Supplementary file of the paper titled "On Nonparametric Image Registration"

Peihua Qiu and Chen Xing
School of Statistics
University of Minnesota

This supplementary file has two parts. The first part provides proofs of Theorems 2 and 3 presented in Section 4 of the paper, and the second part provides some numerical results related to the IBIR procedures considered in Section 5.

Lemma 1 Assume that $h_n^* = o(1)$, $\rho_n = o(1)$, $1/(nh_n^*\rho_n) = o(1)$, and K^* is a Lipschitz-1 continuous density kernel function in a unit circle. Then, we have

$$\left\| \frac{1}{n^2 h_n^{*2} \rho_n} \sum_{(x_i,y_j)\in N_{\widehat{\mathbf{G}}}^{(1)}(x,y)} K^*\left(\frac{x_i - x}{h_n^*}, \frac{y_j - y}{h_n^*}\right) - C_n \right\|_{\Omega(R,T,n)} = O\left(\frac{1}{\rho_n n h_n^*}\right), \tag{S.1}$$

where $C_n = \frac{1}{\rho_n} \int\int_{N_{\rho_n}} K^*(u,v)\ dudv = O(1)$, $N_{\rho_n} = \{(u,v) : \sqrt{u^2+v^2} \le 1, |u| \le \rho_n, v \ge 0\}$, $N_{\widehat{\mathbf{G}}}^{(1)}(x,y)$ is defined in the paragraph immediately before (4) of the paper.

Proof. For any $(x,y) \in \Omega(R,T,n)$, we have

$$\begin{aligned}
&\left| \frac{1}{n^2 h_n^{*2} \rho_n} \sum_{(x_i,y_j)\in N_{\widehat{\mathbf{G}}}^{(1)}(x,y)} K^*\left(\frac{x_i - x}{h_n^*}, \frac{y_j - y}{h_n^*}\right) - C_n \right| \\
= &\left| \frac{1}{n^2 h_n^{*2} \rho_n} \sum_{(x_i,y_j)\in N_{\widehat{\mathbf{G}}}^{(1)}(x,y)} K^*\left(\frac{x_i - x}{h_n^*}, \frac{y_j - y}{h_n^*}\right) - \frac{1}{h_n^{*2}\rho_n} \int\int_{N_{\widehat{\mathbf{G}}}^{(1)}(x,y)} K^*\left(\frac{u - x}{h_n^*}, \frac{v - y}{h_n^*}\right) dudv \right| \\
= &\left| \frac{1}{n^2 h_n^{*2} \rho_n} \sum_{(x_i,y_j)\in N_{\widehat{\mathbf{G}}}^{(1)}(x,y)} K^*\left(\frac{x_i - x}{h_n^*}, \frac{y_j - y}{h_n^*}\right) \right. \\
&\left. - \frac{1}{h_n^{*2}\rho_n} \left(\sum_{(x_i,y_j)\in N_{\widehat{\mathbf{G}}}^{(1)}(x,y)} \int\int_{N_{\widehat{\mathbf{G}}}^{(1)}(x,y)\bigcap \Delta_{ij}} K^*\left(\frac{u - x}{h_n^*}, \frac{v - y}{h_n^*}\right) dudv + O\left(\frac{h_n^*}{n}\right) \right) \right| \\
= &\left| \frac{1}{h_n^{*2} \rho_n} \left(\sum_{(x_i,y_j)\in N_{\widehat{\mathbf{G}}}^{(1)}(x,y)} K^*\left(\frac{x_i - x}{h_n^*}, \frac{y_j - y}{h_n^*}\right) \int\int_{N_{\widehat{\mathbf{G}}}^{(1)}(x,y)\bigcap \Delta_{ij}} dudv + O\left(\frac{h_n^*}{n}\right) \right) \right.
\end{aligned}$$

$$
\begin{aligned}
&-\frac{1}{h_n^{*2}\rho_n}\left(\sum_{(x_i,y_j)\in N_{\widehat{\mathbf{G}}}^{(1)}(x,y)}\int\int_{N_{\widehat{\mathbf{G}}}^{(1)}(x,y)\bigcap\Delta_{ij}} K^*\left(\frac{u-x}{h_n^*},\frac{v-y}{h_n^*}\right)dudv+O\left(\frac{h_n^*}{n}\right)\right)\Bigg| \\
= &\left|\frac{1}{h_n^{*2}\rho_n}\sum_{(x_i,y_j)\in N_{\widehat{\mathbf{G}}}^{(1)}(x,y)}\int\int_{N_{\widehat{\mathbf{G}}}^{(1)}(x,y)\bigcap\Delta_{ij}}\left[K^*\left(\frac{x_i-x}{h_n^*},\frac{y_j-y}{h_n^*}\right)-K^*\left(\frac{u-x}{h_n^*},\frac{v-y}{h_n^*}\right)\right]dudv\right. \\
&\left.+O\left(\frac{1}{nh_n^*\rho_n}\right)\right| \\
\leq &\frac{1}{h_n^{*2}\rho_n}\sum_{(x_i,y_j)\in N_{\widehat{\mathbf{G}}}^{(1)}(x,y)}\int\int_{N_{\widehat{\mathbf{G}}}^{(1)}(x,y)\bigcap\Delta_{ij}}\left|K^*\left(\frac{x_i-x}{h_n^*},\frac{y_j-y}{h_n^*}\right)-K^*\left(\frac{u-x}{h_n^*},\frac{v-y}{h_n^*}\right)\right|dudv \\
&+O\left(\frac{1}{nh_n^*\rho_n}\right) \\
\leq &\frac{1}{h_n^{*2}\rho_n}\frac{\sqrt{2}C_k}{nh_n^*}\sum_{(x_i,y_j)\in N_{\widehat{\mathbf{G}}}^{(1)}(x,y)}\int\int_{N_{\widehat{\mathbf{G}}}^{(1)}(x,y)\bigcap\Delta_{ij}}dudv+O\left(\frac{1}{nh_n^*\rho_n}\right) \\
\leq &\frac{1}{h_n^{*2}\rho_n}\frac{\sqrt{2}C_k}{nh_n^*}\int\int_{N_{\widehat{\mathbf{G}}}^{(1)}(x,y)}dudv+O\left(\frac{1}{nh_n^*\rho_n}\right) \\
= &\frac{1}{h_n^{*2}\rho_n}\frac{\sqrt{2}C_k}{nh_n^*}O\left(h_n^{*2}\rho_n\right)+O\left(\frac{1}{nh_n^*\rho_n}\right) \\
= &O\left(\frac{1}{nh_n^*\rho_n}\right),
\end{aligned}
$$

where $\Delta_{ij}=[x_{i-1},x_i]\times[y_{j-1},y_j]$, for $i,j=1,2,\ldots,n$, $x_0=y_0=0$, and $C_k>0$ is the Lipschitz constant satisfying $|K^*(x',y')-K^*(x'',y'')|\leq C_k\sqrt{(x'-x'')^2+(y'-y'')^2}$ for any points (x',y') and (x'',y'') in the unit circle.

