ALEA, Lat. Am. J. Probab. Math. Stat. 18, 309–324 (2021)
DOI: 10.30757/ALEA.v18-14

Differentiability of semigroups of stochastic differential equations with Hölder-continuous diffusion coefficients

Martin Hutzenthaler and Daniel Pieper

Faculty of Mathematics, University of Duisburg-Essen, 45117 Essen, Germany
E-mail address: martin.hutzenthaler@uni-due.de , daniel.pieper@uni-due.de

Abstract. Differentiability of semigroups is useful for many applications. Here we focus on stochastic differential equations whose diffusion coefficient is the square root of a differentiable function but not differentiable itself. For every $m \in \{0,1,2\}$ we establish an upper bound for a C^m-norm of the semigroup of such a diffusion in terms of the C^m-norms of the drift coefficient and of the squared diffusion coefficient. The constants in our upper bound are often bounded in the dimension. Our estimates are thus suitable for analyzing certain high-dimensional and infinite-dimensional degenerate stochastic differential equations.

1. Introduction

Let $d \in \mathbb{N}$ and let $X = (X_t)_{t\in[0,\infty)}$ be the solution of a stochastic differential equation (SDE)

$$dX_t(i) = b_i(X_t)\,dt + \sqrt{a_i(X_t(i))}\,dW_t(i), \quad i \in \{1,\dots,d\}, \tag{1.1}$$

with values in $[0,1]^d$. We prove existence and continuity of spatial derivatives of the functions $[0,\infty) \times [0,1]^d \ni (t,x) \mapsto (T_t f)(x) := \mathbb{E}[f(X_t) \mid X_0 = x] \in \mathbb{R}$, $f \in C^2([0,1]^d,\mathbb{R})$, under suitable assumptions. We focus on derivatives up to order 2 since these are needed for Itô's formula. More precisely, Theorem 4.1 below shows under suitable assumptions for every $t \in [0,\infty)$ and every $m \in \{0,1,2\}$ that

$$\|T_t f\|_{C^m} \le e^{(m^2\lambda_m+\mu_m)t}\|f\|_{C^m}, \tag{1.2}$$

where λ_m and μ_m depend respectively on the partial derivatives of the drift function and of the squared diffusion function up to order m and where for every $f \in$

Received by the editors January 17th, 2020; accepted November 11th, 2020.
2010 *Mathematics Subject Classification.* 60J35 (Primary); 47D07, 60J60 (Secondary).
Key words and phrases. Markov semigroups, Feller semigroups, diffusion processes, square-root diffusions, stochastic differential equations, C^m-estimate, generators, degenerate operators, differentiability.

$C^m([0,1]^d,\mathbb{R})$ we define

$$\|f\|_{C^m} := \max_{\alpha\in\mathbb{N}_0^d,|\alpha|\leq m} \|\tfrac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1}\cdots\partial x_d^{\alpha_d}} f\|_\infty. \tag{1.3}$$

In particular, note that we do not assume differentiability of the diffusion coefficient but only of the squared diffusion coefficient. The "cost" of allowing square-root diffusions is that we need to assume the diffusion coefficient matrix to be diagonal. We also note that even differentiability of the semigroup is nontrivial since singular diffusion coefficients (that is, degenerate noise) can lead to loss of regularity; see Theorem 1.2 in Hairer et al. (2015).

Partial differentiability of semigroups is used in a number of applications, e.g.:

- inequalities between expectations of diffusions with different coefficient functions, e.g. Theorem 1 in Cox et al. (1996) or Proposition 2.2 in Hutzenthaler and Wakolbinger (2007),
- weak convergence rates for numerical approximations of SDEs, e.g. Theorem 1 in Talay and Tubaro (1990),
- stochastic representations of quasilinear parabolic partial differential equations, e.g. Theorem 3.2 in Peng (1991),

and many more. These results can now also be derived for those SDEs for which we establish differentiability of the semigroup.

In the literature, differentiability of semigroups is well-known in the case of differentiable coefficient functions of suitable order (see, e.g., Theorem 8.4.3 in Gikhman and Skorokhod, 1969) and in the case of one-dimensional SDEs including the case of square-root diffusion coefficients (see, e.g., Dorea, 1976 or Ethier, 1978). Moreover, Ethier (1976) establishes differentiability of semigroups for a class of multidimensional SDEs with square-root diffusion coefficient $\{y\in[0,1]^d\colon \sum_{i=1}^d y_i\leq 1\}\ni x\mapsto (\sqrt{x_i(1-\sum_{j=1}^d x_j)})_{i\in\{1,\dots,d\}}\in\mathbb{R}^d$. In addition, Lemma 4.3 in Epstein and Pop (2019) establishes differentiability of semigroups corresponding to so-called Kimura operators. So differentiability of semigroups corresponding to degenerate SDEs is in principle known in the literature. However, we have not found a result on differentiability of semigroups corresponding to the specific form of the SDE (4.1) beyond the one-dimensional case.

