Referee report on the manuscript

Annealed averages in spin and matrix models

by L Foini and J Kurchan

The authors of this manuscript investigate the differences between statistical mechanical spin models in which interaction parameters constitute quenched (i.e. frozen) disorder, and those where the interaction variables are not frozen but have a status similar to the spin variables (i.e. they represent annealed disorder). In most of section 2 they study various model versions involving spherical spin systems, where after rotation in the space of spin vectors one can eliminate the eigenvectors of the interaction matrix from the problem and find expressions involving only the statistics of the eigenvalues of the interaction matrices. Here, provided the ensemble of interaction matrices is of a certain form, they can demonstrate the appearance of interesting condensation phenomena in the annealed cases, leading to self-induced Mattis-type states. The annealed solution is written in terms of the so-called R-transform (which in turn is defined in terms of the Stieltjes transform). In section 3 the authors show that the calculation of the marginal distributions for individual entries of random matrices whose distributions have certain symmetry properties can be mapped to annealed averages for spherical spin systems as introduced and studied in section 2.

On the whole I found this an interesting paper. However, I would like the others to take into account the suggestions below.

1. There are some typos/imperfections in the notation that should be fixed. For instance:
   (a) eqn (1): $\sum_i \tilde{s}_k^2$ should be $\sum_i \tilde{s}_i^2$
   (b) eqn (1): $\sum_k \lambda_k \tilde{s}_k \tilde{s}_k$ should be $\sum_k \lambda_k \tilde{s}_k^2$
   (c) eqn (2): $E(\lambda)$ should be $E(\boldsymbol{\lambda})$ or $E(\lambda_1, \ldots, \lambda_N)$
   (d) section 2.8, line 6: 'ans' should be 'and'
   (e) eqn (35): $\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2 = \pm 1$ should be $\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2 \in \{-1, 1\}^N$
   (f) eqn (36): remove the large space between $\int$ and $\mathrm{d}q$

2. The authors use the notation $\mathrm{D}\boldsymbol{x} \equiv \prod_{i=1}^N \mathrm{d}x_i$ throughout their paper. This I find confusing. One would normally simply have written this product as $\mathrm{d}\boldsymbol{x}$, and moreover in the disordered systems community the upper case D has traditionally already been used to denote the zero-average and unit-variance Gaussian measure, i.e. $\mathrm{D}\boldsymbol{x} = (2\pi)^{-N/2} \mathrm{e}^{-\frac{1}{2}\boldsymbol{x}^2}$. I see no reason for or benefit of this deviation from standard practice.

3. The spherical spin models are consistently introduced with *hard* constraints on the spin vector length, imposed via delta-functions. However, in the subsequent calculations they actually use *soft* constraints: they use a real-valued Lagrange parameter z, whereas the delta function would have given an imaginary one. In steepest descent expressions for mean-field models one would for $N \to \infty$ ultimately find the same end result, via contour deformation. Here, however, with the unexpected condensation phenomenon in the annealed case, it is not a priori obvious to me that the soft and the hard constrained model can be interchanged.

4. It is not fully clear why the physics of the problem follows for annealed disorder from minimization of the relevant Hamiltonians over $\boldsymbol{\lambda}$ and z. Thse terms that appear in the exponent of the measure that contain eigenvalues only are not coupled to the heat bath

(they have no β), so it is not down to looking for ground states. If steepest descent is the argument, then one should devote some attention to the fact that we here have also $\mathcal{O}(N)$ integration variables. Is it clear that the curvature around the minimum (and possibly higher order terms) do not contribute to the thermodynamics for large N?

5. In section 3.1 the calculations involve normalised vectors $\boldsymbol{\sigma} \in \mathrm{I\!R}^N$. Why do the authors use matrix notation such as the Hermition conjugation symbol, and write e.g. $\boldsymbol{\sigma}_a^\dagger \boldsymbol{\sigma}_b$ instead of simply $\boldsymbol{\sigma}_a \cdot \boldsymbol{\sigma}_b$, and $\boldsymbol{\sigma}_a^\dagger A \boldsymbol{\sigma}_b$ instead of $\boldsymbol{\sigma}_a \cdot AS\boldsymbol{\sigma}_b$? The present notation wrongly suggests to the reader that $\boldsymbol{\sigma}_a$ is a matrix.

6. In the context of comparing annealed to quenched disorder and interpretations of slowly evolving interactions between spins, the authors could perhaps refer also to a set of early papers in which that idea was developed in a more general context, starting with Penney et al (J Phys A26, 1993), and followed by multiple extensions and applications in physics and biology (see e.g. Rabello et al, J Phys A41, 2008 for further references). The quenched and annealed scenarios are just two special cases ($n = 0$ and $n = 1$) of a more general family, with even more extreme self-induced Mattis type status for $n > 1$. One could even ask at which value of n between 0 and 1 the condensation discussed in the present paper happens.

Once the issues above have been dealt with, I would expect this manuscript to be suitable for publication.