High Order Tensor Formulation for Convolutional Sparse Coding

Adel Bibi and Bernard Ghanem
King Abdullah University of Science and Technology (KAUST)

ICCV17

Abstract

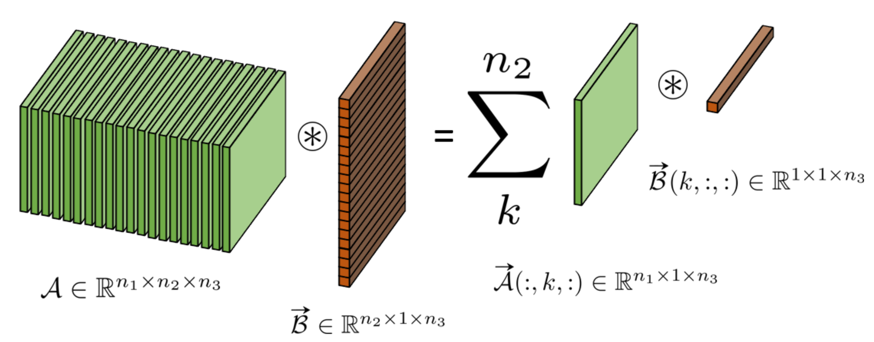
- Convolutional Sparse Coding (CSC) has gained attention for its successful role as a reconstruction and a classification tool.
- Current CSC methods can only reconstruct single feature 2D images independently.
- However, examining correlations among all the data jointly is very important and should be considered for reconstruction applications.
- In this paper, we propose a generic and novel formulation for the CSC problem that can handle an arbitrary order tensor of data.

The CSC problem

$$\min_{\mathbf{x}_k,\mathbf{d}_k \forall k} \frac{1}{2}\sum_n^N \|\mathbf{y}_n - \sum_k^K \mathbf{d}_k * \mathbf{x}_k^n\|_2^2 + \lambda \sum_k^K \|\mathbf{x}_k^n\|_1$$
$$\text{s.t.} \quad \|\mathbf{d}_k\|_2^2 \leq 1 \quad \forall k = 1, ..., K$$

The t-SVD Decomposition

$$circ(\mathcal{A}) = \begin{bmatrix} \mathbf{A}^{(1)} & \mathbf{A}^{(n_3)} & \dots & \mathbf{A}^{(2)} \\ \mathbf{A}^{(2)} & \mathbf{A}^{(1)} & \dots & \mathbf{A}^{(3)} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{A}^{(n_3)} & \mathbf{A}^{(n_3-1)} & \dots & \mathbf{A}^{(1)} \end{bmatrix}$$



$$\mathcal{Z} = \mathcal{A} \circledast \mathcal{B} = fold(circ(\mathcal{A})MatVec(\mathcal{B}))$$

$$\text{such that: } \mathcal{Z}(i,j,:) = \sum_{k=1}^{n_2} \mathcal{A}(i,k,:) \circledast \mathcal{B}(k,j,:) \quad \forall i,j$$

3rd-order Tensor (1-D CSC)

