Online Supplement: A segment-wise dynamic programming algorithm for BSDEs

Christian Bender and Steffen Meyer

July 10, 2023

This online complement contains the proofs of Theorems 4.1 and 4.2.

1 Proof of Theorem 4.2

We first show that

$$\max_{0\le i\le N-1} E\left[|q_i^{N,M}(X_{t_i}) - \overline{y}(t_i, X_{t_i})|^2\right] + \sum_{i=0}^{N-1} E\left[\int_{t_i}^{t_{i+1}} |z_i^{N,M}(X_{t_i}) - \overline{z}(s, X_s)|^2 ds\right]$$
$$\le c \max_{i\in\mathfrak{I}} \left(N^{1-\alpha} \inf_{\psi\in\mathcal{K}_{q,i}} E\left[|\psi(X_{t_i}) - \overline{q}_i^N(X_{t_i})|^2\right] + N^{2-2\alpha}\frac{K_{q,i}}{M_i} + N^{2-2\alpha}\frac{K_{q,i}\log(M_i)}{M_i}\right)$$
$$+ c \max_{0\le i\le N-1}\left(\inf_{\psi\in\mathcal{K}_{q,i}} E\left[|\psi(X_{t_i}) - \overline{q}_i^N(X_{t_i})|^2\right] + \inf_{\psi\in\mathcal{K}_{z,i}} E\left[|\psi(X_{t_i}) - \overline{z}_i^N(X_{t_i})|^2\right]\right.$$
$$\left. + \frac{K_{q,i}}{M_i} + N\frac{K_{z,i}}{M_i} + \frac{K_{q,i}\log(M_i)}{M_i} + N\frac{K_{z,i}\log(M_i)}{M_i}\right)$$
$$+ cN^{-1}$$

under the standing assumptions. Since

$$\max_{0\le i\le N-1} E\left[\left|q_i^{N,M}(X_{t_i}) - \overline{y}(t_i, X_{t_i})\right|^2\right] + \sum_{i=0}^{N-1} E\left[\int_{t_i}^{t_{i+1}} \left|z_i^{N,M}(X_{t_i}) - \overline{z}(s, X_s)\right|^2 ds\right]$$
$$\le 2\Bigg(\max_{0\le i\le N-1} E\left[\left|q_i^{N,M}(X_{t_i}) - \overline{q}_i^N(X_{t_i})\right|^2\right] + \max_{0\le i\le N-1} E\left[\left|\overline{q}_i^N(X_{t_i}) - \overline{y}(t_i, X_{t_i})\right|^2\right]$$
$$+ \sum_{i=0}^{N-1} \Delta E\left[\left|z_i^{N,M}(X_{t_i}) - \overline{z}_i^N(X_{t_i})\right|^2 ds\right] + \sum_{i=0}^{N-1} E\left[\int_{t_i}^{t_{i+1}} \left|\overline{z}_i^N(X_{t_i}) - \overline{z}(s, X_s)\right|^2 ds\right]\Bigg),$$

this follows directly from Theorem 4.1 if we can prove the bounds

$$\max_{0\le i\le N-1} E\left[\left|\overline{q}_i^N(X_{t_i}) - \overline{y}(t_i, X_{t_i})\right|^2\right] + \sum_{i=0}^{N-1} E\left[\int_{t_i}^{t_{i+1}} \left|\overline{z}_i^N(X_{t_i}) - \overline{z}(s, X_s)\right|^2 ds\right] \le c\Delta$$

and $\mathcal{R}^N \leq c\Delta^2$. We start with the bound for $\mathcal{R}^N$ and use Hölder's inequality to get

$$\begin{aligned}\mathcal{R}^N &= \sum_{i=0}^{N-2} E\left[\left(\int_{t_{i+1}}^{t_{i+2}} E_i\left[f\left(s, X_s, Y_s, Z_s\right) - f\left(t_{i+1}, X_{t_{i+1}}, \overline{q}_{i+1}^N, \overline{z}_{i+1}^N\right)\right] ds\right)^2\right] \\ &\leq \sum_{i=0}^{N-2} \Delta E\left[\int_{t_{i+1}}^{t_{i+2}} E_i\left[\left(f\left(s, X_s, Y_s, Z_s\right) - f\left(t_{i+1}, X_{t_{i+1}}, \overline{q}_{i+1}^N, \overline{z}_{i+1}^N\right)\right)^2\right] ds\right].\end{aligned}$$

Then, due to the Lipschitz continuity (respectively Hölder continuity in t) of f, it holds

$$\begin{aligned}\mathcal{R}^N &\leq \sum_{i=0}^{N-2} \Delta E\left[\int_{t_{i+1}}^{t_{i+2}} E_i\left[\left(f\left(s, X_s, Y_s, Z_s\right) - f\left(t_{i+1}, X_{t_{i+1}}, \overline{q}_{i+1}^N, \overline{z}_{i+1}^N\right)\right)^2\right] ds\right] \\ &\leq \sum_{i=0}^{N-2} \Delta E\Bigg[\int_{t_{i+1}}^{t_{i+2}} E_i\Bigg[L_f^2\Big(|s - t_{i+1}|^{\frac{1}{2}} + |X_s - X_{t_{i+1}}| + |Y_s - \overline{q}_{i+1}^N(X_{t_{i+1}})| \\ &\quad + |Z_s - \overline{z}_{i+1}^N(X_{t_{i+1}})|\Big)^2\Bigg] ds\Bigg] \\ &\leq \sum_{i=0}^{N-2} 4L_f^2 \Delta E\Bigg[\int_{t_{i+1}}^{t_{i+2}} |s - t_{i+1}| + E_i\left[(X_s - X_{t_{i+1}})^2\right] + E_i\left[(Y_s - \overline{q}_{i+1}^N(X_{t_{i+1}}))^2\right] \\ &\quad + E_i\left[(Z_s - \overline{z}_{i+1}^N(X_{t_{i+1}}))^2\right] ds\Bigg]\end{aligned}$$

