Appendix

Lemma 12

$$\mathbb{E}[\|\tilde{p}_i - p_i\|_1 | P_i] \leq \overline{C} n^{-1/(2+k)}$$
$$\mathbb{E}[\|\tilde{p}_i - p_i\|_1] \leq \overline{C} n^{-1/(2+k)}$$

where $\overline{C} > 0$ is appropriate constant defined in proof.

Proof. By assumption $\mathfrak{A}3$, the class $\mathcal{I}$ is the set of distribution $\mathcal{H}_k(1)$ with with densities that are 1-smooth Hölder functions, as in (Rigollet & Vert, 2009). Let $D_2^2(\tilde{P}_i, P_i) = \int (\tilde{p}(x) - p(x))^2 \mathrm{d}x$, then $\mathbb{E}[D_2^2(\tilde{P}_i, P_i)|P_i]$ is the integrated mean squared risk for the density estimator $\tilde{p}_i$ for a fixed p_i. Then it is well known that $\mathbb{E}[D_2^2(\tilde{P}_i, P_i)|P_i] \leq c_1^2 b_i^2 + c_2^2 (n_i b_i^k)^{-1}$ for some constants $c_1 c_2 > 0$. Hence, by Jensen's inequality

$$\begin{aligned}\mathbb{E}[D_2(\tilde{P}_i, P_i)|P_i] \leq & (c_1^2 b_i^2 + c_2^2 (n_i b_i^k)^{-1})^{\frac{1}{2}} \\ \leq & c_1 b_i + c_2 (n_i b_i^k)^{-\frac{1}{2}}\end{aligned}$$

Furthermore, since by $\mathfrak{A}3$, P_i's support is compact. So for an appropriate constant $c_0 > 0$, $\int |p_i - \tilde{p}_i| \leq c_0 \sqrt{\int (p_i - \tilde{p}_i)^2}$. Thus:

$$\begin{aligned}\mathbb{E}[D(\tilde{P}_i, P_i)|P_i] \leq & c_0 \mathbb{E}[D_2(\tilde{P}_i, P_i)|P_i] \\ \leq & c_0 (c_1 b_i + c_2 (n_i b_i^k)^{-\frac{1}{2}}) \\ \leq & c_0 (c_1 + c_2) n_i^{-\frac{1}{k+2}}\end{aligned}$$

□

Let

$$\Omega_{M,n} \equiv \left\{ \forall i \in \{0, \ldots, M\},\ D(\tilde{P}_i, P_i) \leq C_* n^{-\frac{1}{k+2}} \right\}.$$

Lemma 13 $\mathbb{P}(\Omega_{M,n}) \geq 1 - (M+1) e^{-\frac{1}{2} n^{\frac{k}{2+k}}}$

Proof. From McDiarmid's inequality for $\epsilon > 0$ we have that $\mathbb{P}(\|\tilde{p}_i - p_i\|_1 - \mathbb{E}[\|\tilde{p}_i - p_i\|_1] > \epsilon) \leq e^{\frac{-n\epsilon^2}{2}}$ (see section 2.4 of (Devroye & Lugosi, 2001)). Hence,

$$\mathbb{P}(\|\tilde{p}_i - p_i\|_1 - \mathbb{E}[\|\tilde{p}_i - p_i\|_1] > n^{-\frac{1}{k+2}}) \leq e^{-\frac{1}{2} n^{-\frac{k}{k+2}}}$$

Thus, by the union bound, and since P_i are i.i.d :

$$\begin{aligned}& 1 - (M+1) e^{-\frac{1}{2} n^{\frac{k}{k+2}}} \\ & \leq \mathbb{P}(\forall i,\ \|\tilde{p}_i - p_i\|_1 \leq \mathbb{E}[\|\tilde{p}_i - p_i\|_1] + n^{-\frac{1}{k+2}}) \\ & \leq \mathbb{P}(\forall i,\ \|\tilde{p}_i - p_i\|_1 \leq (1 + \overline{C}) n^{-\frac{1}{k+2}}) \\ & = \mathbb{P}(\forall i,\ D(\tilde{P}_i, P_i) \leq C_* n^{-\frac{1}{k+2}}) = \mathbb{P}(\Omega_{M,n})\end{aligned}$$

