1 Supplementary material for the referees

In the supplementary material the proof of Lemma 2 as well as the proof of the inequalities (39) and (50) is given.

1.1 Proof of Lemma 2

In the proof of Lemma 2 we need the notion of VC-dimension. Denote by $V_{\mathcal{A}}$ the VC-dimension of a class of subsets $\mathcal{A} \neq \emptyset$ of $\mathbb{R}^d$, which is defined by

$$V_{\mathcal{A}} = \sup\{n \in \mathbb{N} : \ S(\mathcal{A}, n) = 2^n\},$$

where $S(\mathcal{A}, n)$ is the n-th shatter coefficient of $\mathcal{A}$, i.e.

$$S(\mathcal{A}, n) = \max_{\{z_1,\dots,z_n\}\subseteq \mathbb{R}^d} |\{A \cap \{z_1, \dots, z_n\} : \ A \in \mathcal{A}\}|.$$

Proof of Lemma 2. The proof is based on parts of the proof of Lemma 3.2 in Kohler et al. (2003). First, we observe that

$$\mathcal{N}_1(\epsilon_n, \bar{\mathcal{G}}_n, (u_1^n, x_1^n)) \leq \mathcal{N}_1\left(\frac{\epsilon_n}{d_n}, \mathcal{G}_n, (u_1^n, x_1^n)\right), \tag{52}$$

where $\mathcal{G}_n$ is a set of functions defined as

$$\mathcal{G}_n := \Bigg\{ g : [0,1] \times \mathbb{R}^d \to [0, K(0)] : \ g(u,x) = \mathbb{1}_{\{m(u,x) \leq y\}} \cdot K\left(\frac{t-u}{h_{n,1}}\right) \quad ((u,x) \in [0,1] \times \mathbb{R}^d), \ t \in [0,1], \ y \in \mathbb{R} \Bigg\}.$$

Let $g_1, \dots, g_N : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ be an minimal ϵ_n-cover of $\mathcal{G}_n$, i.e. for every $g \in \mathcal{G}_n$ there is a $j = j(g) \in \{1, \dots, N\}$ such that

$$\frac{1}{n}\sum_{i=1}^n |g(u_i, x_i) - g_j(u_i, x_i)| < \epsilon_n.$$

Then $\bar{g}_1, \dots, \bar{g}_N : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$, where

$$\bar{g}_j(u,x) = c_u \cdot g_j(u,x) \quad \text{for all } (u,x) \in \mathbb{R} \times \mathbb{R}^d, \ j = 1, \dots, N,$$

is an δ_n-cover of $\bar{\mathcal{G}}_n$ for $\delta_n = d_n \cdot \epsilon_n$, since

$$\begin{aligned} \frac{1}{n}\sum_{i=1}^n |\bar{g}(u_i, x_i) - \bar{g}_j(u_i, x_i)| &= \frac{1}{n}\sum_{i=1}^n |c_{u_i} \cdot g(u_i, x_i) - c_{u_i} \cdot g_j(u_i, x_i)| \\ &\leq d_n \cdot \frac{1}{n}\sum_{i=1}^n |g(u_i, x_i) - g_j(u_i, x_i)| \\ &< d_n \cdot \epsilon_n. \end{aligned}$$

Hence, we have proven (52). Next, we bound $\mathcal{N}_1\left(\frac{\epsilon_n}{d_n},\mathcal{G}_n,(u_1^n,x_1^n)\right)$. Since the functions are bounded, the proof of Lemma 16.5 in Györfi et al. (2002) implies that

$$\mathcal{N}_1\left(\frac{\epsilon_n}{d_n},\mathcal{G}_n,(u_1^n,x_1^n)\right)\leq \mathcal{N}_1\left(\frac{\epsilon_n}{2d_n},\mathcal{G}_{n,1},u_1^n\right)\cdot \mathcal{N}_1\left(\frac{\epsilon_n}{2d_n\cdot K(0)},\mathcal{G}_{n,2},(u_1^n,x_1^n)\right) \tag{53}$$

for

$$\begin{aligned}
\mathcal{G}_{n,1} &= \left\{g_1:\mathbb{R}\to[0,K(0)]:\ g_1(u)=K\left(\frac{t-u}{h_n}\right)\quad (u\in\mathbb{R}),\ t\in[0,1]\right\},\\
\mathcal{G}_{n,2} &= \left\{g_2:[0,1]\times\mathbb{R}^d\to[0,1]:\ g_2(u,x)=\left(\mathbb{1}_{(-\infty,y]}\circ m\right)(u,x)\ \left((u,x)\in[0,1]\times\mathbb{R}^d\right),\ y\in\mathbb{R}\right\},
\end{aligned}$$

where $\mathbb{1}_{(-\infty,y]}\circ m$ is the composition of the indicator function and the function m. Next, we show

$$\mathcal{N}_1\left(\frac{\epsilon_n}{2\cdot d_n},\mathcal{G}_{n,1},u_1^n\right)\leq 3\cdot\left(\frac{6e\cdot d_n}{\epsilon_n}\right)^8. \tag{54}$$