Lemma 2 Besides the conditions in Lemma 1, if we further assume that $E|\varepsilon_R(x_1,y_1)|^3<\infty$ and $\log^2 n/(nh_n^{*3}\rho_n^{3/2})=O(1)$, then we have

$$
\left\|\frac{1}{n^2h_n^{*2}\rho_n}\sum_{(x_i,y_j)\in N_{\widehat{\mathbf{G}}}^{(1)}(x,y)}\varepsilon_R(x_i,y_j)K^*\left(\frac{x-x_i}{h_n^*},\frac{y-y_j}{h_n^*}\right)\right\|_{\Omega(R,T,n)}=O\left(\frac{\log n}{nh_n^*\sqrt{\rho_n}}\right),\quad a.s. \quad \text{(S.2)}
$$

Proof. (S.2) is a direct conclusion of Proposition 2 in Qiu (2009). Its proof is therefore skipped.

Proposition 1 Besides the conditions in Lemma 2, if we further assume that $E|\varepsilon_M(x_1,y_1)|^3<$

∞, then we have

$$\left\| \frac{\sum_{(x_i,y_j)\in N^{(1)}_{\widehat{\mathbf{G}}}(x,y)} \varepsilon_R(x_i,y_j) K^*\left(\frac{x_i-x}{h_n^*}, \frac{y_j-y}{h_n^*}\right)}{\sum_{(x_i,y_j)\in N^{(1)}_{\widehat{\mathbf{G}}}(x,y)} K^*\left(\frac{x_i-x}{h_n^*}, \frac{y_j-y}{h_n^*}\right)} \right\|_{\Omega(R,T,n)} = O\left(\frac{\log n}{n h_n^* \sqrt{\rho_n}}\right), \quad a.s. \tag{S.3}$$

$$\left\| \frac{\sum_{(x_i,y_j)\in N^{(2)}_{\widehat{\mathbf{G}}}(x,y)} \varepsilon_M(x_i,y_j) K^*\left(\frac{x_i-x}{h_n^*}, \frac{y_j-y}{h_n^*}\right)}{\sum_{(x_i,y_j)\in N^{(2)}_{\widehat{\mathbf{G}}}(x,y)} K^*\left(\frac{x_i-x}{h_n^*}, \frac{y_j-y}{h_n^*}\right)} \right\|_{\Omega(R,T,n)} = O\left(\frac{\log n}{n h_n^* \sqrt{\rho_n}}\right), \quad a.s. \tag{S.4}$$

Proof. (S.3) can be obtained after combining (S.1) and (S.2). (S.4) can be proved in the same way, except that $\{\varepsilon_R(x_i,y_j)\}$ are replaced by $\{\varepsilon_M(x_i,y_j)\}$.

Proof of Theorem 2

First, for any $(x,y)\in\Omega_2(R,T,n)$, we have

$$\begin{aligned}
& a^{(1)}_{\widehat{\mathbf{G}}}(x,y) \\
= & \frac{\sum_{(x_i,y_j)\in N^{(1)}_{\widehat{\mathbf{G}}}(x,y)} Z_R(x_i,y_j) K^*\left(\frac{x_i-x}{h_n^*}, \frac{y_j-y}{h_n^*}\right)}{\sum_{(x_i,y_j)\in N^{(1)}_{\widehat{\mathbf{G}}}(x,y)} K^*\left(\frac{x_i-x}{h_n^*}, \frac{y_j-y}{h_n^*}\right)} \\
= & \frac{\sum_{(x_i,y_j)\in N^{(1)}_{\widehat{\mathbf{G}}}(x,y)} \left(R(x_i,y_j)+\varepsilon_R(x_i,y_j)\right) K^*\left(\frac{x_i-x}{h_n^*}, \frac{y_j-y}{h_n^*}\right)}{\sum_{(x_i,y_j)\in N^{(1)}_{\widehat{\mathbf{G}}}(x,y)} K^*\left(\frac{x_i-x}{h_n^*}, \frac{y_j-y}{h_n^*}\right)} \\
= & \frac{\sum_{(x_i,y_j)\in N^{(1)}_{\widehat{\mathbf{G}}}(x,y)} R(x_i,y_j) K^*\left(\frac{x_i-x}{h_n^*}, \frac{y_j-y}{h_n^*}\right)}{\sum_{(x_i,y_j)\in N^{(1)}_{\widehat{\mathbf{G}}}(x,y)} K^*\left(\frac{x_i-x}{h_n^*}, \frac{y_j-y}{h_n^*}\right)} + \frac{\sum_{(x_i,y_j)\in N^{(1)}_{\widehat{\mathbf{G}}}(x,y)} \varepsilon_R(x_i,y_j) K^*\left(\frac{x_i-x}{h_n^*}, \frac{y_j-y}{h_n^*}\right)}{\sum_{(x_i,y_j)\in N^{(1)}_{\widehat{\mathbf{G}}}(x,y)} K^*\left(\frac{x_i-x}{h_n^*}, \frac{y_j-y}{h_n^*}\right)} \\
= & \frac{\sum_{(x_i,y_j)\in N^{(1)}_{\widehat{\mathbf{G}}}(x,y)} \left[R(x,y)+R'_{\widehat{\mathbf{G}}}(x,y)\sqrt{(x_i-x)^2+(y_j-y)^2}+O\left(h_n^{*2}\right)\right] K^*\left(\frac{x_i-x}{h_n^*}, \frac{y_j-y}{h_n^*}\right)}{\sum_{(x_i,y_j)\in N^{(1)}_{\widehat{\mathbf{G}}}(x,y)} K^*\left(\frac{x_i-x}{h_n^*}, \frac{y_j-y}{h_n^*}\right)} \\
& + \frac{\sum_{(x_i,y_j)\in N^{(1)}_{\widehat{\mathbf{G}}}(x,y)} \varepsilon_R(x_i,y_j) K^*\left(\frac{x_i-x}{h_n^*}, \frac{y_j-y}{h_n^*}\right)}{\sum_{(x_i,y_j)\in N^{(1)}_{\widehat{\mathbf{G}}}(x,y)} K^*\left(\frac{x_i-x}{h_n^*}, \frac{y_j-y}{h_n^*}\right)} \\
= & R(x,y)+R'_{\widehat{\mathbf{G}}}(x,y)C_K h_n^* + O\left(h_n^{*2}\right) + O\left(\frac{\log n}{\sqrt{\rho_n} n h_n^*}\right), \quad a.s. \\
= & R(x,y)+R'_{\mathbf{G}}(x,y)C_K h_n^* + O\left(h_n^{*2}\right), \quad a.s.
\end{aligned} \tag{S.5}$$