□

Lemma 1

$$\begin{aligned}\mathbb{P}\left(\sum_{i=1}^M K_i = 0\right) & \leq \mathbb{P}\left(\sum_{i=1}^M K_i \leq \underline{K}\right) \\ & = \frac{1}{eM} \mathbb{E}\left[\frac{1}{\Phi_P(rh)}\right]\end{aligned}$$

Proof. Since if $\exists i$ s.t. $D(P_0, P_i) \leq rh \implies \sum_i K_i \geq \underline{K}$, we have that $\sum_i K_i < \underline{K} \implies \forall i,\ D(P_0, P_i) > rh$, so $\sum_i K_i < \underline{K} \implies \sum_i I_{\{D(P_0,P_i) \leq rh\}} = 0$. Hence,

$$\begin{aligned}& \mathbb{P}\left(\sum_{i=1}^M K_i < \underline{K}\right) \\ & \leq \mathbb{P}\left(\sum_{i=1}^M I_{\{D(P_0,P_i) \leq rh\}} = 0\right) \\ & = \int \mathbb{P}\left(\sum_{i=1}^M I_{\{D(P_0,P_i) \leq rh\}} = 0 \middle| P_0\right) \mathrm{d}\mathcal{P}(P_0) \\ & = \int \mathbb{P}\left(\forall i\ D(P_0, P_i) > rh \middle| P_0\right) \mathrm{d}\mathcal{P}(P_0) \\ & = \int [1 - \mathcal{P}(P_1 \in \mathcal{B}_D(P_0, rh)|P_0)]^M \mathrm{d}\mathcal{P}(P_0) \qquad (24) \\ & \leq \int \exp[-M\mathcal{P}(P_1 \in \mathcal{B}_D(P_0, rh)|P_0)] \mathrm{d}\mathcal{P}(P_0) \qquad (25) \\ & \leq \int \exp[-M\mathcal{P}(P_1 \in \mathcal{B}_D(P_0, rh)|P_0)] \\ & \quad \times \frac{M\mathcal{P}(P_1 \in \mathcal{B}_D(P_0, rh)|P_0)}{M\mathcal{P}(P_1 \in \mathcal{B}_D(P_0, rh)|P_0)} \mathrm{d}\mathcal{P}(P_0) \\ & \leq \max_{u \geq 0} u \exp(-u) \int \frac{\mathrm{d}\mathcal{P}(P_0)}{M\mathcal{P}(P_1 \in \mathcal{B}_D(P_0, rh)|P_0)} \\ & \leq \frac{1}{e} \int \frac{\mathrm{d}\mathcal{P}(P_0)}{M\mathcal{P}(P_1 \in \mathcal{B}_D(P_0, rh)|P_0)} \qquad (26) \\ & = \frac{1}{eM} \mathbb{E}\left[\frac{1}{\Phi_P(rh)}\right]\end{aligned}$$

where (24) holds since $\{P_i\}$ are drawn iid, (25) holds since for $0 \leq u \leq 1$, $1 \leq M$, $(1-u)^M \leq e^{-Mu}$ and (26) since $\max(u \exp(-u)) = 1/e$. □

Lemma 14

$$|\epsilon_i| \leq \frac{L_K}{h} \left(D(\tilde{P}_0, \tilde{P}_0) + D(\tilde{P}_i, \tilde{P}_i) \right)$$

Proof.

$$\begin{aligned}|\epsilon_i| & = \left| K\left(\frac{D(P_0, P_i)}{h}\right) - K\left(\frac{D(\tilde{P}_0, \tilde{P}_i)}{h}\right) \right| \\ & \leq \frac{L_K}{h} \left| D(P_0, P_i) - D(\tilde{P}_0, \tilde{P}_i) \right|\end{aligned}$$

$$\leq \frac{L_K}{h}(D(P_0,\tilde{P}_0)+D(\tilde{P}_i,P_i))$$

Since

$$\begin{aligned}&D(P_0,P_i)-D(\tilde{P}_0,\tilde{P}_i)\\&\leq D(P_0,\tilde{P}_0)+D(\tilde{P}_0,\tilde{P}_i)+D(\tilde{P}_i,P_i)-D(\tilde{P}_0,\tilde{P}_i)\\&=D(P_0,\tilde{P}_0)+D(\tilde{P}_i,P_i)\end{aligned}$$

and

$$\begin{aligned}&D(\tilde{P}_0,\tilde{P}_i)-D(P_0,P_i)\\&\leq D(\tilde{P}_0,P_0)+D(P_0,P_i)+D(P_i,\tilde{P}_i)-D(P_0,P_i)\\&=D(P_0,\tilde{P}_0)+D(\tilde{P}_i,P_i)\end{aligned}$$

