Appendix

Appendix A: Expected Approximation Errors

The proofs of our results, in Appendix B, rely on a theorem of Detemple, Garcia and Rindisbacher (2004) that characterizes the asymptotic expected approximation error. To state this result we need to introduce the following notation. Recall that $\partial A, \partial B_j$ are $d \times d$ matrices of Jacobians. The tangent process, in this multivariate setup, solves the linear equation

$$d\nabla_{t,x}X_s = \left( \partial A(X_s)ds + \sum_{j=1}^d \partial B_j(X_s)dW^j_s \right) \nabla_{t,x}X_s \quad \text{with} \quad \nabla_{t,x}X_t = I_n. \quad (A-1)$$

With this notation we have,
**Theorem 5**: Let $g \in C^3(\mathbb{R}^d)$ be such that the uniform integrability condition
\[
\lim_{r \to \infty} \limsup_{N} \mathbb{E}_t \left[ 1_{\{|N(g(X^N_T) - g(X_T))| > r \}} N|g(X^N_T) - g(X_T)| \right] = 0 \tag{A-2}
\]
holds ($\mathbb{P}$-a.s.). Then, as $N \to \infty$,
\[
NE_t [g(X^N_T) - g(X_T)] \to \frac{1}{2} K_{t,T}(x) \equiv \frac{1}{2} \mathbb{E}_t [\partial g(X_T)V_1(t,T) + V_2(t,T)] \tag{A-3}
\]
where the random variables $V_1(t,T)$ and $V_2(t,T)$ are
\[
V_1(t,T) = -\nabla_{t,x} X_T \int_t^T (\nabla_{t,x} X_s)^{-1} \left( \partial A(X_s) dX_s + \sum_{j=1}^d [\partial B_j A](X_s) dW_s^j \right) - \sum_{i,j=1}^d ((\partial B_j)(\partial B_j)B_i)(X_s)dW_s^i
\]
\[
+ \nabla_{t,x} X_T \int_t^T (\nabla_{t,x} X_s)^{-1} \left[ \sum_{j=1}^d [\partial B_j \partial B_j A] - \sum_{j,k,l=1}^d [\partial_k (\partial_l AB_{t,j}) B_{k,j}] \right] (X_s)ds
\]
\[
+ \nabla_{t,x} X_T \int_t^T (\nabla_{t,x} X_s)^{-1} \sum_{i,j=1}^d ([\partial [\partial B_j \partial B_j B_i] B_i - \partial B_i \partial B_j \partial B_j B_i] (X_s)) ds
\]
\[
V_2(t,T) = -\int_t^T \sum_{i,j=1}^d \nu_{i,j}(s,T) ds.
\]
where $\nu_{i,j}(s,T) \equiv [h^{i,j}(\nabla_{t,x} X_s)^{-1} [\partial B_j B_i](X_s), W_s^i]$ with $h^{i,j}_t \equiv \mathbb{E}_t [\mathcal{D}_{jl} (\partial g(X_T) \nabla_{t,x} X_T e_i)]$. An explicit expression for $\nu_{i,j}(s,T)$ is in Detemple, Garcia and Rindisbacher (2004).

In order to derive the asymptotic expected approximation error, on the right hand side of (A-3), an expansion, based on the mean value theorem, of the error $g(X^N_T) - g(X_T)$ is used. This gives a linear SDE with four autonomous components. Three of these autonomous terms converge at the rate $1/N$. The last one converges at the rate $1/\sqrt{N}$, but has zero expectation. This difference in the convergence behavior of the autonomous terms in the error expansion explains the presence of two parts in the second order bias $K_{t,T}(x)$. The limits of the first three terms give rise to $V_1(t,T)$. The fourth term leads to $V_2(t,T)$.

The uniform integrability assumption in the theorem enables us to find the asymptotic expected approximation error by taking the expectation of the weak limit of the error expansion.

**Appendix B: Proofs**

In what follows we set $t = 0$, without loss of generality, and write $E_0$ for $E_{0,x}$, $P_0$ for $P_{0,x}$, $\nabla x$ for $\nabla_{0,x}$, and $K_T(x)$ for $K_{0,T}(x)$. 

2
Proof of Theorem 2: Recall that $f(0, x) \equiv \mathbb{E}_0[g(X_T)]$ and that the MCC estimator $\partial_x \tilde{f}(0, x)^{M,N,\tau}$ is given in (26). The error of the MCC estimator, $\tilde{e}(0, x; M, N, \tau) \equiv \partial_x \tilde{f}(0, x)^{M,N,\tau} - \partial_x f(0, x)$, can be expanded as follows

$$
\tilde{e}(0, x; M, N, \tau) = \mathbb{E}_0 \left[ g(X_T) \frac{\Delta_x W_0}{\tau} \right] C(x) - \partial_x f(0, x)
+ \frac{1}{M} \sum_{i=1}^{M} \left( g(X_{T_i}^i) - g(X_T^i) \right) \frac{\Delta_x W_0^i}{\tau} C(x)
+ \frac{1}{M} \sum_{i=1}^{M} \left( g(X_T^i) \frac{\Delta_x W_0^i}{\tau} - \mathbb{E}_0 \left[ g(X_T) \frac{\Delta_x W_0}{\tau} \right] \right) C(x)
$$

where $\Delta_x W_0 = W_\tau - W_0$ and $C$ is the generalized inverse of $B$ defined in (13). In this expansion, $X_{T_i}^i$ is an independent replication of the Euler discretization of the diffusion given in (4), $X_T^i$ (resp. $W_T^i$) is an independent replication of the terminal value of the diffusion (resp. of the Brownian motion driving the SDE) and $\Delta_x W_0^i$ is an independent replication of the increment in the Brownian motion.

Our next three lemmas give the asymptotic errors for the three terms in this expansion.

Lemma 1: Under the assumption of Theorem 2 we have

$$
\frac{1}{\tau} \left( \mathbb{E}_0 \left[ g(X_T) \frac{\Delta_x W_0}{\tau} \right] C(x) - \partial_x f(0, x) \right) \to \frac{1}{2} \partial_v \mathbb{E}_0 [\partial_x f(v, X_v)B(X_v)]_{v=0} C(x)
$$
as $1/\tau \to \infty$.

