A Bias of Nadaraya-Watson Kernel Regression under Smoothness Assumptions

Definition 3. Lipschitz in L_1-normed space.
A function $f : \mathbb{R}^d \to \mathbb{R}$ is said to be Lipschitz continuous in the normed space with Lipschitz constant L if and only if

$$|f(\mathbf{s}_1) - f(\mathbf{s}_2)| \leq |\mathbf{s}_1 - \mathbf{s}_2| L \qquad \forall \mathbf{s}_1, \mathbf{s}_2 \in \mathbb{R}^d.$$

where $|\cdot|$ is the norm in the L_1 space.
Definition 4. Log-Lipschitz in L_1-normed space.
A positive defined function $f : \mathbb{R}^d \to \mathbb{R}$ is said to be log-Lipschitz continuous in the normed space with Lipschitz constant L if and only if

$$|\log f(\mathbf{s}_1) - \log f(\mathbf{s}_2)| \leq |\mathbf{s}_1 - \mathbf{s}_2| L \qquad \forall \mathbf{s}_1, \mathbf{s}_2 \in \mathbb{R}^d.$$

Proposition 1. Log-Lipschitz Distribution
There always exists a Lipschitz continuous function $g : \mathbb{R}^d \to \mathbb{R}$ with constant L as defined in Definition 3 for any distribution with positive-defined Log-Lipschitz density p defined over $\mathbb{R}^d$ with constant L as defined in Definition 4, and

$$p(\mathbf{s}) = \frac{e^{g(\mathbf{s})}}{\int_{-\infty}^{+\infty} e^{g(\mathbf{s})} \, \mathrm{d}\mathbf{s}}.$$

Proof. When $g(\mathbf{s}) = \log p(\mathbf{s})$

$$p(\mathbf{s}) = \frac{e^{\log p(\mathbf{s})}}{\int_{-\infty}^{+\infty} e^{\log p(\mathbf{s})} \, \mathrm{d}\mathbf{s}} = \frac{p(\mathbf{s})}{\int_{-\infty}^{+\infty} p(\mathbf{s}) \, \mathrm{d}\mathbf{s}} = p(\mathbf{s}),$$

and $g(x)$ is Lipschitz with constant L since p is log-Lipschitz

$$|\log p(\mathbf{s}_1) - \log p(\mathbf{s}_2)| \leq |\mathbf{s}_1 - \mathbf{s}_2| L \implies |g(\mathbf{s}_1) - g(\mathbf{s}_2)| \leq |\mathbf{s}_1 - \mathbf{s}_2| L.$$

□

Proposition 2. *Given* $\mathbf{h}, \mathbf{B} \in \mathbb{R}^d$, $L \in \mathbb{R}$

$$\int_0^{\mathbf{B}} \prod_{i=1}^{d} \frac{e^{-\frac{\mathbf{l}_i^2}{2\mathbf{h}_i^2} - \mathbf{l}_i L_\beta}}{\sqrt{2\pi \mathbf{h}_i^2}} \, \mathrm{d}\mathbf{l} = 2^{-d} \prod_{i=1}^{d} e^{\frac{L_\beta^2 h_i^2}{2}} \left(\mathrm{erf}\left(\frac{\mathbf{B}_i + \mathbf{h}_i^2 L_\beta}{\mathbf{h}_i \sqrt{2}} \right) - \mathrm{erf}\left(\frac{\mathbf{h}_i L_\beta}{\sqrt{2}} \right) \right),$$

where erf *is the error-function.*

Proof.

$$\begin{aligned}
\int_0^{\mathbf{B}} \prod_{i=1}^{d} \frac{e^{-\frac{\mathbf{l}_i^2}{2\mathbf{h}_i^2} - \mathbf{l}_i L_\beta}}{\sqrt{2\pi \mathbf{h}_i^2}} \, \mathrm{d}\mathbf{l} &= \prod_{i=1}^{d} \int_0^{\mathbf{B}_i} \frac{e^{-\frac{l_i^2}{2\mathbf{h}_i^2} - l L_\beta}}{\sqrt{2\pi \mathbf{h}_i^2}} \, \mathrm{d}l \qquad \text{(Independent Factorization)} \\
&= \prod_{i=1}^{d} \frac{1}{2} e^{\frac{L_\beta^2 \mathbf{h}_i^2}{2}} \left[1 + \mathrm{erf}\left(\frac{l + \mathbf{h}_i^2 L_\beta}{\mathbf{h}_i \sqrt{2}} \right) \right]_{l=0}^{\mathbf{B}_i} \\
&= 2^{-d} \prod_{i=1}^{d} e^{\frac{L_\beta^2 \mathbf{h}_i^2}{2}} \left(\mathrm{erf}\left(\frac{\mathbf{B}_i + \mathbf{h}_i^2 L_\beta}{\mathbf{h}_i \sqrt{2}} \right) - \mathrm{erf}\left(\frac{\mathbf{h}_i L_\beta}{\sqrt{2}} \right) \right).
\end{aligned}$$

□

Proposition 3.

$$\int_0^B -\frac{e^{-\frac{l^2}{2\mathbf{h}_i^2}-lL_\beta}}{\sqrt{2\pi\mathbf{h}_i^2}} lL_f \,\mathrm{d}l = L_f\mathbf{h}_i \frac{e^{-\frac{B^2}{2\mathbf{h}_i^2}-BL_\beta}-1}{2\pi} + L_f L_\beta \mathbf{h}_i^2 \frac{e^{\frac{L_\beta^2\mathbf{h}_i^2}{2}}}{2}\left(\operatorname{erf}\left(\frac{B+\mathbf{h}_i^2 L_\beta}{\mathbf{h}_i\sqrt{2}}\right)-\operatorname{erf}\left(\frac{\mathbf{h}_i L_\beta}{\sqrt{2}}\right)\right)$$

Proof.

$$\begin{aligned}
\int_0^B -\frac{e^{-\frac{l^2}{2\mathbf{h}_i^2}-lL_\beta}}{\sqrt{2\pi\mathbf{h}_i^2}} lL_f \,\mathrm{d}l &= \frac{L_f\mathbf{h}_i}{\sqrt{2\pi}}\int_0^B -\frac{l}{\mathbf{h}_i^2}e^{-\frac{l^2}{2\mathbf{h}_i^2}-lL_\beta}\,\mathrm{d}l \\
&= \frac{L_f\mathbf{h}_i}{\sqrt{2\pi}}\Bigg(\underbrace{\int_0^B\left(-\frac{l}{\mathbf{h}_i^2}-L_\beta\right)e^{-\frac{l^2}{2\mathbf{h}_i^2}-lL_\beta}\,\mathrm{d}l}_{\text{Integral 1}} + L_\beta\sqrt{2\pi\mathbf{h}_i^2}\underbrace{\int_0^B \frac{e^{-\frac{l^2}{2\mathbf{h}_i^2}-lL_\beta}}{\sqrt{2\pi\mathbf{h}_i^2}}\,\mathrm{d}l}_{\text{Integral 2}}\Bigg)
\end{aligned} \tag{7}$$

The solution of Integral 1 is

$$\int_0^B\left(-\frac{l}{\mathbf{h}_i^2}-L_\beta\right)e^{-\frac{l^2}{2\mathbf{h}_i^2}-lL_\beta}\,\mathrm{d}l = \left[e^{-\frac{l^2}{2\mathbf{h}_i^2}-lL_\beta}\right]_{l=0}^B = e^{-\frac{B^2}{2\mathbf{h}_i^2}-BL_\beta}-1,$$

and based on Proposition 2 the second integral is

$$\int_0^B \frac{e^{-\frac{l^2}{2\mathbf{h}_i^2}-lL_\beta}}{\sqrt{2\pi\mathbf{h}_i^2}}\,\mathrm{d}l = \frac{e^{\frac{L_\beta^2\mathbf{h}_i^2}{2}}}{2}\left(\operatorname{erf}\left(\frac{B+\mathbf{h}_i^2 L_\beta}{\mathbf{h}_i\sqrt{2}}\right)-\operatorname{erf}\left(\frac{\mathbf{h}_i L_\beta}{\sqrt{2}}\right)\right).$$

Plugging back in Equation (7)

$$\begin{aligned}
\int_0^B -\frac{e^{-\frac{l^2}{2\mathbf{h}_i^2}-lL_\beta}}{\sqrt{2\pi\mathbf{h}_i^2}} lL_f \,\mathrm{d}l &= \frac{L_f\mathbf{h}_i}{\sqrt{2\pi}}\left(e^{-\frac{B^2}{2\mathbf{h}_i^2}-BL_\beta}-1+L_\beta\sqrt{2\pi\mathbf{h}_i^2}\frac{e^{\frac{L_\beta^2\mathbf{h}_i^2}{2}}}{2}\left(\operatorname{erf}\left(\frac{B+\mathbf{h}_i^2 L_\beta}{\mathbf{h}_i\sqrt{2}}\right)-\operatorname{erf}\left(\frac{\mathbf{h}_i L_\beta}{\sqrt{2}}\right)\right)\right) \\
&= L_f\mathbf{h}_i\frac{e^{-\frac{B^2}{2\mathbf{h}_i^2}-BL_\beta}-1}{\sqrt{2\pi}} + L_f L_\beta\mathbf{h}_i^2\frac{e^{\frac{L_\beta^2\mathbf{h}_i^2}{2}}}{2}\left(\operatorname{erf}\left(\frac{B+\mathbf{h}_i^2 L_\beta}{\mathbf{h}_i\sqrt{2}}\right)-\operatorname{erf}\left(\frac{\mathbf{h}_i L_\beta}{\sqrt{2}}\right)\right).
\end{aligned}$$

