**AE Report (AoS 1910-054)**

I have now read the paper "Five shades of grey: phase transitions in high-dimensional multiple testing" by Zheng Gao. The paper extends the results on the phase transitions for the sparse normal means problem to the setting of $\chi^2$ setting, where all $\chi^2$s have the same degree freedoms, and the non-central parameters for most of the $\chi^2$ are 0, except for a small fraction of them is elevated. The paper adopts the model in Donoho and Jin (2004) to model the sparsity and non-central parameters for the $\chi^2$, and derive the phase transition.

The main contribution of the paper seems to be

- Extend the phase transition from the normal means problem to the $\chi^2$ setting.
- Relate the theoretical problem to an interesting application problem.

The paper is not particularly well-written. The model and definitions can be better organized, and the results can be better elucidated. I have the following detailed comments.

- References: there are a long list of work on phase transition, and many concepts introduced in the paper are explicitly and implicitly introduced in earlier works. While these concepts are originally for the linear regression setting and in the normal means setting, they continue to be valid for the $\chi^2$-settings.

  For example, the definitions in many places of Section 2.1, as well as Definition 2.1, are not new, but there is no single citation to the existing literature. For the 5 curves in Figure 1, 3 of them are explicitly defined in existing literature, one of them is implicitly introduced in the literatures. These facts have to be carefully clarified.

  Of course, these are only examples. There are numerous other places where the author needs to cite existing literature in a better way.

- Lower bound. The lower bound argument is in the *weak sense* and is therefore not sharp either. For example, in Theorem 3.1 and 3.2, the author only claims *the failure for thresholding procedures* in certain parameter region. However, in the literature, it has already been shown that *any procedure would fail* for the same parameter region, in more difficult settings than considered here (i.e., individual tests are correlated).

  Showing the failure of thresholding procedure only involves elementary calculations (e.g., Mill's ratio), but showing the failure of all methods involves a very difficult analysis of $L^1$-distance (or Hellinger distance). The latter is much more difficult.

- Technical difficulty: the technical difficulty is not seen to be very high, if the focus is only on the thresholding procedure, as the evaluation of the thresholding procedure only involves the Mills' ratio type of approximation and some elementary calculations.