where $R'_{\widehat{\mathbf{G}}}(x,y)$ and $R'_{\mathbf{G}}(x,y)$ are the directional derivatives of R(x,y) along the estimated gradient direction $\widehat{\mathbf{G}}(x,y)$ and the gradient direction $\mathbf{G}(x,y)$, respectively, and

$$C_K = 2\int_{-1}^{1}\int_{0}^{\sqrt{1-u^2}} \sqrt{u^2+v^2}K^*(u,v)\,dvdu.$$

In the second last equation of the above expression, we have used (S.3) and the fact that

$$\frac{\sum_{(x_i,y_j)\in N^{(1)}_{\widehat{\mathbf{G}}}(x,y)} \sqrt{\left(\frac{x_i-x}{h_n^*}\right)^2+\left(\frac{y_j-y}{h_n^*}\right)^2} K^*\left(\frac{x_i-x}{h_n^*},\frac{y_j-y}{h_n^*}\right)}{\sum_{(x_i,y_j)\in N^{(1)}_{\widehat{\mathbf{G}}}(x,y)} K^*\left(\frac{x_i-x}{h_n^*},\frac{y_j-y}{h_n^*}\right)} = C_K + o(1).$$

In the last equation of (S.5), the conditions that $\log^2 n/(nh_n^{*3}\rho_n^{3/2}) = O(1)$ and $\rho_n = o(1)$ have been used, so that $O(\frac{\log n}{\sqrt{\rho_n}nh_n^*}) = o(h_n^{*2})$. Also, we have used the fact that $R'_{\widehat{\mathbf{G}}}(x,y) = R'_{\mathbf{G}}(x,y) + o(1)$, which is a direct conclusion of the results (11) and (12) in Qiu (2009). Similarly, we have

$$\begin{aligned} a^{(2)}_{\widehat{\mathbf{G}}}(x,y) &= \frac{\sum_{(x_i,y_j)\in N^{(2)}_{\widehat{\mathbf{G}}}(x,y)} Z_R(x_i,y_j)K^*\left(\frac{x-x_i}{h_n^*},\frac{y-y_j}{h_n^*}\right)}{\sum_{(x_i,y_j)\in N^{(2)}_{\widehat{\mathbf{G}}}(x,y)} K^*\left(\frac{x-x_i}{h_n^*},\frac{y-y_j}{h_n^*}\right)} \\ &= R(x,y) - R'_{\mathbf{G}}(x,y)C_K h_n^* + O\left(h_n^{*2}\right), \quad a.s. \end{aligned} \tag{S.6}$$

By (S.5) and (S.6), we have

$$\begin{aligned} U_n(x,y) &= \left|a^{(1)}_{\widehat{\mathbf{G}}}(x,y) - a^{(2)}_{\widehat{\mathbf{G}}}(x,y)\right| \\ &= 2C_K\left|R'_{\mathbf{G}}(x,y)\right| h_n^* + O\left(h_n^{*2}\right), \quad a.s. \end{aligned}$$

For any $(x,y) \in D_{2,R}\bigcap\Omega_2(R,T,n)$, $R(x,y)$ is a constant in a neighborhood of (x,y). Therefore, $R'_{\mathbf{G}}(x,y) = 0$. So,

$$\sup_{(x,y)\in D_{2,R}\bigcap\Omega(R,T,n)} U_n(x,y) = O\left(h_n^{*2}\right), \qquad a.s.$$

By the condition that $h_n^{*2}/u_n = o(1)$, when n is large enough, we have

$$D_{2,R}\bigcap\Omega_2(R,T,n) \subseteq \widehat{D}_{2,R}\bigcap\Omega_2(R,T,n). \tag{S.7}$$

On the other hand, for any $(x,y) \in \Omega_R\backslash D_{2,R}$, if $|R'_{\mathbf{G}}(x,y)| \geq \epsilon_n$, then we have

$$U_n(x,y) \geq 2C_K\epsilon_n h_n^* + O\left(h_n^{*2}\right), \qquad a.s.$$

which is uniformly true for all such (x, y). Therefore, when n is large enough, u_n is much smaller than $U_n(x, y)$. Consequently $\widehat{D}_{2,R} \subseteq D_{2,R} \bigcup \mathbf{G}_{2,\epsilon_n}$. So, we have

$$\widehat{D}_{2,R} \bigcap \Omega_2(R, T, n) \subseteq D_{2,R} \bigcap \Omega_2(R, T, n). \tag{S.8}$$

The result (10) in Theorem 2 is then obtained after combining (S.7) and (S.8).

To prove (11) in Theorem 2, by Theorem 1 in Qiu (2009) and similar arguments to those in (S.5), when $(x, y) \in J_n$, we have

$$\begin{aligned} a^{(1)}_{\widehat{\mathbf{G}}}(x, y) &= R_+(x, y) + O(h_n^*) + O\left(\frac{\log n}{\sqrt{\rho_n} n h_n^*}\right), \quad a.s. \\ a^{(2)}_{\widehat{\mathbf{G}}}(x, y) &= R_-(x, y) + O(h_n^*) + O\left(\frac{\log n}{\sqrt{\rho_n} n h_n^*}\right), \quad a.s. \end{aligned}$$

where $R_+(x, y)$ and $R_-(x, y)$ are the one-sided limits of R at (x, y) from two different sides of the edge curve containing (x, y). From the definition of J_n, the jump size of R at (x, y) is at least $\epsilon_n h_n^*$. So, we have

$$\begin{aligned} \inf_{(x,y)\in J_n} U_n(x, y) &= \inf_{(x,y)\in J_n} \left| a^{(1)}_{\widehat{\mathbf{G}}}(x, y) - a^{(2)}_{\widehat{\mathbf{G}}}(x, y) \right| \\ &\geq \epsilon_n h_n^* + O(h_n^*) + O\left(\frac{\log n}{\sqrt{\rho_n} n h_n^*}\right), \qquad a.s. \\ &= \epsilon_n h_n^* + O(h_n^*). \end{aligned}$$

In the above expression, we have used the result that $O\left(\frac{\log n}{\sqrt{\rho_n} n h_n^*}\right) = o(h_n^{*2})$, as in (S.5). By the condition that $u_n/(\epsilon_n h_n^*) = o(1)$, when n is large enough, $(x, y) \in \widehat{\widehat{D}}_{2,R}$, which is uniformly true for $(x, y) \in J_n$. Therefore, $J_n \subseteq \widehat{\widehat{D}}_{2,R}$.