Proof of Lemma 1: The Clark-Ocone formula $g(X_T) = \mathbb{E}_0 [g(X_T)] + \int_0^T \mathbb{E}_v [\mathcal{D}_v g(X_T)] dW_v$ gives

$$
\frac{1}{\tau} \left( \mathbb{E}_0 \left[ g(X_T) \frac{\Delta_x W_0}{\tau} \right] C(x) - \partial_x f(0, x) \right) = \frac{1}{\tau^2} \int_0^\tau \beta(v) dv
$$

with $\beta(v) \equiv \mathbb{E}_0 [\mathbb{E}_v [\mathcal{D}_v g(X_T)] C(x) - \partial_x f(0, x)]$. Given that $\beta(0) = 0$, a second order Taylor approximation of $\beta(v)$ yields $\lim_{1/\tau \to \infty} \frac{1}{\tau^2} \int_0^\tau \beta(v) dv = \frac{1}{2} \partial_v \beta(v)|_{v=0}$. But

$$
\partial_v \beta(v)|_{v=0} = (\partial_v \mathbb{E}_0 [\mathbb{E}_v [\mathcal{D}_v g(X_T)] C(x) - \partial_x f(0, x)]|_{v=0}
= (\partial_v \mathbb{E}_0 [\partial_x f(v, X_v)B(X_v)]|_{v=0} C(x)
$$

where the last equality uses $\mathbb{E}_v [\mathcal{D}_v g(X_T)] = \partial_x f(v, X_v)B(X_v)$. Substituting the expression for $\partial_v \beta(v)|_{v=0}$ in the limit above establishes the result in the lemma.
Lemma 2: Let $K_T(x)$ be given by equation (A-3) in Theorem 5. Under the assumption of Theorem 2 we have, as $M \to \infty$,

$$N_M \frac{1}{M} \sum_{i=1}^{M} \left( g(X_{i,N}^1) - g(X_{i,T}^1) \right) \frac{\Delta \tau M_i W_0^i}{\tau M} - \frac{1}{2} \partial K_T(x) B(x)$$

in probability, when $N_M, 1/\tau M \to \infty$ as $M \to \infty$.

Proof of Lemma 2: We proceed in two steps. In the first step we find a sequential limit. In the second step we establish that joint limits and sequential limits are the same. Define,

$$e_{T}^{M,N,\tau} = \frac{1}{M} \sum_{i=1}^{M} N \left( g(X_{i,N}^1) - g(X_{i,T}^1) \right) \frac{\Delta \tau M_i W_0^i}{\tau} - \frac{1}{2} \partial K_T(x) B(x),$$

and note that $e_{T}^{M,N,\tau} = e_{1,T}^{M,N,\tau} + e_{2,T}^{N,\tau} + e_{3,T}^{N}$, with

$$e_{1,T}^{M,N,\tau} = \frac{1}{M} \sum_{i=1}^{M} N \left( g(X_{i,N}^1) - g(X_{i,T}^1) \right) \frac{\Delta \tau M_i W_0^i}{\tau} - \mathbb{E}_0 \left[ N \left( g(X_{T}^N) - g(X_{T}) \right) \frac{\Delta \tau M_i W_0^i}{\tau} \right]$$

$$e_{2,T}^{N,\tau} = \mathbb{E}_0 \left[ \frac{1}{\tau} \int_0^\tau [D_v - D_0] N \left( g(X_{T}^N) - g(X_{T}) \right) dv \right]$$

$$e_{3,T}^{N} = \mathbb{D}_0 N \mathbb{E}_0 \left[ g(X_{T}^N) - g(X_{T}) \right] - \frac{1}{2} \partial K_T(x) B(x).$$

The expression for $e_{2,T}^{N,\tau}$ is obtained by using the Clark-Ocone formula

$$N \left( g(X_{T}^N) - g(X_{T}) \right) = \mathbb{E}_0 \left[ N \left( g(X_{T}^N) - g(X_{T}) \right) \right] + \int_0^T \mathbb{E}_v \left[ D_v N \left( g(X_{T}^N) - g(X_{T}) \right) \right] dW_v$$

to conclude

$$\mathbb{E}_0 \left[ N \left( g(X_{T}^N) - g(X_{T}) \right) \frac{\Delta \tau M_i W_0^i}{\tau} \right] = \mathbb{E}_0 \left[ \frac{1}{\tau} \int_0^\tau \mathbb{E}_v \left[ D_v N \left( g(X_{T}^N) - g(X_{T}) \right) \right] dv \right].$$

By the weak law of large numbers for i.i.d random sequences the first term vanishes in probability for fixed $N$ and $\tau$, i.e. $\mathbb{P} - \lim_{M \to \infty} e_{1,T}^{M,N,\tau} = 0$ for all $N, \tau$. By the continuous differentiability of the Riemann integral $\lim_{\tau \to \infty} e_{2,T}^{N,\tau} = 0$ for fixed $N$. Finally, by Theorem 5 that shows $N \mathbb{E}_0 [g(X_{T}^N) - g(X_{T})] \to \frac{1}{2} K_T(x)$ and by the chain rule of Malliavin that gives $\mathbb{D}_0 K_T(x) = \partial K_T(x) B(x)$ we obtain $\lim_{N \to \infty} e_{3,T}^{N} = 0$. We conclude that $e_{T}^{M,N,\tau} \to 0$ in probability, if sequentially $M, 1/\tau, N \to \infty$.

Next, we show that this sequential limit corresponds to the joint limit (i.e. when $M, N, 1/\tau \to \infty$ jointly) and thus to the limit along any “diagonal”. Invoking arguments in the proof of Theorem 4.2 in Billingsley (1968) (see also Lemma 6 in Phillips and Moon (1996)), it is necessary and sufficient to show that

$$\lim_{M,N,1/\tau \to \infty} \mathbb{P}_0 \left( \left\| e_{1,T}^{M,N,\tau} \right\| > \epsilon \right) = 0 \quad \text{(B-4)}$$
\[ \limsup_{N,1/\tau \to \infty} \left\| e_{2,T}^{N,\tau} \right\| = 0. \] (B-5)

Condition (B-4) is proved by applying the Markov-type inequality,\(^1\)

\[ P_0 \left( \frac{1}{M} \left\| \sum_{i=1}^{M} H_{i,T}^{N,\tau} \right\| > \epsilon \right) \leq \frac{E_0 \left[ \left\| H_{i,T}^{N,\tau} \right\| 1_{\left\| H_{i,T}^{N,\tau} \right\| > \epsilon M} \right]}{\epsilon M} + M^{\gamma-1} P_0 \left( \left\| H_{i,T}^{N,\tau} \right\| \leq \epsilon M^{\gamma} \right) \] (B-6)

for \( \epsilon > 0 \) and \( \gamma \in [0,1] \) to the random variable

\[ H_{i,T}^{N,\tau} \equiv N \left( g(X_{i,T}^N) - g(X_T^i) \right) \frac{\Delta \tau W_0^i}{\tau} - E_0 \left[ N (g(X_T^N) - g(X_T)) \frac{\Delta \tau W_0}{\tau} \right]. \]

Let \( M, N, 1/\tau \to \infty \). The uniform integrability condition (27) with \( r = M^\gamma \epsilon \) establishes that the first term on the right hand side of (B-6) converges to zero for all \( \epsilon > 0 \). The second term also converges to zero as \( M^\gamma \epsilon \to 0 \) when \( M \to \infty \). We conclude that the left hand side converges to zero as \( M, N, 1/\tau \to \infty \), thereby establishing (B-4).