and we consider the terms in the integrand separately for an arbitrary $s \in [t_{i+1}, t_{i+2}]$:
By choice of the time grid, it obviously holds that $|s - t_{i+1}| \leq \Delta$ and, under the assumptions on b and σ, it follows that $E_i[(X_s - X_{t_{i+1}})^2] \leq c(s - t_{i+1}) \leq c\Delta$ (see e.g. Kloeden and Platen, 1992). Then, by the definition of $\overline{q}_{i+1}^N$ and a zero addition, we get:

$$\begin{aligned}&E_i\left[\left(Y_s - \overline{q}_{i+1}^N(X_{i+1})\right)^2\right] \\ &= E_i\left[\left(Y_s - Y_{t_{i+1}} + Y_{t_{i+1}} + E_{i+1}\left[Y_{t_{i+2}}\right]\right)^2\right] \\ &\leq 4 \max_{0 \leq i \leq N-1} \sup_{s \in [t_{i+1}, t_{i+2}]} E_i\left[\left(Y_s - Y_{t_{i+1}}\right)^2\right].\end{aligned} \tag{1}$$

To estimate the difference $Z_s - \overline{z}_{i+1}^N(X_{t_{i+1}})$, we define for each $i \in \{0, \ldots, N-1\}$ the random variable

$$\tilde{Z}_i := \frac{1}{\Delta} E_i\left[\int_{t_i}^{t_{i+1}} Z_s ds\right],$$

which can be used to express the quadratic difference as

$$
\begin{aligned}
&E_i\left[\left(Z_s-\overline{z}^N_{i+1}(X_{t_{i+1}})\right)^2\right]\\
&=E_i\left[\left(Z_s-E_{i+1}\left[\frac{\Delta W_{i+2}}{\Delta}Y_{t_{i+2}}\right]\right)^2\right]\\
&=E_i\left[\left(Z_s-E_{i+1}\left[\frac{\Delta W_{i+2}}{\Delta}\left(Y_{t_{i+1}}-\int_{t_{i+1}}^{t_{i+2}}f(l,X_l,Y_l,Z_l)dl+\int_{t_{i+1}}^{t_{i+2}}Z_l dW_l\right)\right]\right)^2\right]\\
&=E_i\left[\left(Z_s-\frac{1}{\Delta}E_{i+1}\left[\int_{t_{i+1}}^{t_{i+2}}Z_l dl\right]+\frac{1}{\Delta}E_i\left[\Delta W_{i+2}\int_{t_{i+1}}^{t_{i+2}}f(l,X_l,Y_l,Z_l)dl\right]\right)^2\right]\\
&\le 2E_i\left[\left|Z_s-\tilde{Z}_{i+1}\right|^2\right]+2E_i\left[\left(\frac{1}{\Delta}E_{i+1}\left[\Delta W_{i+2}^2\right]^{\frac{1}{2}}E_{i+1}\left[(C_f\Delta)^2\right]^{\frac{1}{2}}\right)^2\right]\\
&\le 2E_i\left[\left|Z_s-\tilde{Z}_{i+1}\right|^2\right]+c\Delta.
\end{aligned}
\tag{2}
$$

Here, the second equality follows by the Itô-isometry and the measurablility of $Y_{t_{i+1}}$, the following inequality due to the boundedness of f and Hölder's inequality. Plugging in the obtained bounds we have

$$
\begin{aligned}
\mathcal{R}^N &\le \sum_{i=0}^{N-2}4L_f^2\Delta E\Bigg[\int_{t_{i+1}}^{t_{i+2}}|s-t_{i+1}|+E_i\left[\left(X_s-X_{t_{i+1}}\right)^2\right]+E_i\left[\left(Y_s-\overline{q}^N_{i+1}(X_{t_{i+1}})\right)^2\right]\\
&\quad+E_i\left[\left(Z_s-\overline{z}_{i+1}(X_{t_{i+1}})\right)^2\right]ds\Bigg]\\
&\le \sum_{i=0}^{N-2}4L_f^2\Delta E\Bigg[\int_{t_{i+1}}^{t_{i+2}}\Delta+c\Delta+4\max_{0\le j\le N-1}\sup_{l\in[t_{j+1},t_{j+2}]}E\left[\left(Y_l-Y_{t_{j+1}}\right)^2\right]\\
&\quad+2E_i\left[\left(Z_s-\tilde{Z}_{i+1}\right)^2\right]+c\Delta ds\Bigg]\\
&\le \sum_{i=0}^{N-2}4L_f^2\Delta E\Bigg[\Delta\left(c\Delta+4\max_{0\le j\le N-1}\sup_{l\in[t_{j+1},t_{j+2}]}E\left[\left(Y_l-Y_{t_{j+1}}\right)^2\right]\right)\\
&\quad+2\int_{t_{i+1}}^{t_{i+2}}E_i\left[\left(Z_s-\tilde{Z}_{i+1}\right)^2\right]ds\Bigg]\\
&\le T4L_f^2\left(c\Delta^2+\Delta 4\max_{0\le j\le N-1}\sup_{l\in[t_{j+1},t_{j+2}]}E\left[\left(Y_l-Y_{t_{j+1}}\right)^2\right]\right)\\
&\quad+8L_f^2\Delta\sum_{i=0}^{N-2}\int_{t_{i+1}}^{t_{i+2}}E\left[\left(Z_s-\tilde{Z}_{i+1}\right)^2\right]ds.
\end{aligned}
$$