. □

Lemma 15

$$\mathbb{P}\left(\sum_i |\epsilon_i|<\omega\middle|\mathbf{P}\right)\geq 1-\frac{2L_K M\overline{C}}{h\omega}n^{-1/(2+k)}$$

Proof. Markov's inequality states that for r.v. X,Y and constant $\omega>0$,

$$\mathbb{P}(|X|<\omega|Y)\geq 1-\frac{\mathbb{E}\left[|X|\middle|Y\right]}{\omega}$$

Hence,

$$\begin{aligned}\mathbb{P}\left(\sum_i |\epsilon_i|<\omega\middle|\mathbf{P}\right)&\geq 1-\frac{\mathbb{E}\left[\sum_i|\epsilon_i|\middle|\mathbf{P}\right]}{\omega}\\&\geq 1-\frac{L_k}{h\omega}\sum_i\mathbb{E}\left[|\epsilon_i|\middle|\mathbf{P}\right]\\&\geq 1-\frac{2L_K M\overline{C}}{h\omega}n^{-1/(2+k)}\qquad(27)\end{aligned}$$

where (27) holds due to Lemma 12, and Lemma 14. □

Lemma 16

$$\mathbb{E}\left[\sum_{i=1}^M|\epsilon_i|\middle|\mathbf{P}\right]\geq\frac{2L_K M\overline{C}}{h}n^{-1/(2+k)}$$

Proof.

$$\begin{aligned}\mathbb{E}\left[\sum_{i=1}^M|\epsilon_i|\middle|\mathbf{P}\right]&\leq\frac{L_k}{h}\sum_{i=1}^M\mathbb{E}\left[D(P_0,\tilde{P}_0)+D(\tilde{P}_i,P_i)\middle|\mathbf{P}\right]\\&\leq\frac{2L_k M}{h}n^{-1/(2+k)}\end{aligned}$$

□

Lemma 2

$$\mathbb{P}\left(\sum_{i=1}^M\tilde{K}_i=0\right)\leq\mathbb{P}\left(\sum_{i=1}^M\tilde{K}_i\leq\underline{K}\right)\leq\zeta(n,M)$$

Proof. First note that if Ω_M,n holds then with $\mathfrak{A}6$:

$$\begin{aligned}D(\tilde{P}_i,\tilde{P})&\leq D(\tilde{P}_i,P_i)+D(P,\tilde{P})+D(P_i,P)\\&\leq\frac{rh}{2}+D(P_i,P)\end{aligned}$$

Hence,

$$\begin{aligned}&\left[D(P_i,P)\leq\frac{rh}{2}\implies D(\tilde{P}_i,\tilde{P})\leq rh\right]\implies\\&1-\mathbb{P}\left(D(P_i,P)\leq\frac{rh}{2}\right)\geq 1-\mathbb{P}\left(D(\tilde{P}_i,\tilde{P})\leq rh\right).\qquad(28)\end{aligned}$$

Thus,

$$\begin{aligned}&\mathbb{P}\left(\sum_{i=1}^M\tilde{K}_i=0\right)\leq\mathbb{P}\left(\sum_{i=1}^M\tilde{K}_i\leq\underline{K}\right)\\&=\mathbb{P}\left(\Omega_{M,n},\sum_{i=1}^M\tilde{K}_i\leq\underline{K}\right)+\mathbb{P}\left(\Omega_{M,n}^c,\sum_{i=1}^M\tilde{K}_i\leq\underline{K}\right)\\&=\mathbb{P}\left(\Omega_{M,n},\sum_{i=1}^M\tilde{K}_i\leq\underline{K}\right)+\mathbb{P}\left(\Omega_{M,n}^c,\sum_{i=1}^M\tilde{K}_i\leq\underline{K}\right)\\&\leq\mathbb{P}\left(\sum_{i=1}^M\tilde{K}_i\leq\underline{K}\right)+\mathbb{P}\left(\Omega_{M,n}^c\right)\\&\leq\mathbb{P}\left(\sum_{i=1}^M\tilde{K}_i\leq\underline{K}\right)+(M+1)e^{-\frac{1}{2}n^{\frac{k}{2+k}}}\end{aligned}$$

