Vision, Modeling, and Visualization (2019)
H.-J. Schulz, M. Teschner, and M. Wimmer (Eds.)

Stochastic Convolutional Sparse Coding Supplementary Material

J. Xiong, P. Richtarik and W. Heidrich

King Abdullah University of Science and Technology (KAUST), Saudi Arabia

1. Solvers for LASSO and QCQP

Algorithm 1 is used to optimize a LASSO problem in the following form:

$$\underset{\mathbf{z}}{\operatorname{argmin}}\ \frac{1}{2}\|\mathbf{x}-\mathbf{D}\mathbf{z}\|_2^2+\lambda\|\mathbf{z}\|_1 \tag{1}$$

Algorithm 1 ADMM framework for solving LASSO

for $s=1$ to S **do**
// z-update step (quadratic programming)
$\mathbf{z}^{s+1} \leftarrow (\mathbf{D}^\top\mathbf{D}+\rho\mathbb{I})^{-1}(\mathbf{D}^\top\mathbf{x}+\rho(\mathbf{y}^s-\mathbf{q}^s))$
// y-update step (soft thresholding)
$\mathbf{y}^{s+1} \leftarrow (\mathbf{z}^{s+1}+\mathbf{q}^s-\frac{\lambda}{\rho})_+ \;-\; (-\mathbf{z}^{s+1}-\mathbf{q}^s-\frac{\lambda}{\rho})_+$
// scaled dual variables update
$\mathbf{q}^{s+1} \leftarrow \mathbf{q}^s+\mathbf{z}^{s+1}-\mathbf{y}^{s+1}$
end for

S is the total number of ADMM iteration. $\mathbf{y}$ is the introduced slack variable, $\mathbf{q}$ is the scaled dual variable, and ρ is the augmented Lagrangian penalty.

Algorithm 2 Projected Block Coordinate Descent for solving QCQP

$\mathbf{r} \leftarrow \mathbf{x}-\mathbf{Z}\mathbf{d}$
while not converge **do**
for $k=1$ to K **do**
$\mathbf{d}_k^* \leftarrow \mathbf{d}_k+\mathbf{Z}_k^\top\mathbf{r}\ /\ L_k$
$\mathbf{d}_k^* \leftarrow \mathbf{d}_k^*\ /\ \max(\|\mathbf{d}_k^*\|,1)$
$\mathbf{r}\ \leftarrow \mathbf{r}+\mathbf{Z}_k(\mathbf{d}_k-\mathbf{d}_k^*)$
end for
$\mathbf{d} \leftarrow \mathbf{d}^*$
end while

Algorithm 2 is used to optimize a QCQP problem in the following form:

$$\begin{aligned}&\underset{\mathbf{d}}{\operatorname{argmin}}\ \frac{1}{2}\|\mathbf{x}-\mathbf{Z}\mathbf{d}\|_2^2\\ &\text{subject to } \|\mathbf{d}_k\|_2^2 \le 1\ \ \forall k \in \{1,\ldots,K\},\end{aligned} \tag{2}$$

L_k is the Lipschitz constant of $\mathbf{Z}_k^\top\mathbf{Z}_k$.

2. Additional Experiments

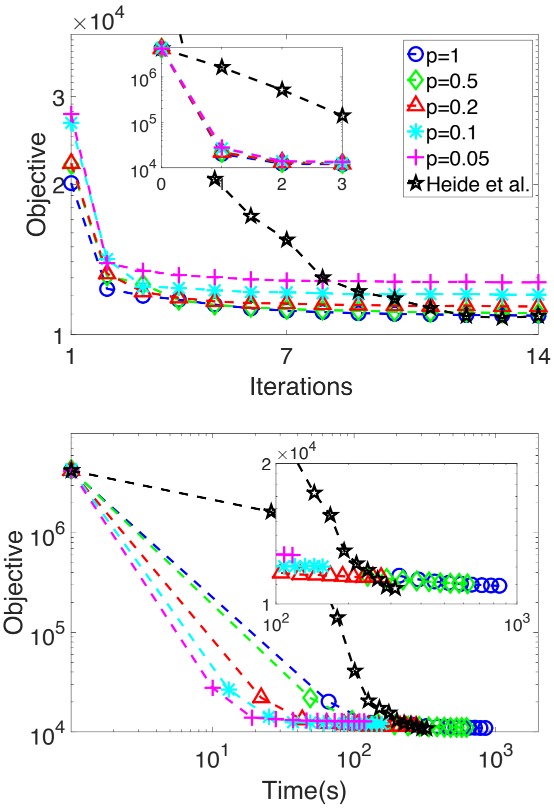
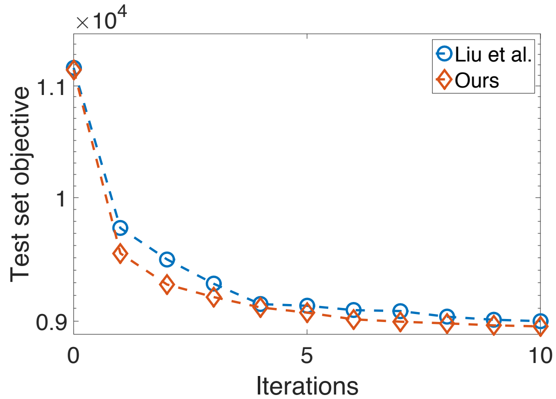


Figure 1: Experimental results obtained on the city dataset. Convergence comparison between the-state-of-art batch method and SBCSC with different subsampling probability.

In order to show the robustness of the proposed algorithms, We also conduct experiments on city dataset. We first compare SBCSC with the state-of-the-art batch-mode algorithm, and the results are shown in Fig. 1. Similar to the results shown in Fig. 1 in the main manuscript, SBCSC, with $p=0.1$, outperforms the compared batch-mode algorithm in terms of runtime performance.

