Supplementary Material

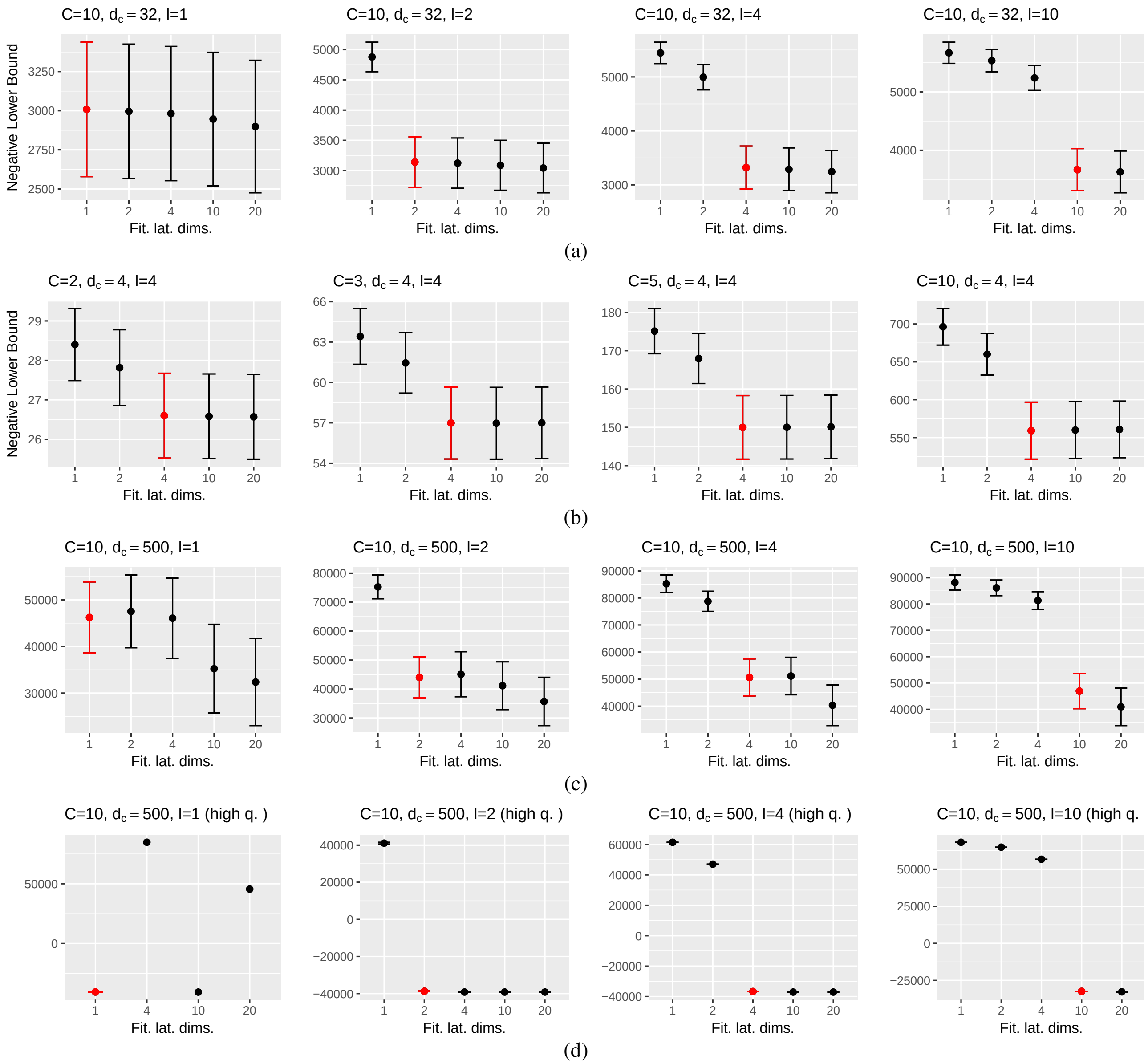
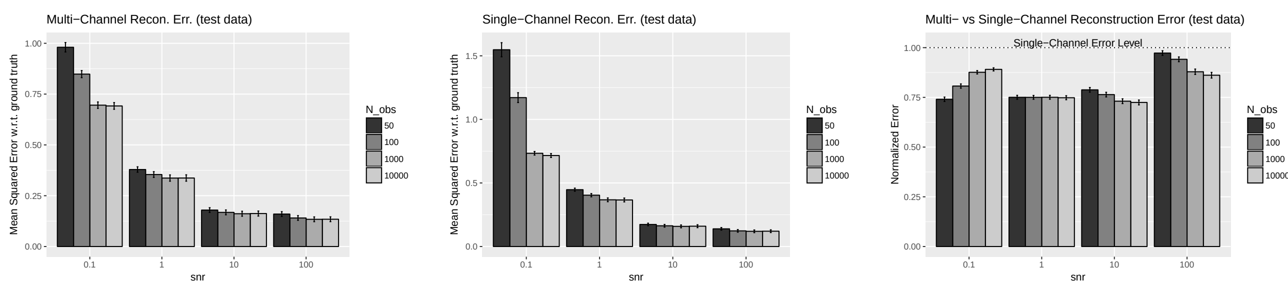


Figure S1: Negative lower bound (NLB) on the synthetic training set computed at convergence for all the scenarios. Each bar shows mean $\pm$ std.err. of $N = 80$ total experiments as a function of the number of fitted latent dimensions. Red bars represents experiments where the number of true and fitted latent dimensions coincide. (a) Experimental setup $C = 10$, $d_c = 32$: NLB stops decreasing when the number of fitted latent dimension coincide with the generated ones; notable gap between the under-fitted and over-fitted experiments (elbow effect). (b) Experimental setup $d_c = 4$, $l = 4$: increasing the number of channels C makes the elbow effect more pronounced. (c) Experimental setup $C = 10$, $d_c = 500$: with high dimensional data ($d_c = 500$) using the lower bound as a model selection criteria to assess the true number of latent dimensions may end up in overestimation. (d) Restricted ($N = 5$ total experiments) high quality experimental setup $C = 10$, $d_c = 500$, $S = 10000$, $snr = 100$: the risk to overestimate the true number of latent dimensions can be mitigated by increasing the snr and S of the observations in the dataset.



(a) (b) (c)

Figure S2: Reconstruction error on synthetic test data reconstructed with the multi-channel model. The reconstruction is better for high snr and high training data sample size. Scenarios where generated by varying one-at-a-time the dataset attributes listed in Tab. 1 for a total of $8\,000$ experiments. (a) Mean squared error from the ground truth test data using the Multi-Channel reconstruction: $\hat{\mathbf{x}}_i = \mathbb{E}_c \left[\mathbb{E}_{q(\mathbf{z}|\mathbf{x}_c, \boldsymbol{\phi}_c)} \left[p\left(\mathbf{x}_i|\mathbf{z}, \boldsymbol{\theta}_i\right) \right] \right]$. (b) Mean squared error from the ground truth test data using the Single-Channel reconstruction: $\hat{\mathbf{x}}_i = \mathbb{E}_{q(\mathbf{z}|\mathbf{x}_i, \boldsymbol{\phi}_i)} \left[p\left(\mathbf{x}_i|\mathbf{z}, \boldsymbol{\theta}_i\right) \right]$. (c) Ratio between Multi- *vs* Single-Channel reconstruction errors: we notice that the error made in ground truth data recovery with multi-channel information is systematically lower than the one obtained with a single-channel decoder.