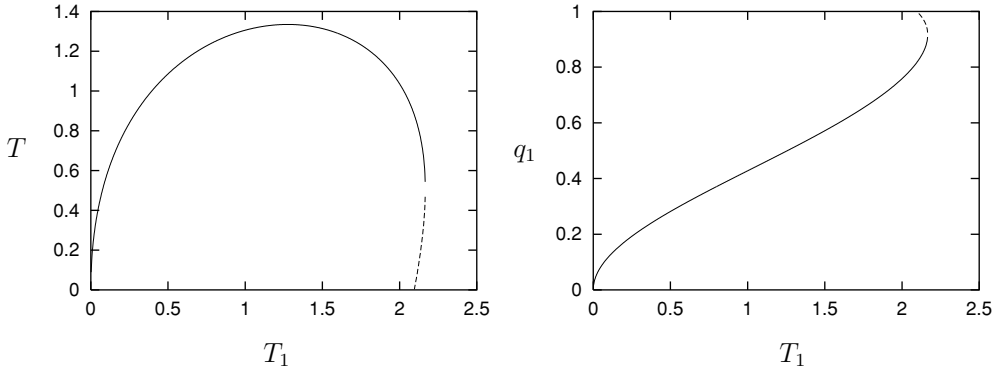

Figure 13. Re-entrance phenomena in the second-order $SG_1 \to SG_2$ transition for $L = 3$ and $m_2 = 1.5$. Left panel: phase diagram in the (T_1, T) plane. Right panel: value of the order parameter q_1 at the transition line as a function of T_1. In both cases the solid part of the curve indicates the physical part of the transition line; the dashed part indicates the non-physical (re-entrant) part.

This is done in appendix C, where we show that the value $T^{\rm 2nd}_{\ell-1,\ell}$ for T_1 which marks the continuous transition from $SG_{\ell-1}$ to SG_ℓ is to be solved from

$$1 = \sqrt{\frac{\tilde{\pi}_\ell}{T_1}} \left\{ 1 + \kappa + \sum_{k=2}^{\ell-1} m_2 \cdots m_{k-1}(m_k - 1) q_k \right\} \tag{99}$$

where $\kappa = \lim_{m_1 \to 0}(1 - q_1)/m_1$, where $q_\ell = 0$, and where for the values of the $\{q_k\}$ with $1 < k < \ell$ one has to substitute the solution of equations (92) and (93). The quantity κ is calculated along the lines of our previous calculation of (86) and (87):

$$\begin{aligned} \kappa &= \lim_{m_1 \to 0} \frac{1}{m_1} \left\langle \cdots \left\langle \frac{\int \mathrm{D}z_1 \cosh^{m_1}(\Xi)[1 - \tanh^2(\Xi)]}{\int \mathrm{D}z_1 \cosh^{m_1}(\Xi)} \right\rangle_2 \cdots \right\rangle_\ell \\ &= \lim_{m_1 \to 0} \frac{1}{m_1} \left\langle \cdots \left\langle \frac{\int \mathrm{D}z_1 \cosh^{m_1-2}(\tilde{\Xi}/m_1)}{\int \mathrm{D}z_1 \cosh^{m_1}(\tilde{\Xi}/m_1)} \right\rangle_2 \cdots \right\rangle_\ell . \end{aligned} \tag{100}$$

We know that $K_1 = \int \mathrm{D}z_1 \cosh^{m_1}(\tilde{\Xi}/m_1) \to \int \mathrm{D}z_1\, \mathrm{e}^{|\tilde{\Xi}|}$ for $m_1 \to 0$. Let us calculate the term in the numerator in leading order.

$$\int \mathrm{D}z_1 \cosh^{m_1-2}(\tilde{\Xi}/m_1) = I_+(\tilde{\Xi}_1) + I_-(\tilde{\Xi}_1)$$

with

$$I_+(\tilde{\Xi}_1) = \int_{\Xi>0} \mathrm{D}z_1 \cosh^{m_1-2}(\tilde{\Xi}/m_1) = \int_{-\frac{\tilde{\Xi}_1}{\tilde{a}_1}}^{\infty} \mathrm{D}z_1 \{\mathrm{e}^{\tilde{\Xi}/m_1}(1 + \mathrm{e}^{-2\tilde{\Xi}/m_1})/2\}^{m_1-2}$$

$$I_-(\tilde{\Xi}_1) = \int_{\Xi<0} \mathrm{D}z_1 \cosh^{m_1-2}(\tilde{\Xi}/m_1) = \int_{-\infty}^{-\frac{\tilde{\Xi}_1}{\tilde{a}_1}} \mathrm{D}z_1 \{\mathrm{e}^{-\tilde{\Xi}/m_1}(1 + \mathrm{e}^{2\tilde{\Xi}/m_1})/2\}^{m_1-2} = I_+(-\tilde{\Xi}_1).$$

Putting $b = 2\tilde{a}_1/m_1$, $b_n = nb$, fixing m_1 to a small number, introducing a cut-off ε and using $(1 + x)^{-2} = \sum_{n=0}^{\infty}(n + 1)(-x)^n$ and the asymptotic properties of the error function [21], we obtain

$$
\begin{aligned}
I_+(\tilde\Xi_1) &= \int_{-\frac{\tilde\Xi_1}{\tilde a_1}}^{\infty} \mathrm{D}z_1\, \mathrm{e}^{-\tilde\Xi(2/m_1-1)}\{(1+\mathrm{e}^{-2\tilde\Xi/m_1})/2\}^{m_1-2} \\
&= \int_{-\frac{\tilde\Xi_1}{\tilde a_1}}^{\infty} \mathrm{D}z_1\, \mathrm{e}^{-\tilde\Xi(2/m_1-1)}\{(1+\mathrm{e}^{-2\tilde\Xi/m_1})/2\}^{-2}(1+\mathcal{O}(m_1)) \\
&= 4\lim_{\varepsilon\to+0}\int_{\varepsilon-\frac{\tilde\Xi_1}{\tilde a_1}}^{\infty} \mathrm{D}z_1\, \mathrm{e}^{-\tilde\Xi(2/m_1-1)}\sum_{n=0}^{\infty}(n+1)(-)^n\,\mathrm{e}^{-2n\tilde\Xi/m_1}(1+\mathcal{O}(m_1)) \\
&= 4\lim_{\varepsilon\to+0}\sum_{n=0}^{\infty}(n+1)(-)^n\int_{\varepsilon-\frac{\tilde\Xi_1}{\tilde a_1}}^{\infty} \mathrm{D}z_1\, \mathrm{e}^{-(bz_1+\frac{2\tilde\Xi_1}{m_1})\left(n+1-\frac{m_1}{2}\right)}(1+\mathcal{O}(m_1)) \\
&= 4\lim_{\varepsilon\to+0}\sum_{n=0}^{\infty}(n+1)(-)^n\,\mathrm{e}^{-\frac{2\tilde\Xi_1}{m_1}\left(n+1-\frac{m_1}{2}\right)}\,\mathrm{e}^{b^2(n+1-\frac{m_1}{2})^2/2} \\
&\quad\times\left\{\frac12-\frac{1}{\pi}\,\mathrm{Erf}\left[\frac{\varepsilon-\frac{\tilde\Xi_1}{\tilde a_1}+b\left(n+1-\frac{m_1}{2}\right)}{\sqrt2}\right]\right\}+\cdots \\
&= 4\lim_{\varepsilon\to+0}\sum_{n=0}^{\infty}(n+1)(-)^n\,\mathrm{e}^{-\frac{2\tilde\Xi_1}{m_1}\left(n+1-\frac{m_1}{2}\right)}\,\mathrm{e}^{b^2(n+1-\frac{m_1}{2})^2/2-(\varepsilon-\frac{\tilde\Xi_1}{\tilde a_1}+b(n+1-\frac{m_1}{2}))^2/2} \\
&\quad\times\frac{1}{\sqrt{2\pi}}\frac{1}{\varepsilon-\frac{\tilde\Xi_1}{\tilde a_1}+b\left(n+1-\frac{m_1}{2}\right)}+\cdots \\
&= 4\lim_{\varepsilon\to+0}\sum_{n=0}^{\infty}(n+1)(-)^n\,\mathrm{e}^{-\frac{2\tilde\Xi_1}{m_1}\left(n+1-\frac{m_1}{2}\right)}\,\mathrm{e}^{-(\varepsilon-\frac{\tilde\Xi_1}{\tilde a_1})b(n+1-\frac{m_1}{2})-(\varepsilon-\frac{\tilde\Xi_1}{\tilde a_1})^2/2} \\
&\quad\times\frac{1}{\sqrt{2\pi}b(n+1)}\frac{1}{\left(\varepsilon-\frac{\tilde\Xi_1}{\tilde a_1}\right)\frac{1}{b(n+1)}+1-\frac{m_1}{2(n+1)}}+\cdots \\
&= \frac{4}{\sqrt{2\pi}b}\lim_{\varepsilon\to+0}\mathrm{e}^{\tilde a_1\varepsilon-(\varepsilon-\frac{\tilde\Xi_1}{\tilde a_1})^2/2}\sum_{n=0}^{\infty}(-)^n(\mathrm{e}^{-b\varepsilon})^{n+1}(1+\mathcal{O}(m_1))+\cdots \\
&= \frac{4}{\sqrt{2\pi}b}\lim_{\varepsilon\to+0}\mathrm{e}^{\tilde a_1\varepsilon-(\varepsilon-\frac{\tilde\Xi_1}{\tilde a_1})^2/2}\frac{\mathrm{e}^{-b\varepsilon}}{1+\mathrm{e}^{-b\varepsilon}}+\cdots \\
&= \frac{2}{\sqrt{2\pi}b}\,\mathrm{e}^{-(\frac{\tilde\Xi_1}{\tilde a_1})^2/2}+\cdots=\frac{m_1}{\sqrt{2\pi}\tilde a_1}\,\mathrm{e}^{-(\frac{\tilde\Xi_1}{\tilde a_1})^2/2}+\cdots.
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
\int \mathrm{D}z_1\cosh^{m_1-2}(\tilde\Xi/m_1) &= I_+(\tilde\Xi_1)+I_-(\tilde\Xi_1)=\frac{2m_1}{\sqrt{2\pi}\tilde a_1}\,\mathrm{e}^{-(\frac{\tilde\Xi_1}{\tilde a_1})^2/2}+\cdots \\
&= \frac{m_1}{\tilde a_1}\sqrt{\frac{2}{\pi}}\,\mathrm{e}^{-\tilde\Xi_1^2/2\tilde a_1^2}+\cdots \qquad (m_1\to0).
\end{aligned}
$$

Hence

$$
\begin{aligned}
\kappa &= \frac{1}{\tilde a_1}\sqrt{\frac{2}{\pi}}\left\langle\cdots\left\langle\frac{\mathrm{e}^{-\tilde\Xi_1^2/2\tilde a_1^2}}{\int \mathrm{D}z_1\,\mathrm{e}^{|\tilde\Xi|}}\right\rangle_2\cdots\right\rangle_\ell \\
&= \left\langle\cdots\left\langle\left[\frac{\tilde a_1^{-1}(2/\pi)^{\frac12}\,\mathrm{e}^{-\tilde\Xi_1^2/2\tilde a_1^2-\tilde a_1^2/2}}{\cosh(\tilde\Xi_1)+\frac12\,\mathrm{e}^{\tilde\Xi_1}\mathrm{Erf}\left[\frac{\tilde a_1^2+\tilde\Xi_1}{\tilde a_1\sqrt2}\right]+\frac12\,\mathrm{e}^{-\tilde\Xi_1}\mathrm{Erf}\left[\frac{\tilde a_1^2-\tilde\Xi_1}{\tilde a_1\sqrt2}\right]}\right]\right\rangle_2\cdots\right\rangle_\ell. \qquad (101)
\end{aligned}
$$

We have now calculated explicitly all ingredients necessary for determining $T^{2\mathrm{nd}}_{\ell-1,\ell}$: one is to solve equation (99), upon insertion of the solution of equations (92) and (93) for $\{q_{k>1}\}$ and of expression (101) for κ.