We then compare SOCSC with the state-of-the-art online algo-



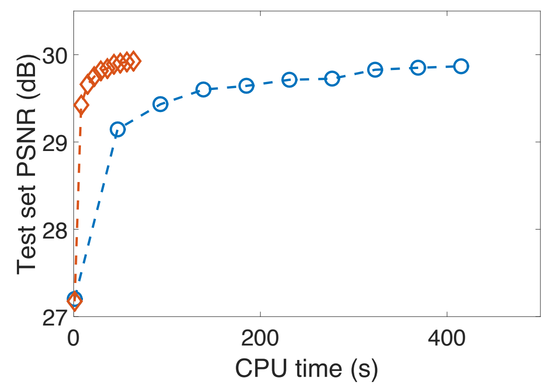


Figure 2: Experimental results obtained on city dataset. Top: Convergence of the test set objectives for our method (SOCSC) and the state-of-the-art online approach. Bottom: Testing PSNR with respect to execution time.

rithm on city dataset, the results of which are shown in Fig. 2. Similar to the results presented in Fig. 3 of the main manuscript, our method obtains comparable outcomes, meanwhile, achieves roughly 6× speedup.

3. Over-complete Dictionary and Large Datasets

We verify the importance of a large dataset when learning the over-complete dictionary. in Fig. 3, we show a visual and quantitative comparisons between the over-complete dictionaries respectively learned by batch-mode algorithm on small dataset (the fruit dataset) and online-mode algorithm on large dataset (1000 images). Most of the filters learned from small dataset have poor structures and reveal very limited representative ability. The numerical results also demonstrate that it even shows a degraded reconstruction performance compared to the under-complete dictionary. Owing to plentiful training samples, our over-complete dictionary not only shows visually decent structures and more representative image features, but also leads to a significant improvement on image reconstruction. Based on all experimental results, it implies that the number of filters and number of training samples are both essential in the CSC model, therefore, the proposed algorithm has prominent advantages over the existing approaches.

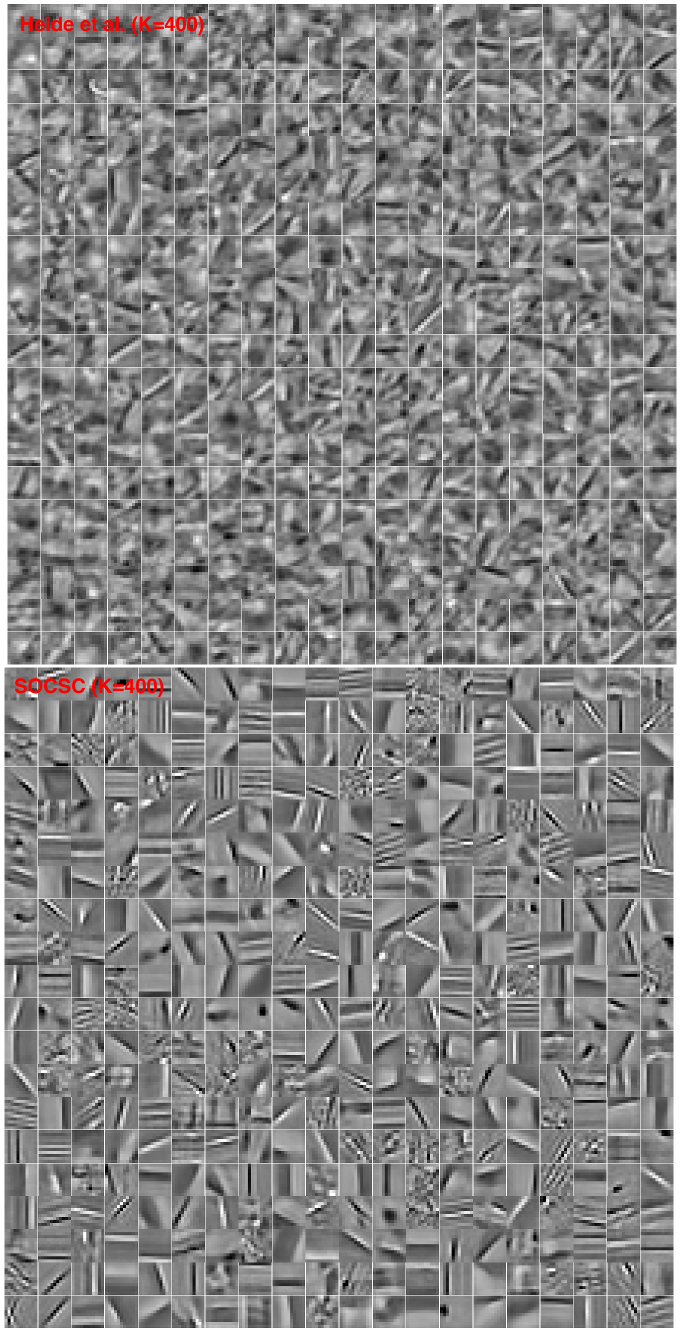


Image	1	2	3	4	5	6	7	8	9	10
PSNR [10]	29.60	28.11	29.47	28.98	28.79	29.21	28.03	29.31	27.42	30.52
PSNR ours	**30.24**	**28.34**	**29.95**	**30.30**	**29.43**	**29.96**	**28.24**	**30.57**	**27.72**	**31.67**

Figure 3: Visual and numerical comparisons between the learned over-complete dictionaries by batch-mode algorithm and by online-mode algorithm. Top: Over-complete dictionary learned by batch CSC model on small dataset, and proposed online CSC model (SOCSC) on large dataset. Bottom: Respective reconstruction quality for these two over-complete dictionaries applied for image inpainting.