Similarly, for any $(x, y) \in J1_n$, we have

$$\begin{aligned} a^{(1)}_{\widehat{\mathbf{G}}}(x, y) &= R(x, y) + R'_+(x, y) C'_K h_n^* + O\left(h_n^{*2}\right) + O\left(\frac{\log n}{\sqrt{\rho_n} n h_n^*}\right), \quad a.s. \\ a^{(2)}_{\widehat{\mathbf{G}}}(x, y) &= R(x, y) + R'_-(x, y) C'_K h_n^* + O\left(h_n^{*2}\right) + O\left(\frac{\log n}{\sqrt{\rho_n} n h_n^*}\right), \quad a.s., \end{aligned}$$

where $C'_K > 0$ is a constant, $R'_+(x, y)$ and $R'_-(x, y)$ are the one-sided limits of R' at (x, y) from two different sides of the roof edge curve containing (x, y). Again, from the definition of $J1_n$, the jump size of R' at (x, y) is at least ϵ_n. So, we have

$$\begin{aligned} \inf_{(x,y)\in J1_n} U_n(x, y) &= \inf_{(x,y)\in J1_n} \left| a^{(1)}_{\widehat{\mathbf{G}}}(x, y) - a^{(2)}_{\widehat{\mathbf{G}}}(x, y) \right| \\ &\geq C'_K \epsilon_n h_n^* + O\left(h_n^{*2}\right), \ a.s., \end{aligned}$$

which is uniformly true for all $(x,y) \in J1_n$. Again, we have used the result that $O\left(\frac{\log n}{\sqrt{\rho_n} n h_n^*}\right) = o\left(h_n^{*2}\right)$. So, by the condition that $u_n/(\epsilon_n h_n^*) = o(1)$, $J1_n \subseteq \widehat{\overline{D}}_{2,R}$. Result (11) in Theorem 2 follows after we combine this result and the result in the previous paragraph.

Result (12) in Theorem 2 can be proved similarly to result (10), as follows. For any $(x,y) \in \Omega_1(R,T,n)$, we have

$$\begin{aligned}
a_{\widehat{\mathbf{n}}}^{(2)}(x,y) &= \frac{\sum_{(x_i,y_j)\in N_{\widehat{\mathbf{n}}}^{(1)}(x,y)} Z_R(x_i,y_j) K^*\left(\frac{x-x_i}{h_n^*}, \frac{y-y_j}{h_n^*}\right)}{\sum_{(x_i,y_j)\in N_{\widehat{\mathbf{n}}}^{(1)}(x,y)} K^*\left(\frac{x-x_i}{h_n^*}, \frac{y-y_j}{h_n^*}\right)}. \\
&= R(x,y) + R'_{\mathbf{n}}(x,y) C''_K h_n^* + O\left(h_n^{*2}\right) + O\left(\frac{\log n}{\sqrt{\rho_n} n h_n^*}\right), \qquad a.s. \\
a_{\widehat{\mathbf{n}}}^{(2)}(x,y) &= \frac{\sum_{(x_i,y_j)\in N_{\widehat{\mathbf{n}}}^{(2)}(x,y)} Z_R(x_i,y_j) K^*\left(\frac{x-x_i}{h_n^*}, \frac{y-y_j}{h_n^*}\right)}{\sum_{(x_i,y_j)\in N_{\widehat{\mathbf{n}}}^{(2)}(x,y)} K^*\left(\frac{x-x_i}{h_n^*}, \frac{y-y_j}{h_n^*}\right)}. \\
&= R(x,y) - R'_{\mathbf{n}}(x,y) C''_K h_n^* + O\left(h_n^{*2}\right) + O\left(\frac{\log n}{\sqrt{\rho_n} n h_n^*}\right), \qquad a.s.,
\end{aligned}$$

where $R'_{\mathbf{n}}(x,y)$ is the directional derivative of R along the normal direction $\mathbf{n}(x,y)$ at (x,y), and $C''_K > 0$ is a constant. So,

$$\begin{aligned}
V_n(x,y) &= \left| a_{\widehat{\mathbf{n}}}^{(1)}(x,y) - a_{\widehat{\mathbf{n}}}^{(2)}(x,y) \right| \\
&= 2C''_K \left| R'_{\mathbf{n}}(x,y) \right| h_n^* + O\left(h_n^{*2}\right) + O\left(\frac{\log n}{\sqrt{\rho_n} n h_n^*}\right), \qquad a.s. \qquad \text{(S.9)}
\end{aligned}$$

For any $(x,y) \in D_{1,R} \bigcap \Omega_1(R,T,n)$, it is obvious that $R'_{\mathbf{n}}(x,y) = 0$. So,

$$\sup_{(x,y)\in D_{1,R} \bigcap \Omega_1(R,T,n)} V_n(x,y) = O\left(h_n^{*2}\right), \qquad a.s.$$

Here, we have used the result that $O\left(\frac{\log n}{\sqrt{\rho_n} n h_n^*}\right) = o\left(h_n^{*2}\right)$. Therefore, when n is large enough, by the condition that $h_n^{*2}/v_n = o(1)$, we have

$$D_{1,R} \bigcap \Omega_1(R,T,n) \subseteq \widehat{D}_{1,R} \bigcap \Omega_1(R,T,n) \qquad \text{(S.10)}$$

On the other hand, for any $(x,y) \in \overline{D}_{2,R} \backslash D_{1,R}$, if $|R'_{\mathbf{n}}(x,y)| \geq \tau_n$, then from (S.9), we have

$$V_n(x,y) \geq 2C''_K \tau_n h_n^* + O\left(h_n^{*2}\right), \qquad a.s.$$

and this is uniformly true for such (x, y). Therefore, when n is large enough, by the condition that $v_n/(\tau_n h_n^*) = o(1)$,

$$\widehat{D}_{1,R} \bigcap \Omega_1(R, T, n) \subseteq D_{1,R} \bigcap \Omega_1(R, T, n). \tag{S.11}$$

Result (12) follows after combining (S.10) and (S.11).