As a simple test one can verify the outcome of (99) for the case $\ell = 2$ which we analysed before. This corresponds to $\tilde{\Xi}_1 = 0$, insertion of which into (101) and subsequently into (99) indeed brings us back to (97), as it should.

8. Discussion

In this paper we have studied self-programming in recurrent neural networks where both neurons (the 'processors') and synaptic interactions ('the programme') evolve in time simultaneously. In contrast to previous models involving coupled dynamics of fast neurons and slow interactions, the interactions in our model do not evolve on a single time-scale; they are divided randomly into a hierarchy of L different groups of prescribed sizes, each with their own characteristic time-scale τ_ℓ and noise level T_ℓ ($\ell = 1, \ldots, L$), describing increasingly non-volatile programming levels.

We have solved our model in equilibrium upon making the replica-symmetric (i.e. ergodic) ansatz within each level of our hierarchy, leading to a theory which resembles, but is not identical to, Parisi's L-level replica symmetry breaking (RSB) solution for spin systems with frozen disorder. In addition to a paramagnetic phase, the phase diagram of our model involves a hierarchy of distinct spin-glass phases SG_ℓ, which differ in terms of the largest time-scale on which the system (spins and couplings) will be found in a frozen state. Our theory involves L replica dimensions m_ℓ, reminiscent of the block sizes in Parisi's RSB scheme, which here represent ratios of the temperatures of subsequent levels in the hierarchy of equilibrating couplings, and which can take any value in the interval $[0, \infty\rangle$.

We have solved our order parameter equations in full detail for the choices $L = 2$ and $L = 3$, leading to extremely rich phase diagrams, with an abundance of first-order transitions especially when the level of stochasticity in the interaction dynamics is chosen to be low, i.e. when one or more of the m_ℓ become large (which can never happen in the Parisi scheme, where always $0 \leqslant m_\ell \leqslant 1$). Increasing L always leads to the creation of new phases, thereby increasing their total number, with associated new order parameters which measure the degree of freezing at larger time-scales. We also studied the asymptotic properties of our model for arbitrary values of L in the limits $m_1 \to \infty$ for fixed T (deterministic dynamics of the level-1 interactions) and $m_1 \to 0$ for fixed T_1 (deterministic dynamics of the neuronal processors). This revealed further non-trivial properties, such as the universal nature of the values of the order parameters (i.e. $q_k = q^\star \approx 0.7911$ for all $k \leqslant \ell$, independent of the control parameters) at the first-order $P \to SG_\ell$ transitions for $m_1 = \infty$, and re-entrance phenomena at the $SG_{\ell-1} \to SG_\ell$ borders in the phase diagrams for $m_1 = 0$ (the latter are likely to be replaced by discrete non-reentrant jumps when replica symmetry breaking is taken into account).

In the present study we have not attempted to validate the predictions of our theory by numerical simulations. With the present CPU resources this would not have been possible, since our model requires L nested equilibrations of disordered sub-systems. Even with fixed couplings it would not have been possible to reach equilibrium with a system size sufficiently large to suppress finite size effects; as a consequence, even for a similar $L = 1$ system it was already found to be impossible to carry out simulations which achieve more than a very rough qualitative agreement with the theory [10]. Experimental verification for a recurrent self-programming network with $L > 1$ will probably require hardware realizations. There is also little scope for comparison with other papers; we are not aware of any other study

involving coupled dynamics of fast neurons and slow synapses where there are multiple separated time-scales for the synapses.

In retrospect, we found ourselves pleasantly surprised by the extent to which the present model allows for analytical solution, in spite of the complications induced by the nested equilibrations. Rather than being restricted by mathematical intractability, our problem was how to select the control parameters for which full phase diagrams are to be shown. This allowed us to gain qualitative and quantitative insight into the structural properties of simple recurrent self-programming systems, in particular regarding the questions of when and how these systems switch from random motion of processors and programme (the paramagnetic phase) to states where processors and some (or all) of the levels of the hierarchy of programme routines are 'locked' into specific fixed-points (the different types of spin-glass phases).

At the same time it will also be clear, however, that both in terms of understanding the physics of realistic self-programming systems and in terms of understanding the mathematical theory by which they are described, this paper represents only a modest step. In realistic self-programming systems one should obviously expect the programming levels not to evolve only on the basis of pair-correlations in processor states, a simple decay term, and Gaussian noise, but one would as a minimum introduce external symmetry-breaking forces into both the processor dynamics and the coupling dynamics, representing data to be processed and programming objectives to be met, respectively. At a theoretical level it would be interesting to investigate the form taken in the present model of replica symmetry breaking, for which we have already found indirect evidence (in the form of re-entrance phenomena) in studying the $m_1 \to 0$ limit, by calculating AT lines [22]. Since in our present model the replica symmetric (RS) theory is already similar to an L-step RSB theory *á la* Parisi, it is not immediately obvious what structure to expect when replica symmetry is broken at one or more levels in our hierarchy. These and other questions and extensions will hopefully be addressed in future studies.

Acknowledgments

One of the authors (TU) would like to acknowledge the Ministry of Education, Culture, Sports, Science and Technology of the Japanese Government for its support of the Overseas Research Fellowship no12-kho-158.

Appendix A. Asymptotic form of $T_{1,2}^{\mathrm{2nd}}$ when $m_1 \gg 1$

In this appendix we derive the asymptotic form of the transition temperature $T_{1,2}^{\mathrm{2nd}}$ for $m_1 \gg 1$. The value of $T_{1,2}^{\mathrm{2nd}}$ is determined by solving the coupled equations

$$1 = \sqrt{\beta\pi_2}[1 + (m_1 - 1)q_1] \tag{A.1}$$

$$q_1 = \frac{1}{K_1}\int \mathrm{D}z_1 \cosh^{m_1}(a_1 z_1)\tanh^2(a_1 z_1) \tag{A.2}$$

where $a_1 = \sqrt{\beta\pi_1 q_1}$. Upon assuming $q_1 \ll 1$ and $m_1 \gg 1$, we can derive from (A.2) the equation for $\tilde{q} \equiv m_1 q_1$ to take the form

$$1 \simeq \sqrt{\beta\pi_1}[1 + \beta\pi_1\tilde{q}(1 + \tilde{q})]. \tag{A.3}$$

From (A.1) and (A.3) we obtain equations for the quantities $w \equiv 1/\sqrt{\beta\pi_2}$ and $\tilde{q}$,

$$\tilde{q} = w - 1 \tag{A.4}$$

$$g(w) \equiv w^3 - aw + b = 0 \tag{A.5}$$

where $a = \frac{\tilde{\pi}_1}{\tilde{\pi}_2}\big(1+\frac{\tilde{\pi}_1}{\tilde{\pi}_2}\big)$, $b = \big(\frac{\tilde{\pi}_1}{\tilde{\pi}_2}\big)^2$ and $\tilde{\pi}_i = m_1\pi_i$, both of which are independent of m_1. From $\tilde{\pi}_1 = \tilde{\pi}_2 + \frac{\epsilon_1}{\mu_1}$ it follows that $g(1) = 1 - \frac{\tilde{\pi}_1}{\tilde{\pi}_2} < 0$. Thus $g(w) = 0$ has the unique solution $w^* > 1$. Therefore we obtain our desired result for $m_1 \to \infty$:

$$q_1 \simeq \frac{\tilde{q}}{m_1} = \frac{1}{m_1}(w^* - 1) > 0 \tag{A.6}$$

$$T \simeq \tilde{\pi}_2 w^* \sqrt{T_1}. \tag{A.7}$$

Appendix B. Building blocks of the general saddle-point equations

In this appendix we calculate the derivatives used in deriving the saddle-point equations for general values of L in section 4.4. In this and subsequent appendices it will be helpful to define a convenient shorthand (further notation is as in section 4.4):

$$M_{\ell>0} = \prod_{k=1}^{\ell} m_k \qquad M_0 = 1$$

$$\psi_{\ell>0} = \langle \cdots \langle \tanh \Xi \rangle_1 \cdots \rangle_\ell \qquad \psi_0 = \tanh \Xi$$

$$J_\ell = K_\ell^{-1} \frac{\partial}{\partial \Xi}(K_\ell \psi_\ell).$$

Appendix B.1. Calculation of $K_\ell^{-1}\partial K_\ell/\partial\Xi$ and $J_\ell = K_\ell^{-1}\partial(K_\ell\psi_\ell)/\partial\Xi$

We first calculate $K_\ell^{-1}\partial K_\ell/\partial\Xi$ for $\ell \geqslant 1$:

$$\begin{aligned} K_\ell^{-1}\frac{\partial K_\ell}{\partial \Xi} &= \frac{m_\ell}{K_\ell}\int \mathrm{D}z_\ell K_{\ell-1}^{m_\ell-1}\frac{\partial K_{\ell-1}}{\partial\Xi} \\ &= \frac{\prod_{k=1}^{\ell} m_k}{K_\ell}\int \mathrm{D}z_\ell K_{\ell-1}^{m_\ell-1}\int \mathrm{D}z_{\ell-1} K_\ell^{m_{\ell-1}-1}\cdots\int \mathrm{D}z_1 \cosh^{m_1-1}\Xi \sinh\Xi \\ &= M_\ell\langle\cdots\langle\tanh\Xi\rangle_1\cdots\rangle_\ell = M_\ell\psi_\ell. \end{aligned} \tag{B.1}$$

Next we calculate $J_\ell = K_\ell^{-1}\partial(K_\ell\psi_\ell)/\partial\Xi$ for $\ell \geqslant 1$. We commence with J_1:

$$\begin{aligned} J_1 &= K_1^{-1}\int \mathrm{D}z_1 \frac{\partial}{\partial\Xi}\big(K_0^{m_1}\psi_0\big) \\ &= K_1^{-1}\int \mathrm{D}z_1\, K_0^{m_1}\big[1+(m_1-1)\tanh^2\Xi\big] \\ &= 1 + (m_1-1)\langle\tanh^2\Xi\rangle_1. \end{aligned} \tag{B.2}$$

We move on to $\ell > 1$, abbreviating $L_\ell = K_\ell^{-1}\partial K_\ell/\partial\Xi$:

$$\begin{aligned} J_\ell &= K_\ell^{-1}\int \mathrm{D}z_\ell \frac{\partial}{\partial\Xi}\big(K_{\ell-1}^{m_\ell-1}K_{\ell-1}\psi_{\ell-1}\big) \\ &= K_\ell^{-1}\int \mathrm{D}z_\ell \left\{(m_\ell-1)K_{\ell-1}^{m_\ell-2}\frac{\partial K_{\ell-1}}{\partial\Xi}K_{\ell-1}\psi_{\ell-1} + K_{\ell-1}^{m_\ell-1}\frac{\partial}{\partial\Xi}(K_{\ell-1}\psi_{\ell-1})\right\} \\ &= K_\ell^{-1}\int \mathrm{D}z_\ell \left\{(m_\ell-1)K_{\ell-1}^{m_\ell}L_{\ell-1}\psi_{\ell-1} + K_{\ell-1}^{m_\ell}J_{\ell-1}\right\} \\ &= (m_\ell-1)\langle L_{\ell-1}\psi_{\ell-1}\rangle_\ell + \langle J_{\ell-1}\rangle_\ell \\ &= M_{\ell-1}(m_\ell-1)\big\langle\psi_{\ell-1}^2\big\rangle_\ell + \langle J_{\ell-1}\rangle_\ell. \end{aligned}$$

Iteration of this relation gives:

$$\begin{aligned} J_\ell &= \sum_{k=2}^{\ell} M_{k-1}(m_k-1)\big\langle\cdots\big\langle\psi_{k-1}^2\big\rangle_k\cdots\big\rangle_\ell + \langle\cdots\langle J_1\rangle_2\cdots\rangle_\ell \\ &= \sum_{k=2}^{\ell} M_{k-1}(m_k-1)\big\langle\cdots\big\langle\psi_{k-1}^2\big\rangle_k\cdots\big\rangle_\ell + 1 + (m_1-1)\big\langle\cdots\langle\tanh^2\Xi\rangle_1\cdots\big\rangle_\ell \\ &= 1 + \sum_{k=1}^{\ell} M_{k-1}(m_k-1)\big\langle\cdots\big\langle\psi_{k-1}^2\big\rangle_k\cdots\big\rangle_\ell. \end{aligned} \tag{B.3}$$