For condition (B-5) note the inequality

\[ \sup_{N \geq N_0, 1/\tau \geq 1/\tau_0} \left\| e_{2,T}^{N,\tau} \right\| \leq e_{2,T}^{N_0,\tau_0} \equiv \sup_{N \geq N_0} \left\| E_0 \left[ \sup_{v \in (0,\tau_0]} D_v Z_T^N - D_0 Z_T^N \right] \right\|, \]

where \( \sup_{v \in (0,\tau_0]} D_v Z_T^N - D_0 Z_T^N \) is non-decreasing in \( \tau_0 \) and null for \( \tau_0 = 0 \). Taking the infimum over \( N_0 > 0, 1/\tau_0 > 0 \) on both sides gives (B-5) as \( \inf_{N_0 > 0, 1/\tau_0 > 0} e_{2,T}^{N_0,\tau_0} = 0. \) \( \blacksquare \)

**Lemma 3:** Under the assumption of Theorem 2 we have

\[ \sqrt{\tau_M} \frac{1}{\sqrt{M}} \sum_{i=1}^{M} \left( g(X_{i,T}) \frac{\Delta \tau W_0^i}{\tau_M} - E_0 \left[ g(X_T^i) \frac{\Delta \tau W_0}{\tau_M} \right] \right) \Rightarrow O_T(x) \] (B-7)

where \( 1/\tau_M \to \infty \) as \( M \to \infty \) and where \( O_T(x) \sim N \left( 0, E_0[g(X_T)^2] \right). \)

**Proof of Lemma 3:** First note that Ito’s lemma implies

\[ (\Delta \tau W_t) (\Delta \tau W_t)' = \int_t^{t+\tau} dW_s (W_s - W_t)' + \int_t^{t+\tau} (W_s - W_t) (dW_s)' + \tau I_d, \]

\(^1\)This inequality holds for a sequence of i.i.d. random vectors \( Z' \) and any \( \epsilon > 0 \), as \( E \left[ \left\| Z \right\| 1_{\left\| Z \right\| > \epsilon M} \right] = E \left[ \left\| Z \right\| \right] - E \left[ \left\| Z \right\| 1_{\left\| Z \right\| \leq \epsilon M} \right] \), by the triangle inequality and the Markov inequality (Kallenberg (1997), Lemma 3.1 page 40),

\[ E \left[ \left\| Z \right\| 1_{\left\| Z \right\| > \epsilon M} \right] \geq E \left[ \left\| Z \right\| \right] - \epsilon M^\gamma \left[ E \left[ \left\| Z \right\| 1_{\left\| Z \right\| \leq \epsilon M} \right] \right] \geq E \left[ \left\| Z \right\| \right] - \epsilon M^\gamma \left[ E \left[ \left\| Z \right\| \right] - E \left[ \left\| Z \right\| 1_{\left\| Z \right\| \leq \epsilon M} \right] \right] \]

\[ \geq P \left( \left\| Z \right\| > \epsilon \right) \epsilon M - \epsilon M^\gamma \left[ P \left( \left\| Z \right\| \leq \epsilon M \right) \right]. \]
while the Clark-Ocone formula gives
\[ g(X_T)^2 = E_0[g(X_T)^2] + \sum_{j=1}^{d} \int_0^T E_v[D_jg(X_T)^2]dW_v^j \]
\[ g(X_T) = E_0[g(X_T)] + \sum_{j=1}^{d} \int_0^T E_v[D_jg(X_T)]dW_v^j. \]

It follows that
\[ m_{2T}^\tau = E_0 [g(X_T)^2 (\Delta_\tau W_0)(\Delta_\tau W_0)'] \]
\[ = E_0 \left[ \int_0^\tau E_v [D_v g(X_T)^2] dW_v \left( \int_0^\tau (W_s - W_0) (dW_s) '\right) \right] + \tau E_0 [g(X_T)^2] I_d \]
\[ = E_0 \left[ \int_0^\tau (W_v - W_0)E_v [D_v g(X_T)^2] dv \right] \]
\[ + \tau E_0 [g(X_T)^2] I_d, \]
and that \( m_{1T}^\tau \equiv E_0 [g(X_T)\Delta_\tau W_0] = E_0 \left[ \int_0^\tau E_v [D_v g(X_T)] dv \right] \). From these expressions we obtain the conditional variance at \( t = 0, \)
\[ \text{VAR}_0 \left[ g(X_T) \frac{\Delta_\tau W_0}{\sqrt{\tau}} \right] = \frac{1}{\tau} m_{2T}^\tau - \tau \left( \frac{1}{\tau} m_{1T}^\tau \right) '\]
\[ = \frac{1}{\tau} E_0 \left[ \int_0^\tau (W_v - W_0)E_v [D_v g(X_T)^2] dv \right] \]
\[ + \frac{1}{\tau} E_0 \left[ \int_0^\tau E_v (D_v g(X_T)^2)' (W_v - W_0)' dv \right] + \tau E_0 [g(X_T)^2] I_d \]
\[ - \tau \left( \frac{1}{\tau} E_0 \left[ \int_0^\tau E_v [D_v g(X_T)] dv \right] \right) '\]
Given that terms one, two and four converge to zero as \( 1/\tau \to \infty \) the limit
\[ \lim_{1/\tau \to \infty} \text{VAR}_0 \left[ g(X_T) \frac{\Delta_\tau W_0}{\sqrt{\tau}} \right] = E_0 [g(X_T)^2] I_d \]
holds. The uniform integrability condition (28), which is sufficient for the Lindeberg Central Limit theorem for independent random variables to hold, can now be invoked to conclude that the weak limit is Gaussian with variance \( E_0 [g(X_T)^2] I_d. \]

Proof of Theorem 2 (continued): Combining Lemmas 1, 2 and 3, shows that\(^2\)
\(^2\) The symbol \( O_p(x) \) stands for "at most of order \( x \) in probability". A sequence of random variables \( Z^N \) is \( O_p(N^k) \), if for every \( \epsilon > 0 \), there exists a real number \( r \) such that \( P(N^{-k}|Z^N| > r) < \epsilon \) for all \( N \). If \( N^{-k}|Z^N| \leq r \) for all \( N \) pointwise, the sequence \( Z^N \) is said to be at most of order \( N^k \). In this case we write \( O(N^k) \). In contrast, a sequence \( Z^N \) is of smaller order than \( N^k \), denoted by \( o(N^k) \), if \( N^{-k}|Z^N| \to 0 \) as \( N \to \infty \).
\[ \partial_x f(0, x) \bigg|_{M,N,\tau_M} - \partial_x f(0, x) = O_p \left( \frac{1}{N_M} + \frac{1}{\sqrt{M\tau_M}} + \left( \frac{1}{\tau_M} \right)^{-1} \right). \]

