We thank all referees for their interest in our work and their comments that will help to clarify our paper.

**R1 - "There is no unified framework for the analysis of weak and strong thresholds."** We respectfully disagree: $\alpha_{\mathrm{FR,IT}}$ is in general not well-defined for an arbitrary channel, which motivates our restriction to the noiseless case. Our analysis is valid for any right-orthogonally invariant data matrix $\mathbf{\Phi}$ with well-defined asymptotic density, and a Gaussian prior (which is limiting, although we believe it can be relaxed and we will work towards this). Concerning $\alpha_{\mathrm{WR,Algo}}$, eq. (11) holds in full generality, i.e. for any data matrix $\mathbf{\Phi}$ as above, and any phase-retrieval probabilistic channel. Eqs.(12),(13) are examples derived from this generic formula. We will clarify on the generality of our results in the revised version. **"The rigorous analysis relies on some Gaussianity, either in the prior or in the data matrix"**. The referee classifies this as a major weakness: while we agree this is a restrictive assumption, it is a fundamental limitation of the interpolation method used for the proof, which will be clarified. We wish to indicate the reviews of **R2** and **R4**, that we thank for underlining the generality of our framework and of our rigorous analysis. **"One weakness is that G-VAMP requires knowledge of the distribution of the true signal".** We would like to emphasize that the algorithm is also well-defined beyond this scope, e.g. it can be used to infer natural images with Fourier matrices. Using a Gaussian prior to infer is actually the minimal assumption on the underlying signal, as it amounts to simply fix its norm: we see this as a strength of our theory, which can predict the G-VAMP performance for any signal, structured or not. We discuss this further in the response to **R4**, and we will clarify this point. We finally thank the referee for pointing out typos, and providing additional references that we will add to the paper.

**R2 -** We thank the referee for her/his appreciation of our work.

**R3 - "The paper is not very clearly stating which results are rigorous + Confusion on the product of Gaussians".** Our analysis in Sections 3-4 relies on Conjecture 2.1, and is thus rigorous whenever the conditions for Theorem 2.2 hold. In this theorem, the matrix $\mathbf{B}$ can be random or deterministic, as long as it satisfies the assumptions of the theorem, which is the case for $\mathbf{B} = \mathbf{W}_2$ i.i.d. Gaussian. We acknowledge this should be clarified in the text, and we hope this will answer the question of the referee. **'The conjectured optimality of G-VAMP".** Indeed we refer to the G-VAMP threshold as "algorithmic", even if a proof for the optimality of G-VAMP is not given. We adopted this notation for consistency with the previous literature on this topic, in which this conjectured optimality is often assumed. We will add a note on this choice on the paper. **"Do the authors prove the existence of a gap ?".** We emphasize that we provide scalar equations (rigorous when in the setting of Theorem 2.2) that can be used to find $\alpha_{\mathrm{FR,Algo}}$. Apart from the inevitable numerical solution, our analysis is thus well-controlled. This discussion will be added in the paper, and we thank the referee for helping clarify this point. **"Can we analyze $\boldsymbol{\alpha}_{\mathbf{WR,IT}}$ ?".** Extending our analysis to $\alpha_{\mathrm{WR,IT}}$ is an interesting open direction, which requires understanding the appearance of a global maximum in the replica-symmetric potential, but not necessarily continuously from the $q = 0$ solution as in the case of $\alpha_{\mathrm{WR,Algo}}$. At the moment we are not able to carry such an analysis, and we will discuss this more extensively in the revised paper. **"On the all-or-nothing transition".** We have observed these transitions for orthogonal/unitary matrices : as stated in the paper, we expect to see this phenomenon for other real matrices. We will discuss this further, as well as the possible dependency on the prior, and add the references provided. We finally thank the referee for the list of typos and comments, all of which will be addressed in the revised version.

**R4 - "I don't expect these results to extend to natural images".** We thank the referee for indicating towards such an analysis, which would be a valuable add to our work. We conducted a simple experiment on a natural image, and the result is given in Fig. 1. Although we are far from a Bayes-optimal setting, the achieved MSE is very close to values of Fig. 2 of the paper, for all values of $\alpha$. In particular, we achieve perfect recovery for $\alpha \geq 2.3$, just above $\alpha_{\mathrm{FR,Algo}} \simeq 2.27$ which was derived for random unitary matrices, i.i.d. data and in the Bayes-optimal setting. As all normalized signals are equivalent under a Gaussian prior (it is a "maximum-entropy" prior), we indeed expect a structured signal to perform exactly as a random one as long as one also infers the signal using a Gaussian prior. This observation is coherent with Fig. 1 amd strengthens the relevance of our theoretical results for real data, and we will discuss this point further in the final version. We point out previous works that investigated the performance of AMP algorithms in phase retrieval [1, 2].

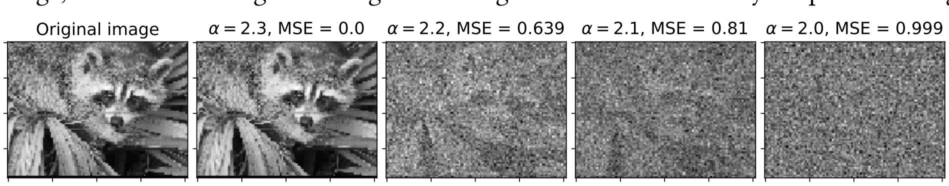


Figure 1: Performance of the G-VAMP algorithm for noiseless phase retrieval. We wish to recover a 77x102 image (on the left), and we use a complex Gaussian prior to infer the signal. The data matrix $\mathbf{\Phi}$ is a randomly subsampled DFT matrix.

[1] Junjie Ma, Ji Xu, and Arian Maleki. Optimization-based amp for phase retrieval: The impact of initialization and l2 regularization. *IEEE Transactions on Information Theory*, 65(6):3600–3629, 2019.

[2] Philip Schniter and Sundeep Rangan. A message-passing approach to phase retrieval of sparse signals. In *Excursions in Harmonic Analysis, Volume 4*, pages 177–204. Springer, 2015.