We thank the reviewers for their suggestions. We have revised the paper for clarity and added experiments on hierarchical models requested by reviewers.

R2, R4: Application to hierarchical models. Closely following the techniques used in (Tucker et al. 2017; Grathwohl et al. 2018; Yin and Zhou 2019), we extend DisARM to hierarchical models. We evaluate $2/3/4$-layer linear models on MNIST, Fashion-MNIST, and Omniglot (Figure 1 shows the Omniglot results) and find that DisARM consistently outperforms ARM and REINFORCE-LOO. RELAX outperforms DisARM, however, the gap between two estimators diminishes for deeper hierarchies and training with DisARM is about twice as fast (wall clock time) as with RELAX.

R3: Section 2.1 clarity. We have clarified this section by moving definitions into the main text and providing motivation and intuition for the choices.

R4: Differences with RELAX. RELAX requires gradients from a (learned) surrogate function. While in principle RELAX is generic because the surrogate can be learned from scratch, the strong performance previously reported (and in this paper) relies on a continuous relaxation of the discrete function and only learns a small deviation from this hard-coded relaxation. Moreover, for discrete VAEs, using the continuous relaxation has the same computational cost as working with the discrete model, but in other problems it can be much slower. For example, with conditional computation, the continuous relaxation requires evaluating the entire model, while pure discrete approaches, such as DisARM, evaluate only the parts of the model selected by the discrete gates.

R5: Extension to categorical variables. Yes, in principle the idea for DisARM can be extended to categorical variables. The authors of ARM released an extension ARSM (Yin et al. 2019) for categorical variables and the same idea of analytic integration can be adapted to ARSM to reduce variance. However, we do not yet know if the analytic integration can be done efficiently in this case.

R3, R4: Application to RL. We agree that this would be interesting, and we expect to see similar improvements compared to ARM. However, this would require extending DisARM to the categorical case. Due to this and the complexity of proper evaluation in RL, we feel applications to RL are beyond the scope of this paper.

R5: Usefulness of antithetic samples. The reviewer is correct that depending on the properties of the function, antithetic samples can result in higher variance compared to the same number of independent samples. This can be resolved by constructing an *interpolated estimator*. Briefly, we define a coupling between Bernoulli variables that is parameterized by $\alpha \in [0, 1]$ that smoothly interpolates between independent and antithetic samples. Furthermore, we construct an unbiased estimator parameterized by α for these coupled variables and such that $\alpha = 0$ corresponds to REINFORCE LOO and $\alpha = 1$ corresponds to DisARM. Because this estimator is unbiased for any choice of $\alpha \in [0, 1]$, we can optimize α to reduce variance as in (Ruiz et al. 2016; Tucker et al. 2017) and thus automatically choose the coupling which is favorable for the function under consideration. We have added an appendix section describing this construction and now mention it in the main text. In preliminary experiments, we did not find significant improvements on the datasets we evaluated.

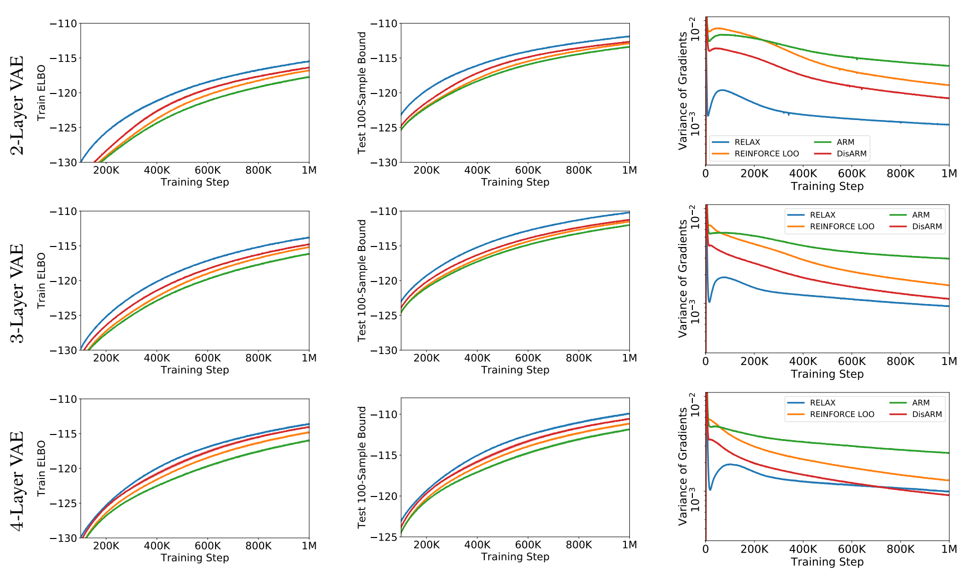

Figure 1: Training 2/3/4-layer Bernoulli VAE on Omniglot using DisARM, RELAX, REINFORCE LOO, and ARM. We report the ELBO on the training set (left), the 100-sample bound on the test set (middle), and the variance of the gradient estimator (right).