With the selection \( 1/\tau_M = M^\delta/\varepsilon_1^2 \) and \( N_M = M^\gamma/\varepsilon_2^2 \) for some \( \varepsilon_1^2, \varepsilon_2^2, \delta, \gamma > 0 \), it is seen that the efficient scheme is obtained for \( (\gamma, \delta) = \inf (\arg \max_{\gamma, \delta} \min \{ \delta, (1-\delta)/2, \gamma \}) = (1/3, 1/3) \). This proves the theorem. \[ \blacksquare \]

**Proof of Theorem 3:** The error \( \tilde{e}_j (0, x; M, N, \tau_j, \alpha_j) \equiv \partial_{x_j} f(0, x) \bigg|_{M,N,\tau_j} - \partial_{x_j} f(0, x) \) of the MCFD estimators \( \partial_{x_j} f(0, x) \) given in (30), can be expanded for arbitrary \( \alpha_j \in \mathbb{R} \), as:

\[ \tilde{e}_j (0, x; M, N, \tau_j, \alpha_j) = R_{1j}(0, x; \alpha_j, \tau_j) + R_{2j}(0, x; M, N, \tau_j, \alpha_j) + R_{3j}(0, x; M, \tau_j, \alpha_j) \]

with

\[ R_{1j}(0, x; \tau_j, \alpha_j) \equiv \frac{E_0 \left[ g \left( X_T(x + \alpha_j \tau_j e_j) \right) - g \left( X_T(x - (1 - \alpha_j) \tau_j e_j) \right) \right]}{\tau_j} - \partial_{x_j} f(0, x) \tag{B-8} \]

\[ R_{2j}(0, x; M, N, \tau_j, \alpha_j) \equiv \frac{1}{M} \sum_{i=1}^{M} \left( \frac{g \left( X^i_T(x + \alpha_j \tau_j e_j) \right) - g \left( X^i_T(x - (1 - \alpha_j) \tau_j e_j) \right)}{\tau_j} \right) \tag{B-9} \]

\[ R_{3j}(0, x; M, \tau_j, \alpha_j) \equiv \frac{1}{M} \sum_{i=1}^{M} \left( \frac{g \left( X^i_T(x + \alpha_j \tau_j e_j) \right) - g \left( X^i_T(x - (1 - \alpha_j) \tau_j e_j) \right)}{\tau_j} \right) \tag{B-10} \]

In this expansion, \( e_j = [0, ..., 1, ..., 0]' \) is the \( j^{th} \) unit vector of dimension \( d \), \( X^i_T(x) \) is an independent replication of the Euler discretization of the SDE based on \( N \) points started at \( X_0^i = x \), and \( X^i_T(x) \) is an independent replication of the continuous diffusion with initial value \( X_0^i = x \).

Our next three lemmas provide the asymptotic errors for \( R_{1j}, R_{2j} \) and \( R_{3j} \).

**Lemma 4:** Let \( R_{1j}(0, x; \tau_j, \alpha_j) \) be defined by (B-8) and suppose that \( f \in C^3(\mathbb{R}^d) \). Then

\[
\begin{align*}
\lim_{1/\tau_j \to \infty} \frac{1}{\tau_j} R_{1j}(0, x; \tau_j, \alpha_j) &= \frac{1}{2} \partial_{x_j}^3 f(0, x) & \text{if } \alpha_j = 1/2 \\
\lim_{1/\tau_j \to \infty} \frac{1}{\tau_j} R_{1j}(0, x; \tau_j, \alpha_j) &= \frac{2\alpha_j - 1}{2} \partial_{x_j}^2 f(0, x) & \text{if } \alpha_j \neq 1/2.
\end{align*}
\]
Proof of Lemma 4: Given the differentiability assumption on \( f \) we can use Taylor series expansions for \( f(0, x + \alpha_j \tau_j e_j) \) and \( f(0, x - (1 - \alpha_j) \tau_j e_j) \) around \( f(0, x) \) to write (see footnote 2 for a definition of \( o(\cdot) \))

\[
R_{1j}(0, x; \tau_j, \alpha_j) = \frac{1}{2} (\alpha_j^2 - (1 - \alpha_j)^2) \partial_{x_j}^2 f(0, x) \tau_j + \frac{1}{6} (\alpha_j^3 + (1 - \alpha_j)^3) \partial_{x_j}^3 f(0, x) \tau_j^2 + o \left( \left( \frac{1}{\tau_j} \right)^{-2} \right).
\]

If \( \alpha_j = 1/2 \) the first term vanishes and \( R_{1j}(0, x; \tau_j, \alpha_j) \) is of order \( (1/j^2) \). If \( \alpha_j \neq 1/2 \) the second term in the expansion is asymptotically negligible and \( R_{1j}(0, x; \tau_j, \alpha_j) \) is of order \( (1/j)^{-1} \). The limits announced are obtained from this Taylor expansion.

Lemma 5: Let \( K_T(x) \) be defined by (A-3), \( R_{2j}(0, x; M, N, \tau_j, \alpha_j) \) by (B-9) and select \( N_M, \tau, M \) such that \( N_M, 1/\tau_j, M \to \infty \) when \( M \to \infty \). Suppose that the assumptions of Theorem 3 hold. Then, for all \( \alpha_j \in [0, 1] \) we have \( N_M R_{2j}(0, x; M, N, \tau_j, \alpha_j) \to \frac{1}{2} \partial_j K_T(x) \) in probability as \( M \to \infty \).

Proof of Lemma 5: We proceed in two steps as in the proof of Lemma 2. Under the assumptions of Theorem 3 we can apply the Clark-Ocone formula to find

\[
R_{2j}(0, x; M, N, \tau_j) = R_{21j}(0, x; M, N, \tau_j) - R_{22j}(0, x; M, N, \tau_j)
\]

with

\[
R_{21j}(0, x; M, N, \tau_j, \alpha_j) \equiv \frac{1}{M} \sum_{i=1}^{M} \frac{g \left( X_{i, T}^M(x + \alpha_j \tau_j e_j) \right) - g \left( X_{i, T}^M(x + \alpha_j \tau_j e_j) \right)}{\tau_j}
\]

\[
R_{22j}(0, x; M, N, \tau_j, \alpha_j) \equiv \frac{1}{M} \sum_{i=1}^{M} \frac{g \left( X_{i, T}^M(x - (1 - \alpha_j) \tau_j e_j) \right) - g \left( X_{i, T}^M(x - (1 - \alpha_j) \tau_j e_j) \right)}{\tau_j}.
\]

The law of large numbers and the arguments in the proof of Theorem 5 in Detemple, Garcia and Rindisbacher (2004) imply

\[
NR_{21j}(0, x; M, N, \tau_j, \alpha_j) \to \frac{1}{2} K_T \left( x + \alpha_j \tau_j e_j \right) / \tau_j,
\]

\[
NR_{22j}(0, x; M, N, \tau_j, \alpha_j) \to \frac{1}{2} K_T \left( x - (1 - \alpha_j) \tau_j e_j \right) / \tau_j
\]

in probability, as \( M, N \to \infty \) sequentially. If next \( 1/j \to \infty \), we obtain \( NR_{2j}(0, x; M, N, \tau_j, \alpha_j) \to \frac{1}{2} \partial_j K_T(x) \). This establishes that \( NR_{2j}(0, x; M, N, \tau_j, \alpha_j) \to \frac{1}{2} \partial_j K_T(x) \) in probability as \( M, N, 1/j \to \infty \) sequentially.