And, using a similar argument to Lemma 1 and (28):

$$\begin{aligned}\mathbb{P}\left(\sum_{i=1}^M\tilde{K}_i\leq\underline{K}\right)&\leq\sum_i I_{\{D(\tilde{P}_0,\tilde{P}_i)\leq rh\}}=0\\&\leq\frac{1}{em}\mathbb{E}\left[\frac{1}{\Phi_P(rh/2)}\right]\end{aligned}$$

□

Lemma 3

$$\mathbb{E}\left[\frac{I_{\{\sum_i K_i\leq\underline{K}\}}}{\sum_i K_i}\right]\leq\frac{1+1/\underline{K}}{M\underline{K}}\mathbb{E}\left[\frac{1}{\Phi_P(rh)}\right]$$

Proof.

$$\mathbb{E}\left[\frac{I_{\{\sum_i K_i\leq\underline{K}\}}}{\sum_i K_i}\right]\leq\mathbb{E}\left[\frac{1+1/\underline{K}}{1+\sum_i K_i}\right]\qquad(29)$$

$$\le \mathbb{E}\left[\frac{1+1/\underline{K}}{1+\underline{K}\sum_i I_{\{D(P_0,P_i)\le rh\}}}\right]$$
$$\le \frac{1+1/\underline{K}}{\underline{K}}\mathbb{E}\left[\frac{1}{1/\underline{K}+\sum_i I_{\{D(P_0,P_i)\le rh\}}}\right]$$
$$\le \frac{1+1/\underline{K}}{\underline{K}}\mathbb{E}\left[\frac{1}{1+\sum_i I_{\{D(P_0,P_i)\le rh\}}}\right] \quad (30)$$
$$\le \frac{1+1/\underline{K}}{\underline{K}}\mathbb{E}\left[\mathbb{E}\left[\frac{1}{1+\sum_i I_{\{D(P_0,P_i)\le rh\}}}\middle|P_0\right]\right]$$
$$\le \frac{1+1/\underline{K}}{M\underline{K}}\mathbb{E}\left[\frac{1}{\Phi_P(rh)}\right] \quad (31)$$

where (29) holds since $\underline{K}\le\sum_i K_i \implies 1+\sum_i K_i \le \sum_i K_i+\sum_i K_i/\underline{K}$, (30) since $\underline{K}<1$, and (31) since for a binomial random variable $B(M,p)$, $\mathbb{E}[\frac{1}{1+B(M,p)}] \le \frac{1}{(M+1)p} \le \frac{1}{Mp}$. □

Lemma 4 $\mathbb{E}\left[\Delta\hat{a}_\alpha I_{\tilde{E}_2} I_{E_2}\right] \le \frac{C_1}{h}\mathbb{E}\left[\frac{1}{\Phi_P(rh)}\right] n^{-1/(2+k)}$ for a $C_1>0$.

Proof.

