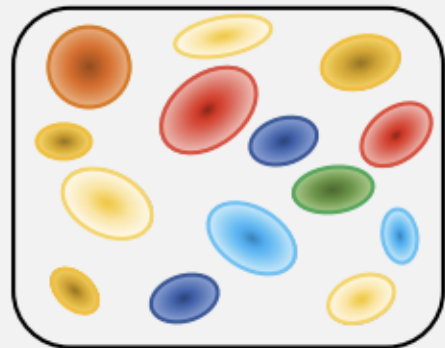
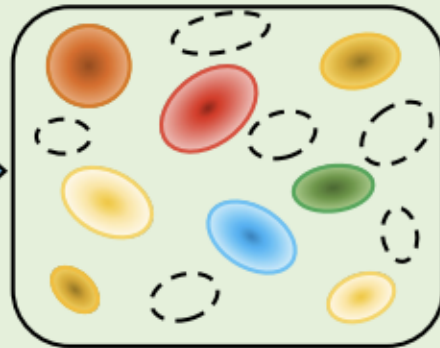
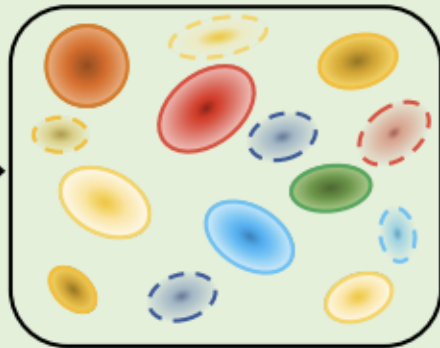
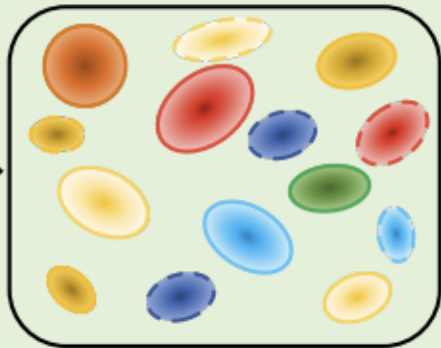


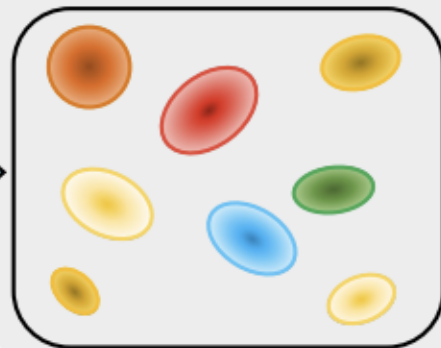
Initialization



"Optimizing-Sparsifying"-Integrated Training



Light Tuning



$$\mathbf{a} \leftarrow \mathbf{a} - \eta \left(\frac{\partial \mathcal{L}(\mathbf{a}, \boldsymbol{\Theta})}{\partial \mathbf{a}} + \delta(\mathbf{a} - \mathbf{z} + \boldsymbol{\lambda}) \right); \boldsymbol{\Theta} \leftarrow \boldsymbol{\Theta} - \eta \frac{\partial \mathcal{L}(\mathbf{a}, \boldsymbol{\Theta})}{\partial \boldsymbol{\Theta}}$$

"Optimizing" Step

$$\mathbf{z} \leftarrow \text{prox}_h(\mathbf{a} + \boldsymbol{\lambda})$$

"Sparsifying" Step

$$\boldsymbol{\lambda} \leftarrow \boldsymbol{\lambda} + \mathbf{a} - \mathbf{z}$$

Multiplier Update