Then, using the the L^2-regularity of BSDEs (see Zhang, 2001), which states

$$\max_{0\leq i\leq N}\sup_{t_i\leq s\leq t_{i+1}} E\left[\left|Y_s - Y_{t_i}\right|^2\right] + \sum_{i=0}^{N-1} E\left[\int_{t_i}^{t_{i+1}} \left|Z_s - \tilde{Z}_i\right|^2 ds\right] \leq c\Delta,$$

it follows

$$\begin{aligned}\mathcal{R}^N &\leq T4L_f^2\left(c\Delta^2 + 4\Delta \max_{0\leq j\leq N-1}\sup_{l\in[t_{j+1},t_{j+2}]} E\left[\left(Y_l - Y_{t_{j+1}}\right)^2\right]\right)\\ &\quad + 8L_f^2\Delta\sum_{i=0}^{N-2}\int_{t_{i+1}}^{t_{i+2}} E\left[\left(Z_s - \tilde{Z}_{i+1}\right)^2\right] ds\\ &\leq c\Delta^2,\end{aligned}$$

what proves the bound for $\mathcal{R}^N$. Note that the inequalities (2) and (1) together with the L^2 regularity of BSDEs (see Zhang, 2001) in particular also imply that

$$\begin{aligned}&\max_{0\leq i\leq N-1} E\left[\left|\overline{q}_i^N(X_{t_i}) - Y_{t_i}\right|^2\right] + \sum_{i=0}^{N-1} E\left[\int_{t_i}^{t_{i+1}} \left|\overline{z}_i^N(X_{t_i}) - Z_s\right|^2 ds\right]\\ &\leq 4\max_{0\leq i\leq N-1}\sup_{l\in[t_{j+1},t_{j+2}]} E\left[\left|Y_l - Y_{t_i}\right|^2\right]\\ &\quad + \sum_{i=0}^{N-1} 2E\left[\int_{t_i}^{t_{i+1}} \left|Z_s - \tilde{Z}_i(X_{t_i})\right|^2 + c\Delta ds\right]\\ &\leq c\Delta,\end{aligned}$$

which shows the other bound.

The final step of the proof of Theorem 4.2 is to obtain the modified representation of the regression error. To this end note that

$$N^{1-\alpha}E\left[|\psi(X_{t_i}) - \overline{q}_i^N(X_{t_i})|^2\right] \leq 2N^{1-\alpha}\left(E\left[|\psi(X_{t_i}) - \overline{y}(t_i, X_{t_i})|^2\right] + E\left[|\overline{y}(t_i, X_{t_i}) - \overline{q}_i^N(X_{t_i})|^2\right]\right)$$

and analogously

$$E\left[|\psi(X_{t_i}) - \overline{z}_i^N(X_{t_i})|^2\right] \leq 2\left(E\left[|\psi(X_{t_i}) - \overline{z}(t_i, X_{t_i})|^2\right] + E\left[|\overline{z}(t_i, X_{t_i}) - \overline{z}_i^N(X_{t_i})|^2\right]\right)$$

for all $\psi \in \mathcal{K}_{q,i}$ or $\psi \in \mathcal{K}_{z,i}$ respectively. Hence it suffices to show that the bounds

$$E\left[|q_i^N(X_{t_i}) - \overline{y}(t_i, X_{t_i})|^2\right] \leq \Delta^2, \quad E\left[|\overline{z}_i^N(X_{t_i}) - \overline{z}(t_i, X_{t_i})|^2\right] \leq \Delta$$

hold true for each $i \in \{0, \ldots, N-1\}$, whenever z is Lipschitz continuous in x and $\frac{1}{2}$-Hölder continuous in t. For the bound regarding y, we directly get by the definition of $\overline{q}_i^N$ and the boundedness assumption on f that

$$E\left[\left|q_i^N(X_{t_i}) - \overline{y}(t_i, X_{t_i})\right|^2\right]$$

$$
\begin{aligned}
&= E\left[\left|E_i\left[\overline{y}(t_i, X_{t_i}) - \int_{t_i}^{t_{i+1}} f(t, X_t, Y_t, Z_t)dt + \int_{t_i}^{t_{i+1}} Z_t dW_t\right] - \overline{y}(t_i, X_{t_i})\right|^2\right] \\
&\leq E\left[\left|\int_{t_i}^{t_{i+1}} C_f dt\right|^2\right] \\
&\leq c\Delta^2.
\end{aligned}
$$

