Supplementary Figure 3. Example of validation loss on SARDU-Net training performed on brain DRI data when different number of hidden layers are used for both selector and predictor sub-networks. The figure refers to the case when sub-protocols of size $D = 8$ out of $M = 32$ (b,TI) measurements are searched. The same training options have been used for all cases shown (mini-batch size of 100 voxels, learning rate of 10^{-3}, no dropout regularisation), and training refers to leave-one-out fold 1 (i.e. training performed on voxels from subjects 2, 3).