□

Proposition 4. *Given* $\mathbf{A}, \mathbf{B} \in \mathbb{R}^d$*,*

$$\int_{\mathbf{A}}^{\mathbf{B}} \left(\prod_{i=1}^{d} f(\mathbf{s}_i)\right)\left(\sum_{k=1}^{d} g(\mathbf{s}_k)\right) \mathrm{d}\mathbf{s} = \sum_{k=1}^{d}\left(\prod_{i\neq k}^{d} \int_{\mathbf{A}_i}^{\mathbf{B}_i} f(x)\,\mathrm{d}x\right)\int_{\mathbf{A}_k}^{\mathbf{B}_k} f(x)g(x)\,\mathrm{d}x.$$

Proof.

$$\begin{aligned}
\int_{\mathbf{A}}^{\mathbf{B}} \left(\prod_{i=1}^{d} f(\mathbf{s}_i)\right)\left(\sum_{k=1}^{d} g(\mathbf{s}_k)\right) \mathrm{d}\mathbf{s} &= \sum_{k=1}^{d}\int_{\mathbf{A}}^{\mathbf{B}}\left(\prod_{i=1}^{d} f(\mathbf{s}_i)\right) g(\mathbf{s}_k)\,\mathrm{d}\mathbf{s} \\
&= \sum_{k=1}^{d}\int_{\mathbf{A}}^{\mathbf{B}} f(\mathbf{s}_1)f(\mathbf{s}_2)\ldots f(\mathbf{s}_d)g(\mathbf{s}_k)\,\mathrm{d}\mathbf{s} \\
&= \sum_{k=1}^{d}\int_{\mathbf{A}_1}^{\mathbf{B}_1} f(x)\,\mathrm{d}x \int_{\mathbf{A}_2}^{\mathbf{B}_2} f(x)\,\mathrm{d}x \ldots \\
&\qquad \int_{\mathbf{A}_n}^{\mathbf{B}_n} f(x)\,\mathrm{d}x \int_{\mathbf{A}_k}^{\mathbf{B}_k} f(x)g(x)\,\mathrm{d}x \\
&= \sum_{k=1}^{d}\left(\prod_{i\neq k}^{d}\int_{\mathbf{A}_i}^{\mathbf{B}_i} f(x)\,\mathrm{d}x\right)\int_{\mathbf{A}_k}^{\mathbf{B}_k} f(x)g(x)\,\mathrm{d}x
\end{aligned}$$

□

Definition 5. *A multivariate Gaussian Kernel with bandwidth* $\mathbf{h} \in \mathbb{R}$ *is defined as*

$$K(\mathbf{s}, \mathbf{z}) = \prod_{i=1}^{d} \frac{e^{-\frac{(\mathbf{s}_i - \mathbf{z}_i)^2}{2\mathbf{h}_i^2}}}{\sqrt{2\pi \mathbf{h}_i^2}}.$$

Definition 6. *Nadaraya-Watson Kernel Regression. We define the Nadaraya-Watson Kernel Regression with Gaussian Kernels (as defined in Definition 5) of a data-set* $\{\mathbf{s}_i, y_i\}_{i=1}^{n}$ *as*

$$\hat{f}_n(\mathbf{s}) = \frac{\sum_{i=1}^{n} K(\mathbf{s}, \mathbf{s}_i) y_i}{\sum_{j=1}^{n} K(\mathbf{s}, \mathbf{s}_j)}.$$

Theorem 2 *Bias of Nadaraya-Watson Kernel Regression under Smooth Assumptions.*

Proof. Substituting the Definition 5 on the left-side

$$\begin{aligned}
\left|\mathbb{E}\left[\lim_{n\to\infty}\hat{f}_n(\mathbf{s})\right] - f(\mathbf{s})\right| &= \left|\mathbb{E}\left[\lim_{n\to\infty}\frac{\sum_{i=1}^{n} K(\mathbf{s},\mathbf{s}_i)y_i}{\sum_{j=1}^{n} K(\mathbf{s},\mathbf{s}_j)}\right] - f(\mathbf{s})\right| \\
&= \left|\mathbb{E}\left[\lim_{n\to\infty}\frac{n^{-1}\sum_{i=1}^{n} K(\mathbf{s},\mathbf{s}_i)y_i}{n^{-1}\sum_{j=1}^{n} K(\mathbf{s},\mathbf{s}_j)}\right] - f(\mathbf{s})\right| \\
&= \left|\mathbb{E}\left[\frac{\int_{-\infty}^{\infty} K(\mathbf{s},\mathbf{z})\big(f(\mathbf{z}) + \epsilon(\mathbf{z})\big)\beta(\mathbf{z})\,\mathrm{d}\mathbf{z}}{\int_{-\infty}^{+\infty} K(\mathbf{s},\mathbf{z})\beta(\mathbf{z})\,\mathrm{d}\mathbf{z}}\right] - f(\mathbf{s})\right| \\
&= \left|\frac{\int_{-\infty}^{+\infty}\int_{-\infty}^{\infty} K(\mathbf{s},\mathbf{z})\big(f(\mathbf{z}) + \int_{-\infty}^{+\infty}\epsilon_{\mathbf{z}}\,\mathrm{d}\epsilon_{\mathbf{z}}\big)\beta(\mathbf{z})\,\mathrm{d}\mathbf{z}}{\int_{-\infty}^{+\infty} K(\mathbf{s},\mathbf{z})\beta(\mathbf{z})\,\mathrm{d}\mathbf{z}} - f(\mathbf{s})\right| \qquad (8) \\
&= \left|\frac{\int_{-\infty}^{\infty} K(\mathbf{s},\mathbf{z})f(\mathbf{z})\beta(\mathbf{z})\,\mathrm{d}\mathbf{z}}{\int_{-\infty}^{+\infty} K(\mathbf{s},\mathbf{z})\beta(\mathbf{z})\,\mathrm{d}\mathbf{z}} - f(\mathbf{s})\right| \\
&= \left|\frac{\int_{-\infty}^{\infty} K(\mathbf{s},\mathbf{z})\big(f(\mathbf{z}) - f(\mathbf{s})\big)\beta(\mathbf{z})\,\mathrm{d}\mathbf{z}}{\int_{-\infty}^{+\infty} K(\mathbf{s},\mathbf{z})\beta(\mathbf{z})\,\mathrm{d}\mathbf{z}}\right| \qquad (9) \\
&= \frac{\left|\int_{-\infty}^{\infty} K(\mathbf{s},\mathbf{z})\big(f(\mathbf{z}) - f(\mathbf{s})\big)\beta(\mathbf{z})\,\mathrm{d}\mathbf{z}\right|}{\left|\int_{-\infty}^{+\infty} K(\mathbf{s},\mathbf{z})\beta(\mathbf{z})\,\mathrm{d}\mathbf{z}\right|}.
\end{aligned}$$

We want to obtain an upper-bound of the bias. Therefore we want to find an upper-bound of the numerator and a lower-bound of the denominator. Note that the magnitude of the noise does not influence the Bias [36].

Lower-bound of the Denominator: The denominator is always positive, so the absolute value operator can be removed,

$$\begin{aligned}
\int_{-\infty}^{+\infty} K(\mathbf{s},\mathbf{z})\beta(\mathbf{z})\,\mathrm{d}\mathbf{z} &= \int_{-\infty}^{+\infty} \prod_{i=1}^{d} \frac{e^{-\frac{(\mathbf{s}_i-\mathbf{z}_i)^2}{2\mathbf{h}_i^2}}}{\sqrt{2\pi\mathbf{h}_i^2}} \beta(\mathbf{z})\,\mathrm{d}\mathbf{z} \\
&= \frac{1}{\int_{-\infty}^{\infty} e^{g(\mathbf{z})}\,\mathrm{d}\mathbf{z}} \int_{-\infty}^{+\infty} \prod_{i=1}^{d} \frac{e^{-\frac{(\mathbf{s}_i-\mathbf{z}_i)^2}{2\mathbf{h}_i^2}}}{\sqrt{2\pi\mathbf{h}_i^2}} e^{g(\mathbf{z})}\,\mathrm{d}\mathbf{z} \\
&= \frac{e^{g(\mathbf{s})}}{\int_{-\infty}^{\infty} e^{g(\mathbf{z})}\,\mathrm{d}\mathbf{z}} \int_{-\infty}^{+\infty} \left(\prod_{i=1}^{d} \frac{e^{-\frac{(\mathbf{s}_i-\mathbf{z}_i)^2}{2\mathbf{h}_i^2}}}{\sqrt{2\pi\mathbf{h}_i^2}}\right) e^{g(\mathbf{z})-g(\mathbf{s})}\,\mathrm{d}\mathbf{z} \\
&= \frac{e^{g(\mathbf{s})}}{\int_{-\infty}^{\infty} e^{g(\mathbf{z})}\,\mathrm{d}\mathbf{z}} \int_{-\infty}^{+\infty} \left(\prod_{i=1}^{d} \frac{e^{-\frac{\mathbf{l}_i^2}{2\mathbf{h}_i^2}}}{\sqrt{2\pi\mathbf{h}_i^2}}\right) e^{g(\mathbf{s}+\mathbf{l})-g(\mathbf{s})}\,\mathrm{d}\mathbf{l} \qquad (\text{let } l = x - z) \\
&\leq \frac{e^{g(\mathbf{s})}}{\int_{-\infty}^{\infty} e^{g(\mathbf{z})}\,\mathrm{d}\mathbf{z}} \int_{-\infty}^{+\infty} \left(\prod_{i=1}^{d} \frac{e^{-\frac{\mathbf{l}^2}{2\mathbf{h}_i^2}}}{\sqrt{2\pi\mathbf{h}_i^2}}\right) e^{-|\mathbf{l}|L_\beta}\,\mathrm{d}\mathbf{l} \\
&= \frac{e^{g(\mathbf{s})}}{\int_{-\infty}^{\infty} e^{g(\mathbf{z})}\,\mathrm{d}\mathbf{z}} \int_{-\infty}^{+\infty} \left(\prod_{i=1}^{d} \frac{e^{-\frac{\mathbf{l}_i^2}{2\mathbf{h}_i^2}}}{\sqrt{2\pi\mathbf{h}_i^2}}\right)\left(\prod_{j=1}^{d} e^{-|\mathbf{l}_j|L_\beta}\right)\mathrm{d}\mathbf{l} \qquad (\text{Lipschitz Inequality}) \\
&= 2^d \frac{e^{g(\mathbf{s})}}{\int_{0}^{\infty} e^{g(\mathbf{z})}\,\mathrm{d}\mathbf{z}} \int_{0}^{+\infty} \left(\prod_{i=1}^{d} \frac{e^{-\frac{\mathbf{l}_i^2}{2\mathbf{h}_i^2}}}{\sqrt{2\pi\mathbf{h}_i^2}}\right)\left(\prod_{j=1}^{d} e^{-\mathbf{l}_jL_\beta}\right)\mathrm{d}\mathbf{l} \qquad (\text{Symmetric Integral}) \\
&= 2^d \frac{e^{g(\mathbf{s})}}{\int_{0}^{\infty} e^{g(\mathbf{z})}\,\mathrm{d}\mathbf{z}} \int_{0}^{+\infty} \prod_{i=1}^{d} \frac{e^{-\frac{\mathbf{l}_i^2}{2\mathbf{h}_i^2} - \mathbf{l}_iL_\beta}}{\sqrt{2\pi\mathbf{h}_i^2}}\,\mathrm{d}\mathbf{l}.
\end{aligned}$$