By Lemma 9.2 und Theorem 9.4 Györfi et al. (2002) we obtain

$$\mathcal{N}_1\left(\frac{\epsilon_n}{2\cdot d_n},\mathcal{G}_{n,1},u_1^n\right)\leq 3\cdot\left(\frac{4e\cdot d_n}{\epsilon_n}\cdot\log\left(\frac{6e\cdot d_n}{\epsilon_n}\right)\right)^{\max\left\{2,V_{\mathcal{G}_{n,1}^+}\right\}}\leq\frac{c_{11}}{2}\cdot\left(\frac{d_n}{\epsilon_n}\right)^{2\cdot\max\left\{2,V_{\mathcal{G}_{n,1}^+}\right\}}$$

for some constant $c_{11}>0$, where $V_{\mathcal{G}_{n,1}^+}$ is the VC-dimension of the class of all subgraphs of $\mathcal{G}_{n,1}$, i.e., of

$$\mathcal{G}_{n,1}^+=\{\{(u,s)\in\mathbb{R}\times\mathbb{R},\ g_1(u)\geq s\}\ :\ \ g_1\in\mathcal{G}_{n,1}\}.$$

Thus, it suffices to bound the VC-dimension of $\mathcal{G}_{n,1}^+$. For this purpose we use the fact that K is left-continuous as well as monotonically decreasing on $\mathbb{R}_+$ and has a compact support, and get for $s>0$

$$K\left(\frac{t-u}{h_n}\right)\geq s\iff\left|\frac{t-u}{h_n}\right|\leq\phi(s)\iff t^2-2ut+u^2-\phi^2(s)\cdot h_n^2\leq 0$$

for $\phi(s)=\sup\{z\in\mathbb{R}:\ K(z)\geq s\}$. Consider the set of functions

$$\begin{aligned}
\tilde{\mathcal{G}}_{n,1}=\{g_{\alpha,\beta,\gamma,\delta}:\mathbb{R}\times\mathbb{R}\to\mathbb{R},\ g_{\alpha,\beta,\gamma,\delta}(u,v) &= \alpha u^2+\beta u+\gamma v^2+\delta,\\
&\quad (u,v)\in\mathbb{R}\times\mathbb{R},\ \alpha,\beta,\gamma,\delta\in\mathbb{R}\}.
\end{aligned}$$

If for a given collection of points $\{(u_i,s_i)\}_{i=1,\dots,n}$, where $s_i>0$ for $i=1,\dots,n$, the set $\{(u,s):g_1(u)\geq s\}$ for $g_1\in\mathcal{G}_{n,1}$ chooses the points $\{(u_{i_1},s_{i_1}),\dots,(u_{i_l},s_{i_l})\}$, i.e.

$$\{(u,s):\ g_1(u)\geq s\}\cap\{(u_i,s_i)\}_{i=1,\dots,n}=\{(u_{i_1},s_{i_1}),\dots,(u_{i_l},s_{i_l})\},$$

then there exist α, β, γ, $\delta\in\mathbb{R}$ such that for $g_{\alpha,\beta,\gamma,\delta}\in\tilde{\mathcal{G}}_n$ the equality

$$\{(u,s):\ g_{\alpha,\beta,\gamma,\delta}(u,s)\geq 0\}\cap\{(u_1,\phi(s_1)),\dots,(u_n,\phi(s_n))\}=\{(u_{i_1},\phi(s_{i_1})),\dots,(u_{i_l},\phi(s_{i_l}))\}$$

holds. Therefore,

$$V_{\mathcal{G}_{n,1}^+} \leq V_{\left\{\{(u,v):\; g_{\alpha,\beta,\gamma,\delta}(u,v)\geq 0\}:\; g\in\tilde{\mathcal{G}}_n\right\}} \leq 4,$$

where we have used Theorem 9.5 from Györfi et al. (2002) in the last inequality. The proof of (54) is complete.