For the bound concerning $\overline{z}$, we get by inequality (2)

$$
\begin{aligned}
E\left[\left|\overline{z}_i^N(X_{t_i}) - \overline{z}(t_i, X_{t_i})\right|^2\right] &\leq 2E\left[\left|\overline{z}(t_i, X_{t_i}) - \tilde{Z}_{i+1}\right|^2\right] + c\Delta \\
&= 2E\left[\left|\frac{1}{\Delta}E_i\left[\int_{t_i}^{t_{i+1}} \overline{z}(t_i, X_{t_i}) - \overline{z}(l, X_l)dl\right]\right|^2\right] + c\Delta \\
&\leq \frac{2}{\Delta^2}E\left[E_i\left[\int_{t_i}^{t_{i+1}} \left|\overline{z}(t_i, X_{t_i}) - \overline{z}(l, X_l)\right| dl\right]^2\right] + c\Delta \\
&\leq \frac{2}{\Delta}\int_{t_i}^{t_{i+1}} E\left[E_i\left[L_z(|t_i - l|^{\frac{1}{2}} + |X_{t_i} - X_l|)dl\right]^2\right] + c\Delta \\
&\leq \frac{2}{\Delta}\int_{t_i}^{t_{i+1}} E\left[c\Delta^{\frac{1}{2}}dl\right]^2 + c\Delta \\
&\leq c\Delta
\end{aligned}
$$

where we used Hölder's inequality in the first step, the continuity assumptions on $\overline{z}$ along with Fubini's theorem in the second inequality and denote the Lipschitz constant of $\overline{z}$ with L_z.

2 Proof of Theorem 4.3

Similarly to the proof of Theorem 4.2, we have

$$
\begin{aligned}
&\max_{0\leq i\leq N-1} E\left[\left|q_i^{N,M}(W_{t_i}) - \overline{y}(t_i, W_{t_i})\right|^2\right] + \sum_{i=0}^{N-1} \Delta E\left[\left|\overline{z}(t_i, W_{t_i}) - z_i^{N,M}(W_{t_i})\right|^2 ds\right] \\
&\leq 2\left(\max_{0\leq i\leq N-1} E\left[\left|q_i^{N,M}(W_{t_i}) - \overline{q}_i^N(W_{t_i})\right|^2\right] + \max_{0\leq i\leq N-1} E\left[\left|\overline{q}_i^N(W_{t_i}) - \overline{y}(t_i, W_{t_i})\right|^2\right]\right. \\
&\quad \left. + \sum_{i=0}^{N-1} \Delta E\left[\left|\overline{z}(t_i, W_{t_i}) - \overline{z}_i^N(W_{t_i})\right|^2 ds\right] + \sum_{i=0}^{N-1} \Delta E\left[\left|z_i^{N,M}(W_s) - \overline{z}_i^N(W_{t_i})\right|^2 ds\right]\right)
\end{aligned}
$$

and it suffices to prove the bounds

$$
\max_{0\leq i\leq N-1} E\left[\left|\overline{q}_i^N(W_{t_i}) - \overline{y}(t_i, W_{t_i})\right|^2\right] + \sum_{i=0}^{N-1} \Delta E\left[\left|\overline{z}(t_i, W_{t_i}) - \overline{z}_i^N(W_{t_i})\right|^2\right] \leq c\Delta^2 \qquad (3)
$$

and $\mathcal{R}^N \leq c\Delta^3$ for the first statement of Theorem 4.3. We focus on the bound on $\mathcal{R}^N$ and derive the bounds in (3) along the way. It suffices to show that

$$\left| E_i \left[f\left(s, W_s, Y_s, Z_s\right) - f\left(t_i, W_{t_i}, \overline{q}_i^N, \overline{z}_i^N\right)\right] \right| \leq c\Delta \tag{4}$$

for any $t_i \in \pi$ and $s \in [t_i, t_{i+1}]$, since then

$$\begin{aligned}
\mathcal{R}^N &= \sum_{i=0}^{N-1} E\left[\left(\int_{t_i}^{t_{i+1}} E_i \left[f\left(s, W_s, Y_s, Z_s\right) - f\left(t_i, W_{t_i}, \overline{q}_i^N, \overline{z}_i^N\right)\right] ds \right)^2 \right] \\
&\leq \sum_{i=0}^{N-1} E\left[\left(\int_{t_i}^{t_{i+1}} c\Delta ds \right)^2 \right] \\
&\leq c\Delta^3.
\end{aligned}$$

In order to prove (4), we set for arbitrary but fixed $t_i \in \pi$ and $s \in [t_i, t_{i+1}]$

$$\begin{aligned}
a &:= \left(t_i, W_{t_i}^{(1)}, \dots, W_{t_i}^{(\mathcal{D})}, \overline{q}_i^N(W_{t_i}), \overline{z}_i^{N,(1)}(W_{t_i}), \dots, \overline{z}_i^{N,(\mathcal{D})}(W_{t_i})\right)^T \\
\tilde{a} &:= \left(s, W_s^{(1)}, \dots, W_s^{(\mathcal{D})}, Y_s, Z_s^{(1)}, \dots, Z_s^{(\mathcal{D})}\right)^T.
\end{aligned}$$