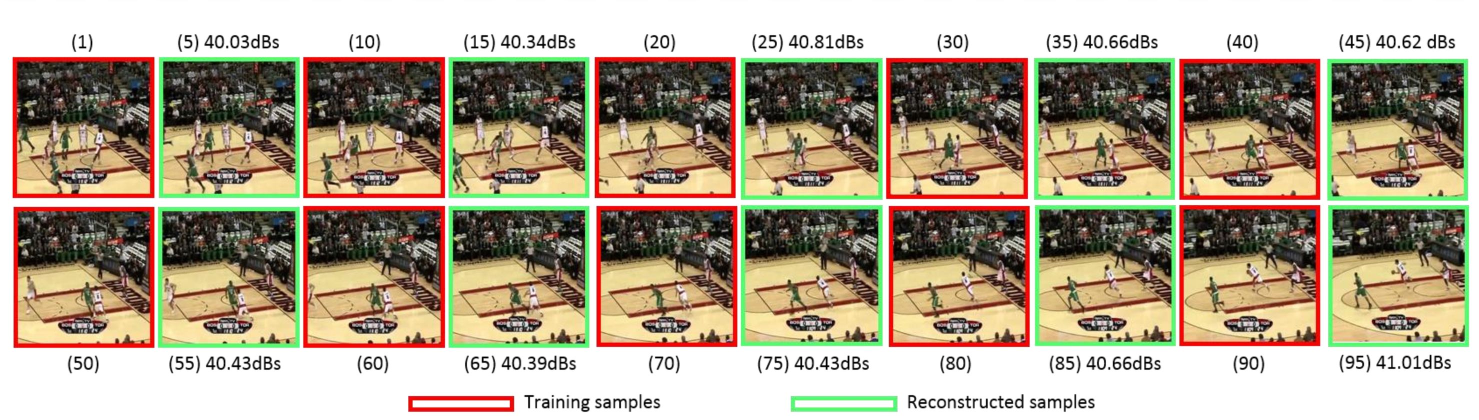
$$\min_{\mathcal{D},\vec{\mathcal{X}}} \quad \frac{1}{2}\sum_n^N \|\vec{\mathcal{Y}}_n - \mathcal{D} \circledast \vec{\mathcal{X}}_n\|_F^2 + \lambda \|\vec{\mathcal{X}}_n\|_{1,1,1}$$
$$\text{s.t.} \quad \|\vec{\mathcal{D}}_k\|_F^2 \leq 1 \ \forall k = 1, \dots, K$$

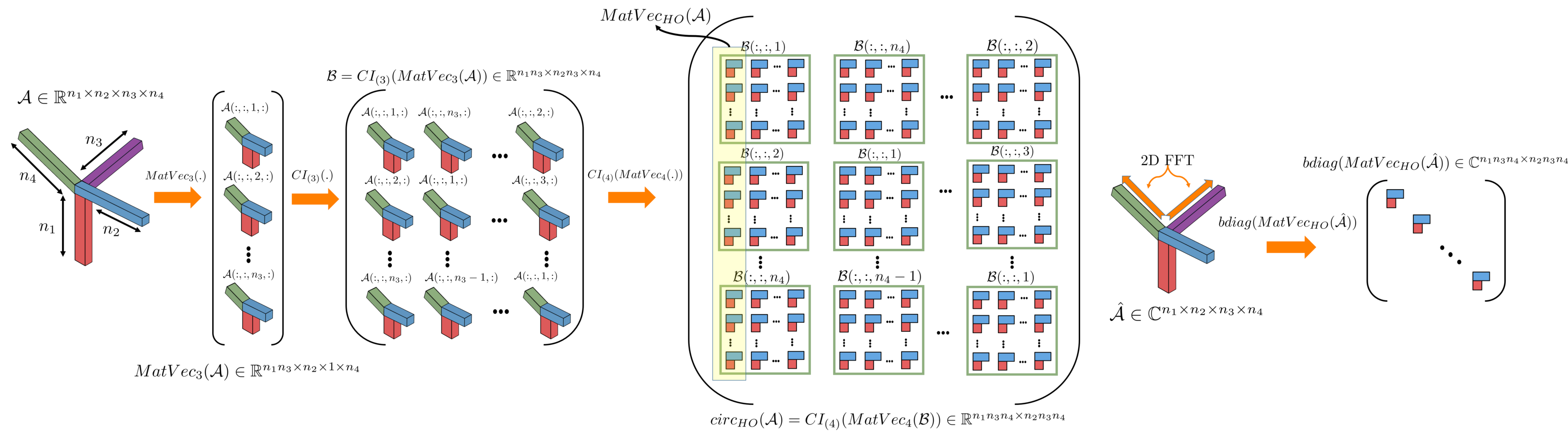
N-order tensor CSC (N-D CSC)

$\vec{\mathcal{Y}}_n \in \mathbb{R}^{n_1 \times 1 \times n_2 \times \dots \times n_d}$.
The K filters $\{\vec{\mathcal{D}}_i\}_{i=1}^K$ $\mathcal{D} \in \mathbb{R}^{n_1 \times K \times n_2 \times \dots \times n_d}$.
The n^{th} sparse code is $\vec{\mathcal{X}}_n \in \mathbb{R}^{K \times 1 \times n_2 \times \dots \times n_d}$.

