Response to the reviews of the manuscript: “Concentration of Maxima and Fundamental Limits in High-Dimensional Testing and Inference” by Zheng Gao and Stilian Stoev submitted to *SpringerBriefs Series in Probability and Statistics*

March 10, 2021

Summary of the revision

We sincerely thank the managing Editor and the five anonymous referees for their very helpful and constructive suggestions. We really appreciate the reviewers’ time, effort, and thoughtful suggestions, which stimulated us to substantially revise and improve the manuscript. We begin with a high-level description of our edits and in the following section we respond to specific points raised by the reviewers.

- **(Brief description)** The revised manuscript has now a brief introductory chapter and 6 substantive chapters, which amount to a total of 124 pages. As with the original submission, Chapter 2 reviews the statistical risks, estimation and testing procedures, as well as probability models for the errors used in the rest of the manuscript. It contains also a high-level summary of our results in the context of the existing literature and some auxiliary facts from probability and extreme value theory. For convenience of the reader, we have added statements of Slepyan’s Lemma and the Sudakov-Fernique inequality (without proofs).

  Chapter 3 provides a panorama of the phase-transition results in the signal detection and support estimation problems for the additive iid Gaussian error model. Its purpose is to emphasize ideas and provide an accessible account, including references to the state-of-the-art in the literature. The chapter contains also some new results on the (sub)optimality of some popular statistics, which appear to have been overlooked in the literature (Theorem 3.1). One main idea behind the exposition is to systematically present the different types of phase transition results as a function of the chosen statistical risk.

  Chapters 4 through 7 contain the core of our original contributions.

  Chapter 4 establishes phase-transitions results for the exact support recovery problem in the additive error model. The main themes of this chapter are *error–dependence* and *concentration of maxima*. We establish that the notion of uniform relative stability – a type of concentration of maxima phenomenon – is key to characterizing the phase-transition in exact support recovery for the general class of thresholding estimators.

  Chapter 5 begins with a study of the finite-sample (Bayes) optimality and sub-optimality of the threshold-based support estimators. Criteria for the latter are established and certain theoretical results indicate that *likelihood thresholding* (as opposed to data thresholding) is in general optimal, when the thresholding estimators are sub-optimal. In the regime when the thresholding estimators are optimal, i.e., for log-concave error densities, our results entail a universal statistical limit in the exact support recovery problem, valid for all estimators (Theorem 5.3). On the other hand, for the general class of thresholding estimators, we obtain minimax characterizations of the phase-transition in exact support recovery, for every fixed error-dependence structure that satisfies mild uniform relative stability condition (Corollary 5.3).

Chapter 6 is nearly unchanged. It is more probabilistic in nature and it provides a complete characterization of the uniform relatively stable (URS) Gaussian triangular arrays. This URS condition is the key property, in the lower bound on exact support recovery under dependence in Chpter 4. These probabilistic results as well as their methods of proof may of independent interest and are accessible to a wider audience of graduate students in statistics.

Chapter 7 focuses on an application to genome-wide association studies (GWAS) in statistical genetics. It is shown that all phase-transition results in the additive error models of Chapter 3, have their close counterparts for the chi-square models arising naturally in GWAS. To connect these asymptotic results to the practical notions in statistical genetics, we establish the connection between the odds-ratio (effect size) in multinomial 2×2 models and the signal-size parameter of the corresponding chi-square association test. This allows us to quantify the statistical power and optimal design questions as well as ultimately explain the role of phase-transitions in the fundamental statistical limits of GWAS.

- **(Organization)** Following a suggestion of a reviewer (**R2**), we have included most proofs in the text right after the corresponding claims (theorems, propositions, lemmas). In the interest of space and to conform with the format of the SpringerBriefs series, we no longer include "Exercise" sections.

  Last but not least, to better balance the exposition, we decided to divide the previous Chapter 4 into two new chapters (now Ch. 4, entitled "Exact Support Recovery Under Dependence", and Ch. 5 "Bayes and Minimax Optimality"). In addition to making all chapters more manageable and of about equal length, it exposes more clearly our contributions. The results contained in Chapters 4 and 5 as well as the more probabilistic Chapter 6 and more applied Chapter 7 are original and new contributions to the literature.

Responses to specific points raised by the reviewers