In fact, differentiability of semigroups of degenerate SDEs is not our main concern. Our main goal is to establish the regularity estimates (1.2) with constants $\lambda_0,\lambda_1,\lambda_2,\mu_0,\mu_1,\mu_2$ that are bounded in the dimension. This dimension-independence of regularity estimates of semigroups of degenerate stochastic differential equations seems to be a new observation. The benefit of such estimates with dimension-independent constants is that it allows us to analyze infinite-dimensional (where $d=\infty$) or high-dimensional (where $d\to\infty$) SDEs. To mention an example application, our main result, Theorem 4.1 below, is applied in Hutzenthaler and Pieper (2020) to a system of interacting diffusions on $D\in\mathbb{N}$ demes to obtain that the partial derivatives of the semigroups are uniformly bounded in $D\in\mathbb{N}$; see Example 4.2 below for details. This then allows to establish a many-demes limit as $D\to\infty$, that is, to generalize Theorem 3.3 in Hutzenthaler (2012) to a class of SDEs with nonlinear squared diffusion coefficients. In addition, by approximation with finite-dimensional SDEs, Theorem 4.1 can also be applied to McKean-Vlasov SDEs (e.g. (1.2) with $g(x)=x(1-x)$ in Dawson and Greven, 1993 or (1.2) in Hutzenthaler, 2012 or (8) in Hutzenthaler et al., 2015).

An important technical insight of this paper is as follows. Results in the literature are often (e.g., Ethier, 1976 or Epstein and Pop, 2019 with the domain suitably replaced) formulated in the norms

$$C^m([0,1]^d,\mathbb{R}) \ni f \mapsto |||f|||_{C^m([0,1]^d,\mathbb{R})} := \sum_{\alpha\in\mathbb{N}_0^d, |\alpha|\le m} \sup_{x\in[0,1]^d} |\partial^\alpha f(x)|. \tag{1.4}$$

This norm, however, introduces unnecessary dimension-dependence due to the sum in (1.4). To give an illustrative example, if the drift coefficient is $[0,1]^d \ni x \mapsto x \in \mathbb{R}^d$, if the diffusion coefficient is zero, and if $f \in C^1(\mathbb{R},\mathbb{R})$, then the solution of the SDE (4.1) is $(x_i e^t)_{t\in[0,\infty), i\in\{1,\ldots,d\}}$ and it holds for all $t\in[0,\infty)$ that

$$\begin{aligned}
&\left|\left|\left|[0,1]^d \ni x \mapsto f\Big(\sum_{i=1}^d x_i e^t\Big) \in \mathbb{R}\right|\right|\right|_{C^1([0,1]^d,\mathbb{R})} \\
&= \sup_{x\in[0,1]^d}\left|f\Big(\sum_{i=1}^d x_i e^t\Big)\right| + \sum_{k=1}^d \sup_{x\in[0,1]^d}\left|f'\Big(\sum_{i=1}^d x_i e^t\Big)e^t\right| \\
&= \sup_{z\in\mathbb{R}}|f(z)| + d\sup_{z\in\mathbb{R}}|f'(z)|e^t.
\end{aligned} \tag{1.5}$$

If the norm $|||\cdot|||_{C^1([0,1]^d,\mathbb{R})}$ is replaced by our norm $\|\cdot\|_{C^1}$ where the sum in (1.4) is replaced by the maximum, then $\|[0,1]^d \ni x \mapsto f\big(\sum_{i=1}^d x_i e^t\big) \in \mathbb{R}\|_{C^1}$ does not depend on the dimension.