Now considering Proposition 2, we obtain

$$\int_{-\infty}^{+\infty} K(\mathbf{s},\mathbf{z})\beta(\mathbf{z})\,\mathrm{d}\mathbf{z} \leq \frac{e^{g(\mathbf{s})}}{\int_{-\infty}^{\infty} e^{g(\mathbf{z})}\,\mathrm{d}\mathbf{z}} \prod_{i=1}^{d} e^{\frac{L_\beta^2 h_i^2}{2}} \left(1 - \mathrm{erf}\left(\frac{\mathbf{h}_iL_\beta}{\sqrt{2}}\right)\right). \tag{10}$$

Upper-bound of the Numerator:

$$
\begin{aligned}
& \Big| \int_{-\infty}^{\infty} K(\mathbf{s},\mathbf{z})\big(f(\mathbf{z}) - f(\mathbf{s})\big)\beta(\mathbf{z})\,\mathrm{d}\mathbf{z} \Big| \\
= & \Big| \int_{-\infty}^{\infty} K(\mathbf{s},\mathbf{s}+\mathbf{l})\big(f(\mathbf{s}+\mathbf{l}) - f(\mathbf{s})\big)\beta(\mathbf{s}+\mathbf{l})\,\mathrm{d}\mathbf{l} \Big| \\
\leq & \int_{-\infty}^{\infty} K(\mathbf{s},\mathbf{s}+\mathbf{l})\big|f(\mathbf{s}+\mathbf{l}) - f(\mathbf{s})\big|\beta(\mathbf{s}+\mathbf{l})\,\mathrm{d}\mathbf{l} \\
= & \frac{1}{\int_{-\infty}^{\infty} e^{g(\mathbf{z})}\,\mathrm{d}\mathbf{z}} \int_{-\infty}^{\infty} K(\mathbf{s},\mathbf{s}+\mathbf{l})\big|f(\mathbf{s}+\mathbf{l}) - f(\mathbf{s})\big| e^{g(\mathbf{s}+\mathbf{l})}\,\mathrm{d}\mathbf{l} \qquad \text{(Jensen's Inequality)} \\
\leq & \frac{1}{\int_{-\infty}^{\infty} e^{g(\mathbf{z})}\,\mathrm{d}\mathbf{z}} \int_{-\infty}^{\infty} K(\mathbf{s},\mathbf{s}+\mathbf{l})\big|\mathbf{l}\big| L_f e^{g(\mathbf{s}+\mathbf{l})}\,\mathrm{d}\mathbf{l} \qquad \text{(Lipschitz Inequality)} \\
= & \frac{1}{\int_{-\infty}^{\infty} e^{g(\mathbf{z})}\,\mathrm{d}\mathbf{z}} \int_{-\infty}^{\infty} K(\mathbf{s},\mathbf{s}+\mathbf{l}) \sum_{i=1}^{n} \big|\mathbf{l}_i\big| L_f e^{g(\mathbf{s}+\mathbf{l})}\,\mathrm{d}\mathbf{l} \\
= & \frac{e^{g(\mathbf{s})}}{\int_{-\infty}^{\infty} e^{g(\mathbf{z})}\,\mathrm{d}\mathbf{z}} \int_{-\infty}^{\infty} K(\mathbf{s},\mathbf{s}+\mathbf{l}) \sum_{i=1}^{n} \big|\mathbf{l}_i\big| L_f e^{g(\mathbf{s}+\mathbf{l}) - g(\mathbf{s})}\,\mathrm{d}\mathbf{l} \\
\leq & \frac{e^{g(\mathbf{s})}}{\int_{-\infty}^{\infty} e^{g(\mathbf{z})}\,\mathrm{d}\mathbf{z}} \int_{-\infty}^{\infty} K(\mathbf{s},\mathbf{s}+\mathbf{l}) \sum_{i=1}^{n} \big|\mathbf{l}_i\big| L_f \prod_{i=1}^{d} e^{|\mathbf{l}_i| L_\beta}\,\mathrm{d}\mathbf{l} \\
= & \frac{e^{g(\mathbf{s})}}{\int_{-\infty}^{\infty} e^{g(\mathbf{z})}\,\mathrm{d}\mathbf{z}} \int_{-\infty}^{\infty} \left(\prod_{i=1}^{d} \frac{e^{-\frac{\mathbf{l}^2}{2\mathbf{h}_i^2}}}{\sqrt{2\pi\mathbf{h}_i^2}} \right) \sum_{i=1}^{n} \big|\mathbf{l}_i\big| L_f \prod_{i=1}^{d} e^{|\mathbf{l}_i| L_\beta}\,\mathrm{d}\mathbf{l} \\
= & 2^d \frac{e^{g(\mathbf{s})}}{\int_{-\infty}^{\infty} e^{g(\mathbf{z})}\,\mathrm{d}\mathbf{z}} \int_{0}^{\infty} \left(\prod_{i=1}^{d} \frac{e^{-\frac{\mathbf{l}^2}{2\mathbf{h}_i^2} + \mathbf{l}_i L_\beta}}{\sqrt{2\pi\mathbf{h}_i^2}} \right) \sum_{i=1}^{n} \mathbf{l}_i L_f\,\mathrm{d}\mathbf{l}.
\end{aligned}
$$

Using Proposition 4,

$$
\begin{aligned}
& 2^d \frac{e^{g(\mathbf{s})}}{\int_{-\infty}^{\infty} e^{g(\mathbf{z})}\,\mathrm{d}\mathbf{z}} \int_{0}^{+\infty} \left(\prod_{i=1}^{d} \frac{e^{-\frac{\mathbf{l}^2}{2\mathbf{h}_i^2} + \mathbf{l}_i L_\beta}}{\sqrt{2\pi\mathbf{h}_i^2}} \right) \sum_{i=1}^{n} \mathbf{l}_i L_f\,\mathrm{d}\mathbf{l} \\
= & 2^d \frac{e^{g(\mathbf{s})}}{\int_{-\infty}^{\infty} e^{g(\mathbf{z})}\,\mathrm{d}\mathbf{z}} \sum_{k=1}^{d} \left(\prod_{i \neq k}^{d} \int_{0}^{+\infty} \frac{e^{-\frac{l^2}{2\mathbf{h}_i^2} + lL_\beta}}{\sqrt{2\pi\mathbf{h}_i^2}}\,\mathrm{d}l \right) \int_{0}^{+\infty} \frac{e^{-\frac{l^2}{2\mathbf{h}_k^2} + lL_\beta}}{\sqrt{2\pi\mathbf{h}_k^2}} lL_f\,\mathrm{d}l \\
= & 2^d \frac{e^{g(\mathbf{s})}}{\int_{-\infty}^{\infty} e^{g(\mathbf{z})}\,\mathrm{d}\mathbf{z}} \sum_{k=1}^{d} \left(\prod_{i \neq k}^{d} \int_{0}^{-\infty} -\frac{e^{-\frac{l^2}{2\mathbf{h}_i^2} - lL_\beta}}{\sqrt{2\pi\mathbf{h}_i^2}}\,\mathrm{d}l \right) \int_{0}^{-\infty} \frac{e^{-\frac{l^2}{2\mathbf{h}_k^2} - lL_\beta}}{\sqrt{2\pi\mathbf{h}_k^2}} lL_f\,\mathrm{d}l.
\end{aligned}
$$

Using Proposition 2,

$$
\begin{aligned}
= & 2^d \frac{e^{g(\mathbf{s})}}{\int_{-\infty}^{\infty} e^{g(\mathbf{z})}\,\mathrm{d}\mathbf{z}} \sum_{k=1}^{d} \left(2^{-d} \prod_{i \neq k}^{d} e^{\frac{L_\beta^2 \mathbf{h}_i^2}{2}} \left(1 + \operatorname{erf}\left(\frac{\mathbf{h}_i L_\beta}{\sqrt{2}} \right) \right) \right) \int_{0}^{-\infty} \frac{e^{-\frac{l^2}{2\mathbf{h}_k^2} - lL_\beta}}{\sqrt{2\pi\mathbf{h}_k^2}} lL_f\,\mathrm{d}l \\
= & \frac{e^{g(\mathbf{s})}}{\int_{-\infty}^{\infty} e^{g(\mathbf{z})}\,\mathrm{d}\mathbf{z}} \sum_{k=1}^{d} \left(\prod_{i \neq k}^{d} e^{\frac{L_\beta^2 \mathbf{h}_i^2}{2}} \left(1 + \operatorname{erf}\left(\frac{\mathbf{h}_i L_\beta}{\sqrt{2}} \right) \right) \right) \int_{0}^{-\infty} \frac{e^{-\frac{l^2}{2\mathbf{h}_k^2} - lL_\beta}}{\sqrt{2\pi\mathbf{h}_k^2}} lL_f\,\mathrm{d}l.
\end{aligned}
$$