Finally we prove that $J_\ell > 0$ for all $\ell \geqslant 1$. We put $q_k = \big\langle\cdots\big\langle\psi_{k-1}^2\big\rangle_k\cdots\big\rangle_\ell$ and use the inequalities $q_1 \geqslant q_2 \geqslant \cdots \geqslant q_\ell \geqslant 0$:

$$\begin{aligned} J_\ell &= \sum_{k=2}^{\ell}(M_k - M_{k-1})q_k + 1 + (m_1-1)q_1 \\ &= \sum_{k=2}^{\ell-1} M_k(q_k - q_{k+1}) + M_\ell q_\ell - m_1 q_2 + 1 + (m_1-1)q_1 \\ &= \sum_{k=2}^{\ell-1} M_k(q_k - q_{k+1}) + M_\ell q_l + m_1(q_1 - q_2) + 1 - q_1 > 0. \end{aligned}$$

Appendix B.2. Calculation of $\partial K_\ell/\partial q_{\ell'}$

In this second part of appendix B we will repeatedly need the following simple relations:

$$\frac{\partial\Xi}{\partial q_1} = \frac{1}{2}\beta\pi_1 z_1\left(\frac{\partial\Xi}{\partial z_1}\right)^{-1} \qquad \frac{\partial\Xi}{\partial q_{\ell>1}} = \frac{1}{2}\beta\pi_\ell\left\{z_\ell\left(\frac{\partial\Xi}{\partial z_\ell}\right)^{-1} - z_{\ell-1}\left(\frac{\partial\Xi}{\partial z_{\ell-1}}\right)^{-1}\right\}. \tag{B.4}$$

We first calculate $\partial K_\ell/\partial q_{\ell'}$ for $1 < \ell' \leqslant \ell$:

$$\begin{aligned} \frac{\partial K_\ell}{\partial q_{\ell'}} &= m_\ell\int \mathrm{D}z_\ell\, K_{\ell-1}^{m_\ell-1}\frac{\partial K_{\ell-1}}{\partial q_{\ell'}} \\ &= \left[\prod_{k=\ell'-1}^{\ell} m_k\right]\int \mathrm{D}z_\ell\, K_{\ell-1}^{m_\ell-1}\int \mathrm{D}z_{\ell-1}K_{\ell-2}^{m_{\ell-1}-1} \\ &\quad \cdots\int \mathrm{D}z_{\ell'}K_{\ell'-1}^{m_{\ell'}-1}\int \mathrm{D}z_{\ell'-1}K_{\ell'-2}^{m_{\ell'-1}-1}\frac{\partial K_{\ell'-2}}{\partial q_{\ell'}} \\ &= \left[\prod_{k=\ell'-1}^{\ell} m_k\right]\int \mathrm{D}z_\ell K_{\ell-1}^{m_\ell-1}\cdots\int \mathrm{D}z_{\ell'}K_{\ell'-1}^{m_{\ell'}-1}\int \mathrm{D}z_{\ell'-1}K_{\ell'-2}^{m_{\ell'-1}-1} \\ &\quad \times\frac{\beta\pi_{\ell'}}{2}\left\{z_{\ell'}\left(\frac{\partial\Xi}{\partial z_{\ell'}}\right)^{-1} - z_{\ell'-1}\left(\frac{\partial\Xi}{\partial z_{\ell'-1}}\right)^{-1}\right\}\frac{\partial K_{\ell'-2}}{\partial\Xi} \\ &= \frac{\beta\pi_{\ell'}}{2}\left[\prod_{k=\ell'-1}^{\ell} m_k\right]\int \mathrm{D}z_\ell K_{\ell-1}^{m_\ell-1}\int \mathrm{D}z_{\ell-1}K_{\ell-2}^{m_{\ell-1}-1} \\ &\quad \cdots\int \mathrm{D}z_{\ell'}\frac{\partial K_{\ell'-1}^{m_{\ell'}-1}}{\partial\Xi}\int \mathrm{D}z_{\ell'-1}K_{\ell'-2}^{m_{\ell'-1}-1}\frac{\partial K_{\ell'-2}}{\partial\Xi} \end{aligned}$$

$$= \frac{\beta\pi_{\ell'}}{2}\left[\prod_{k=\ell'-1}^{\ell} m_k\right](m_{\ell'}-1)K_\ell$$
$$\times\left\langle\left\langle\cdots\left\langle\frac{1}{K_{\ell'-1}}\frac{\partial K_{\ell'-1}}{\partial\Xi}\left\langle\frac{1}{K_{\ell'-2}}\frac{\partial K_{\ell'-2}}{\partial\Xi}\right\rangle_{\ell'-1}\right\rangle_{\ell'}\cdots\right\rangle_{\ell-1}\right\rangle_{\ell}.$$

Hence, upon using $K_\ell^{-1}\partial K_\ell/\partial\Xi = M_\ell\psi_\ell$, we obtain for $1<\ell'\leqslant\ell$ the simple result:

$$K_\ell^{-1}\frac{\partial K_\ell}{\partial q_{\ell'}} = \frac{\beta\pi_{\ell'}}{2}\left[\prod_{k=1}^{\ell} m_k\right]\left[\prod_{j=1}^{\ell'-1} m_j\right](m_{\ell'}-1)\langle\langle\cdots\langle\psi_{\ell'-1}\langle\psi_{\ell'-2}\rangle_{\ell'-1}\rangle_{\ell'}\cdots\rangle_{\ell-1}\rangle_\ell$$
$$= \frac{\beta\pi_{\ell'}}{2}M_\ell M_{\ell'-1}(m_{\ell'}-1)\left\langle\left\langle\cdots\left\langle\psi^2_{\ell'-1}\right\rangle_{\ell'}\cdots\right\rangle_{\ell-1}\right\rangle_\ell. \tag{B.5}$$

For $\ell'=1\leqslant\ell$ we obtain, similarly:

$$K_\ell^{-1}\frac{\partial K_\ell}{\partial q_1} = \frac{m_\ell}{K_\ell}\int \mathrm{D}z_\ell K_{\ell-1}^{m_\ell-1}\frac{\partial K_{\ell-1}}{\partial q_1}$$
$$= \frac{\prod_{k=1}^{\ell} m_k}{K_\ell}\int \mathrm{D}z_\ell K_{\ell-1}^{m_\ell-1}\int \mathrm{D}z_{\ell-1}K_{\ell-2}^{m_{\ell-1}-1}\cdots\int \mathrm{D}z_1 K_0^{m_1-1}\frac{\partial K_0}{\partial q_1}$$
$$= \frac{M_\ell}{K_\ell}\int \mathrm{D}z_\ell K_{\ell-1}^{m_\ell-1}\cdots\int \mathrm{D}z_1 K_0^{m_1-1}\frac{\beta\pi_1}{2}z_1\left(\frac{\partial\Xi}{\partial z_1}\right)^{-1}\frac{\partial K_0}{\partial\Xi}$$
$$= \frac{\beta\pi_1}{2}\frac{M_\ell}{K_\ell}\int \mathrm{D}z_\ell K_{\ell-1}^{m_\ell-1}\cdots\int \mathrm{D}z_1\frac{\partial}{\partial\Xi}\left\{K_0^{m_1-1}\frac{\partial K_0}{\partial\Xi}\right\}$$
$$= \frac{\beta\pi_1}{2}\frac{M_\ell}{K_\ell}\int \mathrm{D}z_\ell K_{\ell-1}^{m_\ell-1}\cdots\int \mathrm{D}z_1\left\{1+(m_1-1)\tanh^2\Xi\right\}K_0^{m_1}$$
$$= \frac{\beta\pi_1}{2}M_\ell\left\{1+(m_1-1)\left\langle\cdots\langle\tanh^2\Xi\rangle_1\cdots\right\rangle_\ell\right\}. \tag{B.6}$$

Next we turn to the case where $\ell'>\ell$, noting that K_ℓ is a function of $\{z_{\ell+1},\ldots,z_L\}$. If $\ell'>\ell+1$ one finds

$$K_\ell^{-1}\frac{\partial K_\ell}{\partial q_{\ell'}} = \frac{\beta\pi_{\ell'}}{2}\left\{z_{\ell'}\left(\frac{\partial\Xi}{\partial z_{\ell'}}\right)^{-1}-z_{\ell'-1}\left(\frac{\partial\Xi}{\partial z_{\ell'-1}}\right)^{-1}\right\}\frac{1}{K_\ell}\frac{\partial K_\ell}{\partial\Xi}$$
$$= \frac{\beta\pi_{\ell'}}{2}\left\{z_{\ell'}\left(\frac{\partial\Xi}{\partial z_{\ell'}}\right)^{-1}-z_{\ell'-1}\left(\frac{\partial\Xi}{\partial z_{\ell'-1}}\right)^{-1}\right\}M_\ell\psi_\ell.$$

For $\ell'=\ell+1$, on the other hand, we obtain

$$\frac{\partial K_\ell}{\partial q_{\ell+1}} = m_\ell\int \mathrm{D}z_\ell K_{\ell-1}^{m_\ell-1}\frac{\partial K_{\ell-1}}{\partial q_{\ell+1}}$$
$$= \frac{m_\ell\beta\pi_{\ell+1}}{2}\int \mathrm{D}z_\ell K_{\ell-1}^{m_\ell-1}\left\{z_{\ell+1}\left(\frac{\partial\Xi}{\partial z_{\ell+1}}\right)^{-1}-z_\ell\left(\frac{\partial\Xi}{\partial z_\ell}\right)^{-1}\right\}\frac{\partial K_{\ell-1}}{\partial\Xi}$$
$$= \frac{m_\ell\beta\pi_{\ell+1}}{2}\left\{z_{\ell+1}\left(\frac{\partial\Xi}{\partial z_{\ell+1}}\right)^{-1}\int \mathrm{D}z_\ell K_{\ell-1}^{m_\ell-1}\frac{\partial K_{\ell-1}}{\partial\Xi}\right.$$

$$- \int \mathrm{D}z_\ell \frac{\partial}{\partial \Xi}\left[K_{\ell-1}^{m_\ell-1}\frac{\partial K_{\ell-1}}{\partial \Xi}\right]\Bigg\}.$$

This latter term will appear inside an integral over $z_{\ell+1}$ whenever $q_{\ell+1} > 0$. Thus, we will at some point have to evaluate integrals of the following form:

$$\int \mathrm{D}z_{\ell+1} g(\Xi)\frac{\partial K_\ell}{\partial q_{\ell+1}} = \frac{m_\ell \beta \pi_{\ell+1}}{2}\int \mathrm{D}z_{\ell+1}\frac{\partial g(\Xi)}{\partial \Xi}\int \mathrm{D}z_\ell K_{\ell-1}^{m_\ell-1}\frac{\partial K_{\ell-1}}{\partial \Xi}. \tag{B.7}$$

We calculate the derivative in the second term of $\partial K_\ell/\partial q_{\ell+1}$ as follows:

$$\begin{aligned}\frac{\partial}{\partial \Xi}\left\{K_{\ell-1}^{m_\ell-1}\frac{\partial K_{\ell-1}}{\partial \Xi}\right\} &= M_{\ell-1}\frac{\partial}{\partial \Xi}\left\{K_{\ell-1}^{m_\ell-1}K_{\ell-1}\psi_{\ell-1}\right\}\\ &= M_{\ell-1}\left\{(m_\ell-1)K_{\ell-1}^{m_\ell-2}\frac{\partial K_{\ell-1}}{\partial \Xi}K_{\ell-1}\psi_{\ell-1} + K_{\ell-1}^{m_\ell-1}\frac{\partial}{\partial \Xi}(K_{\ell-1}\psi_{\ell-1})\right\}\\ &= M_{\ell-1}\left\{(m_\ell-1)K_{\ell-1}^{m_\ell}M_{\ell-1}\psi_{\ell-1}^2 + K_{\ell-1}^{m_\ell-1}\frac{\partial}{\partial \Xi}(K_{\ell-1}\psi_{\ell-1})\right\}.\end{aligned}$$

