Supplementary file of the paper titled "Image Denoising by a Local Clustering Framework" by Partha Sarathi Mukherjee and Peihua Qiu

This file gives proofs of the two theorems in the paper and provides some extra numerical results.

Proof of Theorem 1

By the strong law of large numbers (SLLN), we have $\overline{\xi} \to f(x, y)$ a.s. Suppose, s induces the clustering of sizes n_1 and n_2, i.e. $n_1 + n_2 = nh_n$. For any s, $n_1 \to \infty$ and $n_2 \to \infty$ as $n \to \infty$. Then, in a continuous region of the image, i.e. if $(x, y) \in \Omega_{\bar{J}, \epsilon}$, $(\overline{\xi}_1 - \overline{\xi}) = \frac{-\sigma\phi(t/\sigma)}{\Phi(t/\sigma)} + O(h_n)$ a.s. and $(\overline{\xi}_2 - \overline{\xi}) = \frac{\sigma\phi(t/\sigma)}{1-\Phi(t/\sigma)} + O(h_n)$ a.s. where $t = \sigma\Phi^{-1}(n_1/nh_n)$. Note that the above result is based on the assumption that f has continuous first-order derivatives over $(0, 1) \times (0, 1)$ except the JLCs. In that case, $f(x_i, y_j) = f(x, y) + O(h_n)$ for any $(x_i, y_j) \in O(x, y; h_n)$. Therefore,

$$\frac{1}{n^2 h_n^2} \left(|O_1(x, y; h_n, s)|(\overline{\xi}_1 - \overline{\xi})^2 + |O_2(x, y; h_n, s)|(\overline{\xi}_2 - \overline{\xi})^2 \right) \to \frac{\sigma^2\phi^2(t/\sigma)}{\Phi(t/\sigma)(1 - \Phi(t/\sigma))} \text{ a.s.}$$

Now,

$$\frac{1}{n^2 h_n^2} \left(\sum_{(x_i,y_j)\in O_1(x,y;h_n,s)} (\xi_{ij} - \overline{\xi}_1)^2 + \sum_{(x_i,y_j)\in O_2(x,y;h_n,s)} (\xi_{ij} - \overline{\xi}_2)^2 \right)$$

$$= \frac{1}{n^2 h_n^2} \sum_{(x_i,y_j)\in O(x,y;h_n)} (\xi_{ij} - \overline{\xi})^2 -$$

$$\frac{1}{n^2 h_n^2} \left(|O_1(x, y; h_n, s)|(\overline{\xi}_1 - \overline{\xi})^2 + |O_2(x, y; h_n, s)|(\overline{\xi}_2 - \overline{\xi})^2 \right)$$

$$\to \sigma^2 \left(1 - \frac{\phi^2(t/\sigma)}{\Phi(t/\sigma)(1 - \Phi(t/\sigma))} \right) \text{ a.s.}$$

The first equality is due to the fact that the total sum of squares is equal to the sum of the sum of squares between the groups and the sum of squares within the groups. Since, $\frac{\phi^2(t/\sigma)}{\Phi(t/\sigma)(1-\Phi(t/\sigma))} \neq 0$ for any finite value of t,

$$T(x, y; h_n, s) \to \frac{\left(\frac{\phi^2(t/\sigma)}{\Phi(t/\sigma)(1-\Phi(t/\sigma))} \right)}{1 - \left(\frac{\phi^2(t/\sigma)}{\Phi(t/\sigma)(1-\Phi(t/\sigma))} \right)} \text{ a.s.}$$

It is easy to check that $T(x, y; h_n, s)$ is maximized when $t = 0$, i.e. when $n_1 = n_2 = nh_n/2$. Therefore, $T(x, y; h_n, S_0) \leq u_n$ a.s.

Proceeding similarly, if $(x, y) \in J_{h_n}$, $O(x, y; h_n)$ intersects two Λ_ls, then $T(x, y; h_n, s) \to J^2/4\sigma^2$ a.s. If $J > 4\kappa\sigma^2$, and u_n is as defined in Section 3, then $T(x, y; h_n, S_0) > u_n$ a.s.