$$\mathbb{E}\left[\Delta\hat{a}_\alpha I_{\tilde{E}_2} I_{E_2}\right]$$
$$= \mathbb{E}\left[\left|\frac{\sum_i a_\alpha(\widehat{Q}_i)\tilde{K}_i}{\sum_j \tilde{K}_j} - \frac{\sum_i a_\alpha(\widehat{Q}_i)K_i}{\sum_j K_j}\right| I_{E_2} I_{\tilde{E}_2}\right]$$
$$= \mathbb{E}\left[\left|\sum_i a_\alpha(\widehat{Q}_i)\frac{\tilde{K}_i}{\sum_j \tilde{K}_j} - \frac{K_i}{\sum_j K_j}\right| I_{E_2} I_{\tilde{E}_2}\right]$$
$$\le \varphi_{\max}\mathbb{E}\left[\sum_i\left|\frac{\tilde{K}_i}{\sum_j \tilde{K}_j} - \frac{K_i}{\sum_j K_j}\right| I_{E_2} I_{\tilde{E}_2}\right]$$
$$= \varphi_{\max}\mathbb{E}\left[\sum_i \frac{|K_i(\sum_j \tilde{K}_j) - \tilde{K}_i(\sum_j K_j)|}{(\sum_j \tilde{K}_j)(\sum_j K_j)} I_{E_2} I_{\tilde{E}_2}\right]$$
$$= \varphi_{\max}\mathbb{E}\Bigg[\sum_i \frac{|(\tilde{K}_i-\epsilon_i)(\sum_j \tilde{K}_j) - \tilde{K}_i(\sum_j(\tilde{K}_j-\epsilon_j))|}{(\sum_j \tilde{K}_j)(\sum_j K_j)} \times I_{E_2} I_{\tilde{E}_2}\Bigg]$$
$$= \varphi_{\max}\mathbb{E}\left[\sum_i \frac{|-\epsilon_i(\sum_j \tilde{K}_j) + \tilde{K}_i(\sum_j \epsilon_j)|}{(\sum_j \tilde{K}_j)(\sum_j K_j)} I_{E_2} I_{\tilde{E}_2}\right]$$
$$\le \varphi_{\max}\mathbb{E}\Bigg[\frac{(\sum_i|\epsilon_i|)(\sum_j \tilde{K}_j)}{(\sum_j \tilde{K}_j)(\sum_j K_j)} + \frac{(\sum_i \tilde{K}_i)(\sum_j|\epsilon_j|)}{(\sum_j \tilde{K}_j)(\sum_j K_j)} I_{E_2} I_{\tilde{E}_2}\Bigg]$$
$$\le 2\varphi_{\max}\mathbb{E}\left[\frac{(\sum_j|\epsilon_j|)}{(\sum_j K_j)} I_{E_2}\right]$$
$$\le 2\varphi_{\max}\mathbb{E}\left[\mathbb{E}\left[\sum_j|\epsilon_j|\middle|\mathbf{P}\right]\frac{1}{(\sum_j K_j)} I_{E_2}\right]$$
$$\le 2\varphi_{\max}\frac{2L_k M}{h} n^{-1/(2+k)}\mathbb{E}\left[\frac{1}{(\sum_j K_j)} I_{E_2}\right]$$
$$\le 2\varphi_{\max}\frac{2L_k M}{h} n^{-1/(2+k)}\frac{1+1/\underline{K}}{M\underline{K}}\mathbb{E}\left[\frac{1}{\Phi_P(rh)}\right],$$

having used Lemma 3 and Lemma 16. □

Lemma 5 $\mathbb{E}[|\mu_\alpha^{(i)}|] \le \sqrt{\mathbb{E}[|\mu_\alpha^{(i)}|^2]} \le cm^{-\frac{1}{2}}$

By Jensen's inequality:

$$\mathbb{E}[|\mu_\alpha^{(i)}|] \le \sqrt{\mathbb{E}[|\mu_\alpha^{(i)}|^2]} \le \sqrt{\mathbb{E}[(a_\alpha(\widehat{Q}_i) - a_\alpha(Q_i))^2]}$$
$$\le \sqrt{\mathrm{Var}[a_\alpha(\widehat{Q}_i)]} = \sqrt{\mathrm{Var}\left[\frac{1}{m_i}\sum_{j=1}^{m_i}\varphi_\alpha(Y_{ij})\right]}$$
$$\le \frac{1}{m_i}\sqrt{\sum_{j=1}^{m_i}\mathbb{E}\left[\varphi_\alpha^2(Y_{ij})\right]} \le \frac{1}{m_i}\sqrt{m_i\varphi_{\max}^2}$$
$$\le c_1 m^{-\frac{1}{2}}.$$

Lemma 6 $\mathbb{E}\left[\left|\frac{\sum_i \mu_\alpha^{(i)} K_i}{\sum_i K_i}\right| \bar{I}\right] \le C\sqrt{\frac{1}{mM}\mathbb{E}\left[\frac{1}{\Phi_P(rh)}\right]}$

Proof.