Upon using formula (B.3) we arrive at

$$\begin{aligned}M_{\ell-1}^{-1}\frac{\partial}{\partial \Xi}\left\{K_{\ell-1}^{m_\ell-1}\frac{\partial K_{\ell-1}}{\partial \Xi}\right\} &= (m_\ell-1)K_{\ell-1}^{m_\ell}M_{\ell-1}\psi_{\ell-1}^2\\ &\quad + K_{\ell-1}^{m_\ell}\left[1+\sum_{j=1}^{\ell-1}M_{j-1}(m_j-1)\left\langle\cdots\left\langle\psi_{j-1}^2\right\rangle_j\cdots\right\rangle_{\ell-1}\right].\end{aligned}$$

Hence

$$\begin{aligned}K_\ell^{-1}\int \mathrm{D}z_\ell\frac{\partial}{\partial \Xi}\left\{K_{\ell-1}^{m_\ell-1}\frac{\partial K_{\ell-1}}{\partial \Xi}\right\} &= M_{\ell-1}\left\{1+\sum_{j=1}^{\ell}M_{j-1}(m_j-1)\left\langle\cdots\left\langle\psi_{j-1}^2\right\rangle_j\cdots\right\rangle_\ell\right\}\\ &= M_{\ell-1}J_\ell\end{aligned}$$

$$\begin{aligned}K_\ell^{-1}\frac{\partial K_\ell}{\partial q_{\ell+1}} &= \frac{m_\ell}{K_\ell}\frac{\beta\pi_{\ell+1}}{2}\left\{z_{\ell+1}\left(\frac{\partial \Xi}{\partial z_{\ell+1}}\right)^{-1}K_\ell M_{\ell-1}\langle\psi_{\ell-1}\rangle_\ell - K_\ell M_{\ell-1}J_\ell\right\}\\ &= \frac{\beta\pi_{\ell+1}}{2}M_{\ell-1}m_\ell\left\{z_{\ell+1}\left(\frac{\partial \Xi}{\partial z_{\ell+1}}\right)^{-1}\psi_\ell - J_\ell\right\}.\end{aligned} \tag{B.8}$$

Appendix C. Condition for second-order $SG_{\ell-1} \to SG_\ell$ transitions

In this appendix we derive the condition for the second-order transition $SG_{\ell-1}$ to SG_ℓ, for arbitrary values of $1 \leqslant \ell \leqslant L$. We generalize our previous notational conventions, and write the general saddle-point equations as

$$q_\ell = \varphi_\ell(q_1,\ldots,q_L) \qquad (\ell = 1,\ldots,L).$$

Similarly we define, for $1 \leqslant k \leqslant \ell$:

$$\varphi_k^{(\ell)}(q_1,\ldots,q_L) = \left\langle\cdots\left\langle\psi_{k-1}^2\right\rangle_k\cdots\right\rangle_\ell.$$

Note that, with these conventions, $\varphi_k(q_1,\ldots,q_L) = \left\langle \cdots \left\langle \varphi_k^{(\ell)} \right\rangle_{\ell+1} \cdots \right\rangle_L$. The second-order $SG_{\ell-1} \to SG_\ell$ transition temperature is defined by

$$\left.\frac{\partial \varphi_\ell^{(\ell)}}{\partial q_\ell}\right|_{(q_{\ell-1}>0;\, q_\ell = q_{\ell+1} = \cdots = q_L = 0)} = 1. \tag{C.1}$$

We can write $\varphi_\ell^{(\ell)}(q_1,\ldots,q_L)$ as $\varphi_\ell^{(\ell)} = K_\ell^{-1} \int \mathrm{D}z_\ell K_{\ell-1}^{m_\ell} \psi_{\ell-1}^2 = \left\langle \psi_{\ell-1}^2 \right\rangle_\ell$, with $\psi_{\ell-1} = \langle \cdots \langle \tanh \Xi \rangle_1 \cdots \rangle_{\ell-1}$. From this expression we obtain

$$\begin{aligned}
\frac{\partial \varphi_\ell^{(\ell)}}{\partial q_\ell} &= -\frac{1}{K_\ell}\frac{\partial K_\ell}{\partial q_\ell}\varphi_\ell^{(\ell)} + \frac{1}{K_\ell}\int \mathrm{D}z_\ell \frac{\partial}{\partial q_\ell}\left[K_{\ell-1}^{m_\ell}\psi_{\ell-1}^2\right] \\
&= -\frac{1}{K_\ell}\frac{\partial K_\ell}{\partial q_\ell}\varphi_\ell^{(\ell)} + \frac{m_\ell - 2}{K_\ell}\int \mathrm{D}z_\ell K_{\ell-1}^{m_\ell-3}\frac{\partial K_{\ell-1}}{\partial q_\ell}(K_{\ell-1}\psi_{\ell-1})^2 \\
&\quad + \frac{2}{K_\ell}\int \mathrm{D}z_\ell K_{\ell-1}^{m_\ell-1}\psi_{\ell-1}\frac{\partial}{\partial q_\ell}[K_{\ell-1}\psi_{\ell-1}] \\
&= -\frac{1}{K_\ell}\frac{\partial K_\ell}{\partial q_\ell}\varphi_\ell^{(\ell)} + \frac{m_\ell - 2}{K_\ell}\int \mathrm{D}z_\ell K_{\ell-1}^{m_\ell-3}\frac{\partial K_{\ell-1}}{\partial q_\ell}[K_{\ell-1}\psi_{\ell-1}]^2 \\
&\quad + \frac{2}{K_\ell}\int \mathrm{D}z_\ell K_{\ell-1}^{m_\ell-1}\psi_{\ell-1}\int \mathrm{D}z_{\ell-1}\frac{\partial \Xi}{\partial q_\ell}\frac{\partial}{\partial \Xi}\left[K_{\ell-2}^{m_{\ell-1}}\psi_{\ell-2}\right] \\
&= -\frac{1}{K_\ell}\frac{\partial K_\ell}{\partial q_\ell}\varphi_\ell^{(\ell)} + \frac{m_\ell - 2}{K_\ell}\int \mathrm{D}z_\ell K_{\ell-1}^{m_\ell-3}\frac{\partial K_{\ell-1}}{\partial q_\ell}[K_{\ell-1}\psi_{\ell-1}]^2 \\
&\quad + \frac{\beta\pi_\ell}{K_\ell}\int \mathrm{D}z_\ell \left\{\left(\frac{\partial \Xi}{\partial z_\ell}\right)^{-1}\frac{\partial}{\partial z_\ell}\left(K_{\ell-1}^{m_\ell-1}\psi_{\ell-1}\right)\right\}\int \mathrm{D}z_{\ell-1}\frac{\partial}{\partial \Xi}\left[K_{\ell-2}^{m_{\ell-1}}\psi_{\ell-2}\right] \\
&= -\frac{1}{K_\ell}\frac{\partial K_\ell}{\partial q_\ell}\varphi_\ell^{(\ell)} + \frac{m_\ell - 2}{K_\ell}\int \mathrm{D}z_\ell K_{\ell-1}^{m_\ell-3}\frac{\partial K_{\ell-1}}{\partial q_\ell}[K_{\ell-1}\psi_{\ell-1}]^2 \\
&\quad + \frac{\beta\pi_\ell}{K_\ell}\int \mathrm{D}z_\ell \left\{(m_\ell - 2)K_{\ell-1}^{m_\ell-2}\frac{\partial K_{\ell-1}}{\partial \Xi}\psi_{\ell-1} + K_{\ell-1}^{m_\ell-2}\frac{\partial}{\partial \Xi}(K_{\ell-1}\psi_{\ell-1})\right\} \\
&\quad \times \frac{\partial}{\partial \Xi}[K_{\ell-1}\psi_{\ell-1}] \\
&= -\frac{1}{K_\ell}\frac{\partial K_\ell}{\partial q_\ell}\varphi_\ell^{(\ell)} + (m_\ell - 2)\left\langle K_{\ell-1}^{-1}\frac{\partial K_{\ell-1}}{\partial q_\ell}\psi_{\ell-1}^2\right\rangle_\ell \\
&\quad + \beta\pi_\ell(m_\ell - 2)\left\langle K_{\ell-1}^{-2}\frac{\partial K_{\ell-1}}{\partial \Xi}\psi_{\ell-1}\frac{\partial}{\partial \Xi}(K_{\ell-1}\psi_{\ell-1})\right\rangle_\ell \\
&\quad + \beta\pi_\ell\left\langle K_{\ell-1}^{-2}\left\{\frac{\partial}{\partial \Xi}[K_{\ell-1}\psi_{\ell-1}]\right\}^2\right\rangle_\ell .
\end{aligned}$$

For $\ell > 1$ this reduces to

$$\begin{aligned}
\frac{\partial \varphi_\ell^{(\ell)}}{\partial q_\ell} &= -\frac{1}{K_\ell}\frac{\partial K_\ell}{\partial q_\ell}\varphi_\ell^{(\ell)} + \beta\pi_\ell(m_\ell - 2)\left\langle \frac{1}{K_{\ell-1}}\frac{\partial K_{\ell-1}}{\partial \Xi}\psi_{\ell-1}J_{\ell-1}\right\rangle_\ell + \beta\pi_\ell\left\langle J_{\ell-1}^2\right\rangle_\ell \\
&\quad + (m_\ell - 2)\frac{\beta\pi_\ell}{2K_\ell}m_{\ell-1}\int \mathrm{D}z_\ell\left\{\frac{\partial}{\partial \Xi}\left[K_{\ell-1}^{m_\ell-1}\psi_{\ell-1}^2\right]\right\}\int \mathrm{D}z_{\ell-1}K_{\ell-2}^{m_{\ell-1}-1}\frac{\partial K_{\ell-2}}{\partial \Xi} \\
&= -\frac{1}{K_\ell}\frac{\partial K_\ell}{\partial q_\ell}\varphi_\ell^{(\ell)} + \beta\pi_\ell(m_\ell - 2)\left\langle \frac{1}{K_{\ell-1}}\frac{\partial K_{\ell-1}}{\partial \Xi}\psi_{\ell-1}J_{\ell-1}\right\rangle_\ell + \beta\pi_\ell\left\langle J_{\ell-1}^2\right\rangle_\ell \\
&\quad + (m_\ell - 2)\frac{\beta\pi_\ell}{2}m_{\ell-1}\left\langle\left\{(m_\ell - 3)\frac{1}{K_{\ell-1}}\frac{\partial K_{\ell-1}}{\partial \Xi}\psi_{\ell-1}^2\right.\right.
\end{aligned}$$

$$
\begin{aligned}
&+\ 2\psi_{\ell-1}\frac{1}{K_{\ell-1}}\frac{\partial}{\partial\Xi}(K_{\ell-1}\psi_{\ell-1})\Big\}\left\langle\frac{1}{K_{\ell-2}}\frac{\partial K_{\ell-2}}{\partial\Xi}\right\rangle_{\ell-1}\Big\rangle_{\ell}\\
&=-\frac{1}{K_\ell}\frac{\partial K_\ell}{\partial q_\ell}\varphi_\ell^{(\ell)}+\beta\pi_\ell(m_\ell-2)M_{\ell-1}\left\langle\psi_{\ell-1}^2J_{\ell-1}\right\rangle_\ell+\beta\pi_\ell\left\langle J_{\ell-1}^2\right\rangle_\ell\\
&\quad+(m_\ell-2)\frac{\beta\pi_\ell}{2}m_{\ell-1}\left\langle\left\{(m_\ell-3)M_{\ell-1}\psi_{\ell-1}^3+2\psi_{\ell-1}J_{\ell-1}\right\}\langle M_{\ell-2}\psi_{\ell-2}\rangle_{\ell-1}\right\rangle_\ell\\
&=\frac{\beta\pi_\ell}{2}\Big\{-M_\ell M_{\ell-1}(m_\ell-1)\big(\varphi_\ell^{(\ell)}\big)^2+2(m_\ell-2)m_{\ell-1}M_{\ell-2}\left\langle\psi_{\ell-1}^2J_{\ell-1}\right\rangle_\ell\\
&\quad+(m_\ell-2)m_{\ell-1}M_{\ell-2}(m_\ell-3)M_{\ell-1}\left\langle\psi_{\ell-1}^4\right\rangle_\ell\\
&\quad+2(m_\ell-2)M_{\ell-1}\left\langle\psi_{\ell-1}^2J_{\ell-1}\right\rangle_\ell+2\left\langle J_{\ell-1}^2\right\rangle_\ell\Big\}\\
&=\frac{\beta\pi_\ell}{2}\Big\{-M_\ell M_{\ell-1}(m_\ell-1)\big(\varphi_\ell^{(\ell)}\big)^2+4(m_\ell-2)M_{\ell-1}\left\langle\psi_{\ell-1}^2J_{\ell-1}\right\rangle_\ell\\
&\quad+(m_\ell-2)m_{\ell-1}M_{\ell-2}(m_\ell-3)M_{\ell-1}\left\langle\psi_{\ell-1}^4\right\rangle_\ell+2\left\langle J_{\ell-1}^2\right\rangle_\ell\Big\}
\end{aligned}
\tag{C.2}
$$

