We thank the reviewers for their analysis and suggestions. We would like to reply to specific points:

1. We agree with Reviewer 3 that the paper emphasizes some philosophical aspects; intentionally so. Sadly we can't make all kinds of readers happy; a choice of audience is necessary. As readers ourselves we appreciate when a paper begins by asking: 'What is the question? is it possible to translate it into a mathematical problem? which principles can we use to make such translation?'. There are papers that develop very refined mathematical techniques but leave us unsatisfied, with the lingering question 'why are we doing this?'. This is the reason why we try to emphasize these kinds of questions. But we are sure that part of the NIPS audience will appreciate this emphasis.
2. We also agree with Reviewer 3 that the main formulae of the paper, (17) and (19), probably have little experimental use today. But the point they make is important:
   - the Reviewer says 'the differences between the sample level and population level distribution seem minor' – but we could not have known this, if we hadn't faced the whole problem and derived a formula showing that the difference is minor;
   - the sample-level distribution is actually *less* entropic. By lifting the constraints to the population level, the maximum-entropy principle allows us to explore even *more entropic* sample distributions than those allowed by a direct application at the sample level. Thus this approach agrees with the 'maximum entropy' spirit. – These points should be mentioned, of course, in an amended version of the paper.
   -