Using Proposition 3 to solve the last integral

$$
\begin{aligned}
= \quad & \frac{e^{g(\mathbf{s})}}{\int_{-\infty}^{\infty} e^{g(\mathbf{z})}\,\mathrm{d}\mathbf{z}} \sum_{k=1}^{d} \left(\prod_{i \neq k}^{d} e^{\frac{L_\beta^2 \mathbf{h}_i^2}{2}} \left(1 + \operatorname{erf}\left(\frac{\mathbf{h}_i L_\beta}{\sqrt{2}} \right) \right) \right) \\
& \times L_f \mathbf{h}_k \left(\frac{1}{\sqrt{2\pi}} + L_\beta \mathbf{h}_k \frac{e^{\frac{L_\beta^2 \mathbf{h}_k^2}{2}}}{2} \left(1 + \operatorname{erf}\left(\frac{\mathbf{h}_k L_\beta}{\sqrt{2}} \right) \right) \right).
\end{aligned}
\tag{11}
$$

Conclusion. Unifying equations (11) and (10) we finally obtain

$$
\begin{aligned}
& \left| \mathbb{E}\left[\lim_{n \to \infty} \hat{f}_n(\mathbf{s}) \right] - f(\mathbf{s}) \right| \leq \\
& \frac{L_f \sum_{k=1}^{d} \mathbf{h}_k \left(\prod_{i \neq k}^{d} e^{\frac{L_\beta^2 \mathbf{h}_i^2}{2}} \left(1 + \operatorname{erf}\left(\frac{\mathbf{h}_i L_\beta}{\sqrt{2}} \right) \right) \right) \left(\frac{1}{\sqrt{2\pi}} + L_\beta \mathbf{h}_k \frac{e^{\frac{L_\beta^2 \mathbf{h}_k^2}{2}}}{2} \left(1 + \operatorname{erf}\left(\frac{\mathbf{h}_k L_\beta}{\sqrt{2}} \right) \right) \right)}{\prod_{i=1}^{d} e^{\frac{L_\beta^2 h_i^2}{2}} \left(1 - \operatorname{erf}\left(\frac{\mathbf{h}_i L_\beta}{\sqrt{2}} \right) \right)}.
\end{aligned}
$$

$\square$

B The Nonparametric Bellman Equation

Proofs of Theorem 1 and Theorem 3.

Proposition 5. *In the limit of infinite samples the NPBE defined in Definition 2 with a data-set* $\lim_{n\to\infty} D_n$ *collected under distribution* β *on the state-action space and MDP* $\mathcal{M}$ *converges to*

$$
\begin{aligned}
\hat{V}_\pi(\mathbf{s}) \quad &= \quad \int_{\mathcal{S}\times\mathcal{A}} \varepsilon_\pi(\mathbf{s},\mathbf{z},\mathbf{b})\left(R_{\mathbf{z},\mathbf{b}} + \gamma\int_{\mathcal{S}} \hat{V}_\pi(\mathbf{s}')\phi(\mathbf{s}',\mathbf{z}'_{\mathbf{z},\mathbf{b}})\,\mathrm{d}\mathbf{s}'\right)\beta(\mathbf{z},\mathbf{b})\,\mathrm{d}\mathbf{z}\,\mathrm{d}\mathbf{b}, \\
&\quad\ \textit{with} \quad R_{\mathbf{z},\mathbf{b}} \sim R(\mathbf{z},\mathbf{b}) \quad \forall(\mathbf{z},\mathbf{b})\in\mathcal{S}\times\mathcal{A}, \\
&\quad\ \textit{with} \quad \mathbf{z}'_{\mathbf{z},\mathbf{b}} \sim P(\cdot|\mathbf{z},\mathbf{b}) \quad \forall(\mathbf{z},\mathbf{b})\in\mathcal{S}\times\mathcal{A}.
\end{aligned} \tag{12}
$$

and

$$
\begin{cases}
\varepsilon_\pi(\mathbf{s},\mathbf{z},\mathbf{b}) \coloneqq \int_{\mathcal{A}} \dfrac{\psi(\mathbf{s},\mathbf{z})\varphi(\mathbf{a},\mathbf{b})}{\int_{\mathcal{S},\mathcal{A}}\psi(\mathbf{s},\mathbf{z})\varphi(\mathbf{a},\mathbf{b})\beta(\mathbf{z},\mathbf{b})\,\mathrm{d}\mathbf{z}\,\mathrm{d}\mathbf{b}}\pi(\mathbf{a}|\mathbf{s})\,\mathrm{d}\mathbf{a} & \textit{if } \pi \textit{ is stochastic}, \\[2ex]
\varepsilon_i^\pi(\mathbf{s}) \coloneqq \dfrac{\psi(\mathbf{s},\mathbf{z})\varphi(\pi(\mathbf{s}),\mathbf{b})}{\int_{\mathcal{S},\mathcal{A}}\psi(\mathbf{s},\mathbf{z})\varphi(\pi(\mathbf{s}),\mathbf{b})\beta(\mathbf{z},\mathbf{b})\,\mathrm{d}\mathbf{z}\,\mathrm{d}\mathbf{b}} & \textit{otherwise.}
\end{cases}
$$

Proof.

$$
\begin{aligned}
\hat{V}_\pi(\mathbf{s}) \quad &= \quad \lim_{n\to\infty}\int_{\mathcal{A}} \frac{\sum_{i=1}^n \psi_i(\mathbf{s})\varphi_i(\mathbf{a})\left(r_i + \gamma\int_{\mathcal{S}}\phi_i(\mathbf{s}')\hat{V}_\pi(\mathbf{s}')\,\mathrm{d}\mathbf{s}\right)}{\sum_{i=1}^n \psi_j(\mathbf{s})\varphi_j(\mathbf{a})}\pi(\mathbf{a}|\mathbf{s})\,\mathrm{d}\mathbf{a} \\
&= \quad \int_{\mathcal{A}} \frac{\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^n \psi_i(\mathbf{s})\varphi_i(\mathbf{a})\left(r_i + \gamma\int_{\mathcal{S}}\phi_i(\mathbf{s}')\hat{V}_\pi(\mathbf{s}')\,\mathrm{d}\mathbf{s}\right)}{\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^n \psi_j(\mathbf{s})\varphi_j(\mathbf{a})}\pi(\mathbf{a}|\mathbf{s})\,\mathrm{d}\mathbf{a} \\
&= \quad \int_{\mathcal{A}} \frac{\int_{\mathcal{S}\times\mathcal{A}}\psi(\mathbf{s},\mathbf{z})\varphi(\mathbf{a},\mathbf{b})\left(R(\mathbf{z},\mathbf{b}) + \gamma\int_{\mathcal{S}}\phi(\mathbf{s}',\mathbf{z}')P(\mathbf{z}'|\mathbf{b},\mathbf{z})\hat{V}_\pi(\mathbf{s}')\,\mathrm{d}\mathbf{s}\right)\beta(\mathbf{z},\mathbf{b})\,\mathrm{d}\mathbf{z}\,\mathrm{d}\mathbf{b}}{\int_{\mathcal{S}\times\mathcal{A}}\psi(\mathbf{s},\mathbf{z})\varphi(\mathbf{a},\mathbf{b})\beta(\mathbf{z},\mathbf{b})\,\mathrm{d}\mathbf{z}\,\mathrm{d}\mathbf{b}}\pi(\mathbf{a}|\mathbf{s})\,\mathrm{d}\mathbf{a}.
\end{aligned}
$$

Analogously we can derive the deterministic policy case. □

Proposition 6. *Both for finite samples and infinite samples, when* R *is bounded by* $-R_{max}$ *and* R_{max} *(where* R_{max} *is non-negative defined), then the solution of the NPBE if exists it is bounded between* $\frac{-2R_{max}}{1-\gamma}$ *and* $\frac{2R_{max}}{1-\gamma}$*.*

