



Figure S2. Relative error for the decoder (**a,b**) and the bias-corrected estimator (**c,d**), for different population sizes: Continuous lines, $N=250$, dashed lines, $N=50$. Simulations are based on the model used in the main text, with the same parameters except the number of filters. (**a,c**) Fisher information in the original data. (**b,d**) Fisher information in the shuffled data. Note that in the shuffled data differential correlations are destroyed and the code is not robust, hence a slightly suboptimal decoder (e.g. one trained on finite data) is expected to miss much of the information, and to perform worse as the size of the population increases. This is illustrated in (**b**), and can be contrasted to (**a**) where the error is relatively insensitive to population size. The direct estimator does not suffer from this issue: the estimation error for the shuffled information decreases, rather than increases, with population size (**d**).