whereas for $\ell=1$ one finds

$$
\begin{aligned}
\frac{\partial\varphi_1^{(1)}}{\partial q_1}&=-\frac{1}{K_1}\frac{\partial K_1}{\partial q_1}\varphi_1^{(1)}+\frac{1}{K_1}\int\mathrm{D}z_1\frac{\partial}{\partial q_1}\left[K_0^{m_1}\tanh^2\Xi\right]\\
&=-\frac{1}{K_1}\frac{\partial K_1}{\partial q_1}\varphi_1^{(1)}+\frac{1}{K_1}\int\mathrm{D}z_1\frac{\beta\pi_1}{2}z_1(\partial\Xi/\partial z_1)^{-1}\frac{\partial}{\partial\Xi}\left[K_0^{m_1}\tanh^2\Xi\right]\\
&=-\frac{1}{K_1}\frac{\partial K_1}{\partial q_1}\varphi_1^{(1)}+\frac{\beta\pi_1}{2}\frac{1}{K_1}\int\mathrm{D}z_1\frac{\partial^2}{\partial\Xi^2}\left[K_0^{m_1}\tanh^2\Xi\right].
\end{aligned}
$$

The second derivative of $K_0^{m_1}\tanh^2\Xi$ is given by

$$
\begin{aligned}
\frac{\partial^2}{\partial\Xi^2}\left[K_0^{m_1}\tanh^2\Xi\right]&=\frac{\partial}{\partial\Xi}\left[(m_1-2)\cosh m_1-3\Xi\sinh^3\Xi+2\cosh^{m_1-1}\Xi\sinh\Xi\right]\\
&=\cosh^{m_1}\Xi\left[(m_1-2)(m_1-3)\tanh^4\Xi+(5m_1-8)\tanh^2\Xi+2\right].
\end{aligned}
$$

Inserting this expression, and using (B.6), we now arrive at

$$
\begin{aligned}
\frac{\partial\varphi_1^{(1)}}{\partial q_1}=\frac{\beta\pi_1}{2}\Big\{&-m_1(m_1-1)\left(\varphi_1^{(1)}\right)^2+(m_1-2)(m_1-3)\langle\tanh^4\Xi\rangle_1\\
&+4(m_1-2)\varphi_1^{(1)}+2\Big\}.
\end{aligned}
\tag{C.3}
$$

Finally we note that the property $q_\ell=0$ induces the simplifying relations $\varphi_\ell^{(\ell)}=\psi_{\ell-1}=0$, with which expressions (C.2) and (C.3) leads to

$$
\left.\frac{\partial\varphi_\ell^{(\ell)}}{\partial q_\ell}\right|_{(q_{\ell-1}>0;q_\ell=q_{\ell+1}=\cdots=q_L=0)}=\beta\pi_\ell J_{\ell-1}^2.
$$

Together with (B.3) we can now work out the condition (C.1) for the second-order $SG_{\ell-1}\to SG_\ell$ transition explicitly, giving the final result

$$
\sqrt{\beta\pi_\ell}\left\{1+\sum_{k=1}^{\ell-1}M_{k-1}(m_k-1)q_k\right\}=1.
\tag{C.4}
$$

Appendix D. Derivation of partial derivatives for $L = 2$ and $L = 3$

In this appendix we will find it convenient to use the following abbreviations:

$$\varphi_1(q_1,q_2,q_2) = \left\langle\left\langle\langle\tanh^2\Xi\rangle_1\right\rangle_2\right\rangle_3$$
$$\varphi_2(q_1,q_2,q_3) = \left\langle\left\langle\langle\tanh\Xi\rangle_1^2\right\rangle_2\right\rangle_3$$
$$\varphi_3(q_1,q_1,q_3) = \left\langle\langle\langle\tanh\Xi\rangle_1\rangle_2^2\right\rangle_3$$

(these are the right-hand sides of the $L = 3$ saddle-point equations for $\{q_1, q_2, q_3\}$),

$$\varphi_1^{(1)}(q_1,q_2,q_3) = \langle\tanh^2\Xi\rangle_1 \qquad \varphi_1^{(2)}(q_1,q_2,q_3) = \left\langle\langle\tanh^2\Xi\rangle_1\right\rangle_2$$
$$\varphi_2^{(1)}(q_1,q_2,q_3) = \langle\tanh\Xi\rangle_1^2 \qquad \varphi_2^{(2)}(q_1,q_2,q_3) = \left\langle\langle\tanh\Xi\rangle_1^2\right\rangle_2$$
$$\varphi_3^{(2)}(q_1,q_2,q_3) = \langle\langle\tanh\Xi\rangle_1\rangle_2^2$$

and also $\psi_\ell = \langle\cdots\langle\tanh\Xi\rangle_1\cdots\rangle_\ell$. Note that $\left\langle\varphi_k^{(1)}\right\rangle_2 = \varphi_k^{(2)}$, $\varphi_k = \left\langle\varphi_k^{(2)}\right\rangle_3$ $(k = 1, 2, 3)$, $\psi_1^2 = \varphi_2^{(1)}$ and $\psi_2^2 = \varphi_3^{(2)}$. Further notation is as in sections 4 and 5 and appendix C.

Appendix D.1. Building blocks of the partial derivatives

First, we calculate the derivative $\partial\varphi_1^{(1)}(q_1,q_2,q_3)/\partial q_1$. From (C.3) we obtain

$$\frac{\partial\varphi_1^{(1)}}{\partial q_1} = \frac{\beta\pi_1}{2}\left\{2 - m_1(m_1-1)\left(\varphi_1^{(1)}\right)^2 + (m_1-2)(m_1-3)\langle\tanh^4\Xi\rangle_1 + 4(m_1-2)\varphi_1^{(1)}\right\}. \tag{D.1}$$

Next is $\partial\varphi_2^{(2)}(q_1,q_2,q_3)/\partial q_2$. Equation (C.2) with $k = 2$ can be simplifid by using (B.3):

$$\begin{aligned}\frac{\partial\varphi_2^{(2)}}{\partial q_2} &= \frac{\beta\pi_2}{2}\Big\{-M_2m_1(m_2-1)\left(\varphi_2^{(2)}\right)^2 + m_1^2(m_2-2)(m_2-3)\left\langle\psi_1^4\right\rangle_2 \\ &\quad + 4m_1(m_2-2)\left\langle\psi_1^2J_1\right\rangle_2 + 2\left\langle J_1^2\right\rangle_2\Big\} \\ &= \frac{\beta\pi_2}{2}\Big\{-m_1^2m_2(m_2-1)\left(\varphi_2^{(2)}\right)^2 + m_1^2(m_2-2)(m_2-3)\left\langle\psi_1^4\right\rangle_2 \\ &\quad + 4m_1(m_2-2)\varphi_2^{(2)} + 4m_1(m_1-1)(m_2-2)\left\langle\psi_1^2\langle\tanh^2\Xi\rangle_1\right\rangle_2 \\ &\quad + 2 + 4(m_1-1)\varphi_1^{(2)} + 2(m_1-1)^2\left\langle\langle\tanh^2\Xi\rangle_1^2\right\rangle_2\Big\}.\end{aligned} \tag{D.2}$$

The third partial derivative which we will need is

$$\begin{aligned}\frac{\partial\varphi_1^{(2)}}{\partial q_1} &= -\frac{1}{K_2}\frac{\partial K_2}{\partial q_1}\varphi_1^{(2)} + \frac{1}{K_2}\int Dz_2\frac{\partial}{\partial q_1}\left(K_1^{m_2}\langle\tanh^2\Xi\rangle_1\right) \\ &= -\frac{1}{K_2}\frac{\partial K_2}{\partial q_1}\varphi_1^{(2)} + \frac{m_2}{K_2}\int Dz_2K_1^{m_2-1}\frac{\partial K_1}{\partial q_1}\langle\tanh^2\Xi\rangle_1 + \left\langle\frac{\partial}{\partial q_1}\langle\tanh^2\Xi\rangle_1\right\rangle \\ &= -\frac{1}{K_2}\frac{\partial K_2}{\partial q_1}\varphi_1^{(2)} + m_2\left\langle\frac{1}{K_1}\frac{\partial K_1}{\partial q_1}\langle\tanh^2\Xi\rangle_1\right\rangle_2 + \left\langle\frac{\partial\varphi_1^{(1)}}{\partial q_1}\right\rangle_2.\end{aligned}$$

The last term in this expression is found to be

$$\begin{aligned}\left\langle\frac{\partial\varphi_1^{(1)}}{\partial q_1}\right\rangle_2 &= \frac{\beta\pi_1}{2}\Big\{-m_1(m_1-1)\left\langle\langle\tanh^2\Xi\rangle_1^2\right\rangle_2 \\ &\quad + (m_1-2)(m_1-3)\left\langle\langle\tanh^4\Xi\rangle_1\right\rangle_2 + 4(m_1-2)\varphi_1^{(2)} + 2\Big\}.\end{aligned}$$

Thus

$$\frac{\partial\varphi_1^{(2)}}{\partial q_1} = \frac{\beta\pi_1}{2}\Big\{4(m_1-2)\varphi_1^{(2)} + 2 - M_2\varphi_1^{(2)} - (m_1-1)M_2\left(\varphi_1^{(2)}\right)^2 + m_1m_2\left\langle\left[1+(m_1-1)\langle\tanh^2\Xi\rangle_1\right]\langle\tanh^2\Xi\rangle_1\right\rangle_2 - m_1(m_1-1)\left\langle\langle\tanh^2\Xi\rangle_1^2\right\rangle_2 + (m_1-2)(m_1-3)\left\langle\langle\tanh^4\Xi\rangle_1\right\rangle_2 + (m_1-2)(m_1-3)\left\langle\langle\tanh^4\Xi\rangle_1\right\rangle_2 + 4(m_1-2)\varphi_1^{(2)} + 2\Big\}. \tag{D.3}$$