The remainder of the proof parallels the proof of Lemma 2. To show that the same limit holds if \( M, N, 1/j \to \infty \) jointly, and therefore along any “diagonal”, it is sufficient to show that

\[
\limsup_{M, N, 1/j \to \infty} \mathbf{P}_0 \left( \left| \frac{1}{M} \sum_{i=1}^{M} \frac{N \nabla_{x_j}^2 \alpha_j g \left( X_{i, T}^M(x) \right) - \mathbf{E}_0 \left[ N \nabla_{x_j}^2 \alpha_j g \left( X_{T}^M(x) \right) \right] }{\tau_j} \right| > \epsilon \right) = 0 \quad (B-11)
\]
for all $\epsilon > 0$ and

$$\limsup_{N,1/\tau_j \to \infty} |E_0 \left[ N \nabla_{x_j}^{T_j} g \left( X_T^N(x) \right) \right] - \partial_{x_j} N E_0 \left[ g(X_T^N(x)) - g(X_T(x)) \right]| = 0. \quad (B-12)$$

As in the proof of Lemma 2, (B-11) follows from the Markov-type inequality (B-6) and the uniform integrability condition (31). Similarly (B-12) is satisfied, because

$$\partial_{x_j} N E_0 \left[ g(X_T^N(x)) - g(X_T(x)) \right] = N E_0 \left[ \nabla_{x_j} (g(X_T^N(x)) - g(X_T(x))) \right]$$

and under the assumptions of Theorem 5,

$$\limsup_{1/\tau_j,N \to \infty} E_0 \left[ \nabla_{x_j}^{T_j} N (g(X_T^N(x)) - g(X_T(x))) \right] = \limsup_{N \to \infty} E_0 \left[ \nabla_{x_j} N (g(X_T^N(x)) - g(X_T(x))) \right].$$

We conclude that the joint limit corresponds to the sequential limit. $lacksquare$

**Lemma 6:** Let $R_{3j}(0, x; M, \tau)$ be defined by (B-10) and select $\tau_{j,M}$ such that $1/\tau_{j,M} \to \infty$ when $M \to \infty$. Suppose that the assumptions of Theorem 3 hold. For $j = 1, \ldots, d$ and for all $\alpha_j \in [0,1]$ we have the limit, \( \left( \sqrt{M} R_{3j} (0, x; M, \tau_{j, M}) \right)_{j=1, \ldots, d} \to Q_T \) as $M \to \infty$, where $Q_T \sim N \left( 0, E_0 \left[ \int_0^T L_v L_v' \right] \right)$ with $L_v = E_v [D_v (\partial g(X_T) D_v X_T C(X_v))]$.

**Proof of Lemma 6:** Let us first find the asymptotic variance of

$$H_{T_j}^{T_j} = \frac{g \left( X_T(x + \alpha_j \tau_j e_j) \right) - g \left( X_T(x - (1 - \alpha_j) \tau_j e_j) \right)}{\tau_j}.$$

As $g(X_T(x)) \in \mathbb{D}^{1,2}$, we can apply the Clark-Ocone formula to obtain

$$H_{T_j}^{T_j} - E_0[H_{T_j}^{T_j}] = \int_0^T E_v[D_v H_{T_j}^{T_j}] dW_v = \int_0^T (L_v^{T_j})' dW_v$$

and therefore, $\text{VAR}_0[H_{T_j}^{T_j}] = E_0 \left[ \left( H_{T_j}^{T_j} - H_{T_j}^{T_j} \right)_T \right] = \int_0^T E_0 [(L_v^{T_j})' (L_v^{T_j})] dv$.

Define $f(v, X_v(x)) \equiv E_{v,X_v(x)}[g(X_T(X_v(x)))].$ Commutativity of the Malliavin derivative and the conditional expectation gives

$$(L_v^{T_j})' = D_v E_{v}[H_{T_j}^{T_j}] = D_v \left( \frac{f(v, X_v(x + \alpha_j e_j \tau_j)) - f(v, X_v(x - (1 - \alpha_j) e_j \tau_j))}{\tau_j} \right)$$

$$= \frac{1}{\tau_j} \partial_x f(v, X_v(x + \alpha_j e_j \tau_j)) B(X_v(x + \alpha_j e_j \tau_j))$$

$$- \frac{1}{\tau_j} \partial_x f(v, X_v(x - (1 - \alpha_j) e_j \tau_j)) B(X_v(x - (1 - \alpha_j) e_j \tau_j)).$$

Using $X(x - (1 - \alpha_j) e_j \tau_j) - X(x + \alpha_j e_j \tau_j) \to 0$ $P_0$-a.s., as $\tau_j \to 0$ we obtain

$$(L_v^{T_j})' \left( \frac{\partial_x f(v, X_v(x + \alpha_j e_j \tau_j)) - \partial_x f(v, X_v(x - (1 - \alpha_j) e_j \tau_j))}{\tau_j} \right) B(X_v) \to 0$$
\( P_0 \)-a.s., as \( \tau_j \to 0 \), or, equivalently, \( (L^T_v)' - \partial^2_{x,x_j} f(v, X_v)B(X_v) \to 0 \) \( P_0 \)-a.s., as \( \tau_j \to 0 \). From \( \partial^2_{x,x_j} f(v, X_v)B(X_v) = D_v \partial x_j f(v, X_v) \), we conclude \( (L^T_v)' \to D_v \partial x_j f(v, X_v) \), \( P_0 \)-a.s. as \( \tau_j \to 0 \), and hence \( [H^\tau_j, H^\gamma]_T \to {\int}_0^T L_{jv} L_{jv} dv \) \( P_0 \)-a.s., with \( L_{jv} = D_v \partial x_j f(v, X_v) = E_v [D_v (\partial g(X_T) \nabla x_j X_T)] \).

By assumption (32), \( L^T_v \) is uniformly integrable so that almost sure convergence implies convergence of the mean. We conclude

\[
\lim_{\tau_j \to 0} \text{VAR}_0 [H^T_T] = E_0 \left[ \int_0^T L_{jv} L_{jv} ds \right].
\]

This analysis holds for all \( j = 1, \ldots, d \). With \( L_v = [L_{jv}]_{j=1,\ldots,d} \) we obtain the asymptotic variance

\[
\text{VAR}_0 [H^T_T]_{j=1,\ldots,d} = E_0 \left[ \int_0^T L_v L_v dv \right].
\]

The asymptotic distribution follows from the Lindeberg Central Limit theorem for independent random variables. The uniform integrability condition (32) ensures that the Lindeberg condition for application of this theorem is satisfied.