Then, a Taylor expansion of f yields

$$\begin{aligned}
&E_i \left[f\left(s, W_s, Y_s, Z_s\right) - f\left(t_i, W_{t_i}, \overline{q}_i^N, \overline{z}_i^N\right)\right] \\
&= E_i \left[\nabla f(a)^T (a - \tilde{a}) + \frac{1}{2} \int_0^1 (1 - \Theta)(a - \tilde{a})^T \operatorname{Hess}_f(a + \Theta(\tilde{a} - a))(a - \tilde{a}) d\Theta \right],
\end{aligned}$$

where ∇f denotes the gradient and Hess_f the Hessian matrix of f. Using that f has bounded derivatives and a is $\mathcal{F}_i$-measurable we obtain

$$\begin{aligned}
&E_i \left[f\left(s, W_s, Y_s, Z_s\right) - f\left(t_i, W_{t_i}, \overline{q}_i^N, \overline{z}_i^N\right)\right] \\
&\leq \nabla f(a)^T E_i \left[(a - \tilde{a})\right] + \frac{1}{2} \sup_{\Theta \in [0,1]} \left| E_i \left[(a - \tilde{a})^T H_f(a + \Theta(\tilde{a} - a))(a - \tilde{a}) \right] \right| \\
&\leq C_f \sum_{l=1}^{2\mathcal{D}+2} \left| E_i \left[a^{(l)} - \tilde{a}^{(l)} \right] \right| + \frac{1}{2} C_f E_i \left[\sum_{l,k=1}^{2\mathcal{D}+2} \left| a^{(l)} - \tilde{a}^{(l)} \right| \left| a^{(k)} - \tilde{a}^{(k)} \right| \right] \\
&\leq C_f \sum_{l=1}^{2\mathcal{D}+2} \left| E_i \left[a^{(l)} - \tilde{a}^{(l)} \right] \right| + \frac{1}{2} C_f \sum_{l,k=1}^{2\mathcal{D}+2} E_i \left[\left| a^{(l)} - \tilde{a}^{(l)} \right|^2 \right]^{\frac{1}{2}} E_i \left[\left| a^{(k)} - \tilde{a}^{(k)} \right|^2 \right]^{\frac{1}{2}}
\end{aligned}$$

and it suffices to show that it holds $|E_i[(a^{(l)} - \tilde{a}^{(l)})^p]| \leq c\Delta$ for $l \in \{1, \dots, 2\mathcal{D}+2\}$ and $p \in \{1, 2\}$. This is trivial for $l = 1, \dots, \mathcal{D}+1$, since W is a Brownian motion and the step width of the time grid is Δ. For the remaining values of l, we either have $a^{(l)} - \tilde{a}^{(l)} = \overline{y}(s, W_s) - \overline{q}_i^N(W_{t_i})$ or $a^{(l)} - \tilde{a}^{(l)} = \overline{z}^{(d)}(s, W_s) - \overline{z}_i^{N,(d)}(W_{t_i})$ for a $d \in \{1, \dots, \mathcal{D}\}$. We first consider the terms $\overline{y}(s, W_s) - \overline{q}_i^N(W_{t_i})$.

By the definition of $\overline{q}_i^N$, we get for $p \in \{1,2\}$ with Hölder's inequality that

$$
\begin{aligned}
&E_i\left[\left(\overline{y}(s,W_s) - \overline{q}_i^N(W_{t_i})\right)^p\right] \\
&= E_i\left[\left(\overline{y}(s,W_s) - E_i\left[\overline{y}(t_{i+1},W_{t_{i+1}})\right]\right)^p\right] \\
&= E_i\left[\left(\overline{y}(s,W_s) - E_i\left[\overline{y}(t_i,W_{t_i}) - \int_{t_i}^{t_{i+1}} f(l,W_l,Y_l,Z_l)dl + \int_{t_i}^{t_{i+1}} Z_l dW_l\right]\right)^p\right] \\
&= E_i\left[\left(\overline{y}(s,W_s) - E_i\left[\overline{y}(t_i,W_{t_i}) - \int_{t_i}^{t_{i+1}} f(l,W_l,Y_l,Z_l)dl\right]\right)^p\right] \\
&= pE_i\left[(\overline{y}(s,W_s) - \overline{y}(t_i,W_{t_i}))^p\right] + pE_i\left[\left(\int_{t_i}^{t_{i+1}} f(l,W_l,Y_l,Z_l)dl\right)^p\right] \\
&\leq pE_i\left[(\overline{y}(s,W_s) - \overline{y}(t_i,W_{t_i}))^p\right] + p\Delta^p C_f^p
\end{aligned}
$$

where we used that f is uniformly bounded by C_f in the last step. Note that for $s = t_i$, this shows in particular that

$$
E_i\left[\left(\overline{y}(s,W_{t_i}) - \overline{q}_i^N(W_{t_i})\right)^2\right] \leq c\Delta^2
$$

which is the first part of the bound in (3). We now set $\tilde{a}_y := (s, W_s^{(1)}, \ldots, W_s^{(\mathfrak{D})})^T$ and $a_y := (t_i, W_{t_i}^{(1)}, \ldots, W_{t_i}^{(\mathfrak{D})})^T$. Then, for $p = 2$, a Taylor expansion on y yields

$$
\begin{aligned}
E_i\left[(\overline{y}(s,W_s) - \overline{y}(t_i,W_{t_i}))^2\right] &= E_i\left[\left(\int_0^1 (1-\Theta)\nabla\overline{y}(a_y + \Theta(\tilde{a}_y - a_y))(a_y - \tilde{a}_y)d\Theta\right)^2\right] \\
&\leq E_i\left[\left(\sup_{\Theta\in[0,1]} \nabla\overline{y}(a_y + \Theta(\tilde{a}_y - a_y))(a_y - \tilde{a}_y)\right)^2\right] \\
&\leq C_y^2 E_i\left[|a_y - \tilde{a}_y|^2\right] \leq c\Delta.
\end{aligned}
$$