Definition 3.1 (High Order t-products):

$$\mathcal{D} \circledast_{HO} \vec{\mathcal{X}} = fold_{HO}(circ_{HO}(\mathcal{D})MatVec_{HO}(\vec{\mathcal{X}}))$$





$$\min_{\mathcal{D},\vec{\mathcal{X}}} \quad \frac{1}{2}\sum_n^N \|\vec{\mathcal{Y}}_n - \mathcal{D} \circledast_{HO} \vec{\mathcal{X}}_n\|_F^2 + \lambda \|\vec{\mathcal{X}}_n\|_{\underbrace{1,\dots,1}_{d+1}}$$
$$\text{s.t.} \quad \|\vec{\mathcal{D}}_k\|_F^2 \leq 1 \ \forall k = 1, \dots, K$$

$$circ_{HO}(\mathcal{D}) = (\mathbf{F}_{n_d} \otimes \mathbf{F}_{n_{d-1}} \otimes \dots \otimes \mathbf{F}_{n_2} \otimes \mathbf{I}_{n_1})$$
$$bdiag(MatVec_{HO}(\hat{\mathcal{D}}))(\mathbf{F}_{n_d} \otimes \mathbf{F}_{n_{d-1}} \otimes \dots \otimes \mathbf{F}_{n_2} \otimes \mathbf{I}_K)^{\mathbf{H}}$$

ADMM Solver

Subproblem (1): Sparse Coding.

$$\underset{\vec{\mathcal{X}},\vec{\mathcal{Z}}}{\arg\min} \quad \frac{1}{2}\|\vec{\mathcal{Y}} - \mathcal{D} \circledast_{HO} \vec{\mathcal{X}}\|_F^2 + \lambda\|\vec{\mathcal{Z}}\|_{\underbrace{1,\dots,1}_{d+1}}$$
$$\text{s.t.} \quad \vec{\mathcal{X}} = \vec{\mathcal{Z}}$$

$$\frac{1}{2}\|\vec{\mathcal{Y}} - \mathcal{D} \circledast_{HO} \vec{\mathcal{X}}\|_F^2 = \frac{1}{2}\|MatVec_{HO}(\hat{\vec{\mathcal{Y}}}) - bdiag\Big(MatVec_{HO}(\hat{\mathcal{D}})\Big)MatVec(\hat{\vec{\mathcal{X}}}))\|_F^2$$

$$\hat{\mathcal{X}}^{(i)} \leftarrow \underset{\hat{\mathcal{X}}^{(i)}}{\arg\min} \ \frac{1}{2}\|\hat{\mathcal{D}}^{(i)}\hat{\mathcal{X}}^{(i)} - \hat{\mathcal{Y}}^{(i)}\|_2^2 + \frac{\rho_1}{2}\|\hat{\mathcal{X}}^{(i)} - \hat{\mathcal{Z}}^{(i)}\|_2^2 + <\hat{\mathcal{U}}^{(i)}, \hat{\mathcal{X}}^{(i)}>$$
$$\hat{\mathcal{X}}^{(i)} \leftarrow (\hat{\mathcal{D}}^{(i)\mathbf{H}}\hat{\mathcal{D}}^{(i)} + \rho_1 \mathbf{I}_K)^{-1}\Big(\hat{\mathcal{D}}^{(i)\mathbf{H}}\hat{\mathcal{Y}}^{(i)} + \rho_1\hat{\mathcal{Z}}^{(i)} - \hat{\mathcal{U}}^{(i)}\Big)$$

$$\vec{\mathcal{Z}} \leftarrow \underset{\vec{\mathcal{Z}}}{\arg\min} \frac{\lambda}{\rho_1}\|\vec{\mathcal{Z}}\|_{\underbrace{1,\dots,1}_{d+1}} + \frac{\rho_1}{2}\|\vec{\mathcal{Z}} - \overbrace{\Big(\vec{\mathcal{X}} + \frac{\vec{\mathcal{U}}}{\rho_1}\Big)}^{\vec{\mathcal{A}}}\|_F^2$$

$$\vec{\mathcal{U}} \leftarrow \vec{\mathcal{U}} + \rho_1(\vec{\mathcal{X}} - \vec{\mathcal{Z}})$$

Table 1: Reconstruction error using TCSC trained with $\lambda = 20$ and an average sparsity of 64% across all 29 test samples for different values of K (number of filters in the dictionary).