Next, we observe that for

$$\mathcal{G}_{n,3} = \left\{g_3 : \mathbb{R} \to [0,1] : \quad g_3(w) = \mathbb{1}_{(-\infty,y]}(w) \quad (w\in\mathbb{R}), \quad y \in \mathbb{R}\right\}$$

it holds

$$\mathcal{N}_1\left(\frac{\epsilon_n}{2d_n \cdot K(0)}, \mathcal{G}_{n,2}, (u_1^n, x_1^n)\right) = \mathcal{N}_1\left(\frac{\epsilon_n}{2d_n \cdot 2K(0)}, \mathcal{G}_{n,3}, v_1^n\right),$$

where $v_i \in v_1^n$ is defined as $v_i = m(u_i, x_i)$ for $i = 1, \ldots, n$.

Finally, we bound $\mathcal{N}_1\big(\frac{\epsilon_n}{2d_n\cdot K(0)}, \mathcal{G}_{n,3}, v_1^n\big)$ using the n-th shatter coefficient $S(\mathcal{A}, n)$ of the set $\mathcal{A}$. Since $G_{n,3}$ is a set of indicator functions $\mathbb{1}_A$ with $A \in \mathcal{A} = \{(-\infty, y] :\; y \in \mathbb{R}\}$, we have

$$\mathcal{N}_1\left(\frac{\epsilon_n}{2d_n \cdot K(0)}, \mathcal{G}_{n,3}, v_1^n\right) \leq S(\mathcal{A}, n) \leq n+1 \leq 2n$$

for $n \in \mathbb{N}$, where the last two inequalities follow from Theorem 9.3 and Example 9.1 in Györfi et al. (2002). The assertion is implied by (53), (54) and the last result. □

1.2 Proof of inequality (39)

Inequality (39) is implied by

$$\sup_{t\in[0,1]} \int_{\mathbb{R}^d} I_{\{x\notin K_n\}} f(t,x)dx = \sup_{t\in[0,1]} \mathbf{P}(X_t \notin K_n) \leq \mathbf{P}(\exists t\in[0,1] : X_t \notin K_n) \leq c_{32}(\beta_n + \eta_n)$$

for some constant $c_{32} > 0$ and $n \in \mathbb{N}$ sufficiently large, where the last step holds by assumption (16), and by the fact that on C_n we have

$$\begin{aligned}
&\sup_{t\in[0,1]} \int_{\mathbb{R}^d} I_{\{x\in K_n\,:\,\hat{q}_{Y_t,\alpha}-3\beta_n-3\eta_n\leq m_n(t,x)\leq \hat{q}_{Y_t,\alpha}+3\beta_n+3\eta_n\}} \cdot f(t,x)dx \\
&\leq \sup_{t\in[0,1]} \int_{\mathbb{R}^d} I_{\{x\in K_n\,:\,q_{Y_t,\alpha}-4\beta_n-4\eta_n\leq m(t,x)\leq q_{Y_t,\alpha}+4\beta_n+4\eta_n\}} \cdot f(t,x)dx \\
&\leq \sup_{t\in[0,1]} \mathbf{P}\big(q_{Y_t,\alpha} - 4\beta_n - 4\eta_n \leq m(t,X_t) \leq q_{Y_t,\alpha} + 4\beta_n + 4\eta_n\big) \\
&\leq \sup_{t\in[0,1]} \sup_{x\in F_{t,n}} g(t,x) \cdot |8\beta_n + 8\eta_n| \\
&\leq c_{33} \cdot (\beta_n + \eta_n),
\end{aligned}$$

for $F_{t,n} = [q_{Y_t,\alpha} - 4\beta_n - 4\eta_n, q_{Y_t,\alpha} + 4\beta_n + 4\eta_n]$ and some constant $c_{33} > 0$, because of (19). □

1.3 Proof of (50).

Analogously to (48) one can show for any $t \in [0,1]$

$$G_{Y_t}\left(q_{Y_t,\alpha} + \frac{1}{2}\cdot(\beta_n+\eta_n)\right) - \alpha = G_{Y_t}\left(q_{Y_t,\alpha} + \frac{1}{2}\cdot(\beta_n+\eta_n)\right) - G_{Y_t}(q_{Y_t,\alpha})$$
$$> \frac{c_1}{2}\cdot(\beta_n+\eta_n) \tag{55}$$

for some constant $c_1 > 0$ and $n \in \mathbb{N}$ large enough. Using (47) and (55) as well as the assumptions (23) and (24), we get on the event C_n