Proof of Theorem 2

If (x, y) is in a continuous region of an image, i.e., if $(x, y) \in \Omega_{\bar{J}, \epsilon}$, then for any $(x_i, y_j) \in O(x, y; h_n)$, $f(x_i, y_j) = f(x, y) + O(h_n)$ a.s. Therefore, $\left(\frac{\|\widetilde{O}(x_i, y_j) - \widetilde{O}(x, y)\|_2^2}{|\widetilde{O}(x_i, y_j)|} \right) = O(h_n^2)$ a.s., assuming that $\widetilde{h}_n \to \infty$ as $n \to \infty$, and $\frac{\widetilde{h}_n}{h_n} = o(1)$. In that case, if $B_n = O(h_n^{1/2})$, $\widetilde{W}_{ij} = 1 + O(h_n^{3/2})$ a.s. Also, from Theorem 1, $T(x, y; h_n, S_0) \leq u_n$ a.s. Therefore,

$$\widehat{f}(x, y) = \frac{\sum\limits_{(x_i, y_j) \in O(x, y; h_n, S_0)} \widetilde{W}_{ij} \xi_{ij}}{\sum\limits_{(x_i, y_j) \in O(x, y; h_n, S_0)} \widetilde{W}_{ij}} = f(x, y) + O(h_n^{1/2}) \quad \text{a.s.}$$

If (x, y) is a non-singular point close to a JLC of an image, i.e., when $O(x, y; h_n)$ intersects with two Λ_ls, for sufficiently small h_n, and if the minimum jump size J of the JLC within $O(x, y; h_n)$ is larger than $4\kappa\sigma^2$, then by Theorem 1, $T(x, y; h_n, S_0) > u_n$ a.s. In this case, for any $(x_i, y_j) \in O_1(x, y; h_n)$, $f(x_i, y_j) = f(x, y) + O(h_n)$ a.s., and proceeding as before, we see that $\widetilde{W}_{ij} = 1 + O(h_n^{3/2})$ a.s. So,

$$\widehat{f}(x, y) = \frac{\sum\limits_{(x_i, y_j) \in O_1(x, y; h_n, S_0)} \widetilde{W}_{ij} \xi_{ij}}{\sum\limits_{(x_i, y_j) \in O_1(x, y; h_n, S_0)} \widetilde{W}_{ij}} = f(x, y) + O(h_n^{1/2}) \quad \text{a.s.}$$

If (x, y) is a non-singular point close to a JLC of an image, i.e., when $O(x, y; h_n)$ intersects with two Λ_ls, for sufficiently small h_n, and if the minimum jump size J of the JLC within $O(x, y; h_n)$ is smaller than or equal to $4\kappa\sigma^2$, then for any (x_i, y_j) which is inside the same Λ_l (say, Λ_{l_1}) that contains (x, y), $f(x_i, y_j) = f(x, y) + O(h_n)$ a.s., and for any (x_i, y_j) which is inside the other Λ_l (say, Λ_{l_2}), there exists $\delta^* > 0$ such that $f(x_i, y_j) = f(x, y) + \delta^* + O(h_n)$. The last results follows from the definition of the singular points and the assumption that f has continuous first-order derivatives over $(0, 1) \times (0, 1)$ except the JLCs. Then, proceeding similarly, for $(x_i, y_j) \in \Lambda_{l_1}$, $\widetilde{W}_{ij} = 1 + O(h_n^{3/2})$ a.s. and for $(x_i, y_j) \in \Lambda_{l_2}$, $\widetilde{W}_{ij} = o(1)$ a.s. Therefore,

$$\widehat{f}(x, y) = \frac{\sum\limits_{(x_i, y_j) \in O_1(x, y; h_n, S_0)} \widetilde{W}_{ij} \xi_{ij}}{\sum\limits_{(x_i, y_j) \in O_1(x, y; h_n, S_0)} \widetilde{W}_{ij}} = f(x, y) + O(h_n^{1/2}) \quad \text{a.s.}$$

In this case, if we replace $O_1(x, y; h_n)$ by $O(x, y; h_n)$, i.e., even if $T(x, y; h_n, S_0) \leq u_n$, $\widehat{f}(x, y) = f(x, y) + O(h_n^{1/2})$ a.s. From this proof, we see that even if we do not consider the clustering step, point-wise asymptotic convergence of $\widehat{f}(x, y)$ at a non-singular point (x, y) still holds.

2

Comparisons of the denoising methods when the noise is uniformly distributed

In practice, the noise distribution may not be normal. Here, we consider cases when the noise distribution is $Uniform[-N_R, N_R]$, where N_R is a constant. Let us still use the three test images discussed in Section 4 of the paper, and choose $N_R = 0.15$ in the example with the artificial image and $N_R = 25$ in the examples with the fingerprint and MRI images. The image denoising methods are executed in the same way as in the Gaussian noise cases. The corresponding results are presented in Table S.1 and Figure S.1. From Table S.1, we see that NEW is better than its three competitors in most cases with the artificial and fingerprint images. In cases with the MRI brain image, NEW outperforms TV, performs similarly to ONLM, and performs better than ASSK in terms of MISE. Similar to cases with the Gaussian noise, ASSK seems to preserve edges better with the price of a weaker noise removal ability. This is confirmed by Figure 1, and is consistent with the results in Table 3 and Figure 6 in the main article.