K	20	40	60	80	100
dBs	70.63	94.45	98.35	104.43	104.70
Avg Sparsity(%)	63.93	64.00	66.66	60.73	65.43

Subproblem (2): Dictionary Learning.

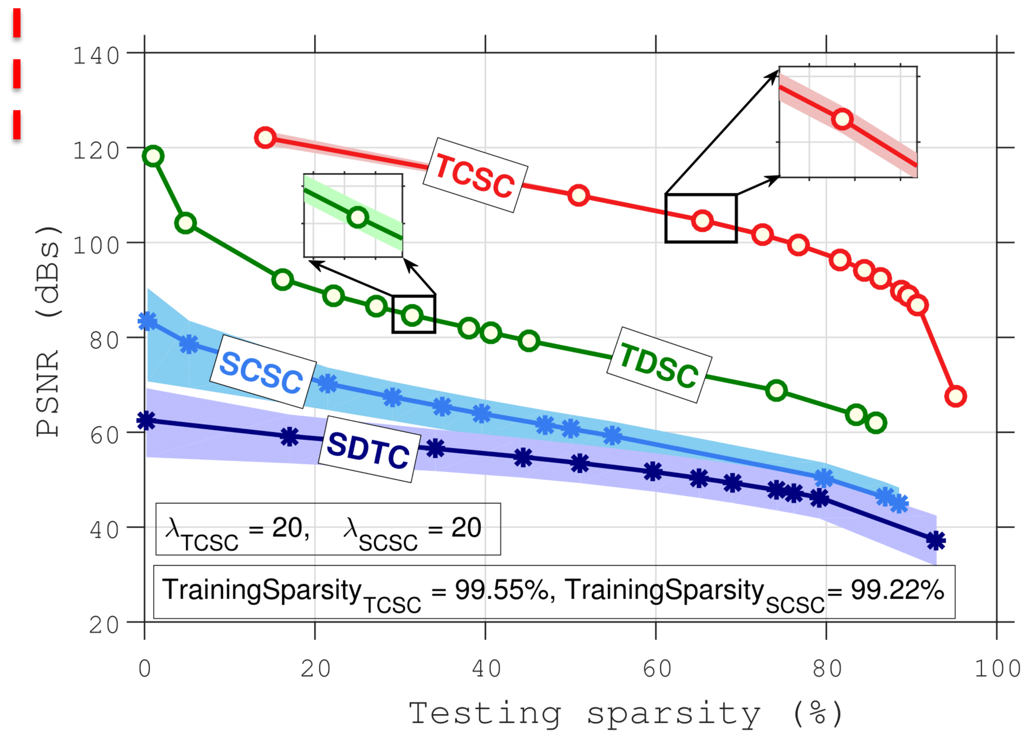
$$\underset{\mathcal{D},\vec{\mathcal{Z}}}{\arg\min} \quad \frac{1}{2}\|\vec{\mathcal{Y}} - \mathcal{D} \circledast_{HO} \vec{\mathcal{X}}\|_F^2$$
$$\text{s.t.} \quad \vec{\mathcal{D}} = \vec{\mathcal{T}}, \quad \|\mathcal{T}_k\|_F^2 \leq 1 \forall k$$