Proof. Starting with the finite samples case. Suppose by absurd proposition that if the NPBE admits a solution $\hat{V}_\pi$ then $\sup_{\mathbf{s}}|\hat{V}_\pi(\mathbf{s})| = \frac{R_{\max}}{1-\gamma} + \epsilon$ with $\epsilon > 0$ strictly positive (and eventually $+\infty$). It immediately follows that $\sup_{\mathbf{s}_1,\mathbf{s}_2}|\hat{V}_\pi(\mathbf{s}_1) - \hat{V}_\pi(\mathbf{s}_2)| = \frac{2R_{\max}}{1-\gamma} + 2\epsilon$. Expanding this term

$$
\begin{aligned}
\sup_{\mathbf{s}_1,\mathbf{s}_2}|\hat{V}_\pi(\mathbf{s}_1) - \hat{V}_\pi(\mathbf{s}_2)| \quad &= \quad \sup_{\mathbf{s}_1,\mathbf{s}_2}\left|\left(\boldsymbol{\varepsilon}_\pi^T(\mathbf{s}_1) - \boldsymbol{\varepsilon}_\pi^T(\mathbf{s}_2)\right)\left(\mathbf{r} + \gamma\int_{\mathcal{S}}\boldsymbol{\phi}(\mathbf{s}')\hat{V}_\pi(\mathbf{s}')\,\mathrm{d}\mathbf{s}'\right)\right| \\
&\leq \quad \sup_{\mathbf{s}_1,\mathbf{s}_2}\left|\boldsymbol{\varepsilon}_\pi^T(\mathbf{s}_1) - \boldsymbol{\varepsilon}_\pi^T(\mathbf{s}_2)\right|\left|\mathbf{r} + \gamma\int_{\mathcal{S}}\boldsymbol{\phi}(\mathbf{s}')\hat{V}_\pi(\mathbf{s}')\,\mathrm{d}\mathbf{s}'\right| \\
&\leq \quad \sup_{\mathbf{s}_1,\mathbf{s}_2}\left(\left|\boldsymbol{\varepsilon}_\pi^T(\mathbf{s}_1)\right| + \left|\boldsymbol{\varepsilon}_\pi^T(\mathbf{s}_2)\right|\right)\left|\mathbf{r} + \gamma\int_{\mathcal{S}}\boldsymbol{\phi}(\mathbf{s}')\hat{V}_\pi(\mathbf{s}')\,\mathrm{d}\mathbf{s}'\right|.
\end{aligned}
$$

Notice that $\boldsymbol{\varepsilon}_\pi^T(\mathbf{s})$ is a stochastic vector (non-negative definite and sums up to 0),

$$\begin{aligned}
&\sup_{\mathbf{s}_1,\mathbf{s}_2}\left(\left|\boldsymbol{\varepsilon}_\pi^T(\mathbf{s}_1)\right|+\left|\boldsymbol{\varepsilon}_\pi^T(\mathbf{s}_2)\right|\right)\left|\mathbf{r}+\gamma\int_{\mathcal{S}}\boldsymbol{\phi}(\mathbf{s}')\hat{V}_\pi(\mathbf{s}')\,\mathrm{d}\mathbf{s}'\right| \\
\leq\ & 2R_{\max}+\gamma\sup_{\mathbf{s}_1,\mathbf{s}_2}\left(\left|\boldsymbol{\varepsilon}_\pi^T(\mathbf{s}_1)\right|+\left|\boldsymbol{\varepsilon}_\pi^T(\mathbf{s}_2)\right|\right)\left|\int_{\mathcal{S}}\boldsymbol{\phi}(\mathbf{s}')\hat{V}_\pi(\mathbf{s}')\,\mathrm{d}\mathbf{s}'\right| \\
\leq\ & 2R_{\max}+\gamma\left(\frac{R_{\max}}{1-\gamma}+\epsilon\right)\sup_{\mathbf{s}_1,\mathbf{s}_2}\left(\left|\boldsymbol{\varepsilon}_\pi^T(\mathbf{s}_1)\right|+\left|\boldsymbol{\varepsilon}_\pi^T(\mathbf{s}_2)\right|\right)\int_{\mathcal{S}}\boldsymbol{\phi}(\mathbf{s}')\,\mathrm{d}\mathbf{s}' \\
=\ & 2R_{\max}+\gamma\left(\frac{R_{\max}}{1-\gamma}+\epsilon\right)\sup_{\mathbf{s}_1,\mathbf{s}_2}\left(\left|\boldsymbol{\varepsilon}_\pi^T(\mathbf{s}_1)\right|+\left|\boldsymbol{\varepsilon}_\pi^T(\mathbf{s}_2)\right|\right)\mathbf{1} \\
\leq\ & 2R_{\max}+\gamma\left(\frac{2R_{\max}}{1-\gamma}+2\epsilon\right),
\end{aligned}$$

which implies that

$$\begin{aligned}
\sup_{\mathbf{s}_1,\mathbf{s}_2}|\hat{V}_\pi(\mathbf{s}_1)-\hat{V}_\pi(\mathbf{s}_2)| &\leq 2R_{\max}+\gamma\left(\frac{2R_{\max}}{1-\gamma}+2\epsilon\right) \\
\implies 2\frac{R_{\max}}{1-\gamma}+2\epsilon &\leq 2R_{\max}+\gamma\left(\frac{2R_{\max}}{1-\gamma}+2\epsilon\right) \\
\implies 0\leq\epsilon\big(\gamma(1-\gamma)-1\big). &
\end{aligned} \tag{13}$$

Since $\gamma(1-\gamma)-1$ is always negative (we defined $0\leq\gamma<1$), then there are no positive values for ϵ which satisfy the inequality, which is in clear contradiction with the absurd premise. For the infinite samples case we can do similar reasoning noting that ϕ,β,P are probability measures. □

Proposition 7. *If R is bounded by R_{max} and if $f^*:\mathcal{S}\to\mathbb{R}$ satisfies the NPBE, then there is no other function $f:\mathcal{S}\to\mathbb{R}$ for which $\exists\mathbf{z}\in\mathcal{S}$ and $|f^*(\mathbf{z})-f(\mathbf{z})|>0$.*

Proof. Suppose, by absurd assumption, that a function $g:\mathcal{S}\to\mathbb{R}$ exists such that $f(\mathbf{s})+g(\mathbf{s})$ satisfies Equation (12) for every $\mathbf{s}\in\mathcal{S}$ and a function $G\in\mathbb{R}^+$ exists for which $|g(\mathbf{z})|>G$. Note that the existence of $f:\mathcal{S}\to\mathbb{R}$ as a solution for the NPBE implies the existence of

$$\int_{\mathcal{S}}\boldsymbol{\varepsilon}_\pi^T(\mathbf{s})\boldsymbol{\phi}(\mathbf{s}')f^*(\mathbf{s}')\,\mathrm{d}\mathbf{s}'\in\mathbb{R}, \tag{14}$$

and similarly, the existence of $f(\mathbf{s})\in\mathbb{R}$ with $f(\mathbf{s})=f^*(\mathbf{s})+g(\mathbf{s})$ as a solution of the NPBE implies that

$$\int_{\mathcal{S}}\boldsymbol{\varepsilon}_\pi^T(\mathbf{s})\boldsymbol{\phi}(\mathbf{s}')f^*(\mathbf{s}')+g(\mathbf{s}')\,\mathrm{d}\mathbf{s}'\in\mathbb{R}. \tag{15}$$

Note that the existence of the integral in Equations (14) and (15) implies

$$\int_{\mathcal{S}}\boldsymbol{\varepsilon}_\pi^T(\mathbf{s})\boldsymbol{\phi}(\mathbf{s}')g(\mathbf{s}')\,\mathrm{d}\mathbf{s}'\in\mathbb{R}. \tag{16}$$

Note that

$$\begin{aligned}
|f^*(\mathbf{s})-f(\mathbf{s})| &= \left|f^*(\mathbf{s})-\boldsymbol{\varepsilon}_\pi^T(\mathbf{s})\left(\mathbf{r}+\gamma\int_{\mathcal{S}}\boldsymbol{\phi}(\mathbf{s}')\big(f(\mathbf{s}')+g(\mathbf{s}')\big)\,\mathrm{d}\mathbf{s}'\right)\right| \\
&= \left|\boldsymbol{\varepsilon}_\pi^T(\mathbf{s})\left(\mathbf{r}+\gamma\int_{\mathcal{S}}\boldsymbol{\phi}(\mathbf{s}')g(\mathbf{s}')\,\mathrm{d}\mathbf{s}'\right)-\boldsymbol{\varepsilon}_\pi^T(\mathbf{s})\left(\mathbf{r}+\gamma\int_{\mathcal{S}}\boldsymbol{\phi}(\mathbf{s}')\big(f^*(\mathbf{s}')+g(\mathbf{s}')\big)\,\mathrm{d}\mathbf{s}'\right)\right| \\
&= \gamma\left|\boldsymbol{\varepsilon}_\pi^T(\mathbf{s})\int_{\mathcal{S}}\boldsymbol{\phi}(\mathbf{s}')g(\mathbf{s}')\,\mathrm{d}\mathbf{s}'\right| \\
\implies |g(\mathbf{s})| &= \gamma\left|\boldsymbol{\varepsilon}_\pi^T(\mathbf{s})\int_{\mathcal{S}}\boldsymbol{\phi}(\mathbf{s}')g(\mathbf{s}')\,\mathrm{d}\mathbf{s}'\right|.
\end{aligned}$$

Using Jensen's inequality

$$|g(\mathbf{s})| \leq \gamma\boldsymbol{\varepsilon}_\pi^T(\mathbf{s})\int_{\mathcal{S}}\boldsymbol{\phi}(\mathbf{s}')|g(\mathbf{s}')|\,\mathrm{d}\mathbf{s}'.$$

Note that since both f^* and f are bounded by $\frac{R_{\max}}{1-\gamma}$ then $|g(\mathbf{s})| \leq \frac{2R_{\max}}{1-\gamma}$, thus

$$\begin{aligned}
|g(\mathbf{s})| &\leq \gamma \boldsymbol{\varepsilon}_\pi^T(\mathbf{s}) \int_\mathcal{S} \boldsymbol{\phi}(\mathbf{s}')|g(\mathbf{s}')| \,\mathrm{d}\mathbf{s}' \\
&\leq \gamma 2 \frac{R_{\max}}{1-\gamma} \boldsymbol{\varepsilon}_\pi^T(\mathbf{s}) \int_\mathcal{S} \boldsymbol{\phi}(\mathbf{s}') \,\mathrm{d}\mathbf{s}' \\
&= \gamma 2 \frac{R_{\max}}{1-\gamma} \\
\implies |g(\mathbf{s})| &\leq \gamma \frac{2R_{\max}}{1-\gamma} \\
\implies |g(\mathbf{s})| &\leq \gamma^2 \frac{2R_{\max}}{1-\gamma} \qquad \text{using (17)} \\
\implies |g(\mathbf{s})| &\leq \gamma^3 \frac{2R_{\max}}{1-\gamma} \qquad \text{using (17)} \\
&\dots \\
\implies |g(\mathbf{s})| &\leq 0,
\end{aligned} \tag{17}$$

which is in clear disagreement with the assumption made. Again here a similar procedure shows the same result for the infinite case. □

