Figure 1: (Top) Evolution of the training loss components. (Bottom) Nonlinear iterations per residual layer of the trained network.
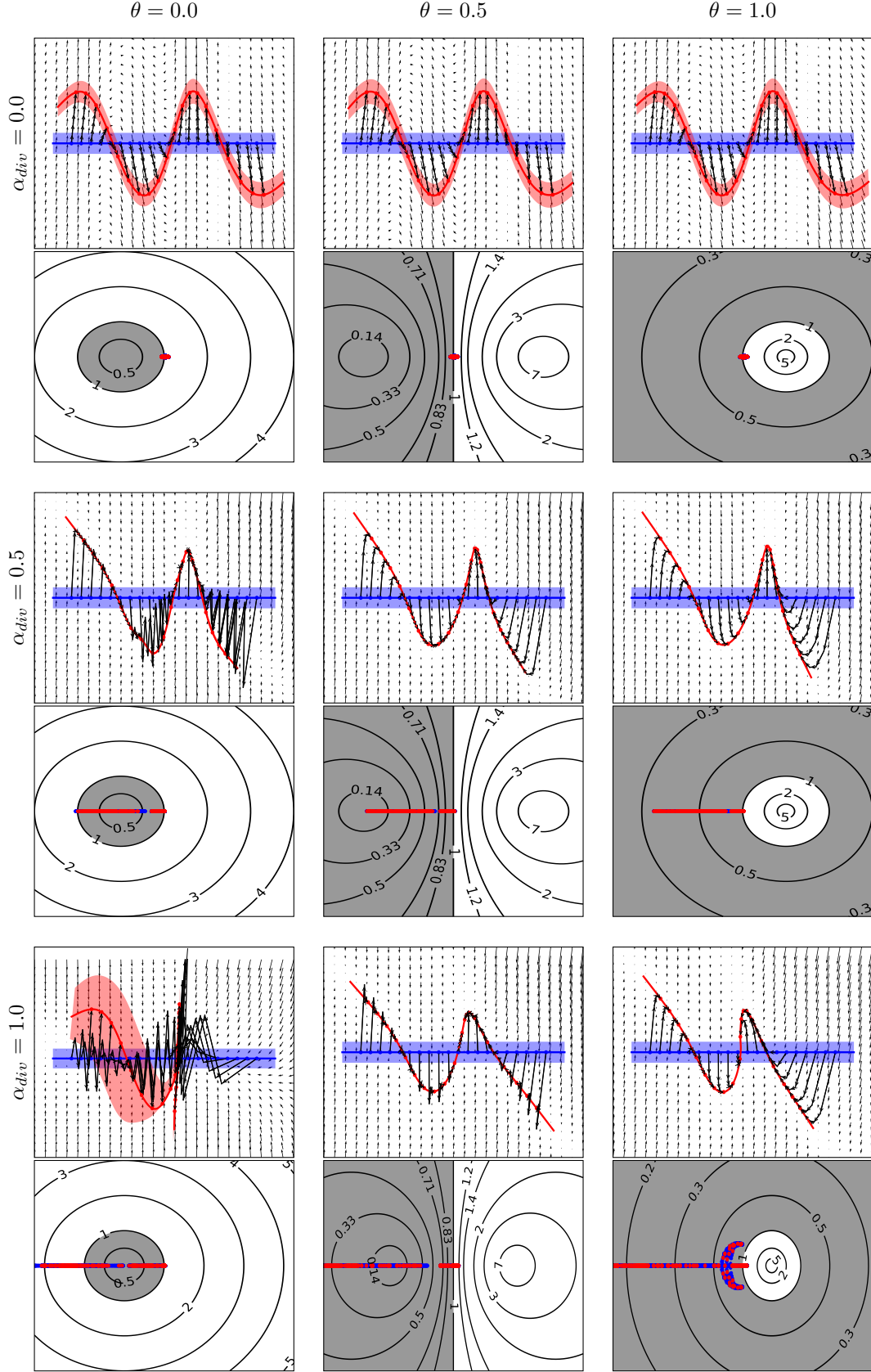
Figure 2: (Top) Learned vector fields $F(\gamma, y)$ in Example 1. Blue line is the initial state, red curve is the final state after $T = 5$ steps, solid black lines are the trajectories of the training data. (Bottom) Eigenvalues of $\frac{\partial F(\gamma, y)}{\partial y}$ evaluated along the learned trajectories at times $t = 0, ..., T$. Red and blue dots are used for the train and test datasets respectively. Stability regions of implicit layers are highlighted with grey color and the contours depict the values of the stability function.