Here we used that the derivatives of $\overline{y}$ are bounded by a constant C_y and that it holds for the entries of $a_y - \tilde{a}_y$

$$
\begin{aligned}
E_i\left[|a_y^{(d)} - \tilde{a}_y^{(d)}|^2\right] &= \begin{cases} E_i\left[|s - t_i|^2\right] & d = 1 \\ E_i\left[W_s^{(d-1)} - W_{t_i}^{(d-1)}|^2\right] & d > 1 \end{cases} \\
&\leq \begin{cases} \Delta^2 & d = 1 \\ \Delta & d > 1 \end{cases},
\end{aligned}
$$

since W is a Brownian motion. In the case $p = 1$, we have to continue the Taylor expansion an additional step and get similarly:

$$
\begin{aligned}
&E_i\left[(\overline{y}(s,W_s) - \overline{y}(t_i,W_{t_i}))\right] \\
&= E_i\left[\nabla\overline{y}(a_y)^T(a_y - \tilde{a}_y) + \frac{1}{2}\int_0^1 (1-\Theta)(a_y - \tilde{a}_y)^T \operatorname{Hess}_{\overline{y}}(a_y + \Theta(\tilde{a}_y - a_y))(a_y - \tilde{a}_y)d\Theta\right] \\
&\leq E_i\left[\nabla\overline{y}(a_y)^T(a_y - \tilde{a}_y)\right] + \frac{1}{2}E_i\left[\sup_{\Theta\in[0,1]} (a_y - \tilde{a}_y)^T \operatorname{Hess}_{\overline{y}}(a_y + \Theta(\tilde{a}_y - a_y))(a_y - \tilde{a}_y)\right].
\end{aligned}
$$

Now $\nabla \overline{y}(a_y)$ is $\mathcal{F}_i$-measurable and $E_i[(a_y - \tilde{a}_y)] = (s - t_i, 0, \ldots, 0)^T$, since $W_s - W_{t_i}$ is independent of $\mathcal{F}_{t_i}$ and has expectation 0. Additionally, using that $\overline{y}$ has bounded derivatives, we conclude

$$E_i\left[(\overline{y}(s, W_s) - \overline{y}(t_i, W_{t_i}))\right] \leq C_y \Delta + \frac{1}{2} \sum_{d,l=1}^{1+\mathcal{D}} C_f E_i\left[|(a_y^{(d)} - \tilde{a}_y^{(d)})|^2\right]^{\frac{1}{2}} E_i\left[|(a_y^{(l)} - \tilde{a}_y^{(l)})|^2\right]^{\frac{1}{2}}$$
$$\leq c\Delta.$$

It remains to show that $E_i[\big(z^{(d)}(s, W_s) - \overline{z}_i^N(W_{t_i})^{(d)}\big)^p]$ is bounded by a multiple of Δ for $p \in \{0, 1\}$. For this, we first rewrite the $d - th$ component of $\overline{z}_i^N$ by a Taylor expansion on y as

$$\begin{aligned}
\overline{z}_i^{N,(d)} &= E_i\left[\frac{\Delta W_{i+1}^{(d)}}{\Delta} \overline{y}(t_{i+1}, W_{t_{i+1}})\right] \\
&= E_i\Bigg[\frac{\Delta W_{i+1}^{(d)}}{\Delta}\Bigg(\overline{y}(t_i, W_{t_i}) + \sum_{e=1}^{\mathcal{D}} \frac{\partial}{\partial x^{(e)}} \overline{y}(t_i, W_{t_i})(\Delta W_{i+1}^{(e)}) + \frac{\partial}{\partial t}\overline{y}(t_i, W_{t_i})\Delta \\
&\quad + \frac{1}{2}\sum_{e,l=1}^{\mathcal{D}} \frac{\partial^2}{\partial x^{(e)}\partial x^{(l)}} \overline{y}(t_i, W_{t_i})(\Delta W_{i+1}^{(e)})(\Delta W_{i+1}^{(l)}) + \frac{1}{2}\sum_{e=1}^{\mathcal{D}} \frac{\partial^2}{\partial t \partial x^{(e)}} \overline{y}(t_i, W_{t_i})(\Delta W_{i+1}^{(e)})\Delta \\
&\quad + \frac{1}{2}\frac{\partial^2}{\partial^2 t}\overline{y}(t_i, W_{t_i})\Delta^2 + \frac{1}{6}\int_0^1 (1-\Theta)\Bigg(\frac{\partial^3}{\partial^3 t}\overline{y}(t_i + \Theta\Delta, W_{t_i} + \Theta\Delta W_{i+1})\Delta^3 \\
&\quad + \sum_{e,l,k=1}^{\mathcal{D}} \frac{\partial^3}{\partial x^{(e)}\partial x^{(l)}\partial x^{(k)}} \overline{y}(t_i + \Theta\Delta, W_{t_i} + \Theta\Delta W_{i+1})(\Delta W_{i+1}^{(e)})(\Delta W_{i+1}^{(l)})(\Delta W_{i+1}^{(k)}) \\
&\quad + \sum_{e,l=1}^{\mathcal{D}} \frac{\partial^3}{\partial t\partial x^{(e)}\partial x^{(l)}} \overline{y}(t_i + \Theta\Delta, W_{t_i} + \Theta\Delta W_{i+1})(\Delta W_{i+1}^{(e)})(\Delta W_{i+1}^{(l)})\Delta \\
&\quad + \sum_{e=1}^{\mathcal{D}} \frac{\partial^3}{\partial^2 t\partial x^{(e)}} \overline{y}(t_i + \Theta\Delta, W_{t_i} + \Theta\Delta W_{i+1})(\Delta W_{i+1}^{(e)})\Delta^2\Bigg) d\Theta\Bigg)\Bigg].
\end{aligned}$$