Similarly we calculate the fourth relevant partial derivative:

$$\begin{aligned}\frac{\partial\varphi_1^{(2)}}{\partial q_2} &= -\frac{1}{K_2}\frac{\partial K_2}{\partial q_2}\varphi_1^{(2)} + \frac{1}{K_2}\int Dz_2\frac{\partial}{\partial q_2}\left[K_1^{m_2-1}\int Dz_1\cosh^{m_1}\Xi\tanh^2\Xi\right]\\ &= -\frac{1}{K_2}\frac{\partial K_2}{\partial q_2}\varphi_1^{(2)} + (m_2-1)\frac{1}{K_2}\int Dz_2K_1^{m_2-2}\frac{\partial K_1}{\partial q_2}K_1\langle\tanh^2\Xi\rangle_1\\ &\quad + \frac{1}{K_2}\int Dz_2K_1^{m_2-1}\int Dz_1\frac{\partial}{\partial q_2}[\cosh^{m_1}\Xi\tanh^2\Xi]\\ &= -\frac{\beta\pi_2}{2}m_1^2m_2(m_2-1)\left\langle\psi_1^2\right\rangle_2\varphi_1^{(2)} + (m_2-1)\frac{\beta m_1\pi_2}{2K_2}\int Dz_2\\ &\quad\times\frac{\partial}{\partial\Xi}[K_1^{m_2-1}\langle\tanh^2\Xi\rangle_1]\int Dz_1\cosh^{m_1-1}\Xi\frac{\partial K_0}{\partial\Xi}\\ &\quad + \frac{1}{K_2}\int Dz_2K_1^{m_2-1}\frac{\beta\pi_2}{2}\int Dz_1\left(\frac{z_2}{a_2}-\frac{z_1}{a_1}\right)\frac{\partial}{\partial\Xi}(\cosh^{m_1}\Xi\tanh^2\Xi)\\ &= \frac{\beta\pi_2}{2}\left\{\frac{1}{K_2}\int Dz_2\frac{\partial K_1^{m_2-1}}{\partial\Xi}\frac{\partial}{\partial\Xi}\left[K_1\langle\tanh^2\Xi\rangle_1\right] - m_1^2m_2(m_2-1)\varphi_2^{(2)}\varphi_1^{(2)}\right.\\ &\quad\left. + (m_2-1)\frac{m_1}{K_2}\int Dz_2K_1\psi_1\frac{\partial}{\partial\Xi}\left[K_1^{m_2-1}\langle\tanh^2\Xi\rangle_1\right]\right\}.\end{aligned}$$

In this expression we need to calculate the following two objects:

$$\begin{aligned}\frac{\partial}{\partial\Xi}\left[K_1^{m_2-1}\langle\tanh^2\Xi\rangle_1\right] &= \frac{\partial}{\partial\Xi}\left[K_1^{m_2-2}K_1\langle\tanh^2\Xi\rangle_1\right]\\ &= (m_2-2)\frac{\partial K_1}{\partial\Xi}K_1^{m_2-3}K_1\langle\tanh^2\Xi\rangle_1 + K_1^{m_2-2}\frac{\partial}{\partial\Xi}\left[K_1\langle\tanh^2\Xi\rangle_1\right]\\ &= K_2^{m_1}\left\{(m_2-2)m_1\psi_1K_1^{-1}\langle\tanh^2\Xi\rangle_1 + K_1^{-2}\frac{\partial}{\partial\Xi}\left[K_1\langle\tanh^2\Xi\rangle_1\right]\right\}\end{aligned}$$

$$\frac{\partial}{\partial\Xi}\left[K_1\langle\tanh^2\Xi\rangle_1\right] = K_1\left[(m_1-2)\langle\tanh^3\Xi\rangle_1 + 2\langle\tanh\Xi\rangle_1\right].$$

Insertion of these intermediate results into our previous expression for $\partial\varphi_1^{(2)}/\partial q_2$ gives

$$\begin{aligned}\frac{\partial\varphi_1^{(2)}}{\partial q_2} &= \frac{\beta\pi_2}{2}m_1(m_2-1)\Big\{-m_1m_2\varphi_2^{(2)}\varphi_1^{(2)} + m_1(m_2-2)\left\langle\psi_1^2\langle\tanh^2\Xi\rangle_1\right\rangle_2\\ &\quad + 2\left\langle\psi_1\left((m_1-2)\langle\tanh^3\Xi\rangle_1 + 2\langle\tanh\Xi\rangle_1\right)\right\rangle_2\Big\}\\ &= \frac{\beta\pi_2}{2}m_1(m_2-1)\Big\{-m_1m_2\varphi_2^{(2)}\varphi_1^{(2)} + m_1(m_2-2)\left\langle\psi_1^2\langle\tanh^2\Xi\rangle_1\right\rangle_2\\ &\quad + 2(m_1-2)\left\langle\psi_1\langle\tanh^3\Xi\rangle_1\right\rangle_2 + 4\varphi_2^{(2)}\Big\}.\end{aligned} \tag{D.4}$$

From this last result one immediately obtains also the remaining partial derivative

$$\frac{\partial\varphi_2^{(2)}}{\partial q_1} = \frac{\alpha_1}{\alpha_2}\frac{\partial\varphi_1^{(2)}}{\partial q_2} = \frac{\pi_1(m_1-1)}{\pi_2 m_1(m_2-1)}\frac{\partial\varphi_1^{(2)}}{\partial q_2}$$
$$= \frac{\beta\pi_1}{2}(m_1-1)\Big\{-m_1m_2\varphi_1^{(2)}\varphi_2^{(2)} + m_1(m_2-2)\left\langle\psi_1^2\langle\tanh^2\Xi\rangle_1\right\rangle_2$$
$$+ 2(m_1-2)\left\langle\psi_1\langle\tanh^3\Xi\rangle_1\right\rangle_2 + 4\varphi_2^{(2)}\Big\} \tag{D.5}$$

where $\alpha_l = \frac{\beta\pi_\ell}{2}M_{\ell-1}(m_\ell-1)$.

Appendix D.2. Partial derivatives of the saddle-point equations

Here we calculate all partial derivatives of the form $\partial\varphi_\ell(q_1,q_2,q_3)/\partial q_{\ell'}$, with $\ell,\ell'\in\{1,2,3\}$, which play a role in our analysis of the $L\in\{2,3\}$ saddle-point equations.

Calculation of $\partial\varphi_3/\partial q_3$*:* Starting from the general formula (C.2)

$$\frac{\partial\varphi_3}{\partial q_3} = \frac{\beta\pi_3}{2}\left\{-M_3M_2(m_3-1)\varphi_3^2 + (m_3-2)m_2m_1(m_3-3)M_2\left\langle\psi_2^4\right\rangle_3\right.$$
$$\left.+4(m_3-2)M_2\left\langle\psi_2^2J_2\right\rangle_3 + 2\left\langle J_2^2\right\rangle_3\right\}$$

we derive

$$\frac{\partial\varphi_3}{\partial q_3} = \frac{\beta\pi_3}{2}\Big\{-M_3M_2(m_3-1)\varphi_3^2 + (m_3-2)m_2m_1(m_3-3)M_2\big\langle\psi_2^4\big\rangle_3$$
$$+4(m_3-2)M_2\Big(\big\langle\psi_2^2\big\rangle_3 + (m_1-1)\big\langle\psi_2^2\big\langle\langle\tanh^2\Xi\rangle_1\big\rangle_2\Big\rangle_3$$
$$+m_1(m_2-1)\big\langle\psi_2^2\big\langle\psi_1^2\big\rangle_2\big\rangle_3\Big)$$
$$+2\Big\langle\Big(1+(m_1-1)\big\langle\big\langle\psi_0^2\big\rangle_1\big\rangle_2 + m_1(m_2-1)\big\langle\psi_1^2\big\rangle_2\Big)^2\Big\rangle_3\Big\}$$
$$= \frac{\beta\pi_3}{2}\Big\{-M_3M_2(m_3-1)\varphi_3^2 + (m_3-2)m_2m_1(m_3-3)M_2\big\langle\psi_2^4\big\rangle_3$$
$$+4(m_3-2)M_2\Big(\varphi_3 + (m_1-1)\langle\psi_2^2\big\langle\langle\tanh^2\Xi\rangle_1\rangle_2\big\rangle_3$$
$$+m_1(m_2-1)\big\langle\psi_2^2\big\langle\psi_1^2\big\rangle_2\big\rangle_3\Big) + 2 + 4(m_1-1)\varphi_1 + 4m_1(m_2-1)\varphi_2$$
$$+2(m_1-1)^2\Big\langle\Big(\big\langle\langle\tanh^2\Xi\rangle_1\big\rangle_2\Big)^2\Big\rangle_3 + 2m_1^2(m_2-1)^2\Big\langle\Big(\big\langle\psi_1^2\big\rangle_2\Big)^2\Big\rangle_3$$
$$+4m_1(m_1-1)(m_2-1)\big\langle\langle\langle\tanh^2\Xi\rangle_1\rangle_2\big\langle\psi_1^2\big\rangle_2\big\rangle_3\Big\}. \tag{D.6}$$

Calculation of $\partial\varphi_3/\partial q_1$*:*

$$\frac{\partial\varphi_3}{\partial q_1} = -\frac{1}{K_3}\frac{\partial K_3}{\partial q_1}\varphi_3 + \frac{1}{K_3}\int Dz_3\frac{\partial}{\partial q_1}\left(K_2^{m_3}\psi_2^2\right)$$
$$= -\frac{1}{K_3}\frac{\partial K_3}{\partial q_1}\varphi_3 + \frac{m_3-2}{K_3}\int Dz_3K_2^{m_3-3}\frac{\partial K_2}{\partial q_1}(K_2\psi_2)^2$$
$$+\frac{2}{K_3}\int Dz_3K_2^{m_3-1}\psi_2\frac{\partial}{\partial q_1}(K_2\psi_2)$$
$$= -\frac{1}{K_3}\frac{\partial K_3}{\partial q_1}\varphi_3 + \frac{m_3-2}{K_3}\int Dz_3K_2^{m_3-3}\frac{\partial K_2}{\partial q_1}(K_2\psi_2)^2$$
$$+\frac{2(m_2-1)}{K_3}\int Dz_3K_2^{m_3-1}\psi_2\int Dz_2K_1^{m_2-2}\frac{\partial K_1}{\partial q_1}K_1\psi_1$$

$$
\begin{aligned}
&+\frac{2}{K_3}\int \mathrm{D}z_3 K_2^{m_3-1}\psi_2 \int \mathrm{D}z_2 K_1^{m_2-1}\frac{\partial}{\partial q_1}(K_1\psi_1)\\
&=-\frac{1}{K_3}\frac{\partial K_3}{\partial q_1}\varphi_3+\frac{m_3-2}{K_3}\int \mathrm{D}z_3 K_2^{m_3-3}\frac{\partial K_2}{\partial q_1}(K_2\psi_2)^2\\
&+\frac{2(m_2-1)}{K_3}\int \mathrm{D}z_3 K_2^{m_3-1}\psi_2 \int \mathrm{D}z_2 K_1^{m_2-2}\frac{\partial K_1}{\partial q_1}K_1\psi_1\\
&+\frac{\beta\pi_2}{2}\frac{2}{K_3}\int \mathrm{D}z_3 K_2^{m_3-1}\psi_2 \int \mathrm{D}z_2 K_1^{m_2-1}\int \mathrm{D}z_1 z_1(\partial\Xi/\partial z_1)^{-1}\frac{\partial}{\partial\Xi}\left(K_0^{m_1}\psi_0\right)\\
&=-\frac{1}{K_3}\frac{\partial K_3}{\partial q_1}\varphi_3+(m_3-2)\left\langle\frac{1}{K_2}\frac{\partial K_2}{\partial q_1}\psi_2^2\right\rangle_3+2(m_2-1)\left\langle\psi_2\left\langle\frac{1}{K_1}\frac{\partial K_1}{\partial q_1}\psi_1\right\rangle_2\right\rangle_3\\
&+\frac{\beta\pi_1}{2}\frac{2}{K_3}\int \mathrm{D}z_3 K_2^{m_3-1}\psi_2 \int \mathrm{D}z_2 K_1^{m_2-1}\int \mathrm{D}z_1\frac{\partial^2}{\partial\Xi^2}\left(K_0^{m_1}\psi_0\right)\\
&=-\frac{1}{K_3}\frac{\partial K_3}{\partial q_1}\varphi_3+(m_3-2)\left\langle\frac{1}{K_2}\frac{\partial K_2}{\partial q_1}\psi_2^2\right\rangle_3+2(m_2-1)\left\langle\psi_2\left\langle\frac{1}{K_1}\frac{\partial K_1}{\partial q_1}\psi_1\right\rangle_2\right\rangle_3\\
&+\frac{\beta\pi_1}{2}2\left\langle\psi_2\left\langle\frac{1}{K_1}\frac{\partial^2}{\partial\Xi^2}(K_1\langle\psi_0\rangle_1)\right\rangle_2\right\rangle_3 .
\end{aligned}
$$