Proof of Theorem 3

For any point $(x, y) \in (\Omega_{R,\mathbf{T}} \cap \overline{D}_{2,R}) \backslash (D_{1,R} \bigcup P \bigcup J \bigcup J1)$ and a point $(x', y') \in O(x, y, h_n)$, we have

$$\begin{aligned}
& WMSD\left((x, y), (x', y'); h_n\right) \\
= & \frac{\sum_{\sqrt{s^2+t^2} \le h_n} \left[Z_R(x+s, y+t) - Z_M(x'+s, y'+t)\right]^2 K\left(\frac{s}{h_n}, \frac{t}{h_n}\right)}{\sum_{\sqrt{s^2+t^2} \le h_n} K\left(\frac{s}{h_n}, \frac{t}{h_n}\right)} \\
= & \frac{\sum_{\sqrt{s^2+t^2} \le h_n} \eta_{RM}^2(s,t) K\left(\frac{s}{h_n}, \frac{t}{h_n}\right)}{\sum_{\sqrt{s^2+t^2} \le h_n} K\left(\frac{s}{h_n}, \frac{t}{h_n}\right)} \\
& + \frac{\sum_{\sqrt{s^2+t^2} \le h_n} \left(\varepsilon_R(x+s, y+t) - \varepsilon_M(x'+s, y'+t)\right)^2 K\left(\frac{s}{h_n}, \frac{t}{h_n}\right)}{\sum_{\sqrt{s^2+t^2} \le h_n} K\left(\frac{s}{h_n}, \frac{t}{h_n}\right)} \\
& + \frac{2\sum_{\sqrt{s^2+t^2} \le h_n} \left(\varepsilon_R(x+s, y+t) - \varepsilon_M(x'+s, y'+t)\right) \eta_{RM}(s,t) K\left(\frac{s}{h_n}, \frac{t}{h_n}\right)}{\sum_{\sqrt{s^2+t^2} \le h_n} K\left(\frac{s}{h_n}, \frac{t}{h_n}\right)} \\
=: & A_1 + A_2 + A_3,
\end{aligned} \tag{S.12}$$

where $\eta_{RM}(s,t) = R(x+s, y+t) - M(x'+s, y'+t)$.

By Proposition 2 in Qiu (2009), we have

$$A_3 = O\left(\frac{\log n}{n h_n}\right), \qquad a.s. \tag{S.13}$$

For handling A_2, define

$$\xi_{s,t} = \left(\varepsilon_R(x+s, y+t) - \varepsilon_M(x'+s, y'+t)\right)^2 - \left(\sigma_R^2 + \sigma_M^2\right).$$

Then, $E(\xi_{s,t}) = 0$ for any s and t. So, similar to (S.13), under the conditions that $E(\varepsilon_R(x_1, y_1))^6 < \infty$, $E(\varepsilon_M(x_1, y_1))^6 < \infty$, and K is Lipschitz-1 continuous, we have

$$A_2 = \frac{\sum_{\sqrt{s^2+t^2} \le h_n} \left[\xi_{s,t} + (\sigma_R^2 + \sigma_M^2)\right] K\left(\frac{s}{h_n}, \frac{t}{h_n}\right)}{\sum_{\sqrt{s^2+t^2} \le h_n} K\left(\frac{s}{h_n}, \frac{t}{h_n}\right)}$$

$$
\begin{aligned}
&= \sigma_R^2 + \sigma_M^2 + \frac{\sum_{\sqrt{s^2+t^2}\le h_n} \xi_{s,t} K\left(\frac{s}{h_n}, \frac{t}{h_n}\right)}{\sum_{\sqrt{s^2+t^2}\le h_n} K\left(\frac{s}{h_n}, \frac{t}{h_n}\right)} \\
&= \sigma_R^2 + \sigma_M^2 + O\left(\frac{\log n}{nh_n}\right), \qquad a.s. \qquad \text{(S.14)}
\end{aligned}
$$

When $(x', y') = \mathbf{T}(x, y)$, $M(x', y') = R(x, y)$. By the Taylor's expansions of $R(x+s, y+t)$ at (x, y) and $M(x' + s, y' + t)$ at (x', y'), we have

$$
\begin{aligned}
A_1 &= \frac{\sum_{\sqrt{s^2+t^2}\le h_n} \left(R(x+s, y+t) - M(x'+s, y'+t)\right)^2 K\left(\frac{s}{h_n}, \frac{t}{h_n}\right)}{\sum_{\sqrt{s^2+t^2}\le h_n} K\left(\frac{s}{h_n}, \frac{t}{h_n}\right)} \\
&= \frac{\sum_{\sqrt{s^2+t^2}\le h_n} \left\{\left[R_x'(x, y) - M_x'(x', y')\right] s + \left[R_y'(x, y) - M_y'(x', y')\right] t\right\}^2 K\left(\frac{s}{h_n}, \frac{t}{h_n}\right)}{\sum_{\sqrt{s^2+t^2}\le h_n} K\left(\frac{s}{h_n}, \frac{t}{h_n}\right)} + o\left(h_n^2\right) \\
&= O\left(h_n^2\right). \qquad \text{(S.15)}
\end{aligned}
$$

By combining (S.12)–(S.15), we have

$$
WMSD\left((x, y), \mathbf{T}(x, y); h_n\right) = \sigma_R^2 + \sigma_M^2 + O\left(h_n^2\right) + O\left(\frac{\log n}{nh_n}\right) \qquad a.s., \qquad \text{(S.16)}
$$