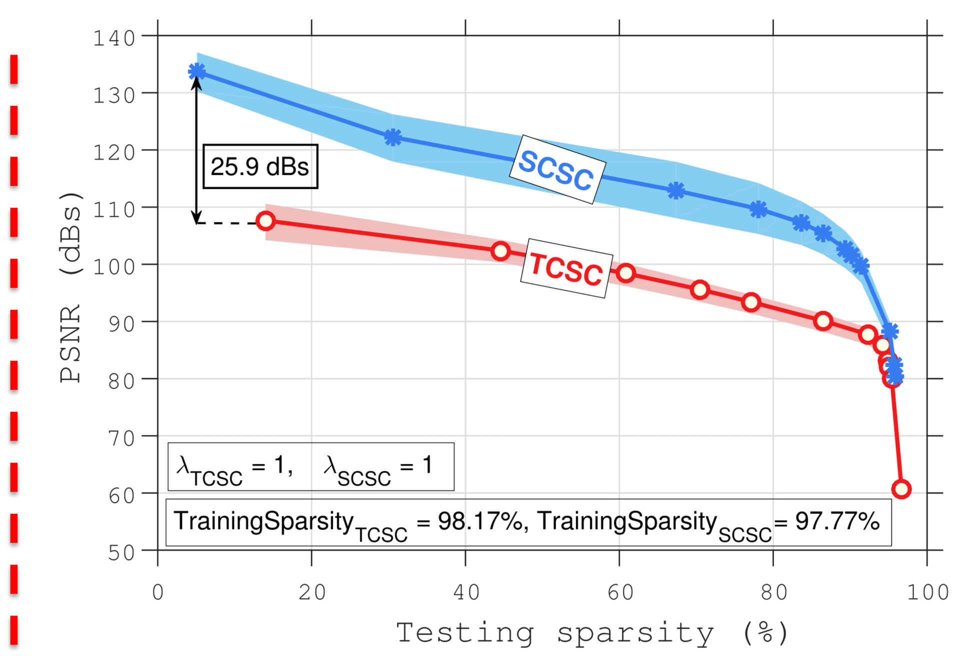
$$\hat{\mathcal{D}}^{(i)} \leftarrow \underset{\hat{\mathcal{D}}^{(i)}}{\arg\min} \frac{1}{2}\|\hat{\mathcal{D}}^{(i)}\hat{\mathcal{X}}^{(i)} - \hat{\mathcal{Y}}^{(i)}\|_2^2 + \frac{\rho}{2}\|\hat{\mathcal{D}}^{(i)} - \hat{\mathcal{T}}^{(i)}\|_2^2 + <\hat{\mathcal{G}}^{(i)}, \hat{\mathcal{D}}^{(i)}>$$
$$\hat{\mathcal{D}}^{(i)} \leftarrow \Big(\hat{\mathcal{Y}}^{(i)}\hat{\mathcal{X}}^{(i)\top} + \rho_2\hat{\mathcal{T}}^{(i)} - \hat{\mathcal{G}}^{(i)}\Big)(\hat{\mathcal{X}}^{(i)}\hat{\mathcal{X}}^{(i)\top} + \rho_2\mathbf{I}_K)^{-1}$$