The second derivative $\partial^2(K_1\langle\psi_0\rangle_1)/\partial\Xi^2$ in this expression is calculated as

$$
\begin{aligned}
\frac{\partial^2}{\partial\Xi^2}(K_1\langle\psi_0\rangle_1)&=\int \mathrm{D}z_1\frac{\partial^2}{\partial\Xi^2}\left[\cosh^{m_1-1}\Xi\sinh\Xi\right]\\
&=\int \mathrm{D}z_1\left[(m_1-1)(m_1-2)\cosh^{m_1-3}\Xi\sinh^3\Xi\right.\\
&\quad\left.+(3m_1-2)\cosh^{m_1-1}\Xi\sinh\Xi\right]\\
&=K_1\left((m_1-1)(m_1-2)\langle\tanh^3\Xi\rangle_1+(3m_1-2)\psi_1\right).
\end{aligned}
$$

Insertion into our previous expression for $\partial\varphi_3/\partial q_1$ gives

$$
\begin{aligned}
\frac{\partial\varphi_3}{\partial q_1}&=-\frac{1}{K_3}\frac{\partial K_3}{\partial q_1}\varphi_3+(m_3-2)\left\langle\frac{1}{K_2}\frac{\partial K_2}{\partial q_1}\psi_2^2\right\rangle_3+2(m_2-1)\left\langle\psi_2\left\langle\frac{1}{K_1}\frac{\partial K_1}{\partial q_1}\psi_1\right\rangle_2\right\rangle_3\\
&\quad+\frac{\beta\pi_1}{2}\left\{2(m_1-1)(m_1-2)\left\langle\psi_2\left\langle\langle\tanh^3\Xi\rangle_1\right\rangle_2\right\rangle_3+2(3m_1-2)\langle\psi_2\langle\psi_1\rangle_2\rangle_3\right\}.
\end{aligned}
$$

Upon using our previous result (B.6) of appendix B, this subsequently translates into

$$
\begin{aligned}
\frac{\partial\varphi_3}{\partial q_1}&=\frac{\beta\pi_1}{2}\Big\{-M_3\varphi_3-M_3(m_1-1)\varphi_1\varphi_3+(m_3-2)M_2\left\langle\psi_2^2\right\rangle_3\\
&\quad+(m_1-1)(m_3-2)M_2\left\langle\varphi_1^{(2)}\psi_2^2\right\rangle_3\\
&\quad+2m_1(m_2-1)\langle\psi_2\langle\psi_1\rangle_2\rangle_3+2m_1(m_1-1)(m_2-1)\langle\psi_2\langle\langle\tanh^2\Xi\rangle_1\psi_1\rangle_2\rangle_3\\
&\quad+2(m_1-1)(m_1-2)\left\langle\psi_2\left\langle\langle\tanh^3\Xi\rangle_1\right\rangle_2\right\rangle_3+2(3m_1-2)\langle\psi_2\langle\psi_1\rangle_2\rangle_3\Big\}\\
&=\frac{\beta\pi_1}{2}(m_1-1)\Big\{4\varphi_3-M_3\varphi_1\varphi_3+(m_3-2)M_2\left\langle\varphi_1^{(2)}\psi_2^2\right\rangle_3\\
&\quad+2m_1(m_2-1)\left\langle\psi_2\left\langle\langle\tanh^2\Xi\rangle_1\psi_1\right\rangle_2\right\rangle_3\\
&\quad+2(m_1-2)\left\langle\psi_2\left\langle\langle\tanh^3\Xi\rangle_1\right\rangle_2\right\rangle_3\Big\}.
\end{aligned}
\tag{D.7}
$$

Calculation of $\partial\varphi_3/\partial q_2$*:*

$$
\frac{\partial\varphi_3}{\partial q_2}=-\frac{1}{K_3}\frac{\partial K_3}{\partial q_2}\varphi_3+\frac{1}{K_3}\int \mathrm{D}z_3\frac{\partial}{\partial q_2}\left(K_2^{m_3}\psi_2^2\right)
$$

$$
\begin{aligned}
&= -\frac{1}{K_3}\frac{\partial K_3}{\partial q_2}\varphi_3 + \frac{m_3-2}{K_3}\int \mathrm{D}z_3 K_2^{m_3-3}\frac{\partial K_2}{\partial q_2}(K_2\psi_2)^2 \\
&\quad + \frac{2}{K_3}\int \mathrm{D}z_3 K_2^{m_3-1}\psi_2 \int \mathrm{D}z_2 \frac{\partial}{\partial q_2}\left(K_1^{m_2-1}K_1\psi_1\right) \\
&= -\frac{1}{K_3}\frac{\partial K_3}{\partial q_2}\varphi_3 + \frac{m_3-2}{K_3}\int \mathrm{D}z_3 K_2^{m_3-3}\frac{\partial K_2}{\partial q_2}(K_2\psi_2)^2 \\
&\quad + \frac{2(m_2-1)}{K_3}\int \mathrm{D}z_3 K_2^{m_3-1}\psi_2 \int \mathrm{D}z_2 K_1^{m_2-2}\frac{\partial K_1}{\partial q_2}K_1\psi_1 \\
&\quad + \frac{\beta\pi_2}{2}\frac{2}{K_3}\int \mathrm{D}z_3 K_2^{m_3-1}\psi_2 \int \mathrm{D}z_2 \frac{\partial K_1^{m_2-1}}{\partial \Xi}\int \mathrm{D}z_1 \frac{\partial}{\partial \Xi}\left(K_0^{m_1}\psi_0\right) \\
&= -\frac{1}{K_3}\frac{\partial K_3}{\partial q_2}\varphi_3 + \frac{m_3-2}{K_3}\int \mathrm{D}z_3 K_2^{m_3-3}\frac{\partial K_2}{\partial q_2}(K_2\psi_2)^2 \\
&\quad + \frac{2(m_2-1)}{K_3}\int \mathrm{D}z_3 K_2^{m_3-1}\psi_2 \int \mathrm{D}z_2 K_1^{m_2-2}\frac{\partial K_1}{\partial q_2}K_1\psi_1 \\
&\quad + \frac{\beta\pi_2}{2}\frac{2}{K_3}\int \mathrm{D}z_3 K_2^{m_3-1}\psi_2 \int \mathrm{D}z_2 (m_2-1)K_1^{m_2-2}\frac{\partial K_1}{\partial \Xi}K_1 J_1 \\
&= -\frac{1}{K_3}\frac{\partial K_3}{\partial q_2}\varphi_3 + (m_3-2)\left\langle \frac{1}{K_2}\frac{\partial K_2}{\partial q_2}\psi_2^2\right\rangle_3 \\
&\quad + 2(m_2-1)\left\langle \psi_2\left\langle \frac{1}{K_1}\frac{\partial K_1}{\partial q_2}\psi_1\right\rangle_2\right\rangle_3 + \frac{\beta\pi_2}{2}2m_1(m_2-1)\langle\psi_2\langle\psi_1 J_1\rangle_2\rangle_3 .
\end{aligned}
$$

Here we require the following object:

$$
\begin{aligned}
\left\langle \frac{1}{K_1}\frac{\partial K_1}{\partial q_2}\psi_1\right\rangle_2 &= \frac{1}{K_2}\int \mathrm{D}z_2 K_1^{m_2-1}\psi_1\frac{\partial K_1}{\partial q_2} \\
&= \frac{1}{K_2}m_1\frac{\beta\pi_2}{2}\int \mathrm{D}z_2\left\{\frac{\partial}{\partial \Xi}\left[K_1^{m_2-1}\psi_1\right]\right\}\int \mathrm{D}z_1 K_0^{m_1-1}\frac{\partial K_0}{\partial \Xi} \\
&= \frac{1}{K_2}m_1\frac{\beta\pi_2}{2}\int \mathrm{D}z_2 K_1^{m_2}\left[(m_2-2)\frac{1}{K_1}\frac{\partial K_1}{\partial \Xi}\frac{1}{K_1}\psi_1 + \frac{1}{K_1}\frac{\partial}{\partial \Xi}(K_1\psi_1)\frac{1}{K_1}\right] \\
&\quad \times \int \mathrm{D}z_1 K_0^{m_1}\tanh\Xi \\
&= m_1\frac{\beta\pi_2}{2}\left\langle\left[(m_2-2)\frac{1}{K_1}\frac{\partial K_1}{\partial \Xi}\psi_1 + \frac{1}{K_1}\frac{\partial}{\partial \Xi}(K_1\psi_1)\right]\psi_1\right\rangle_2 \\
&= m_1\frac{\beta\pi_2}{2}\left[m_1(m_2-2)\left\langle\psi_1^3\right\rangle_2 + \langle J_1\psi_1\rangle_2\right].
\end{aligned}
$$

Insertion into our earlier expression for $\partial\varphi_3/\partial q_2$, followed by usage of (B.5)–(B.8) whenever appropriate, allows us to write

$$
\begin{aligned}
\frac{\partial\varphi_3}{\partial q_2} &= -\frac{1}{K_3}\frac{\partial K_3}{\partial q_2}\varphi_3 + (m_3-2)\left\langle \frac{1}{K_2}\frac{\partial K_2}{\partial q_2}\psi_2^2\right\rangle_3 \\
&\quad + 2m_1(m_2-1)\frac{\beta\pi_2}{2}\left\langle\psi_2\left\{m_1(m_2-2)\left\langle\psi_1^3\right\rangle_2 + \langle J_1\psi_1\rangle_2\right\}\right\rangle_3 \\
&\quad + \frac{\beta\pi_2}{2}2m_1(m_2-1)\langle\psi_2\langle\psi_1 J_1\rangle_2\rangle_3 \\
&= -\frac{1}{K_3}\frac{\partial K_3}{\partial q_2}\varphi_3 + (m_3-2)\left\langle \frac{1}{K_2}\frac{\partial K_2}{\partial q_2}\psi_2^2\right\rangle_3
\end{aligned}
$$

$$
\begin{aligned}
&+2m_1^2(m_2-1)(m_2-2)\frac{\beta\pi_2}{2}\left\langle\psi_2\left\langle\psi_1^2\right\rangle_2\right\rangle_3\\
&+\frac{\beta\pi_2}{2}4m_1(m_2-1)\left\{\langle\psi_2\langle\psi_1\rangle_3\rangle_3+(m_1-1)\left\langle\psi_2\left\langle\psi_1\langle\tanh^2\Xi\rangle_1\right\rangle_2\right\rangle_3\right\}\\
&=\frac{\beta\pi_2}{2}(m_2-1)\left\{-M_3m_1\varphi_2\varphi_3+M_2m_1(m_3-2)\left\langle\left\langle\psi_1^2\right\rangle_2\psi_2^2\right\rangle_3\right.\\
&+2m_1^2(m_2-2)\left\langle\psi_2\left\langle\psi_1^3\right\rangle_2\right\rangle_3\\
&\left.+4m_1\varphi_3+4m_1(m_1-1)\left\langle\psi_2\left\langle\psi_1\langle\tanh^2\Xi\rangle_1\right\rangle_2\right\rangle_3\right\}. \qquad \text{(D.8)}
\end{aligned}
$$