**Proof of Theorem 3 (continued):** Given the results of Lemmas 4, 5 and 6, we see that

\[
\partial x f(0, x)^M_{N,M,\tau_j,M} - \partial x f(0, x) = \begin{cases} O_P \left( \frac{1}{N_M} + \frac{1}{\sqrt{M}} + \left( \frac{1}{\tau_j,M} \right)^{-2} \right) & \text{if } \alpha_j = 1/2 \\
O_P \left( \frac{1}{N_M} + \frac{1}{\sqrt{M}} + \left( \frac{1}{\tau_j,M} \right)^{-1} \right) & \text{if } \alpha_j \neq 1/2. \end{cases}
\]

Choose \( 1/\tau_j,M = M^{\delta_j}/\varepsilon_j^\kappa \) and \( N_M = M^\gamma/\varepsilon_j^\delta \) for some \( \varepsilon_j^\kappa, \varepsilon_j^\delta, \delta_j, \gamma > 0 \) and \( \kappa \in \{ f c d, f d \} \).

If \( \alpha_j = 1/2 \) the efficient scheme is obtained for \( (\gamma, \delta_j) = \inf \{ \arg \max_{\gamma, \delta_j} \min \{ 2\delta_j, 1/2, \gamma \} \} = (1/2, 1/4) \). For \( \alpha_j \neq 1/2 \), \( (\gamma, \delta_j) = \inf \{ \arg \max_{\gamma, \delta_j} \min \{ \delta_j, 1/2, \gamma \} \} = (1/2, 1/2) \) provides the efficient scheme. This proves the theorem.

**Proof of Theorem 4:** The proof of Theorem 4 parallels the proof of Theorem 3. For the discontinuous functions under consideration, \( \sum_{j=1}^\infty \gamma_j 1_{B_j}(x) \in \mathbb{D}^{1,2} \), Lemma 6 is replaced by:

**Lemma 7:** Let \( R_{3j} (0, x; M, \tau_j) \) be defined by (B-10) and select \( \tau_j,M \) such that \( 1/\tau_j,M \to \infty \) when \( M \to \infty \). Suppose that the assumptions of Theorem 4 hold. For \( j = 1, \ldots, d \) and for all \( \alpha_j \in [0, 1] \) we have the limit, \( \sqrt{M\tau_j,M} R_{3j} (0, x; M, \tau_j,M) \Rightarrow Q^T_1 \) as \( M \to \infty \), where \( Q^T_1 \sim N (0, V_0(x)) \) with

\[
V_0(x) = \sum_{k=1}^\infty \gamma_k^2 (2\alpha_j - 1) \partial x_j P_0 (\{ X_T(x) \in B_k \})
\]

\[
-2\alpha_j \sum_{k,l=1}^\infty \gamma_k \gamma_l \partial x_j P_0 (\{ X_T(x) \in B_k \} \cap \{ X_T(x') \in B_l \}) |_{x' = x}
\]

\[
+2(1 - \alpha_j) \sum_{k,l=1}^\infty \gamma_k \gamma_l \partial x_j P_0 (\{ X_T(x) \in B_k \} \cap \{ X_T(x') \in B_l \}) |_{x' = x},
\]
To derive the asymptotic limits of the variance and the covariance we therefore need the limits of $g^c(x)$ and $g^d(x)$.

Using $g^c(x)g^d(x) = 0$, we obtain $\text{VAR}_0 \left[ H_T \right] = \text{VAR}_0 \left[ H_{T}^{c,\tau_j} \right] + \text{VAR}_0 \left[ H_{T}^{d,\tau_j} \right]$.

From the assumption $g^c(X_T) \in \mathbb{D}^{1,2}$ and Lemma 6, we get $\text{VAR}_0 \left[ H_{T}^{c,\tau_j} \right] = O((1/\tau_j)^{-1/2})$ when $1/\tau_j \to \infty$, where $O(\cdot)$ is defined in footnote 2.

Next, we show that $\lim_{\tau_i,\tau_j \to 0} \left( \text{COV}_0[H_T^{d,\tau_i}, H_T^{d,\tau_j}] - E_0 \left[ H_{T}^{d,\tau_i} H_{T}^{d,\tau_j} \right] \right) = 0$ for all $i, j$.

As $\text{COV}_0[H_T^{d,\tau_i}, H_T^{d,\tau_j}] = E_0 \left[ H_{T}^{d,\tau_i} H_{T}^{d,\tau_j} \right] - E_0 \left[ H_{T}^{d,\tau_i} \right] E_0 \left[ H_{T}^{d,\tau_j} \right]$, and $\frac{1}{\sqrt{\tau_j}} E_0[H_T^{c,\tau_j}] \to \partial_x f(0, x)$ we have $E_0[H_T^{d,\tau_i}]E_0[H_T^{d,\tau_j}] = O \left( ((1/\tau_i)(1/\tau_j))^{-1/2} \right)$, and therefore $E_0[H_T^{d,\tau_i}]E_0[H_T^{d,\tau_j}] \to 0$, $P_0$-a.s., as $1/\tau_i, 1/\tau_j \to \infty$.

Hence,

$$\text{VAR}_0[H_T^{d,\tau_j}] - E_0 \left[ \left( H_{T}^{d,\tau_j} \right)^2 \right] \to 0, P_0 - a.s., \quad \frac{1}{\tau_j} \to \infty \quad \text{(B-13)}$$

$$\text{COV}_0[H_T^{d,\tau_i}, H_T^{d,\tau_j}] - E_0 \left[ H_{T}^{d,\tau_i} H_{T}^{d,\tau_j} \right] \to 0, P_0 - a.s., \quad \frac{1}{\tau_i}, \frac{1}{\tau_j} \to \infty. \quad \text{(B-14)}$$

To derive the asymptotic limits of the variance and the covariance we therefore need the limits of $E \left[ \left( H_{T}^{d,\tau_j} \right)^2 \right]$ and $E \left[ H_{T}^{d,\tau_i} H_{T}^{d,\tau_j} \right]$.