1.1. *Notation.* We write $\mathbb{N}_0 := \{0,1,2,\ldots\}$ and $\mathbb{N} := \mathbb{N}_0 \setminus \{0\}$. For every topological space $(E,\mathcal{E})$ we denote by $\mathcal{B}(E)$ the Borel σ-algebra on $(E,\mathcal{E})$. For every $d\in\mathbb{N}$ and every $m\in\mathbb{N}_0$ we denote by $C^m([0,1]^d,\mathbb{R})$ the set of functions $f\colon [0,1]^d\to\mathbb{R}$ whose partial derivatives of order 0 through m exist and are continuous on $[0,1]^d$. For every $d\in\mathbb{N}$ and every $f\colon[0,1]^d\to\mathbb{R}$ we define $\|f\|_\infty := \sup_{x\in[0,1]^d}|f(x)| \in [0,\infty]$. For every $d\in\mathbb{N}$ and every multiindex $\alpha=(\alpha_1,\ldots,\alpha_d)\in\mathbb{N}_0^d$ of length $|\alpha| := \sum_{k=1}^d \alpha_k$ we write $\partial^\alpha := \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1}\cdots\partial x_d^{\alpha_d}}$. For every $d\in\mathbb{N}$, every $m\in\mathbb{N}_0$, and every $f\in C^m([0,1]^d,\mathbb{R})$ we define $\|f\|_{C^m} := \max_{\alpha\in\mathbb{N}_0^d, |\alpha|\le m}\|\partial^\alpha f\|_\infty$. For every $d\in\mathbb{N}$, every $x=(x_k)_{k\in\{1,\ldots,d\}}\in[0,1]^d$, and every $i,j\in\{1,\ldots,d\}$ we write $\hat{x}_i := (x_k)_{k\in\{1,\ldots,d\}\setminus\{i\}}$ and $\hat{x}_{ij} := (x_k)_{k\in\{1,\ldots,d\}\setminus\{i,j\}}$.

2. Drift part

In this section, we prove (1.2) for $m\in\{0,1,2\}$ and an analogous result for the $\|\cdot\|_{C^3}$-norm under suitable assumptions in the case where the diffusion coefficient is zero. The case of non-zero diffusion coefficients is analyzed in Section 3.

Lemma 2.1 (C^m-esimate for drift part)**.** *Let $d\in\mathbb{N}$, let $b_1,\ldots,b_d\in C^3([0,1]^d,\mathbb{R})$ satisfy for all $i\in\{1,\ldots,d\}$ and all $x=(x_1,\ldots,x_d)\in[0,1]^d$ with $x_i\in\{0,1\}$ that $(-1)^{x_i}b_i(x)\ge 0$, for every $m\in\{1,2,3\}$ we define $\lambda_m := \max_{\alpha\in\mathbb{N}_0^d, 0<|\alpha|\le m}\sum_{i=1}^d\|\partial^\alpha b_i\|_\infty$, let $y=(y_1,\ldots,y_d)\colon[0,\infty)\times[0,1]^d\to[0,1]^d$ satisfy for all $i\in\{1,\ldots,d\}$, all $t\in[0,\infty)$, and all $x=(x_1,\ldots,x_d)\in[0,1]^d$ that*

$$y_i(t,x) = x_i + \int_0^t b_i(y(s,x))\,ds, \tag{2.1}$$

let $c_1 = 1$, $c_2 = 4$, $c_3 = 13$, *and let* $\{T_t^1 \colon t \in [0,\infty)\}$ *satisfy for all* $t \in [0,\infty)$, *all* $f \in C([0,1]^d, \mathbb{R})$, *and all* $x \in [0,1]^d$ *that* $(T_t^1 f)(x) = (f \circ y)(t,x)$. *Then it holds for all* $m \in \{1,2,3\}$, *all* $f \in C^m([0,1]^d, \mathbb{R})$, *and all* $t \in [0,\infty)$ *that* $T_t^1 f \in C^m([0,1]^d, \mathbb{R})$ *and*

$$\|T_t^1 f\|_{C^m} \le e^{c_m \lambda_m t} \|f\|_{C^m}. \tag{2.2}$$

Proof: For the rest of the proof fix $m \in \{1,2,3\}$ and $f \in C^m([0,1]^d, \mathbb{R})$. The theory of ordinary differential equations yields for all $t \in [0,\infty)$ that $y(t,\,\cdot\,) \in C^m([0,1]^d, [0,1]^d)$ (see, e.g. Corollary V.4.1 in Hartman, 2002) and this together with $f \in C^m([0,1]^d, \mathbb{R})$ implies that $T_t^m f \in C^1([0,1]^d, \mathbb{R})$.

Case $m = 1$**:** The dominated convergence theorem and (2.1) imply for all $i, j \in \{1,\ldots,d\}$, all $t \in [0,\infty)$, and all $x \in [0,1]^d$ that

$$\frac{\partial y_i}{\partial x_j}(t,x) = \mathbb{1}_{i=j} + \int_0^t \sum_{k=1}^d \frac{\partial b_i}{\partial y_k}(y(s,x)) \frac{\partial y_k}{\partial x_j}(s,x)\, ds. \tag{2.3}$$