Calculation of $\partial\varphi_1/\partial q_3$ and $\partial\varphi_2/\partial q_3$: These two partial derivatives are found to be relatively easy to calculate:

$$
\begin{aligned}
\frac{\partial\varphi_1}{\partial q_3}&=\frac{\alpha_3}{\alpha_1}\frac{\partial\varphi_3}{\partial q_1}=\frac{\pi_3M_2(m_3-1)}{\pi_1(m_1-1)}\frac{\partial\varphi_3}{\partial q_1}\\
&=\frac{\beta\pi_3}{2}M_2(m_3-1)\left\{4\varphi_3-M_3\varphi_1\varphi_3+(m_3-2)M_2\left\langle\varphi_1^{(2)}\psi_2^2\right\rangle_3\right.\\
&\left.+2m_1(m_2-1)\left\langle\psi_2\left\langle\langle\tanh^2\Xi\rangle_1\psi_1\right\rangle_2\right\rangle_3+2(m_1-2)\left\langle\psi_2\left\langle\langle\tanh^3\Xi\rangle_1\right\rangle_2\right\rangle_3\right\} \qquad \text{(D.9)}\\
\frac{\partial\varphi_2}{\partial q_3}&=\frac{\alpha_3}{\alpha_2}\frac{\partial\varphi_3}{\partial q_2}=\frac{\pi_3M_2(m_3-1)}{\pi_2M_1(m_2-1)}\frac{\partial\varphi_3}{\partial q_2}\\
&=\frac{\beta\pi_3}{2}m_2(m_3-1)\left\{-M_3m_1\varphi_2\varphi_3+M_2m_1(m_3-2)\left\langle\left\langle\psi_1^2\right\rangle_2\psi_2^2\right\rangle_3+4m_1\varphi_3\right.\\
&\left.+2m_1^2(m_2-2)\left\langle\psi_2\left\langle\psi_1^3\right\rangle_2\right\rangle_3+4m_1(m_1-1)\left\langle\psi_2\left\langle\psi_1\langle\tanh^2\Xi\rangle_1\right\rangle_2\right\rangle_3\right\}. \qquad \text{(D.10)}
\end{aligned}
$$

Calculation of $\partial\varphi_2/\partial q_1$ and $\partial\varphi_1/\partial q_2$:

$$
\begin{aligned}
\frac{\partial\varphi_1}{\partial q_2}&=-\frac{1}{K_3}\frac{\partial K_3}{\partial q_2}\varphi_1+\frac{1}{K_3}\int Dz_3\frac{\partial}{\partial q_2}\left(K_2^{m_3}\varphi_1^{(2)}\right)\\
&=-\frac{1}{K_3}\frac{\partial K_3}{\partial q_2}\varphi_1+m_3\frac{1}{K_3}\int Dz_3\frac{1}{K_2}\frac{\partial K_2}{\partial q_2}K_2^{m_3}\varphi_1^{(2)}+\frac{1}{K_3}\int Dz_3K_2^{m_3}\frac{\partial\varphi_1^{(2)}}{\partial q_2}\\
&=-\frac{\beta\pi_2}{2}M_3m_1(m_2-1)\varphi_1\varphi_2+m_3\left\langle\frac{1}{K_2}\frac{\partial K_2}{\partial q_2}\varphi_1^{(2)}\right\rangle_3+\left\langle\frac{\partial\varphi_1^{(2)}}{\partial q_2}\right\rangle_3\\
&=\frac{\beta\pi_2}{2}M_3m_1(m_2-1)\left(-\varphi_1\varphi_2+\left\langle\varphi_1^{(2)}\varphi_2^{(2)}\right\rangle_3\right)\\
&\quad+\left\langle\frac{\beta\pi_2}{2}m_1(m_2-1)\left\{-m_1m_2\varphi_1^{(2)}\varphi_2^{(2)}+m_1(m_2-2)\left\langle\psi_1^2\langle\tanh^2\Xi\rangle_1\right\rangle_2\right.\right.\\
&\quad\left.\left.+2(m_1-2)\left\langle\psi_1\langle\tanh^3\Xi\rangle_1\right\rangle_2+4\varphi_2^{(2)}\right\}\right\rangle_3\\
&=\frac{\beta\pi_2}{2}m_1(m_2-1)\left\{-M_3\varphi_1\varphi_2+m_1m_2(m_3-1)\left\langle\varphi_1^{(2)}\varphi_2^{(2)}\right\rangle_3+4\varphi_2\right.\\
&\left.+m_1(m_2-2)\left\langle\left\langle\psi_1^2\langle\tanh^2\Xi\rangle_1\right\rangle_2\right\rangle_3+2(m_1-2)\left\langle\left\langle\psi_1\langle\tanh^3\Xi\rangle_1\right\rangle_2\right\rangle_3\right\}. \qquad \text{(D.11)}
\end{aligned}
$$

From this result, in turn, immediately follows

$$\begin{aligned}\frac{\partial\varphi_2}{\partial q_1} &= \frac{\alpha_1}{\alpha_2}\frac{\partial\varphi_1}{\partial q_2} = \frac{\pi_1(m_1-1)}{\pi_2 m_1(m_2-1)}\frac{\partial\varphi_1}{\partial q_2}\\ &= \frac{\beta\pi_1}{2}(m_1-1)\Big\{-m_1m_2m_3\varphi_1\varphi_2 + m_1m_2(m_3-1)\big\langle\varphi_1^{(2)}\varphi_2^{(2)}\big\rangle_3 + 4\varphi_2\\ &\quad + m_1(m_2-2)\left\langle\left\langle\psi_1^2\langle\tanh^2\Xi\rangle_1\right\rangle_2\right\rangle_3 + 2(m_1-2)\left\langle\left\langle\psi_1\langle\tanh^3\Xi\rangle_1\right\rangle_2\right\rangle_3\Big\}. \qquad \text{(D.12)}\end{aligned}$$

Calculation of $\partial\varphi_1/\partial q_1$ and $\partial\varphi_2/\partial q_2$: Finally, let us turn to the remaining two partial derivatives, $\partial\varphi_1/\partial q_1$ and $\partial\varphi_2/\partial q_2$:

$$\begin{aligned}\frac{\partial\varphi_1}{\partial q_1} &= -\frac{1}{K_3}\frac{\partial K_3}{\partial q_1}\varphi_1 + \frac{1}{K_3}\int Dz_3\frac{\partial}{\partial q_1}\big[K_2^{m_3}\varphi_1^{(2)}\big]\\ &= -\frac{1}{K_3}\frac{\partial K_3}{\partial q_1}\varphi_1 + \frac{m_3}{K_3}\int Dz_3 K_2^{m_3-1}\frac{\partial K_2}{\partial q_1}\varphi_1^{(2)} + \left\langle\frac{\partial\varphi_1^{(2)}}{\partial q_1}\right\rangle_3\\ &= \frac{\beta\pi_1}{2}\Big\{-M_3\{1+(m_1-1)\varphi_1\}\varphi_1 + M_3\left\langle\left\{1+(m_1-1)\varphi_1^{(2)}\right\}\varphi_1^{(2)}\right\rangle_3\\ &\quad -(m_1-1)M_2\left\langle\left(\varphi_1^{(2)}\right)^2\right\rangle_3 + m_1(m_1-1)(m_2-1)\left\langle\left\langle\langle\tanh^2\Xi\rangle_1^2\right\rangle_2\right\rangle_3\\ &\quad +(m_1-2)(m_1-3)\left\langle\left\langle\langle\tanh^4\Xi\rangle_1\right\rangle_2\right\rangle_3 + 4(m_1-2)\left\langle\varphi_1^{(2)}\right\rangle_3 + 2\Big\}\\ &= \frac{\beta\pi_1}{2}\Big\{-(m_1-1)M_3(\varphi_1)^2 + (m_1-1)M_2(m_3-1)\left\langle\left(\varphi_1^{(2)}\right)^2\right\rangle_3\\ &\quad + m_1(m_1-1)(m_2-1)\left\langle\left\langle\langle\tanh^2\Xi\rangle_1^2\right\rangle_2\right\rangle_3\\ &\quad +(m_1-2)(m_1-3)\left\langle\left\langle\langle\tanh^4\Xi\rangle_1\right\rangle_2\right\rangle_3 + 4(m_1-2)\varphi_1 + 2\Big\} \qquad \text{(D.13)}\end{aligned}$$

where we use (D.3), and, similarly, using (D.2)

$$\begin{aligned}\frac{\partial\varphi_2}{\partial q_2} &= -\frac{1}{K_3}\frac{\partial K_3}{\partial q_2}\varphi_2 + \frac{1}{K_3}\int Dz_3\frac{\partial}{\partial q_2}\left(K_2^{m_3}\varphi_2^{(2)}\right)\\ &= -\frac{\beta\pi_2}{2}M_3m_1(m_2-1)(\varphi_2)^2 + m_3\left\langle\frac{\beta\pi_2}{2}M_2m_1(m_2-1)\varphi_2^{(2)}\varphi_2^{(2)}\right\rangle_3\\ &\quad +\Big\langle\frac{\beta\pi_1}{2}\Big\{-m_1^2m_2(m_2-1)\left(\varphi_2^{(2)}\right)^2 + m_1^2(m_2-2)(m_2-3)\langle\psi_1^4\rangle_2\\ &\quad +4m_1(m_2-2)\varphi_2^{(2)} + 4m_1(m_1-1)(m_2-2)\left\langle\psi_1^2\langle\tanh^2\Xi\rangle_1\right\rangle_2\\ &\quad +2+4(m_1-1)\varphi_1^{(2)} + 2(m_1-1)^2\left\langle\langle\tanh^2\Xi\rangle_1^2\right\rangle_2\Big\}\Big\rangle_3\\ &= \frac{\beta\pi_2}{2}\Big\{-M_3m_1(m_2-1)(\varphi_2)^2 + m_1^2m_2(m_2-1)(m_3-1)\left\langle\left(\varphi_2^{(2)}\right)^2\right\rangle_3\\ &\quad +m_1^2(m_2-2)(m_2-3)\left\langle\left\langle\psi_1^4\right\rangle_2\right\rangle_3\\ &\quad +4m_1(m_2-2)\varphi_2 + 4m_1(m_1-1)(m_2-2)\left\langle\left\langle\psi_1^2\langle\tanh^2\Xi\rangle_1\right\rangle_2\right\rangle_3\\ &\quad +2+4(m_1-1)\varphi_1 + 2(m_1-1)^2\left\langle\left\langle\langle\tanh^2\Xi\rangle_1^2\right\rangle_2\right\rangle_3\Big\}. \qquad \text{(D.14)}\end{aligned}$$

References

[1] Shinomoto S 1987 *J. Phys. A: Math. Gen.* **20** L1305
[2] Dong D W and Hopfield J J 1992 *Network* **3** 267
[3] Coolen A C C, Penney R W and Sherrington D 1993 *Phys. Rev.* B **48** 16 116
[4] Penney R W, Coolen A C C and Sherrington D 1993 *J. Phys. A: Math. Gen.* **26** 3681
[5] Dotsenko V, Franz S and Mézard M 1994 *J. Phys. A: Math. Gen.* **27** 2351
[6] Penney R W and Sherrington D 1994 *J. Phys. A: Math. Gen.* **27** 4027
[7] Feldman D E and Dotsenko V S 1994 *J. Phys. A: Math. Gen.* **27** 4401
[8] Caticha N 1994 *J. Phys. A: Math. Gen.* **27** 5501
[9] Jongen G, Bollé D and Coolen A C C 1998 *J. Phys. A: Math. Gen.* **31** L737
[10] Jongen G, Anemüller J, Bollé D, Coolen A C C and Pérez-Vicente C J 2000 *J. Phys. A: Math. Gen.* **34** 3957
[11] Sherrington D and Kirkpatrick S 1975 *Phys. Rev. Lett.* **35** 1792
[12] Mézard M, Parisi G and Virasoro M A 1987 *Spin Glass Theory and Beyond* (Singapore: World Scientific)
[13] Van Mourik J and Coolen A C C 2001 *J. Phys. A: Math. Gen.* **34** L111
[14] Torres J J, Garrido P L and Marro J 1997 *J. Phys. A: Math. Gen.* **30** 7801
[15] Heerema M and Van Leeuwen W A 2000 *J. Phys. A: Math. Gen.* **33** 1781
[16] Hertz J and Prügel-Bennet A 1996 *Network* **2** 357
[17] Bressloff P C 1999 *Phys. Rev.* E **60** 2160
[18] Hertz J and Solinas S 2001 *Neurocomputing* **38** 915
[19] Lattanzi G, Nardulli G, Pasquariello G and Stramaglia S 1997 *Phys. Rev.* E **56** 4567
[20] Mannarelli M, Nardulli G and Stramaglia S 2001 *Phys. Rev.* E **64** 052904
[21] Abramowitz M and Stegun I A (ed) 1970 *Handbook of Mathematical Functions* (New York: Dover)
[22] de Almeida J R L and Thouless D J 1978 *J. Phys. A: Math. Gen.* **11** 983