We thank the reviewers for providing useful feedback. To address a shared comment, we first explain our contribution in uncertainty quantification, compared to multiple imputation (MI) methods. While previous MI work can quantify sample variance, few papers explicitly explore the issue of *calibration*: does MI sample variance predict imputation accuracy? Our paper does address calibration: imputation accuracy correlates with our uncertainty metric. Moreover, our method allows for fast large-scale computation and uses only a single easy-to-choose parameter: rank. In contrast, Bayesian MI methods are less user-friendly: they are sensitive to the choice of prior and the selection of (often many) tuning parameters. We compare our method with one of the fastest MI methods, MIPCA (Josse et al. 2011), on synthetic data (here, Figure 1), which shows the MI sample variance does not predict imputation accuracy well. Worse, on the Movielens 1M data, even a single imputation of MIPCA cannot finish in 3h (and 20 imputations are usually used to quantify uncertainty); our method takes 38m. The author of the famous MI method MICE warns against its use on data sets with many columns due to both speed and quality consideration (Van Buuren 2018 Chapter 9.1). On our simulated low rank continuous matrix (500×200), a single imputation of MICE ran 4m, while our method only ran 7s. Hence our new proposed uncertainty measure outperforms competitors in both accuracy and speed.

Reviewer #2 asks us to clarify the relation between our work, multiple imputation (MI), and matrix completion (MC). We use the term MC to mean the task of imputing missing values in a tabular data set (a broader definition than R2 uses), and MI to mean methods that provide multiple imputed values for each missing entry. While many so-called MC methods assume a deterministic signal (and so require deterministic assumptions e.g. incoherence), probabilistic models including MI are also widely used for tabular imputation (and so require probabilistic assumptions e.g. subGaussianity, as in our Cor. 1). Our method is a new probabilistic single imputation approach for tabular imputation. It can also be used for MI if desired.

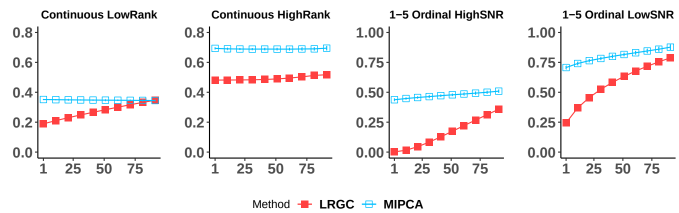

Figure 1: Imputation error (NRMSE for continuous and MAE for ordinal) on the subset of $m\%$ entries for which method's associated uncertainty metric indicates highest reliability. For MIPCA, we use 20 imputations; low sample variance corresponds to high reliability.

Reviewer #2 doubts the novelty of our methods. This paper presents the first imputation method using the Gaussian copula that scales to the recommendation system setting. While there exist methods for MI using the Gaussian copula, they all use MCMC and require large computation to achieve desirable imputation accuracy. Zhao & Udell (2019) proposed a frequentist algorithm for imputation using Gaussian copula, which is faster than the MCMC algorithms but still scales as the cube of the number of columns. Our methodological innovation is to use low rank factorization to reduce the complexity, which yields new probabilistic error guarantees (Thm 3).

Response to Reviewer #3 We appreciate the positive comments and have added the suggested references. Unfortunately, the honest confidence intervals proposed by Carpentier et al. (2017) depend on (possibly huge) hidden constants.

Reviewer #4 suggests comparing to DataWig (and MICE, see above). Datawig is an imputation method aimed at non-numerical (e.g. text) missing data without a direct way to quantify uncertainty, so we cannot see how to compare.

Reviewer #4 asks how we selected hyper-parameters in experiments. As detailed in the supplement, we selected hyper-parameters for all methods using grid-search. We also show our method is not sensitive to its hyper-parameter.

We thank **Reviewer #4** for the suggested references and will add them in the revision, and for the suggestion to call our task (of predicting imputation accuracy) "calibration". But it seems these references do not address the task of calibration for ordinal or real-valued matrices. Our reference [3] does not discuss uncertainty quantification. The reference [A] does not discuss calibration. The reference [B] can calibrate imputation accuracy for Boolean matrices but does not generalize to ordinal or real-valued matrices. We will clarify that "expensive MCMC" applies only to our references [28, 6, 25, 11], the Bayesian approaches most similar to ours, rather than to all Bayesian imputation.

Minor comments from Reviewer #4: (1) In "...deterministic distribution...", you're right that we meant "closed form", not "deterministic". (2) We agree that many low rank models have an underlying generative model, and addressed this point immediately following the sentence quoted by the reviewer. (3) The reviewer's suggestion ("Experiments") results from a misinterpretation of Figure 1 of the paper. The figure shows that MMMF performs worse on the items that MMMF (not LRGC) takes as most reliable. In other words, the reliability metric for MMMF is *negatively* correlated with true reliability. This result indicates sample variance does not correlate with imputation uncertainty.

Response to Reviewer #5. We appreciate the positive comments and will improve the presentation clarity. (1) Analyzing the estimation error is challenging because EM algorithms are only guaranteed to converge to a local maximizer rather than the global maximizer and our objective likelihood function is nonconvex. (2) "Our measure...low NRMSE" is a design principle rather than theoretical property. (3) Empirically, we found the measure using $D_\alpha(i,j)/|\hat{x}^i_j|$ correlates less well with imputation error. (4) We use top-m% entries because the imputation error on a single point yields noisy plots. (5) Our imputation and uncertainty quantification methods extend to mixed data quite easily.