To find the limits of $E_0 \left[ \left( H_{T}^{d,\tau_j} \right)^2 \right]$ and $E_0 \left[ H_{T}^{d,\tau_i} H_{T}^{d,\tau_j} \right]$, note that

$$H_{T}^{d,\tau_i} H_{T}^{d,\tau_j} = \frac{1}{\sqrt{\tau_i \tau_j}} g^d \left( X_T(x + \alpha_i \tau_i e_i) \right) g^d \left( X_T(x + \alpha_j \tau_j e_j) \right) \quad \text{(B-15)}$$

$$+ \frac{1}{\sqrt{\tau_i \tau_j}} g^d \left( X_T(x - (1 - \alpha_i) \tau_i e_i) \right) g^d \left( X_T(x - (1 - \alpha_j) \tau_j e_j) \right)$$

$$- \frac{1}{\sqrt{\tau_i \tau_j}} g^d \left( X_T(x + \alpha_i \tau_i e_i) \right) g^d \left( X_T(x - (1 - \alpha_j) \tau_j e_j) \right)$$

$$- \frac{1}{\sqrt{\tau_i \tau_j}} g^d \left( X_T(x - (1 - \alpha_i) \tau_i e_i) \right) g^d \left( X_T(x + \alpha_j \tau_j e_j) \right) \quad \text{(B-16)}$$

where $g^d \left( X_T(x) \right) g^d \left( X_T(x') \right) = \sum_{k,d=1}^{\infty} \gamma_k \gamma_d 1_{B_k}(X_T(x))1_{B_d}(X_T(x'))$. With the definitions $p_k(x) \equiv P_0(X_T(x) \in B_k)$ and $q_k(x, x') \equiv P_0 (\{X_T(x) \in B_k\} \cap \{X_T(x') \in B_l\})$ and using $p_k(x) = q_{kk}(x, x)$.
and \( q_{kl}(x,x) = 0 \) for \( k \neq l \), we obtain
\[
E_0 \left[ \left( H_T^{d,\tau_j} \right)^2 \right] = \sum_{k=1}^{\infty} \gamma_k^2 \left( \frac{p_k(x + \alpha_j e_j \tau_j) - p_k(x)}{\tau_j} + \frac{p_k(x - (1 - \alpha_j) e_j \tau_j) - p_k(x)}{\tau_j} \right) - 2 \sum_{k=1}^{\infty} \gamma_k^2 \left( \frac{q_{kk}(x + \alpha_j e_j \tau_j, x - (1 - \alpha_j) e_j \tau_j) - q_{kk}(x,x)}{\tau_j} \right) - 2 \sum_{k,l=1, k \neq l}^{\infty} \gamma_k \gamma_l \left( \frac{q_{kl}(x + \alpha_j e_j \tau_j, x - (1 - \alpha_j) e_j \tau_j) - q_{kl}(x,x)}{\tau_j} \right)
\]
and therefore with \( V_0^{\tau_j}(x) \equiv \text{VAR}_0[H_T^{d,\tau_j}] \),
\[
V_0^{\tau_j}(x) \to \sum_{k=1}^{\infty} \gamma_k^2 (2\alpha_j - 1) \partial_{x_j} p_k(x) - 2 \sum_{k,l=1}^{\infty} \gamma_k \gamma_l \left( \alpha_j \partial_{x_j} q_{kl}(x, x') |_{x' = x} - (1 - \alpha_j) \partial_{x_j} q_{kl}(x, x') |_{x' = x} \right),
\]
P_0\text{-a.s., as } 1/\tau_j \to \infty.
Next, we show that \( E_0[H_T^{d,\tau_i} H_T^{d,\tau_j}] \to 0 \) P_0\text{-a.s. when } i \neq j \text{ and } 1/\tau_i, 1/\tau_j \to \infty. Using (B-15) we obtain
\[
E_0[H_T^{d,\tau_i} H_T^{d,\tau_j}] = \sum_{k,l=1}^{\infty} \gamma_k \gamma_l h_{kl}(x, \alpha, \tau_i, \tau_j)
\]
where
\[
h_{kl}(x, \alpha, \tau_i, \tau_j) \equiv \frac{q_{kl}(x + \alpha_i e_i \tau_i, x + \alpha_j e_j \tau_j) - q_{kl}(x,x)}{\sqrt{\tau_i \tau_j}} + \frac{q_{kl}(x - (1 - \alpha_i) e_i \tau_i, x - (1 - \alpha_j) e_j \tau_j) - q_{kl}(x,x)}{\sqrt{\tau_i \tau_j}} - \frac{q_{kl}(x + \alpha_i e_i \tau_i, x - (1 - \alpha_j) e_j \tau_j) - q_{kl}(x,x)}{\sqrt{\tau_i \tau_j}} - \frac{q_{kl}(x - (1 - \alpha_i) e_i \tau_i, x + \alpha_j e_j \tau_j) - q_{kl}(x,x)}{\sqrt{\tau_i \tau_j}}.
\]
It follows that
\[
h_{kl}(x, \alpha, \tau_i, \tau_j) = \alpha_i \partial_{x_i} q_{kl}(x,x) \sqrt{\frac{\tau_i}{\tau_j}} + \alpha_j \partial_{x_j} q_{kl}(x,x') |_{x' = x} \sqrt{\frac{\tau_j}{\tau_i}} - (1 - \alpha_i) \partial_{x_i} q_{kl}(x,x) \sqrt{\frac{\tau_i}{\tau_j}} - (1 - \alpha_j) \partial_{x_j} q_{kl}(x,x') |_{x' = x} \sqrt{\frac{\tau_j}{\tau_i}} - \alpha_i \partial_{x_i} q_{kl}(x,x) \sqrt{\frac{\tau_i}{\tau_j}} + (1 - \alpha_j) \partial_{x_j} q_{kl}(x,x') |_{x' = x} \sqrt{\frac{\tau_j}{\tau_i}} + (1 - \alpha_i) \partial_{x_i} q_{kl}(x,x) \sqrt{\frac{\tau_i}{\tau_j}} - \alpha_j \partial_{x_j} q_{kl}(x,x') |_{x' = x} \sqrt{\frac{\tau_j}{\tau_i}} + o(1) = o(1).
\]
Hence $h_{kl}(x, \alpha, \tau_i, \tau_j) \to 0, P_0$-a.s. as $1/\tau_i, 1/\tau_j \to \infty$. This shows that $H^{d, \tau_i}_T$ and $H^{d, \tau_j}_T$ are asymptotically uncorrelated, for $i \neq j$.

The asymptotic distribution follows from the Lindeberg Central Limit theorem for independent random variables: the uniform integrability condition (32) ensures that Lindeberg’s condition is satisfied. ■

**Proof of Theorem 4 (continued):** The results of Lemmas 4, 5 and 7 give

$$
\hat{\partial_x f}(0, x)_{M,N,\tau_j,M} - \partial_x f(0, x) = \begin{cases} 
O_P \left( \frac{1}{N^M} + \frac{1}{\sqrt{M \tau_j,M}} + \left( \frac{1}{\tau_j,M} \right)^{-2} \right) & \text{if } \alpha_j = 1/2 \\
O_P \left( \frac{1}{N^M} + \frac{1}{\sqrt{M \tau_j,M}} + \left( \frac{1}{\tau_j,M} \right)^{-1} \right) & \text{if } \alpha_j \neq 1/2.
\end{cases}
$$

Choose $1/\tau_j,M = M^\beta/\varepsilon_j^\gamma_1$ and $N_M = M^\gamma/\varepsilon_j^\gamma_2$ for some $\varepsilon_j^\gamma_1, \varepsilon_j^\gamma_2, \delta_j, \gamma > 0$ and $\kappa \in \{fcd, fd\}$. For $\alpha_j = 1/2$, the efficient scheme is attained for $(\gamma, \delta_j) = \inf \{ \arg \max_{\delta_j, \gamma} \min \{2 \delta_j, (1 - \delta_j)/2, \gamma \} \} = (2/5, 1/5)$. For $\alpha_j \neq 1/2$, $(\gamma, \delta_j) = \inf \{ \arg \max_{\gamma, \delta_j} \min \{ \delta_j, (1 - \delta_j)/2, \gamma \} \} = (1/3, 1/3)$ provides the efficient scheme. This proves the theorem. ■