$$\mathbb{E}\left[\left|\frac{\sum_i \mu_\alpha^{(i)} K_i}{\sum_i K_i}\right| \bar{I}\right] = \mathbb{E}\left[\mathbb{E}\left[\left|\frac{\sum_i \mu_\alpha^{(i)} K_i}{\sum_i K_i}\right| \bar{I}\middle|\mathbf{P}\right]\right]$$
$$\le \mathbb{E}\left[\sqrt{\mathbb{E}\left[\left|\frac{\sum_i \mu_\alpha^{(i)} K_i}{\sum_i K_i}\right|^2\middle|\mathbf{P}\right]}\bar{I}\right]$$
$$\le \mathbb{E}\left[\frac{1}{\sum_i K_i}\sqrt{\mathbb{E}\left[\sum_i(\mu_\alpha^{(i)}K_i)^2\middle|\mathbf{P}\right]}\bar{I}\right]$$
$$\le \mathbb{E}\left[c_1 m^{-\frac{1}{2}}\frac{\sqrt{\sum_i K_i^2}}{\sum_i K_i}\bar{I}\right] \le c_1 m^{-\frac{1}{2}}\sqrt{\mathbb{E}\left[\frac{\sum_i K_i^2}{(\sum_i K_i)^2}\bar{I}\right]}$$
$$\le c_1 m^{-\frac{1}{2}}\sqrt{\mathbb{E}\left[\frac{\sum_i K_i}{(\sum_i K_i)^2}\bar{I}\right]} = c_1 m^{-\frac{1}{2}}\sqrt{\mathbb{E}\left[\frac{\bar{I}}{\sum_i K_i}\right]}$$
$$\le c_1 m^{-\frac{1}{2}}\sqrt{\frac{1+1/\underline{K}}{M\underline{K}}\mathbb{E}\left[\frac{1}{\Phi_P(rh)}\right]},$$

where we used Lemma 5, Jensen's inequality, $K_i<1$, and Lemma 3 □

Lemma 8 If $\frac{1}{Mh^d} = \Omega(\sqrt{\frac{1}{nMh^d}})$ and $\frac{1}{Mh^d} = \Omega(\frac{n^{-\frac{1}{2+k}}}{h^{d+1}})$, then $R(M,n) = O(h^\beta + \frac{1}{Mh^d})$ and choosing h optimally leads to $R(M,n) = O(M^{-\frac{\beta}{\beta+d}})$.

Proof. Suppose: i) $\frac{1}{Mh^d} = \Omega(\sqrt{\frac{1}{nMh^d}})$, ii) $\frac{1}{Mh^d} = \Omega(\frac{n^{-\frac{1}{2+k}}}{h^{d+1}})$, then $R(M,n) = O(h^\beta + \frac{1}{Mh^d})$. The optimal choice of h is $h = O(M^{\frac{-1}{\beta+d}})$, leading to $R(M,n) = O(M^{-\frac{\beta}{\beta+d}})$. Note that i) implies that $n = \Omega(M^{\frac{\beta}{\beta+d}})$ and ii) implies $n = \Omega(M^{\frac{(\beta+d+1)(k+2)}{\beta+d}})$; since $M^{\frac{(\beta+d+1)(k+2)}{\beta+d}} = \Omega(M^{\frac{\beta}{\beta+d}})$, i) and ii) imply $n = \Omega(M^{\frac{(\beta+d+1)(k+2)}{\beta+d}})$. Furthermore, $n = \Omega(M^{\frac{(\beta+d+1)(k+2)}{\beta+d}})$ implies $h = \Theta(M^{\frac{-1}{\beta+d}}) = \Omega(n^{-\frac{1}{(\beta+d+1)(k+2)}})$ and $n^{-\frac{1}{(\beta+d+1)(k+2)}} = \Omega(n^{-\frac{1}{k+2}})$, thus assumption $\mathfrak{A}6$ is not violated. Moreover, $e^{\frac{1}{2}M^{\frac{k(\beta+d+1)}{\beta+d}}} = \Omega(M^{\frac{2\beta+d}{\beta+d}}) \implies M^{-\frac{\beta}{\beta+d}} = \Omega(Me^{-\frac{1}{2}M^{\frac{k(\beta+d+1)}{\beta+d}}}) \implies (M+1)e^{-\frac{1}{2}n^{\frac{k}{k+2}}} = O(M^{-\frac{\beta}{\beta+d}})$, where the last implication follows by $n = \Omega(M^{\frac{(\beta+d+1)(k+2)}{\beta+d}})$. Hence, $R(M,n,m) = O(R(M,n)) = O(M^{\frac{-\beta}{\beta+d}})$. Note that since M is slow growing in this case, it makes sense that the rate be driven by it. □