On the other hand, when $(x', y') \in O(x, y, r_n) \backslash O(\mathbf{T}(x, y), h_n)$,

$$
\begin{aligned}
A_1 &= \frac{\sum_{\sqrt{s^2+t^2}\le h_n} \left(R(x+s, y+t) - M(x'+s, y'+t)\right)^2 K\left(\frac{s}{h_n}, \frac{t}{h_n}\right)}{\sum_{\sqrt{s^2+t^2}\le h_n} K\left(\frac{s}{h_n}, \frac{t}{h_n}\right)} \\
&\ge \frac{\sum_{(s,t)\in\Lambda_n} \left(R(x+s, y+t) - M(x'+s, y'+t)\right)^2 K\left(\frac{s}{h_n}, \frac{t}{h_n}\right)}{\sum_{\sqrt{s^2+t^2}\le h_n} K\left(\frac{s}{h_n}, \frac{t}{h_n}\right)} \\
&\ge \tau_n^2 \frac{\sum_{(s,t)\in\Lambda_n} K\left(\frac{s}{h_n}, \frac{t}{h_n}\right)}{\sum_{\sqrt{s^2+t^2}\le h_n} K\left(\frac{s}{h_n}, \frac{t}{h_n}\right)} \\
&= O\left(\frac{q_n \tau_n^2}{h_n^2}\right), \qquad \text{(S.17)}
\end{aligned}
$$

where Λ_n is a subregion of $\{(s, t) : \sqrt{s^2 + t^2} \le h_n\}$ in which $|R(x+s, y+t) - M(x'+s, y'+t)|$ is at least τ_n. So, by combining (S.12)-(S.14) and (S.17), we have

$$
\begin{aligned}
&\inf_{(x',y')\in(O(x,y,r_n)\backslash O(\mathbf{T}(x,y),h_n))} \left|WMSD\left((x, y), (x', y'); h_n\right)\right| \\
&\ge \sigma_R^2 + \sigma_M^2 + O\left(\frac{q_n \tau_n^2}{h_n^2}\right) + O\left(\frac{\log n}{nh_n}\right), \qquad a.s., \qquad \text{(S.18)}
\end{aligned}
$$

Therefore, by (S.16) and (S.18), under the conditions that $h_n^4/(q_n\tau_n^2) = o(1)$ and $h_n \log n/(nq_n\tau_n^2) = o(1)$, we have

$$\inf_{(x',y')\in(O(x,y,r_n)\setminus O(\mathbf{T}(x,y),h_n))} |WMSD\left((x,y),(x',y');h_n\right)| > WMSD\left((x,y),\mathbf{T}(x,y);h_n\right), \qquad a.s.,$$

A direction conclusion of this result is that

$$d_E(\widehat{\mathbf{T}}(x,y),\mathbf{T}(x,y)) = O(h_n), \qquad a.s.$$

Result (13) in Theorem 3 is then proved.

Some Extra Numerical Results

Table S.1 presents the RRMS and CC values shown in Figure 4 of the paper, along with their standard errors, in the IR problem shown in Figure 1 when the pointwise noise at different pixels is independent.

In the example of Figures 4-7 in the printed paper, the pointwise noise is assumed independent at different pixels. In practice, the noise could be spatially correlated. In the next example, we consider adding spatially correlated noise to the two images shown in Figure 1(a)-(b), and the spatial correlation is described as follows. In model (2) of the paper, for any two pixels (x,y) and (x',y'), the error term ε_R is assumed to follow the spatial correlation model

$$\mathrm{Cov}\left(\varepsilon_R(x,y),\varepsilon_R(x',y')\right) = \begin{cases} \sigma_R^2 & \text{if } (x,y)=(x',y') \\ 0.7\sigma_R^2 & \text{if } |x-x'|=1 \text{ or } |y-y'|=1 \\ 0 & \text{otherwise.} \end{cases}$$

Namely, the correlation between $\varepsilon_R(x,y)$ and the random error at an immediately neighboring pixel of (x,y) is 0.7, and $\varepsilon_R(x,y)$ is uncorrelated with the error terms at other pixels. The error term ε_M is assumed to have the same spatial correlation property. All other aspects of the experiment are the same as those in the example of Figures 4-7 in the printed paper. The computed RRMS and CC values of all six procedures are shown in Figure S.1 and presented in Table S.2. From the figure and the table, it can be seen that NEW outperforms all other five procedures in a quite large margin in this example as well in terms of both RRMS and CC. In the case when $n = 256$ and $\sigma_R = 5$, one set of restored reference images and the corresponding difference images of the six procedures are shown in Figure S.2. Similar conclusions can be made from this figure to those from Figure 5 in the printed paper.

