Figure R1: mDLAG test performance (leave-group-out R^2 evaluated on 75 held-out trials) versus the number of available training trials. Red dashed line: test performance of a group factor analysis (GFA) model fit to all 225 training trials. All models were fit to the example dataset shown in Fig. 5.
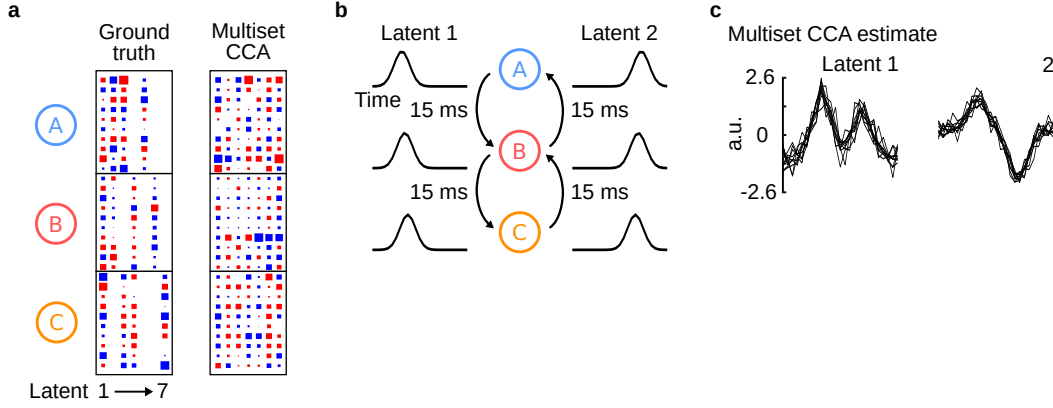

Figure R2: Multiset CCA performance across simulations. (**a**) Left: Ground truth loading matrix in Simulation 1 (copied from Fig. 3a). Right: Multiset canonical correlation analysis (CCA) estimate. Same conventions as in Fig. 2b. (**b**) Schematic of ground truth in Simulation 2 (copied from Fig. 4a). (**c**) Multiset CCA estimates. Same conventions as in Fig. 4c.
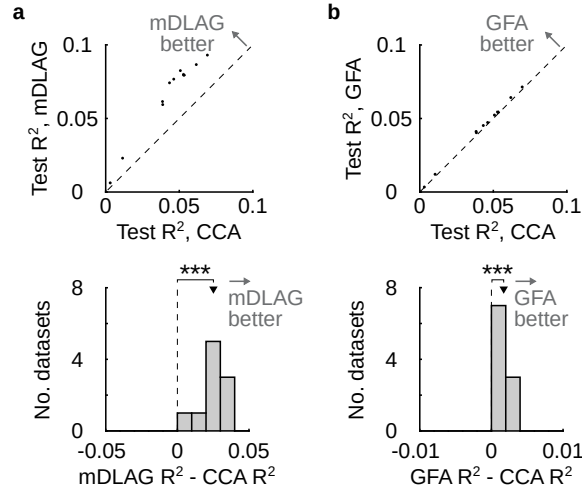

Figure R3: Multiset CCA performance across Neuropixels recordings. (**a**) Top: mDLAG performance versus multiset CCA performance (leave-group-out R^2 evaluated on 75 test trials). Each data point represents one Neuropixels dataset. Bottom: Distribution of differences in performance between mDLAG and multiset CCA. mDLAG significantly outperformed multiset CCA across datasets (***: one-sided paired sign test; $p = 9.8 \times 10^{-4}$). (**b**) Top: Group factor analysis (GFA) performance versus multiset CCA performance. Bottom: GFA significantly outperformed multiset CCA across datasets (***: one-sided paired sign test; $p = 9.8 \times 10^{-4}$). Same conventions as in (a).