Proof of Theorem 1

Proof. Saying that $\hat{V}_\pi^*$ is a solution for Equation (12) is equivalent to saying

$$\hat{V}_\pi^*(\mathbf{s}) - \boldsymbol{\varepsilon}^\pi(\mathbf{s}) \left(\mathbf{r} + \gamma \int_\mathcal{S} \boldsymbol{\phi}(\mathbf{s}') \hat{V}_\pi^*(\mathbf{s}') \,\mathrm{d}\mathbf{s}' \right) = 0 \qquad \forall \mathbf{s} \in \mathcal{S}.$$

We can verify that by simple algebraic manipulation

$$\begin{aligned}
& \hat{V}_\pi^*(\mathbf{s}) - \boldsymbol{\varepsilon}_\pi^T(\mathbf{s}) \left(\mathbf{r} + \gamma \int_\mathcal{S} \boldsymbol{\phi}(\mathbf{s}') \hat{V}_\pi^*(\mathbf{s}') \,\mathrm{d}\mathbf{s}' \right) \\
=\ & \boldsymbol{\varepsilon}_\pi^T(\mathbf{s}) \boldsymbol{\Lambda}_\pi^{-1} \mathbf{r} - \boldsymbol{\varepsilon}^\pi(\mathbf{s}) \left(\mathbf{r} + \gamma \int_\mathcal{S} \boldsymbol{\phi}(\mathbf{s}') \boldsymbol{\varepsilon}_\pi^T(\mathbf{s}') \boldsymbol{\Lambda}_\pi^{-1} \mathbf{r} \,\mathrm{d}\mathbf{s}' \right) \\
=\ & \boldsymbol{\varepsilon}_\pi^T(\mathbf{s}) \left(\boldsymbol{\Lambda}_\pi^{-1} \mathbf{r} - \mathbf{r} - \gamma \int_\mathcal{S} \boldsymbol{\phi}(\mathbf{s}') \boldsymbol{\varepsilon}_\pi^T(\mathbf{s}') \boldsymbol{\Lambda}_\pi^{-1} \mathbf{r} \,\mathrm{d}\mathbf{s}' \right) \\
=\ & \boldsymbol{\varepsilon}_\pi^T(\mathbf{s}) \left(\left(I - \gamma \int_\mathcal{S} \boldsymbol{\phi}(\mathbf{s}') \boldsymbol{\varepsilon}_\pi^T(\mathbf{s}') \,\mathrm{d}\mathbf{s}' \right) \boldsymbol{\Lambda}_\pi^{-1} \mathbf{r} - \mathbf{r} \right) \\
=\ & \boldsymbol{\varepsilon}_\pi^T(\mathbf{s}) \left(\boldsymbol{\Lambda}_\pi \boldsymbol{\Lambda}_\pi^{-1} \mathbf{r} - \mathbf{r} \right) \\
=\ & 0.
\end{aligned} \tag{18}$$

Since equation (12) has (at least) one solution, Proposition 7 guarantees that the solution ($\hat{V}_\pi^*$) is unique. □

Proof of Theorem 3.

Proof. We perform the derivation for the stochastic policy, however the same derivation applies for the deterministic case almost identically. Expanding $\left| \mathbb{E}_D[\overline{V}_D(\mathbf{s})] - V^*(\mathbf{s}) \right|$ using the NPBE and the classic Bellman equation,

$$\begin{aligned}
\left| \mathop{\mathbb{E}}_D[\overline{V}_D(\mathbf{s})] - V^*(\mathbf{s}) \right| = & \Bigg| \mathop{\mathbb{E}}_D \left[\int_{\mathcal{S}\times\mathcal{A}} \varepsilon_\pi(\mathbf{s}, \mathbf{z}, \mathbf{b}) \left(R_{\mathbf{z},\mathbf{b}} + \gamma \int_\mathcal{S} V_D(\mathbf{s}') \phi(\mathbf{s}', \mathbf{z}'_{\mathbf{z},\mathbf{b}}) \,\mathrm{d}\mathbf{s} \right) \beta(\mathbf{z}, \mathbf{b}) \,\mathrm{d}\mathbf{z} \,\mathrm{d}\mathbf{b} \right] \\
& - \int_\mathcal{A} \left(\overline{R}(\mathbf{s}, \mathbf{a}) + \gamma \int_\mathcal{S} V^*(\mathbf{s}') P(\mathbf{s}'|\mathbf{s}, \mathbf{a}) \,\mathrm{d}\mathbf{s}' \right) \pi(\mathbf{a}|\mathbf{s}) \,\mathrm{d}\mathbf{a} \Bigg|.
\end{aligned} \tag{19}$$

As can be easily verified, $\varepsilon_\pi(\mathbf{s},\mathbf{z},\mathbf{b})\beta(\mathbf{z},\mathbf{b})$ is a density distribution over $\mathbf{z},\mathbf{b}$. Hence Equation (19) can be rewritten

$$\begin{aligned}
&\left|\mathop{\mathbb{E}}_{D}\left[\int_{\mathcal{S}\times\mathcal{A}}\varepsilon_\pi(\mathbf{s},\mathbf{z},\mathbf{b})\Big(R_{\mathbf{z},\mathbf{b}}+\gamma\int_{\mathcal{S}}V_D(\mathbf{s}')\phi(\mathbf{s}',\mathbf{z}'_{\mathbf{z},\mathbf{b}})\,\mathrm{d}\mathbf{s}'\Big)\beta(\mathbf{z},\mathbf{b})\,\mathrm{d}\mathbf{z}\,\mathrm{d}\mathbf{b}\right]\right.\\
&\left.-\int_{\mathcal{A}}\Big(\overline{R}(\mathbf{s},\mathbf{a})+\gamma\int_{\mathcal{S}}V^*(\mathbf{s}')P(\mathbf{s}'|\mathbf{s},\mathbf{a})\,\mathrm{d}\mathbf{s}'\Big)\pi(\mathbf{a}|\mathbf{s})\,\mathrm{d}\mathbf{a}\right|\\
= &\left|\mathop{\mathbb{E}}_{D}\left[\int_{\mathcal{A}}\frac{\int_{\mathcal{S}\times\mathcal{A}}\psi(\mathbf{s},\mathbf{z})\varphi(\mathbf{a},\mathbf{b})\big(R_{\mathbf{z},\mathbf{b}}-\overline{R}(\mathbf{s},\mathbf{a})\big)\beta(\mathbf{z},\mathbf{b})\,\mathrm{d}\mathbf{z}\,\mathrm{d}\mathbf{b}}{\int_{\mathcal{S},\mathcal{A}}\psi(\mathbf{s},\mathbf{z})\varphi(\mathbf{a},\mathbf{b})\beta(\mathbf{z},\mathbf{b})\,\mathrm{d}\mathbf{z}\,\mathrm{d}\mathbf{b}}\pi(\mathbf{a}|\mathbf{s})\,\mathrm{d}\mathbf{a}\right]\right.\\
+ &\left.\gamma\int_{\mathcal{A}}\mathop{\mathbb{E}}_{D}\left[\frac{\int_{\mathcal{S}\times\mathcal{A}}\psi(\mathbf{s},\mathbf{z})\varphi(\mathbf{a},\mathbf{b})\big(\int_{\mathcal{S}}V_D(\mathbf{s}')\phi(\mathbf{s}',\mathbf{z}'_{\mathbf{z},\mathbf{b}})\,\mathrm{d}\mathbf{s}'-\int_{\mathcal{S}}V^*(\mathbf{s}')P(\mathbf{s}'|\mathbf{s},\mathbf{a})\,\mathrm{d}\mathbf{s}'\big)\beta(\mathbf{z},\mathbf{b})\,\mathrm{d}\mathbf{z}\,\mathrm{d}\mathbf{b}}{\int_{\mathcal{S},\mathcal{A}}\psi(\mathbf{s},\mathbf{z})\varphi(\mathbf{a},\mathbf{b})\beta(\mathbf{z},\mathbf{b})\,\mathrm{d}\mathbf{z}\,\mathrm{d}\mathbf{b}}\right]\pi(\mathbf{a}|\mathbf{s})\,\mathrm{d}\mathbf{a}\right|\\
\leq &\left|\mathop{\mathbb{E}}_{D}\left[\int_{\mathcal{A}}\underbrace{\frac{\int_{\mathcal{S}\times\mathcal{A}}\psi(\mathbf{s},\mathbf{z})\varphi(\mathbf{a},\mathbf{b})\big(R_{\mathbf{z},\mathbf{b}}-\overline{R}(\mathbf{s},\mathbf{a})\big)\beta(\mathbf{z},\mathbf{b})\,\mathrm{d}\mathbf{z}\,\mathrm{d}\mathbf{b}}{\int_{\mathcal{S},\mathcal{A}}\psi(\mathbf{s},\mathbf{z})\varphi(\mathbf{a},\mathbf{b})\beta(\mathbf{z},\mathbf{b})\,\mathrm{d}\mathbf{z}\,\mathrm{d}\mathbf{b}}}_{\mathrm{A}}\pi(\mathbf{a}|\mathbf{s})\,\mathrm{d}\mathbf{a}\right]\right|\\
+ &\gamma\left|\underbrace{\int_{\mathcal{A}}\mathop{\mathbb{E}}_{D}\left[\frac{\int_{\mathcal{S}\times\mathcal{A}}\psi(\mathbf{s},\mathbf{z})\varphi(\mathbf{a},\mathbf{b})\big(\int_{\mathcal{S}}V_D(\mathbf{s}')\phi(\mathbf{s}',\mathbf{z}'_{\mathbf{z},\mathbf{b}})\,\mathrm{d}\mathbf{s}'-\int_{\mathcal{S}}V^*(\mathbf{s}')P(\mathbf{s}'|\mathbf{s},\mathbf{a})\,\mathrm{d}\mathbf{s}'\big)\beta(\mathbf{z},\mathbf{b})\,\mathrm{d}\mathbf{z}\,\mathrm{d}\mathbf{b}}{\int_{\mathcal{S},\mathcal{A}}\psi(\mathbf{s},\mathbf{z})\varphi(\mathbf{a},\mathbf{b})\beta(\mathbf{z},\mathbf{b})\,\mathrm{d}\mathbf{z}\,\mathrm{d}\mathbf{b}}\right]}_{\mathrm{B}}\pi(\mathbf{a}|\mathbf{s})\,\mathrm{d}\mathbf{a}\right|\\
\leq &\ \mathrm{A}_{\mathrm{Bias}}+\gamma\mathrm{B}_{\mathrm{Bias}}.
\end{aligned}\tag{20}$$