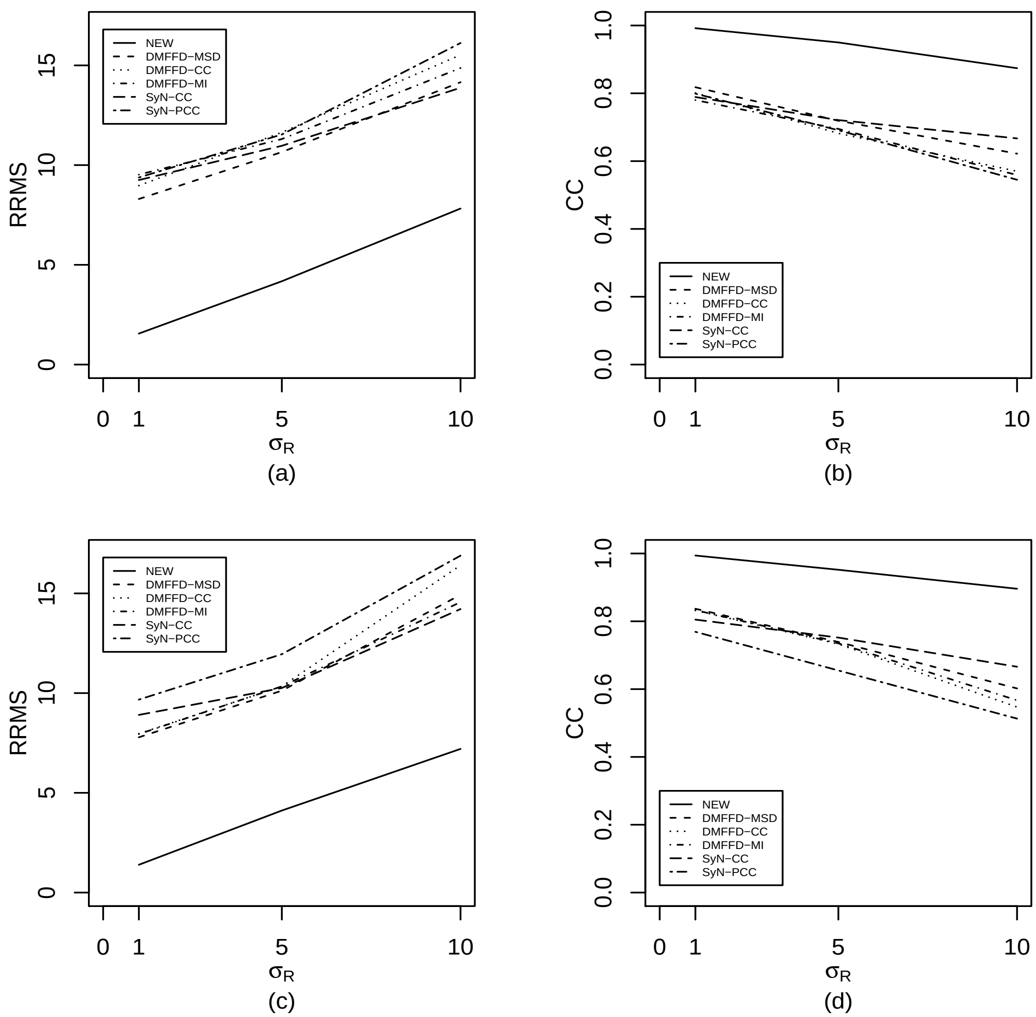

Figure S.1: RRMS and CC values of the six IBIR methods NEW, DMFFD-MSD, DMFFD-CC, DMFFD-MI, SyN-CC, and SyN-PCC in cases when the pointwise noise is spatially correlated. (a) RRMS values when $n = 128$, (b) CC values when $n = 128$, (c) RRMS values when $n = 256$, and (d) CC values when $n = 256$.

Figure S.3 presents the estimated $\mathbf{T}$ of the proposed method NEW in the satellite image of Chicago example. Figure S.4 presents the estimated $\mathbf{T}$ of the proposed method NEW in the Teddy bear example.

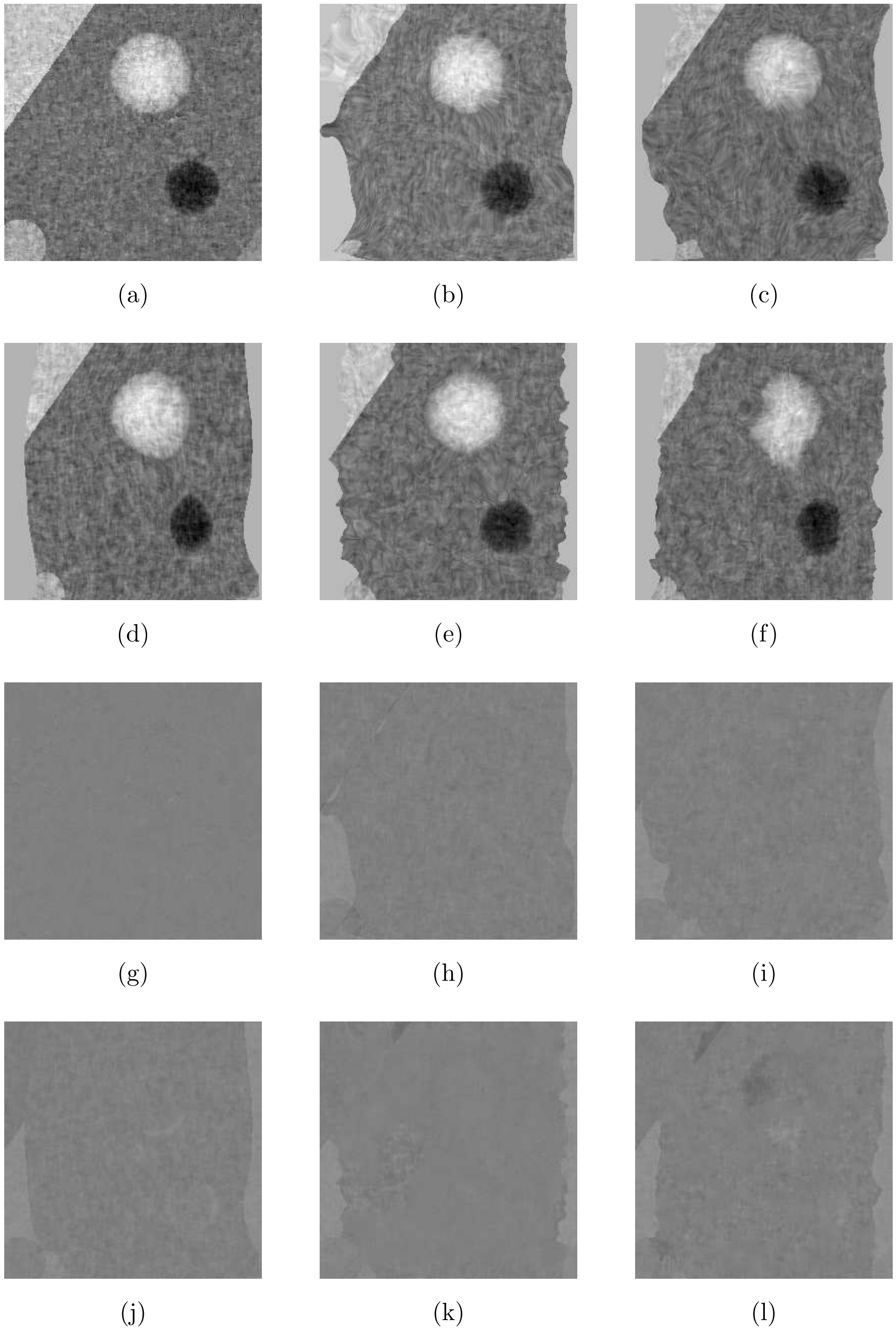

Figure S.2: Restored reference images (plots (a)–(f)), defined to be $Z_M(\widehat{\mathbf{T}}(x, y))$, and the corresponding difference images (plots (g)–(l)), defined to be $Z_R(x, y) - Z_M(\widehat{\mathbf{T}}(x, y))$, of the procedures NEW, DMFFD-MSD, DMFFD-CC, DMFFD-MI, SyN-CC, and SyN-PCC, respectively, in the case when $n = 256$, $\sigma_R = 5$, and the pointwise noise is spatially correlated.

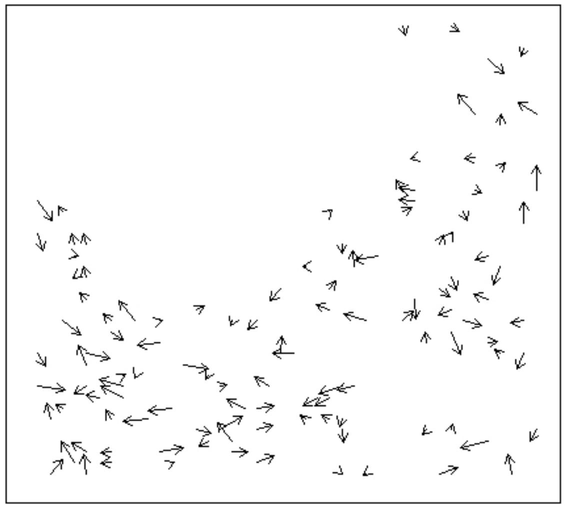

Figure S.3: Estimated **T** of the proposed method NEW in the satellite image of Chicago example.