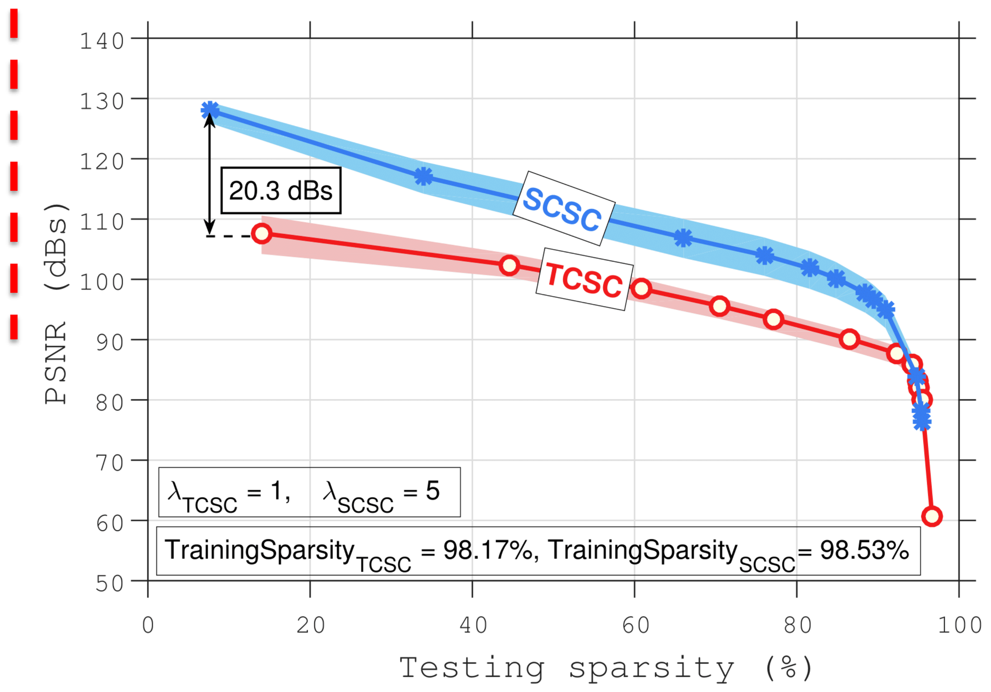
$$\hat{\mathcal{T}} \leftarrow \underset{\hat{\mathcal{T}}}{\arg\min} \quad \frac{\rho_2}{2}\|\hat{\mathcal{T}} - \Big(\hat{\mathcal{D}} + \frac{1}{\rho_2}\hat{\mathcal{G}}\Big)\|_F^2$$
$$\text{s.t.} \quad \|\hat{\vec{\mathcal{T}}}_i \times_p \Psi \times_q \Gamma\|_F^2 \leq 1 \quad \forall i = 1, \dots, K$$

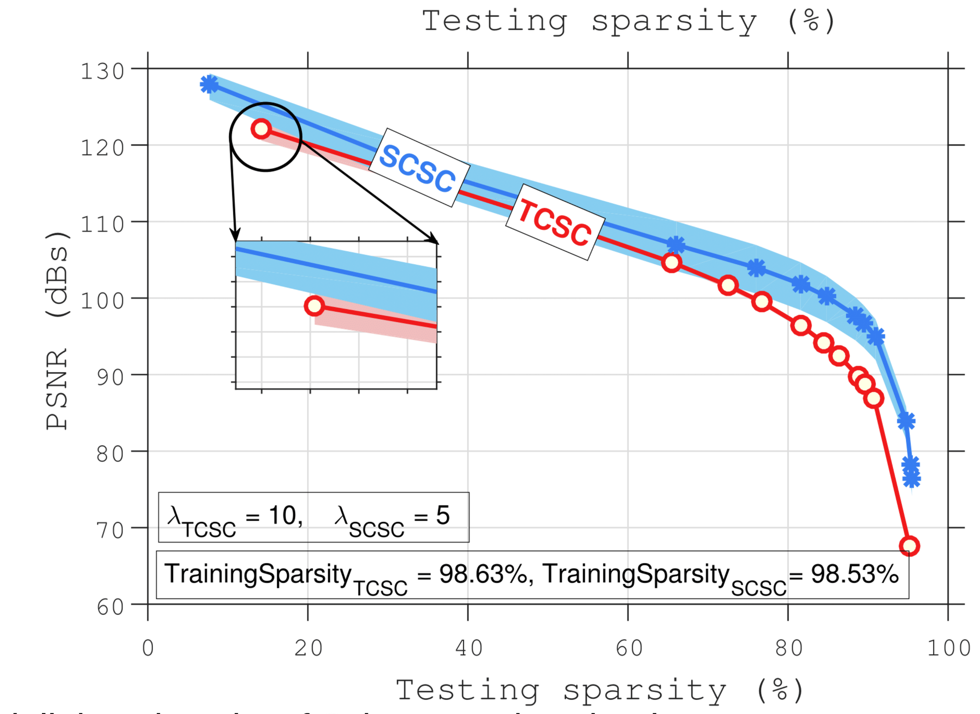
$$\hat{\mathcal{G}} \leftarrow \hat{\mathcal{G}} + \rho_2(\hat{\mathcal{D}} - \hat{\mathcal{T}})$$

Experiments









Acknowledgements: Research reported in this paper was supported by competitive funding from King Abdullah University of Science and Technology