Since the derivatives of $\overline{y}$ are $\mathcal{F}_i$-measurable when evaluated in (t_i, W_{t_i}) and the components of $W_{t_{i+1}} - W_{t_i}$ are independent with mean 0 each, most terms in the right hand side of the equality above vanish and we get

$$\begin{aligned}
\overline{z}_i^{N,(d)} &= E_i\left[\frac{\Delta W_{i+1}^{(d)}}{\Delta}\frac{\partial}{\partial x^{(d)}}\overline{y}(t_i, W_{t_i})\Delta W_{i+1}^{(d)}\right] \\
&\quad + \frac{1}{2}E_i\left[\frac{\Delta W_{i+1}^{(d)}}{\Delta}\frac{\partial^2}{\partial t\partial x^{(d)}}\overline{y}(t_i, W_{t_i})\Delta W_{i+1}^{(d)}\Delta\right] \\
&\quad + \frac{1}{6}E_i\Bigg[\frac{\Delta W_{i+1}^{(d)}}{\Delta}\int_0^1 (1-\Theta)\Bigg(\frac{\partial^3}{\partial^3 t}\overline{y}(t_i + \Theta\Delta, W_{t_i} + \Theta\Delta W_{i+1})\Delta^3 \\
&\quad + \sum_{e,l,k=1}^{\mathcal{D}} \frac{\partial^3}{\partial x^{(e)}\partial x^{(l)}\partial x^{(k)}}\overline{y}(t_i + \Theta\Delta, W_{t_i} + \Theta\Delta W_{i+1})(\Delta W_{i+1}^{(e)})(\Delta W_{i+1}^{(l)})(\Delta W_{i+1}^{(k)})
\end{aligned}$$

$$
\begin{aligned}
&+\sum_{e,l=1}^{\mathcal{D}} \frac{\partial^3}{\partial t \partial x^{(e)} \partial x^{(l)}} \overline{y}(t_i + \Theta\Delta, W_{t_i} + \Theta\Delta W_{i+1})(\Delta W_{i+1}^{(e)})(\Delta W_{i+1}^{(l)})\Delta \\
&+\sum_{e=1}^{\mathcal{D}} \frac{\partial^3}{\partial^2 t \partial x^{(e)}} \overline{y}(t_i + \Theta\Delta, W_{t_i} + \Theta\Delta W_{i+1})(\Delta W_{i+1}^{(e)})\Delta^2 \Big) d\Theta \Bigg] \\
&= E_i \left[\frac{\Delta W_{i+1}^{(d)}}{\Delta} \frac{\partial}{\partial x^{(d)}} \overline{y}(t_i, W_{t_i}) \Delta W_{i+1}^{(d)} \right] + R_T \\
&= \frac{\partial}{\partial x^{(d)}} \overline{y}(t_i, W_{t_i}) + R_T.
\end{aligned}
$$

It is straightforward to check that $|R_T| \leq c\Delta$. Then, since $\overline{z}(s, W_s) = \nabla_x \overline{y}(s, W_s)$, where $\nabla_x \overline{y}$ denotes the vector of first degree partial derivatives of $\overline{y}$ with respect to $x^{(1)}, \ldots, x^{(\mathcal{D})}$, it follows

$$
\begin{aligned}
&E_i \left[\left(\overline{z}^{(d)}(s, W_s) - \overline{z}_i^{N,(d)} \right)^p \right] \\
&= E_i \left[\left(\frac{\partial}{\partial x^{(d)}} \overline{y}(s, W_s) - \frac{\partial}{\partial x^{(d)}} \overline{y}(t_i, W_{t_i}) - R_T \right)^p \right] \\
&\leq p E_i \left[\left(\frac{\partial}{\partial x^{(d)}} \overline{y}(s, W_s) - \frac{\partial}{\partial x^{(d)}} \overline{y}(t_i, W_{t_i}) \right)^p \right] + p E_i \left[|R|^p \right] \\
&\leq p E_i \left[\left(\frac{\partial}{\partial x^{(d)}} \overline{y}(s, W_s) - \frac{\partial}{\partial x^{(d)}} \overline{y}(t_i, W_{t_i}) \right)^p \right] + c\Delta^p.
\end{aligned} \tag{5}
$$

The term $E_i[\frac{\partial}{\partial x^{(d)}} \overline{y}(s, W_s) - \frac{\partial}{\partial x^{(d)}} \overline{y}(t_i, W_{t_i}))^p]$ is for $p \in \{1, 2\}$ bounded by $c\Delta$ for a constant c not depending on Δ, which follows by the same calculations used for the term $E_i[(\overline{y}(s, W_s) - \overline{y}(t_i, W_{t_i}))^p]$ where we have to replace $\overline{y}$ by its first partial derivative $\frac{\partial}{\partial x^{(d)}} \overline{y}$. Note that the Taylor expansion than uses the derivatives of $\overline{y}$ up to degree 3, which still all exist are bounded by assumption. This finishes the proof of (4) and hence the bound on $\mathcal{R}_i^N$. Also, note that (5) for $s = t_i$ shows in particular that

$$
E_i \left[\left(\overline{z}^{(d)}(t_i, W_{t_i}) - \overline{z}_i^{N,(d)} \right)^p \right] \leq c\Delta
$$

which completes the proof of the bound in (3).