Table S.1: In each entry, the first line presents the estimated MISE value based on 100 simulations and the corresponding standard error (in parenthesis), the second line presents the value of EP and its standard error (in parenthesis), and the third line presents the searched procedure parameter values. This table is about the cases with uniformly distributed noise. The best method in each case is indicated by *italicized numbers*.

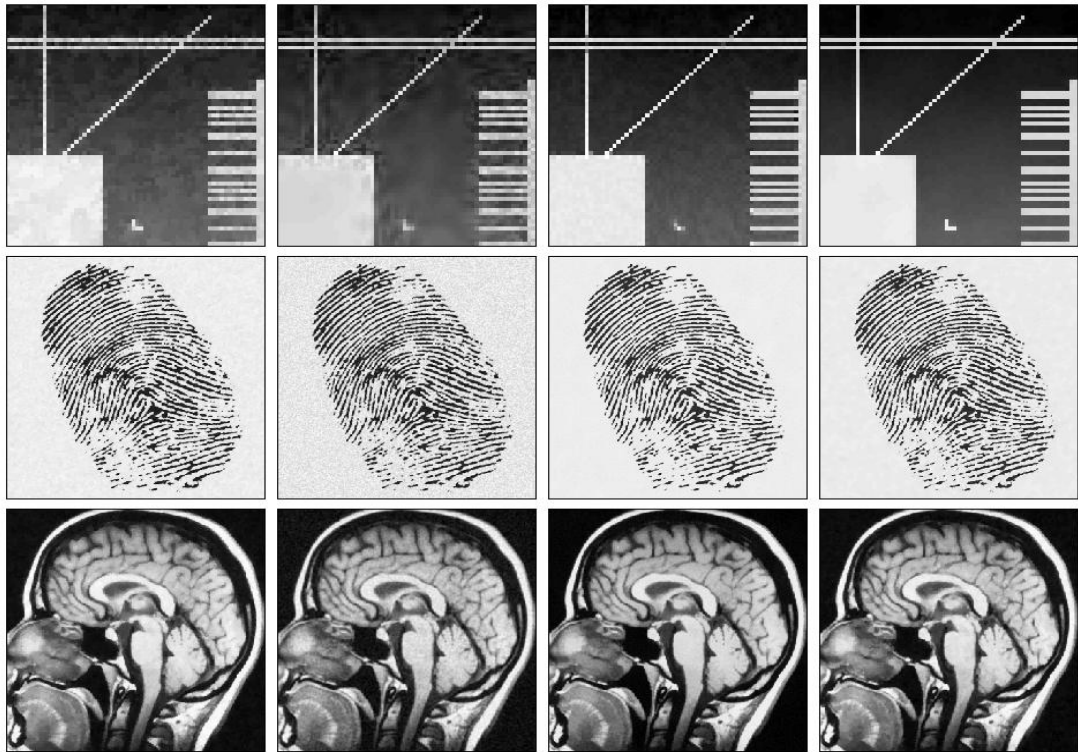| Image	TV	ASSK	ONLM	NEW
Artificial	0.0028 (0.0001)	0.0026 (0.0001)	0.0014 (0.0001)	*0.0002 (0.0000)*
	0.0401 (0.0067)	0.1191 (0.0086)	0.0366 (0.0060)	*0.0040 (0.0031)*
	25.0	0.18, 5	10, 1, 0.10	*0.0938, 2.0, 3.0*
Fingerprint	75.1 (0.4)	157.3 (0.5)	81.4 (0.6)	*75.3 (0.4)*
	0.0393 (0.0007)	0.1048 (0.0009)	0.0145 (0.0009)	*0.0145 (0.0008)*
	0.12	0.0040, 4	10, 1, 20	*0.0043, 2.0, 0.8*
Brain	61.6 (0.4)	61.8 (0.4)	*51.9 (0.3)*	*52.1 (0.3)*
	0.1538 (0.0012)	0.0175 (0.0014)	*0.0968 (0.0013)*	*0.0938 (0.0014)*
	0.10	0.0050, 5	*10, 1, 15*	*0.0117, 10.0, 2.5*

Figure S.1: The columns from the left to the right present the denoised images by TV, ASSK, ONLM and NEW, respectively, in cases with uniformly distributed noise.