Appendix: Application scenario CIFAR-10

We perform a structured analysis for the CIFAR-10 data set similar to the MNIST data set to investigate whether we can make similar observations. For performance analysis with the PCP, we compute (P1) for time step $1,000$ (focus on range $[0.35, 0.55]$), (P2) for all time steps (focus on range $[0.50, 0.72]$), and (P3) for the first $1,000$ time steps, see Figure 3(a). When brushing on high (green) and low (red) accuracy, we make similar observations as for the MNIST data set. One striking and surprising difference is that the high accuracy values (axes 3 and 4) actually exhibit an anti-correlation to the standard deviation values (axis 5). This can be confirmed over all time steps by looking at the Loss-accuracy Graph in Figure 3(c). Here, the green lines have consistently higher accuracy, but do also form a wider band (indicating larger standard deviation). Another surprising observation can be made when looking at the Performance Heat Map in Figure 3(b): The observed pattern is strikingly similar to the one for the MNIST data set.
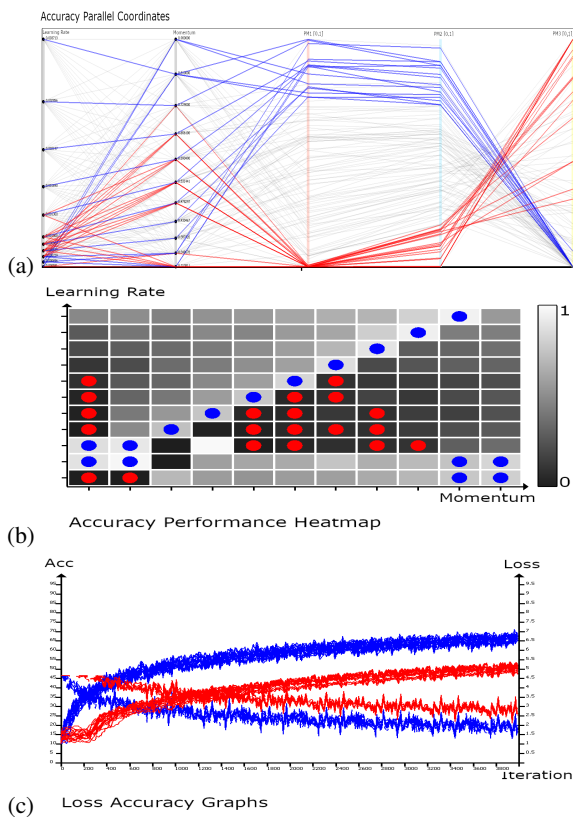

(a) Accuracy Parallel Coordinates

(b) Learning Rate / Momentum

(c) Loss Accuracy Graphs

Figure 3: *Performance analysis of CIFAR-10 data set: Coordinated views of (a) PCP, (b) Performance Heat map, (c) Loss-accuracy Graph.*

For analyzing the influence of the hyper-parameters on topological structures, we generate the Ensemble Similarity Plot for the last time step and for layer Conv3 using a t-SNE embedding, see Figure 4(a). Unlike the MNIST data set, here we immediately see smooth color transition from cyan to red. Figure 4(b) exhibits very low correlation values in the Filter Correlation Heat Map for the red selection, while Figure 4(c) exhibits higher correlations for the

cyan selection. When compared to MNIST, filter correlations are generally lower for CIFAR-10, which may indicate lower redundancy. The Filter Correlation Graph computed for threshold 0.7 in Figure 4(d) shows that the training runs colored red have low correlations after an initial phase, while the ones colored cyan exhibit high correlations until very late in the training process.
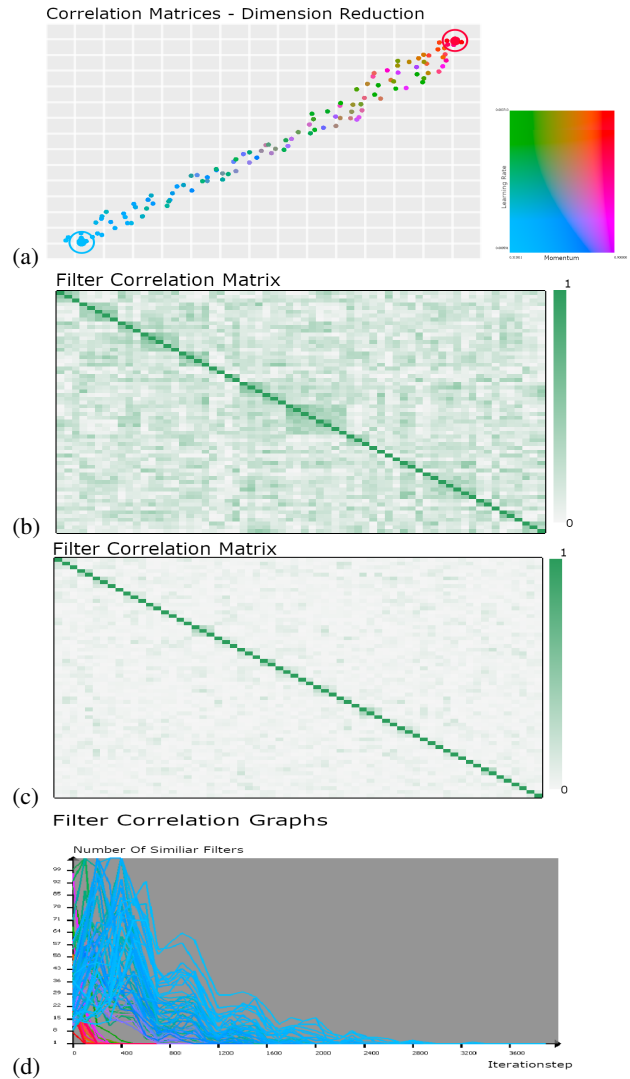

(a) Correlation Matrices - Dimension Reduction

(b) Filter Correlation Matrix

(c) Filter Correlation Matrix

(d) Filter Correlation Graphs / Number Of Similiar Filters

Figure 4: *Analyzing hyper-parameter influence on topological structures for third layer (Conv3) of CIFAR-10 data set: Coordinated views of (a) Ensemble Similarity Plot, (b,c) Filter Correlation Heat Map of selections in (a), and (d) Filter Correlation Graph.*