It remains to show that, whenever $\overline{y}$ bounded and $s + 1$ times differentiable with bounded derivatives, the functions $\overline{q}_i^N$ and $\overline{z}_i^N$ are bounded as well and are respectively $s + 1$ and s times differentiable with bounded derivatives. For this, we can simply use that the components of ΔW_{i+1} are independent and Gaussian-distributed with mean 0 and variance Δ each. Hence we have

$$
\begin{aligned}
\overline{q}_i^N(x) &= E_i \left[\overline{y}(t_{i+1}, W_{t_{i+1}}) \middle| W_{t_i} = x \right] = E_i \left[\overline{y}(t_{i+1}, \Delta W_{i+1} + W_{t_i}) | W_{t_i} = x \right] \\
&= \int_{\mathbb{R}^{\mathcal{D}}} \overline{y}(t_{i+1}, \tilde{x} + x) \frac{1}{(\sqrt{2\pi\Delta})^{\mathcal{D}}} e^{-\frac{1}{2} \sum_{d=1}^{\mathcal{D}} \frac{(\tilde{x}_i)^2}{\Delta}} d\tilde{x}.
\end{aligned}
$$

Then, since $\overline{y}$ is differentiable in x with bounded derivative, we can partial differentiate under the integral and get

$$\begin{aligned}\frac{\partial}{\partial x^{(d)}}\overline{q}_i^N(x) &= \int_{\mathbb{R}^{\mathcal{D}}} \frac{\partial}{\partial x^{(d)}}\overline{y}(t_{i+1}, \tilde{x}+x)\frac{1}{(\sqrt{2\pi\Delta})^{\mathcal{D}}}e^{-\frac{1}{2}\sum_{d=1}^{\mathcal{D}}\frac{(\tilde{x}_i)^2}{\Delta}}d\tilde{x} \\ &= E\left[\frac{\partial}{\partial x^{(d)}}\overline{y}(t_i, W_{t_{i+1}})\middle| W_{t_i}=x\right]\end{aligned}$$

for each $d \in \{1, \dots, \mathcal{D}\}$, which shows that $\overline{q}_i^N$ is continuous differentiable with bounded derivative. The same argumentation can be applied for the higher order derivatives.
Next, we consider the first coordinate of $\overline{z}_i^N$ as the derivatives of the others follow analogously. By the definition of $\overline{z}_i^N$ and Fubini's theorem, we have

$$\begin{aligned}\overline{z}_i^{N,(1)}(x) &= E_i\left[\frac{\Delta W_{i+1}^{(1)}}{\Delta}\overline{y}(t_{i+1}, W_{t_{i+1}})\middle| W_{t_i}=x\right] \\ &= \int_{\mathbb{R}}\cdots\int_{\mathbb{R}}\prod_{d=2}^{\mathcal{D}}\frac{1}{\sqrt{2\pi\Delta}}e^{-\frac{\left(\tilde{x}^{(d)}\right)^2}{2\Delta}}\int_{\mathbb{R}}\frac{\tilde{x}^{(1)}}{\Delta}\frac{1}{\sqrt{2\pi\Delta}}e^{-\frac{\left(\tilde{x}^{(1)}\right)^2}{2\Delta}}\overline{y}(t_{i+1}, x+\tilde{x})d\tilde{x}^{(1)}d\tilde{x}^{(2)}\dots d\tilde{x}^{(\mathcal{D})}.\end{aligned}$$

Since $\overline{y}$ is bounded by assumption, integration by parts leads to

$$\begin{aligned}\overline{z}_i^{N,(1)}(x) &= \int_{\mathbb{R}}\cdots\int_{\mathbb{R}}\prod_{d=2}^{\mathcal{D}}\frac{1}{\sqrt{2\pi\Delta}}e^{-\frac{\left(\tilde{x}^{(d)}\right)^2}{2\Delta}}\int_{\mathbb{R}}\frac{1}{\sqrt{2\pi\Delta}}e^{-\frac{\left(\tilde{x}^{(1)}\right)^2}{2\Delta}}\frac{\partial}{\partial x^{(1)}}\overline{y}(t_{i+1}, x+\tilde{x})d\tilde{x}^{(1)}d\tilde{x}^{(2)}\dots d\tilde{x}^{(\mathcal{D})} \\ &= E_i\left[\frac{\partial}{\partial x^{(1)}}\overline{y}(t_{i+1}, x+\tilde{x})\middle| W_{t_i}=x\right] = \frac{\partial}{\partial x^{(1)}}\overline{q}^N(x).\end{aligned}$$

Hence the statement for $\overline{z}_i^N$ follows by the one for $\overline{q}_i^N$.

References

P. E. Kloeden and E. Platen. Stochastic differential equations. In *Numerical Solution of Stochastic Differential Equations*, pages 103–160. Springer, 1992.

J. Zhang. *Some fine properties of backward stochastic differential equations.* PhD thesis, PhD thesis, Purdue University, 2001.