The author(s) thank the three reviewers for their feedback. Here are our replies point-by-point.

Reviewer 1 ("Assigned_Reviewer_1")

1. We agree that the introduction part did not give enough neuroscientific introductory context and literature. The maximum-entropy method at the "sample level" has actually been repeatedly used and ~~analysed~~discussed in computational neuroscience since the work of Bohte et al (2000) ~~to the present~~. A more complete bibliography is given at the end of this response and ~~can be easily extended~~will be added to the manuscript. The method is also used by non-experts, ~~and our point would be~~motivating the need to clarify for them some implicit assumptions ~~implicit the method also for non-experts~~ they largely seem to be unaware of.

2. The reviewer is again right. Our discussion starts from the ~~*constatation*~~observation that most ~~of the neuroscientific literature~~authors ~~implicitly or explicitly~~ generalize conclusions drawn by maximum entropy ~~s the statistical properties of a~~on sample of recorded neurons, ~~analysed via maximum-entropy,~~ to whole areas the neurons are recorded from, eg retina or motor cortex. ~~This generalization is a complex argument lying outside our competence; we cannot justify or deprecate it in this manuscript. Our~~The purpose of our contribution is to ~~warn neuroscientists~~launch a warning ~~who make this generalization~~: the typical sample-level approach (Eq. 17) ~~method commonly used in the literature~~ is ~~inconsistent~~incompatible with this generalization. We ~~are~~also propose~~ing~~ a simple and more consistent ~~approach~~alternative (Eq. 15).

The case of absolute ignorance (I") is not used in the rest of the manuscript and was confusing; we would omit it in a revised version.

We invite the reviewer to read also our responses 1 & 2 to the second reviewer.

3. We again agree: our principal equations are:

(15) to be used when a larger population is assumed and its size roughly known;

(17) *the equation used in the literature*, to be used when no larger population is assumed (practically, never the case);

(19) to be used when the the size of the larger population is ~~unknown~~uncertain, but can be assigned a probability distribution.

Equations (15) and (19) are novel to our knowledge. Unfortunately we failed to emphasize the role of these three equations and give clear guidance to their use. We ~~would~~are happy to amend this in a revised version.

4. The captions should indeed give a clearer explanation of the figures. We would amend this in a revised version.

Reviewer 2 ("Assigned_Reviewer_2")

1. We agree with the reviewer, we did not give enough context and literature. Our starting point is indeed that

- the *sample-level* maximum-entropy method is increasingly being used in the neuroscientific literature (references below, to be added to the manuscript);

- this literature often generalizes results obtained by this method from a recorded sample to a whole area, or even to draw general principles on the role of correlations in the cortex;

- worries have already been expressed and mathematically formulated (eg Ganmor et al 2011, Roudi et al PlosCB 2009) about the sample-size dependence of maximum-entropy results.

Our manuscript ~~would like to explain~~explains and re-interprets these worries starting from simple probabilistic~~ty~~ considerations ~~why these worries are motivated, why the results depend on the sample size, and what~~and provides alternative formulae ~~can~~to be used: eq~~.ns~~ (15) when the size N of the full population is known, Eq. (19) when it is uncertain.

2. ~~The reviewer is right, w~~We should have indeed mentioned and referenced to an important fact: the sample approach is routinely used in the literature. ~~We agree that it~~It does~~is~~ not provide an exact solution~~option~~ if ~~we assume sampling from~~the observed samples is embedded into a larger population. Our manuscript ~~wanted to~~ explains why ~~it is not an option~~, and proposes ~~more valid~~two ~~alternatives~~amendments (eqns 15, 19). We will better emphasize this line of argument in the revised manuscript.

The phrase "left for future work" was unfortunate. Eq~~.n~~ (19)~~,~~ for ~~unknown~~uncertain population size~~, can actually be~~is actually an end result, ready to use~~d~~ and numerically unproblematic to compute~~d as is~~. What we ~~L~~left for future work was ~~is~~ the ~~analysis~~quantification of its ~~mathematical properties~~actual difference from the distributions (15) and (17). In a revised version of the paper we ~~would~~plan to add to each existing figure also an example of the distribution ~~obtained by eqn~~ (19).

We invite the reviewer to read also our responses 1 & 2 to the first reviewer.

3. ~~It is true, we give much space to the old formulae of sampling without replacement. This was on purpose, see lines 42-45: "The relations we present are well-known in survey sampling and in the pedagogic problem of drawing from an urn without replacement, yet they are somewhat hard to find explicitly written in the neuroscientific literature, so they may be of interest on their own". We should have stressed more that our manuscript wanted to "advertize" the importance of these formulae, not much used in neuroscience but very central to it.~~
While large space was dedicated to deriving the known result Eq. (17), we did also deliver novel and important extensions to it, Eqs (15) and (19). We realize now that the failed in clearly separating these, and in highlighting the importance and role of the latter. Probably our statement "The relations we present are well-known [...], yet they are somewhat hard to find explicitly written in the neuroscientific literature..." was shadowing these achievements. We will reformulate accordingly (see also our responses to your previous two questions).

4. Equation (12) and its discussion (lines 113-129) are an excursion irrelevant to the rest of the paper, to the discussion of maximum-entropy and to the main equations we derive. We would eliminate this whole passage in a revised version, in favor for a better explanation of existing results.

5. We would have also liked to present a Bayesian analysis of the problem, but the page limit of the NIPS format prevented this~~; the manuscript could become confusing~~. In a revised version we would add a measure of discrepancy, eg Kullback-Leibler divergence, between the "sample" and "population" approaches to quantify the difference between Eqs. (15), (17) and (19).

Reviewer 3 ("Assigned_Reviewer_5")

We cordially thank the reviewer for the ~~appreciative~~supportive words. We would modify the manuscript to address the concerns of the other reviewers but leaving the theoretical points. We want to add that our figures do use actual recorded neural activity data from a previous manuscripts of ours, but because of the double-blind review procedure we ~~could not give more details~~omitted this information, which we plan to add back in the revised version.~~ and made comparison with recorded frequency distributions in the reviewed version.ÿ~~

Sample references of maximum-entropy use in neuroscientific literature:

Martignon & al, Biol Cybern (1995)
Bohte & al, Neural Comp (2000)
Schneidman & al, Nature (2006)
Shlens & al, J Neurosci (2006)
Tkačik & al, arXiv:q-bio/0611072 (2006)
Macke & al, Neural Comp (2009)
Roudi & al, PLoS CB (2009)
Roudi & al, Front Comp Neurosci (2009)
Roudi & al, Phys Rev E (2009)
Tkačik & al, arXiv:09125409 (2009)
Macke & al, Tech Rep (2009)
Gerwinn & al, Front Comput Neurosci (2010)
Barreiro & al, arXiv:10112797 (2010)
Macke & al, NIPS (2011)
Macke & al, Phys Rev Lett (2011)
Ganmor & al, Proc Natl Acad Sci (2011)
Granot-Atedgi & al, PLoS CB (2013)
Macke & al, Entropy (2013)
Tkačik & al, Proc Natl Acad Sci (2014)
Mora & al, Phys Rev Lett (2015)
Shimazaki & al, Sci Rep (2015)