$$\begin{aligned}
L_n &= \left\{\inf_{t\in[0,1]} \hat{G}^{(IS)}_{Y_t}\left(q_{Y_t,\alpha} + \frac{1}{2}\cdot(\beta_n+\eta_n)\right) \geq \alpha\right\} \\
&\supseteq \left\{\inf_{t\in[0,1]}\left(\hat{G}^{(IS)}_{Y_t}\left(q_{Y_t,\alpha} + \frac{1}{2}\cdot(\beta_n+\eta_n)\right) - \mathbf{E}^\star_{t_1,\dots,t_n}\left\{\hat{G}^{(IS)}_{Y_t}\left(q_{Y_t,\alpha} + \frac{1}{2}\cdot(\beta_n+\eta_n)\right)\right\}\right)\right. \\
&\qquad + \inf_{t\in[0,1]}\left(\mathbf{E}^\star_{t_1,\dots,t_n}\left\{\hat{G}^{(IS)}_{Y_t}\left(q_{Y_t,\alpha} + \frac{1}{2}\cdot(\beta_n+\eta_n)\right)\right\} - G_{Y_t}\left(q_{Y_t,\alpha} + \frac{1}{2}\cdot(\beta_n+\eta_n)\right)\right) \\
&\qquad \left. + \inf_{t\in[0,1]}\left(G_{Y_t}\left(q_{Y_t,\alpha} + \frac{1}{2}\cdot(\beta_n+\eta_n)\right) - \alpha\right) \geq 0\right\} \\
&= \left\{-\sup_{t\in[0,1]}\left(\mathbf{E}^\star_{t_1,\dots,t_n}\left\{\hat{G}^{(IS)}_{Y_t}\left(q_{Y_t,\alpha} + \frac{1}{2}\cdot(\beta_n+\eta_n)\right)\right\} - \hat{G}^{(IS)}_{Y_t}\left(q_{Y_t,\alpha} + \frac{1}{2}\cdot(\beta_n+\eta_n)\right)\right)\right. \\
&\qquad - \sup_{t\in[0,1]}\left(G_{Y_t}\left(q_{Y_t,\alpha} + \frac{1}{2}\cdot(\beta_n+\eta_n)\right) - \mathbf{E}^\star_{t_1,\dots,t_n}\left\{\hat{G}^{(IS)}_{Y_t}\left(q_{Y_t,\alpha} + \frac{1}{2}\cdot(\beta_n+\eta_n)\right)\right\}\right) \\
&\qquad \left. + \inf_{t\in[0,1]}\left(G_{Y_t}\left(q_{Y_t,\alpha} + \frac{1}{2}\cdot(\beta_n+\eta_n)\right) - G_{Y_t}(q_{Y_t,\alpha})\right) \geq 0\right\} \\
&\supseteq \left\{-\sup_{t\in[0,1]}\left(\mathbf{E}^\star_{t_1,\dots,t_n}\left\{\hat{G}^{(IS)}_{Y_t}\left(q_{Y_t,\alpha} + \frac{1}{2}\cdot(\beta_n+\eta_n)\right)\right\} - \hat{G}^{(IS)}_{Y_t}\left(q_{Y_t,\alpha} + \frac{1}{2}\cdot(\beta_n+\eta_n)\right)\right)\right. \\
&\qquad \left. \geq -\frac{c_1}{2}\cdot(\beta_n+\eta_n) + C_2\cdot\beta^p\cdot h^p_{n,1}\right\} \\
&\supseteq \left\{\sup_{t\in[0,1],\ y\in\mathbb{R}}\left(\mathbf{E}^\star_{t_1,\dots,t_n}\left\{\hat{G}^{(IS)}_{Y_t}(y)\right\} - \hat{G}^{(IS)}_{Y_t}(y)\right) \leq \frac{c_1}{2}\cdot(\beta_n+\eta_n) - C_2\cdot\beta^p\cdot h^p_{n,1}\right\} \\
&\supseteq \left\{\sup_{t\in[0,1],\ y\in\mathbb{R}}\left|\mathbf{E}^\star_{t_1,\dots,t_n}\left\{\hat{G}^{(IS)}_{Y_t}(y)\right\} - \hat{G}^{(IS)}_{Y_t}(y)\right| \leq c_{24}\cdot(\beta_n+\eta_n)\cdot\sqrt{\frac{\log(n)}{nh_{n,1}}}\right\}
\end{aligned}$$

for some constant $c_{24} > 1$ and $n \in \mathbb{N}$ sufficiently large. □