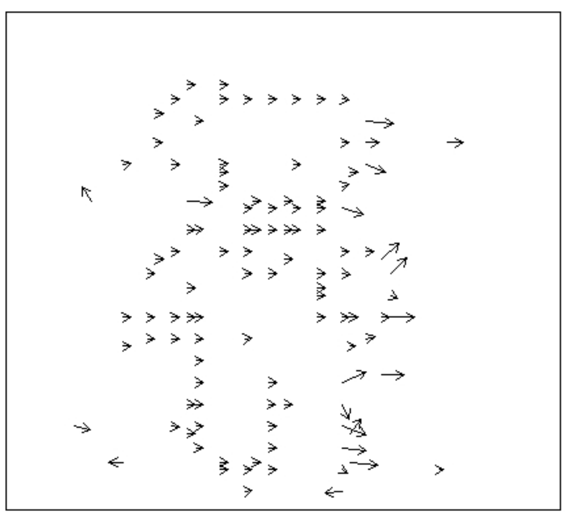

Figure S.4: Estimated **T** of the proposed method NEW in the Teddy bear example.

Table S.1: Performance measures RRMS (first line) and CC (second line) of the six IBIR methods along with their standard errors (in parentheses) in the IR problem shown in Figure 1 when the pointwise noise at different pixels is independent.

n	σ_R	NEW	DMFFD-MSD	DMFFD-CC	DMFFD-MI	SyN-CC	SyN-PCC
128	0	1.284 (0)	8.136 (0)	9.631 (0)	9.582 (0)	9.105 (0)	9.100 (0)
		0.994 (0)	0.830 (0)	0.767 (0)	0.784 (0)	0.796 (0)	0.824 (0)
	1	1.565 (0.024)	8.039 (0.604)	8.528 (0.566)	9.780 (1.641)	8.759 (0.401)	9.157 (0.399)
		0.992 (0)	0.826 (0.022)	0.814 (0.016)	0.768 (0.060)	0.806 (0.010)	0.807 (0.011)
	5	4.767 (0.051)	9.836 (1.454)	10.256 (1.282)	10.638 (1.278)	9.940 (1.181)	10.426 (1.270)
		0.935 (0.001)	0.741 (0.045)	0.723 (0.035)	0.693 (0.045)	0.748 (0.031)	0.729 (0.033)
	10	8.333 (0.053)	14.038 (1.164)	14.294 (1.579)	14.322 (1.211)	13.485 (1.558)	14.650 (1.481)
		0.857 (0.002)	0.585 (0.025)	0.576 (0.031)	0.552 (0.027)	0.631 (0.038)	0.563 (0.029)
256	0	0.916 (0)	7.264 (0)	5.840 (0)	5.091 (0)	8.817 (0)	9.927 (0)
		0.997 (0)	0.859 (0)	0.907 (0)	0.923 (0)	0.811 (0)	0.784 (0)
	1	1.399 (0.020)	7.499 (0.659)	7.599 (0.383)	8.723 (1.569)	8.669 (0.241)	9.442 (0.370)
		0.993 (0)	0.846 (0.021)	0.843 (0.012)	0.808 (0.052)	0.810 (0.006)	0.785 (0.017)
	5	4.466 (0.017)	9.522 (0.837)	9.257 (0.781)	9.110 (1.512)	10.023 (0.943)	11.273 (0.936)
		0.942 (0)	0.756 (0.026)	0.768 (0.023)	0.776 (0.052)	0.746 (0.025)	0.675 (0.027)
	10	8.071 (0.040)	13.858 (0.779)	13.785 (0.872)	12.857 (0.219)	13.669 (1.188)	15.114 (1.120)
		0.865 (0)	0.596 (0.015)	0.603 (0.014)	0.621 (0.021)	0.623 (0.026)	0.532 (0.022)

Table S.2: Performance measures RRMS (first line) and CC (second line) of the six IBIR methods along with their standard errors (in parentheses) in the IR problem shown in Figure 1 when the pointwise noise at different pixels is spatially correlated.

n	σ_R	NEW	DMFFD-MSD	DMFFD-CC	DMFFD-MI	SyN-CC	SyN-PCC
128	1	1.554 (0.035)	8.300 (0.486)	8.973 (0.908)	9.523 (2.015)	9.251 (0.296)	9.379 (0.255)
		0.992 (0)	0.818 (0.021)	0.801 (0.029)	0.780 (0.069)	0.789 (0.009)	0.799 (0.008)
	5	4.177 (0.034)	10.654 (0.910)	11.640 (1.535)	11.306 (2.446)	10.968 (1.254)	11.535 (1.207)
		0.950 (0.002)	0.719 (0.028)	0.682 (0.042)	0.695 (0.075)	0.721 (0.031)	0.691 (0.033)
	10	7.822 (0.089)	14.156 (1.692)	15.525 (1.925)	14.867 (1.982)	13.879 (2.519)	16.122 (1.909)
		0.874 (0.005)	0.622 (0.043)	0.570 (0.046)	0.559 (0.047)	0.667 (0.057)	0.545 (0.037)
256	1	1.393 (0.023)	7.784 (0.781)	7.966 (0.704)	7.945 (2.638)	8.907 (0.357)	9.674 (0.388)
		0.994 (0)	0.837 (0.029)	0.832 (0.023)	0.833 (0.083)	0.805 (0.010)	0.769 (0.018)
	5	4.112 (0.045)	10.122 (0.874)	10.339 (0.956)	10.340 (2.157)	10.255 (1.215)	11.951 (1.057)
		0.952 (0.001)	0.739 (0.021)	0.733 (0.031)	0.735 (0.070)	0.752 (0.033)	0.655 (0.023)
	10	7.202 (0.052)	14.930 (2.543)	16.406 (2.706)	14.562 (1.651)	14.213 (3.195)	16.891 (2.461)
		0.896 (0.003)	0.602 (0.047)	0.547 (0.051)	0.567 (0.046)	0.666 (0.062)	0.513 (0.039)

References

Qiu, P. (2009), "Jump-preserving surface reconstruction from noisy data," *Annals of the Institute of Statistical Mathematics*, **61**, 715–751.