It is evident that the term A is the Nadaraya-Watson kernel regression from Equation (8) - the noise has been washed out in Equation (9) -, therefore Theorem 2 applies

$$\mathrm{A}_{\mathrm{Bias}} = \frac{L_R\sum_{k=1}^{d}\mathbf{h}_k\left(\prod_{i\neq k}^{d}e^{\frac{L_\beta^2\mathbf{h}_i^2}{2}}\left(1+\mathrm{erf}\left(\frac{\mathbf{h}_iL_\beta}{\sqrt{2}}\right)\right)\right)\left(\frac{1}{\sqrt{2\pi}}+L_\beta\mathbf{h}_k\frac{e^{\frac{L_\beta^2\mathbf{h}_k^2}{2}}}{2}\left(1+\mathrm{erf}\left(\frac{\mathbf{h}_kL_\beta}{\sqrt{2}}\right)\right)\right)}{\prod_{i=1}^{d}e^{\frac{L_\beta^2h_i^2}{2}}\left(1-\mathrm{erf}\left(\frac{\mathbf{h}_iL_\beta}{\sqrt{2}}\right)\right)},$$

where $\mathbf{h}=[\mathbf{h}_\psi,\mathbf{h}_\varphi]$ and $d=d_s+d_a$.
Returning to the estimate of $\mathrm{B}_{\mathrm{Bias}}$

$$\begin{aligned}
&\left|\int_{\mathcal{A}}\mathop{\mathbb{E}}_{D}\left[\frac{\int_{\mathcal{S}\times\mathcal{A}}\psi(\mathbf{s},\mathbf{z})\varphi(\mathbf{a},\mathbf{b})\big(\int_{\mathcal{S}}V_D(\mathbf{s}')\phi(\mathbf{s}',\mathbf{z}'_{\mathbf{z},\mathbf{b}})\,\mathrm{d}\mathbf{s}'-\int_{\mathcal{S}}V^*(\mathbf{s}')P(\mathbf{s}'|\mathbf{s},\mathbf{a})\,\mathrm{d}\mathbf{s}'\big)\beta(\mathbf{z},\mathbf{b})\,\mathrm{d}\mathbf{z}\,\mathrm{d}\mathbf{b}}{\int_{\mathcal{S},\mathcal{A}}\psi(\mathbf{s},\mathbf{z})\varphi(\mathbf{a},\mathbf{b})\beta(\mathbf{z},\mathbf{b})\,\mathrm{d}\mathbf{z}\,\mathrm{d}\mathbf{b}}\right]\pi(\mathbf{a}|\mathbf{s})\,\mathrm{d}\mathbf{a}\right|\\
= &\left|\int_{\mathcal{A}}\frac{\int_{\mathcal{S}\times\mathcal{A}}\psi(\mathbf{s},\mathbf{z})\varphi(\mathbf{a},\mathbf{b})\big(\int_{\mathcal{S}}\mathbb{E}\left[V_D(\mathbf{s}')\phi(\mathbf{s}',\mathbf{z}'_{\mathbf{z},\mathbf{b}})\right]\mathrm{d}\mathbf{s}'-\int_{\mathcal{S}}V^*(\mathbf{s}')P(\mathbf{s}'|\mathbf{s},\mathbf{a})\,\mathrm{d}\mathbf{s}'\big)\beta(\mathbf{z},\mathbf{b})\,\mathrm{d}\mathbf{z}\,\mathrm{d}\mathbf{b}}{\int_{\mathcal{S},\mathcal{A}}\psi(\mathbf{s},\mathbf{z})\varphi(\mathbf{a},\mathbf{b})\beta(\mathbf{z},\mathbf{b})\,\mathrm{d}\mathbf{z}\,\mathrm{d}\mathbf{b}}\pi(\mathbf{a}|\mathbf{s})\,\mathrm{d}\mathbf{a}\right|
\end{aligned}$$

One my ask whether the terms in $\mathbb{E}[V_D(\mathbf{s}')\phi(\mathbf{s}',\mathbf{z}'_{\mathbf{z},\mathbf{b}})]$ are uncorrelated. The answer it is affirmative, since, even if V_D depends by $\mathbf{z}_{\mathbf{z},\mathbf{b}}$ (integral in Equation (12)), this corresponds only in the variation of a single point in the integral, and therefore the overall estimate does not change. This argument, however, does not immediately hold for the case of an infinitesimal bandwidth, and therefore we provide the results for that case separately.

For Finite Bandwidth:

$$
\begin{aligned}
&\left|\int_{\mathcal{A}} \frac{\int_{\mathcal{S}\times\mathcal{A}} \psi(\mathbf{s},\mathbf{z})\varphi(\mathbf{a},\mathbf{b})\big(\int_{\mathcal{S}} \mathbb{E}\left[V_D(\mathbf{s}')\phi(\mathbf{s}',\mathbf{z}'_{\mathbf{z},\mathbf{b}})\right]\mathrm{d}\mathbf{s}' - \int_{\mathcal{S}} V^*(\mathbf{s}')P(\mathbf{s}'|\mathbf{s},\mathbf{a})\,\mathrm{d}\mathbf{s}'\big)\beta(\mathbf{z},\mathbf{b})\,\mathrm{d}\mathbf{z}\,\mathrm{d}\mathbf{b}}{\int_{\mathcal{S},\mathcal{A}} \psi(\mathbf{s},\mathbf{z})\varphi(\mathbf{a},\mathbf{b})\beta(\mathbf{z},\mathbf{b})\,\mathrm{d}\mathbf{z}\,\mathrm{d}\mathbf{b}}\pi(\mathbf{a}|\mathbf{s})\,\mathrm{d}\mathbf{a}\right| \\
\leq\ & \max_{\mathbf{s},\mathbf{a}} \left|\frac{\int_{\mathcal{S}\times\mathcal{A}} \psi(\mathbf{s},\mathbf{z})\varphi(\mathbf{a},\mathbf{b})\big(\int_{\mathcal{S}\times\mathcal{S}} \overline{V}(\mathbf{z}')\phi(\mathbf{z}',\mathbf{s}')P(\mathbf{s}'|\mathbf{s},\mathbf{a})\,\mathrm{d}\mathbf{s}'\,\mathrm{d}\mathbf{z}' - \int_{\mathcal{S}} V^*(\mathbf{s}')P(\mathbf{s}'|\mathbf{s},\mathbf{a})\,\mathrm{d}\mathbf{s}'\big)\beta(\mathbf{z},\mathbf{b})\,\mathrm{d}\mathbf{z}\,\mathrm{d}\mathbf{b}}{\int_{\mathcal{S},\mathcal{A}} \psi(\mathbf{s},\mathbf{z})\varphi(\mathbf{a},\mathbf{b})\beta(\mathbf{z},\mathbf{b})\,\mathrm{d}\mathbf{z}\,\mathrm{d}\mathbf{b}}\right| \\
=\ & \max_{\mathbf{s},\mathbf{a}} \left|\frac{\int_{\mathcal{S}\times\mathcal{A}} \psi(\mathbf{s},\mathbf{z})\varphi(\mathbf{a},\mathbf{b})\big(\int_{\mathcal{S}}\int_{\mathcal{S}} \big(\overline{V}(\mathbf{z}')\phi(\mathbf{z}',\mathbf{s}') - V^*(\mathbf{s}')\big)P(\mathbf{s}'|\mathbf{s},\mathbf{a})\,\mathrm{d}\mathbf{s}'\,\mathrm{d}\mathbf{z}'\big)\beta(\mathbf{z},\mathbf{b})\,\mathrm{d}\mathbf{z}\,\mathrm{d}\mathbf{b}}{\int_{\mathcal{S},\mathcal{A}} \psi(\mathbf{s},\mathbf{z})\varphi(\mathbf{a},\mathbf{b})\beta(\mathbf{z},\mathbf{b})\,\mathrm{d}\mathbf{z}\,\mathrm{d}\mathbf{b}}\right| \\
\leq\ & \max_{\mathbf{s},\mathbf{a},\mathbf{s}'} \left|\frac{\int_{\mathcal{S}\times\mathcal{A}} \psi(\mathbf{s},\mathbf{z})\varphi(\mathbf{a},\mathbf{b})\big(\int_{\mathcal{S}} \overline{V}(\mathbf{z}')\phi(\mathbf{z}',\mathbf{s}') - V^*(\mathbf{s}')\,\mathrm{d}\mathbf{z}'\big)\beta(\mathbf{z},\mathbf{b})\,\mathrm{d}\mathbf{z}\,\mathrm{d}\mathbf{b}}{\int_{\mathcal{S},\mathcal{A}} \psi(\mathbf{s},\mathbf{z})\varphi(\mathbf{a},\mathbf{b})\beta(\mathbf{z},\mathbf{b})\,\mathrm{d}\mathbf{z}\,\mathrm{d}\mathbf{b}}\right| \\
=\ & \max_{\mathbf{s},\mathbf{a},\mathbf{s}'} \left|\frac{\int_{\mathcal{S}\times\mathcal{A}} \psi(\mathbf{s},\mathbf{z})\varphi(\mathbf{a},\mathbf{b})\beta(\mathbf{z},\mathbf{b})\,\mathrm{d}\mathbf{z}\,\mathrm{d}\mathbf{b}}{\int_{\mathcal{S},\mathcal{A}} \psi(\mathbf{s},\mathbf{z})\varphi(\mathbf{a},\mathbf{b})\beta(\mathbf{z},\mathbf{b})\,\mathrm{d}\mathbf{z}\,\mathrm{d}\mathbf{b}}\left(\int_{\mathcal{S}} \overline{V}(\mathbf{z}')\phi(\mathbf{z}',\mathbf{s}') - V^*(\mathbf{s}')\,\mathrm{d}\mathbf{z}'\right)\right| \\
=\ & \max_{\mathbf{s},\mathbf{a},\mathbf{s}'} \left|\int_{\mathcal{S}} \overline{V}(\mathbf{z}')\phi(\mathbf{z}',\mathbf{s}') - V^*(\mathbf{s}')\,\mathrm{d}\mathbf{z}'\right| \\
=\ & \max_{\mathbf{s},\mathbf{a},\mathbf{s}'} \left|\int_{\mathcal{S}} \overline{V}(\mathbf{s}'+\mathbf{l})\phi(\mathbf{s}+\mathbf{l},\mathbf{s}') - V^*(\mathbf{s}')\,\mathrm{d}\mathbf{l}\right|. \qquad (21)
\end{aligned}
$$