**Appendix C: Abstract Integration by Parts**

This Appendix shows how to use an integration by parts argument to handle a discontinuous payoff function when the transition density is unknown. Consider a diffusion with volatility coefficient $B \in C^1(\mathbb{R}^n \times \mathbb{R}^d)$ such that $\text{rank}(B(x)) = n$ for all $x \in \mathbb{R}^n$ and define the adapted shift on the Wiener space, $\Omega = C^0([0, T]; \mathbb{R}^d)$,

$$
\theta^\lambda_v(\omega)_v = \omega_v + \int_t^v \nu_{t,s}(\omega) \lambda ds.
$$

This perturbation of the state space depends on a parameter $\lambda \in \mathbb{R}^n$ and a progressively measurable $\mathbb{R}^n \times \mathbb{R}^d$-valued matrix process $\nu_{t,s}$. To obtain our integration by parts formula we pass to a new measure $P_{t,x}^\lambda$ under which $W_v + \int_t^v \nu_{t,s} \lambda ds, \ v \geq t$, is a standard Brownian motion process. The existence of this probability measure requires, by Girsanov’s theorem,

$$
E_{t,x} \left[ \mathcal{E} \left( - \int_t^v (\nu_{t,s} \lambda)' dW_s \right) \right] = 1, \quad \text{for all } v \geq t.
$$

Under this condition we can define $\frac{dP_{t,x}^\lambda}{dP_{t,x}} \equiv \mathcal{E} \left( - \int_t^T (\nu_{t,s} \lambda)' dW_s \right)$.

---

3 The presentation is based on Bismut’s approach to Malliavin calculus (see Bichteler, Gravereaux and Jacod (1987) for a comparison of this approach with Malliavin’s original approach).
Taking derivatives with respect to the parameter \( \lambda \) on both sides of the equality
\[
\mathbb{E}_{t,x} \left[ \frac{d \mathbb{P}_{t,x}^\lambda}{d \mathbb{P}_{t,x}^T} h(X_T(\theta^\lambda)) \right] = \mathbb{E}_{t,x} [h(X_T)]
\] (C-17)
and evaluating the resulting expressions at \( \lambda = 0 \), gives, for \( h \in \mathcal{C}^1(\mathbb{R}^d) \)
\[
\mathbb{E}_{t,x} \left[ \left( \frac{d \mathbb{P}_{t,x}^\lambda}{d \mathbb{P}_{t,x}^T} h(X_T(\theta^\lambda)) \right) \right] = 0.
\]
Using
\[
\frac{d \mathbb{P}_{t,x}^\lambda}{d \mathbb{P}_{t,x}^T} = \frac{d \mathbb{P}_{t,x}^\lambda}{d \mathbb{P}_{t,x}^T} \partial_\lambda \left( \int_t^T (\nu_{t,s}^\lambda) dW_s - \frac{1}{2} \int_t^T ||\nu_{t,s}^\lambda||^2 ds \right)
\]
and
\[
\frac{d \mathbb{P}_{t,x}^\lambda}{d \mathbb{P}_{t,x}^T} \partial_\lambda \left( \int_t^T (\nu_{t,s}^\lambda) dW_s - \frac{1}{2} \int_t^T ||\nu_{t,s}^\lambda||^2 ds \right) = 1
\]
gives
\[
\frac{d \mathbb{P}_{t,x}^\lambda}{d \mathbb{P}_{t,x}^T} \frac{d \mathbb{P}_{t,x}^{\lambda}}{d \mathbb{P}_{t,x}^T} = - \int_t^T dW_s \nu_{t,s}.
\]
Substituting in (C-17) yields the abstract integration by parts formula,
\[
\mathbb{E}_{t,x} \left[ \left( \int_t^T (dW_s)^\lambda \right) h(X_T(\theta^\lambda)) \right] = \mathbb{E}_{t,x} \left[ \partial h(X_T) \left( \partial_\lambda X_T(\theta^\lambda) \right) \right].
\] (C-18)
To identify \( \left( \partial_\lambda X_T(\theta^\lambda) \right) \) differentiate the integral representation
\[
X_T(\theta^\lambda) = x + \int_0^T A(X_s(\theta^\lambda)) ds + \sum_{j=1}^d \int_0^T B_j(X_s(\theta^\lambda)) (dW_s^j + \nu_{t,s}^\lambda ds)
\]
with respect to \( \lambda \), to obtain
\[
\left( \partial_\lambda X_T(\theta^\lambda) \right) = \int_0^T \left( \partial A(X_s) ds + \sum_{j=1}^d \partial B_j(X_s) dW_s^j \right) \left( \partial_\lambda X_s(\theta^\lambda) \right) + \int_0^T B(X_s) \nu_{t,s} ds.
\]
The solution of this linear SDE is
\[
\left( \partial_\lambda X_T(\theta^\lambda) \right) = \nabla_{t,x} X_T(X_t) \int_t^T (\nabla_{t,x} X_s(X_t))^{-1} B(X_s) \nu_{t,s} ds
\]
where \( \nabla_{t,x} X_T(X_t) \) is the derivative process. \( \nabla_{t,x} X_T(X_t) \) is the tangent process.

For \( C \) defined in (13), selecting the process \( \nu_{t,s}^{(\alpha)} \equiv C(X_s) \nabla_{t,x} X_s(X_t) \alpha_{t,s} \) for some progressively measurable process \( \alpha_{t,s} \), such that \( \int_0^T \alpha_{t,s} ds = I_n \), and substituting in (C-18) gives (18), that is
\[
\mathbb{E}_{t,x} \left[ \left( \int_t^T (dW_s)^\lambda C(X_s) \nabla_{t,x} X_s(X_t) \right) h(X_T) \right] = \mathbb{E}_{t,x} \left[ \partial h(X_T) \nabla_{t,x} X_T(X_t) \right].
\]
To establish the same result when \( h \in \mathbb{L}^2 \) is not continuous, we use the fact that any square integrable function can be approximated by a sequence of infinitely differentiable functions with compact support. This enables us to proceed by first using the approximating sequence and then passing to the limit. See Fournié et. al. (1999) for detailed arguments.

---

4Note that \( \left( \partial_\lambda X_T(\theta^\lambda) \right) \) exists whenever the stochastic flow is differentiable with respect to the initial condition. A sufficient condition for this is \( X_T \in \mathcal{D}^{1,2} \), which is satisfied when the diffusion is in the domain of the Malliavin derivative operator. See Bichteler, Gravereaux, Jacod (1987) for more on the differentiability of perturbed diffusions.
References