Lemma 9 If $\frac{n^{-\frac{1}{2+k}}}{h^{d+1}} = \Omega(\sqrt{\frac{1}{nMh^d}})$ and $\frac{n^{-\frac{1}{2+k}}}{h^{d+1}} = \Omega(\frac{1}{Mh^d})$, then $R(M,n) = O(h^\beta + n^{\frac{-1}{2+k}}h^{-(d+1)})$ and choosing h optimally leads to $R(M,n) = O(n^{-\frac{\beta}{(k+2)(\beta+d+1)}})$.

Proof. Suppose: i) $\frac{n^{-\frac{1}{2+k}}}{h^{d+1}} = \Omega(\sqrt{\frac{1}{nMh^d}})$, ii) $\frac{n^{-\frac{1}{2+k}}}{h^{d+1}} = \Omega(\frac{1}{Mh^d})$, then $R(M,n) = O(h^\beta + n^{\frac{-1}{2+k}}h^{-(d+1)})$. The optimal choice of h is $h = \Theta(n^{-\frac{1}{(k+2)(\beta+d+1)}})$, leading to $R(M,n) = O(n^{-\frac{\beta}{(k+2)(\beta+d+1)}})$. Here, i) implies $n^{\frac{k(\beta+d+1)+d+2}{(k+2)(\beta+d+1)}}M = \Omega(1)$, which is always true, and ii) implies $M = \Omega(n^{\frac{\beta+d}{(k+2)(\beta+d+1)}})$; hence, i) and ii) implies $M = \Omega(n^{\frac{\beta+d}{(k+2)(\beta+d+1)}})$. Moreover, $h = \Theta(n^{-\frac{1}{(k+2)(\beta+d+1)}})$ and $n^{-\frac{1}{(k+2)(\beta+d+1)}} = \Omega(n^{-\frac{1}{k+2}})$ so $\mathfrak{A}6$ holds. Note that this case invoked when $M = \Omega(n^{\frac{\beta+d}{(k+2)(\beta+d+1)}})$; thus, in order to prevent M from growing too fast in this case, and not having $(M+1)e^{-\frac{1}{2}n^{\frac{k}{k+2}}} = O(n^{-\frac{\beta}{(k+2)(\beta+d+1)}})$, assumption $\mathfrak{A}5$ was slightly extended as follows: $M = O(n^{-\frac{\beta}{(k+2)(\beta+d+1)}}e^{n^{\frac{k}{k+2}}})$. Then, $R(M,n,m) = O(n^{-\frac{\beta}{(k+2)(\beta+d+1)}})$. This rate is again intuitive since n is slow growing in this case, so it should drive the rate. □

Lemma 10 If one is chooses h optimally it can not be that $\sqrt{\frac{1}{nMh^d}} = \Omega(\frac{n^{-\frac{1}{2+k}}}{h^{d+1}})$ and $\sqrt{\frac{1}{nMh^d}} = \Omega(\frac{1}{Mh^d})$.

Proof. If $\sqrt{\frac{1}{nMh^d}} = \Omega(\frac{n^{-\frac{1}{2+k}}}{h^{d+1}})$ and $\sqrt{\frac{1}{nMh^d}} = \Omega(\frac{1}{Mh^d})$, then $R(M,n) = O(h^\beta + \sqrt{\frac{1}{nMh^d}})$. Here, the optimal choice of h is $h = \Theta((nM)^{-1/(2\beta+d)})$. Then $\sqrt{\frac{1}{nMh^d}} = \Omega(\frac{n^{-\frac{1}{2+k}}}{h^{d+1}}) \implies \sqrt{nMh^d} = O(n^{\frac{1}{2+k}}h^{d+1}) \implies n^{\frac{k}{2+k}}M = O(h^{d+2}) \implies n^{\frac{k}{2+k}}M = O((nM)^{-\frac{d+2}{2\beta+d}})$, contradiction. □