Note that

$$
\phi(\mathbf{s}'+\mathbf{l},\mathbf{s}') = \prod_{i=1}^{d_s} \frac{e^{-\frac{l_i^2}{2h_{\phi,i}^2}}}{\sqrt{2\pi h_{\phi,i}^2}},
$$

thus

$$
\begin{aligned}
& \max_{\mathbf{s},\mathbf{a},\mathbf{s}'} \left|\int_{\mathcal{S}} \overline{V}(\mathbf{s}'+\mathbf{l})\phi(\mathbf{s}+\mathbf{l},\mathbf{s}') - V^*(\mathbf{s}')\,\mathrm{d}\mathbf{l}\right| \\
\leq\ & \max_{\mathbf{s},\mathbf{a},\mathbf{s}'} \left|\overline{V}(\mathbf{s}') - V^*(\mathbf{s}')\right| + \int_{\mathcal{S}} L_V\left(\sum_{i=1}^{d_s}|l_i|\right)\prod_{i=1}^{d_s} \frac{e^{-\frac{l_i^2}{2h_{\phi,i}^2}}}{\sqrt{2\pi h_{\phi,i}^2}}\,\mathrm{d}.\mathbf{l}
\end{aligned}
$$

Using Proposition 4

$$
\begin{aligned}
& \max_{\mathbf{s},\mathbf{a},\mathbf{s}'} \left|\overline{V}(\mathbf{s}') - V^*(\mathbf{s}')\right| + L_V\int_{\mathcal{S}} \left(\sum_{i=1}^{d_s}|l_i|\right)\prod_{i=1}^{d_s} \frac{e^{-\frac{l_i^2}{2h_{\phi,i}^2}}}{\sqrt{2\pi h_{\phi,i}^2}}\,\mathrm{d}\mathbf{l} \\
=\ & \max_{\mathbf{s},\mathbf{a},\mathbf{s}'} \left|\overline{V}(\mathbf{s}') - V^*(\mathbf{s}')\right| + L_V\sum_{k=1}^{d_s}\left(\prod_{i\neq k}^{d_s}\int_{-\infty}^{+\infty} \frac{e^{-\frac{l_i^2}{2h_{\phi,i}^2}}}{\sqrt{2\pi h_{\phi,i}^2}}\,\mathrm{d}l_i\right)\int_{-\infty}^{+\infty} |l_k| \frac{e^{-\frac{l_k^2}{2h_{\phi,k}^2}}}{\sqrt{2\pi h_{\phi,k}^2}}\,\mathrm{d}l_k \\
=\ & \max_{\mathbf{s},\mathbf{a},\mathbf{s}'} \left|\overline{V}(\mathbf{s}') - V^*(\mathbf{s}')\right| + L_V 2\sum_{k=1}^{d_s}\int_{0}^{+\infty} l_k \frac{e^{-\frac{l_k^2}{2h_{\phi,k}^2}}}{\sqrt{2\pi h_{\phi,k}^2}}\,\mathrm{d}l_k \\
=\ & \max_{\mathbf{s},\mathbf{a},\mathbf{s}'} \left|\overline{V}(\mathbf{s}') - V^*(\mathbf{s}')\right| + L_V\sum_{k=1}^{d_s}\frac{h_{\phi,k}}{\sqrt{2\pi}} \qquad (22)
\end{aligned}
$$

which means that when $\mathbf{h}$ not infinitesimal

$$
\left|\overline{V}(\mathbf{s}) - V^*(\mathbf{s})\right| \leq \mathrm{A}_{\text{Bias}} + \gamma\left(\max_{\mathbf{s},\mathbf{a},\mathbf{s}'} \left|\overline{V}(\mathbf{s}') - V^*(\mathbf{s}')\right| + L_V\sum_{k=1}^{d_s}\frac{h_{\phi,k}}{\sqrt{2\pi}}\right).
$$

It is however known that $\left|\overline{V}(\mathbf{s}) - V^*(\mathbf{s})\right| \leq 2\frac{R_{\max}}{1-\gamma}$, thus

$$\begin{aligned}
&\left|\overline{V}(\mathbf{s}) - V^*(\mathbf{s})\right| \leq \mathrm{A}_{\text{Bias}} + \gamma\left(\max_{\mathbf{s},\mathbf{a},\mathbf{s}'}\left|\overline{V}(\mathbf{s}') - V^*(\mathbf{s}')\right| + L_V \sum_{k=1}^{d_s} \frac{h_{\phi,k}}{\sqrt{2\pi}}\right) && (23)\\
&\left|\overline{V}(\mathbf{s}) - V^*(\mathbf{s})\right| \leq \mathrm{A}_{\text{Bias}} + \gamma\left(2\frac{R_{\max}}{1-\gamma} + L_V \sum_{k=1}^{d_s} \frac{h_{\phi,k}}{\sqrt{2\pi}}\right)\\
\implies\quad &\left|\overline{V}(\mathbf{s}) - V^*(\mathbf{s})\right| \leq \mathrm{A}_{\text{Bias}} + \gamma\left(\mathrm{A}_{\text{Bias}} + \gamma\left(2\frac{R_{\max}}{1-\gamma} + L_V \sum_{k=1}^{d_s} \frac{h_{\phi,k}}{\sqrt{2\pi}}\right) + L_V \sum_{k=1}^{d_s} \frac{h_{\phi,k}}{\sqrt{2\pi}}\right) \quad \text{using Equation (23)}\\
\implies\quad &\left|\overline{V}(\mathbf{s}) - V^*(\mathbf{s})\right| \leq \sum_{t=0}^{\infty} \gamma^t\left(\mathrm{A}_{\text{Bias}} + \gamma L_V \sum_{k=1}^{d_s} \frac{h_{\phi,k}}{\sqrt{2\pi}}\right) \quad \text{using Equation (23)}\\
\implies\quad &\left|\overline{V}(\mathbf{s}) - V^*(\mathbf{s})\right| \leq \frac{1}{1-\gamma}\left(\mathrm{A}_{\text{Bias}} + \gamma L_V \sum_{k=1}^{d_s} \frac{h_{\phi,k}}{\sqrt{2\pi}}\right) \quad \text{using Equation (23).}
\end{aligned} \tag{24}$$

For Infinitesimal Bandwidth: In the case of an infinitesimal bandwidth note that, even if V_D and ϕ are correlated the overall integral reduces only on a single point, and the same argument made in the case of finite bandwidth applies,

$$\int_{\mathcal{S}} \mathbb{E}\left[V_D(\mathbf{s}')\phi(\mathbf{s}', \mathbf{z}'_{\mathbf{z},\mathbf{b}})\right] \mathrm{d}\mathbf{s}' = \mathbb{E}\left[\int_{\mathcal{S}} V_D(\mathbf{s}')\phi(\mathbf{s}', \mathbf{z}'_{\mathbf{z},\mathbf{b}}\, \mathrm{d}\mathbf{s}')\, \mathrm{d}\mathbf{s}'\right] = \mathbb{E}\left[V_D(\mathbf{z}'_{\mathbf{z},\mathbf{b}})\right] = \int_{\mathcal{S}} \overline{V}_D(\mathbf{s}')P(\mathbf{s}'|\mathbf{s},\mathbf{a})\, \mathrm{d}\mathbf{s}'.$$

It follows that, proceeding similarly to Equation (21), we obtain

$$\left|\mathop{\mathbb{E}}_{D}[\overline{V}_D(\mathbf{s})] - V^*(\mathbf{s})\right| \leq \max_{\mathbf{s},\mathbf{a},\mathbf{s}'}\left|\overline{V}(\mathbf{s}') - V^*(\mathbf{s}')\right|, \tag{25}$$

which yields

$$\left|\overline{V}(\mathbf{s}) - V^*(\mathbf{s})\right| \leq \frac{1}{1-\gamma}\mathrm{A}_{\text{Bias}}. \tag{26}$$

□