It follows for all $j \in \{1,\ldots,d\}$, all $t \in [0,\infty)$, and all $x \in [0,1]^d$ that

$$\begin{aligned} \sum_{i=1}^d \left| \frac{\partial y_i}{\partial x_j}(t,x) \right| &\le 1 + \int_0^t \sum_{k=1}^d \left(\sum_{i=1}^d \left| \frac{\partial b_i}{\partial y_k}(y(s,x)) \right| \right) \left| \frac{\partial y_k}{\partial x_j}(s,x) \right| ds \\ &\le 1 + \int_0^t \left(\max_{\alpha \in \mathbb{N}_0^d, |\alpha|=1} \sum_{i=1}^d \|\partial^\alpha b_i\|_\infty \right) \left(\sum_{k=1}^d \left| \frac{\partial y_k}{\partial x_j}(s,x) \right| \right) ds \\ &= 1 + \int_0^t \lambda_1 \sum_{k=1}^d \left| \frac{\partial y_k}{\partial x_j}(s,x) \right| ds. \end{aligned} \tag{2.4}$$

This and Gronwall's inequality yield for all $j \in \{1,\ldots,d\}$, all $t \in [0,\infty)$, and all $x \in [0,1]^d$ that

$$\sum_{i=1}^d \left| \frac{\partial y_i}{\partial x_j}(t,x) \right| \le e^{\lambda_1 t}. \tag{2.5}$$

It follows from the chain rule and from (2.5) for all $j \in \{1,\ldots,d\}$, all $t \in [0,\infty)$, and all $x \in [0,1]^d$ that

$$\left| \frac{\partial (f \circ y)}{\partial x_j}(t,x) \right| = \left| \sum_{i=1}^d \frac{\partial f}{\partial y_i}(y(t,x)) \frac{\partial y_i}{\partial x_j}(t,x) \right| \le \|f\|_{C^1} \sum_{i=1}^d \left| \frac{\partial y_i}{\partial x_j}(t,x) \right| \le e^{\lambda_1 t} \|f\|_{C^1}. \tag{2.6}$$

Together with the fact that $\sup_{t \in [0,\infty)} \|T_t^1 f\|_\infty \le \|f\|_\infty$, this implies for all $t \in [0,\infty)$ that

$$\|T_t^1 f\|_{C^1} = \max \left\{ \|T_t^1 f\|_\infty, \max_{j \in \{1,\ldots,d\}} \sup_{x \in [0,1]^d} \left| \frac{\partial (f \circ y)}{\partial x_j}(t,x) \right| \right\} \le e^{\lambda_1 t} \|f\|_{C^1}. \tag{2.7}$$

Case $m = 2$: The dominated convergence theorem and (2.1) imply for all $i, j, k \in \{1, \ldots, d\}$, all $t \in [0, \infty)$, and all $x \in [0,1]^d$ that

$$\begin{aligned}
&\frac{\partial^2 y_i}{\partial x_k \partial x_j}(t,x) \\
&= \int_0^t \sum_{l,m=1}^d \frac{\partial^2 b_i}{\partial y_m \partial y_l}(y(s,x)) \frac{\partial y_m}{\partial x_k}(s,x) \frac{\partial y_l}{\partial x_j}(s,x) + \sum_{l=1}^d \frac{\partial b_i}{\partial y_l}(y(s,x)) \frac{\partial^2 y_l}{\partial x_k \partial x_j}(s,x)\, ds.
\end{aligned} \tag{2.8}$$

This, (2.5), and $\lambda_1 \leq \lambda_2$ imply for all $j, k \in \{1, \ldots, d\}$, all $t \in [0, \infty)$, and all $x \in [0,1]^d$ that

$$\begin{aligned}
&\sum_{i=1}^d \left| \frac{\partial^2 y_i}{\partial x_k \partial x_j}(t,x) \right| \\
&\leq \int_0^t \sum_{l,m=1}^d \left(\sum_{i=1}^d \left| \frac{\partial^2 b_i}{\partial y_m \partial y_l}(y(s,x)) \right| \right) \left| \frac{\partial y_m}{\partial x_k}(s,x) \right| \left| \frac{\partial y_l}{\partial x_j}(s,x) \right| \\
&\quad + \sum_{l=1}^d \left(\sum_{i=1}^d \left| \frac{\partial b_i}{\partial y_l}(y(s,x)) \right| \right) \left| \frac{\partial^2 y_l}{\partial x_k \partial x_j}(s,x) \right| ds \\
&\leq \int_0^t \left(\max_{\alpha \in \mathbb{N}_0^d, |\alpha|=2} \sum_{i=1}^d \|\partial^\alpha b_i\|_\infty \right) \left(\sum_{m=1}^d \left| \frac{\partial y_m}{\partial x_k}(s,x) \right| \right) \left(\sum_{l=1}^d \left| \frac{\partial y_l}{\partial x_j}(s,x) \right| \right) \\
&\quad + \left(\max_{\alpha \in \mathbb{N}_0^d, |\alpha|=1} \sum_{i=1}^d \|\partial^\alpha b_i\|_\infty \right) \left(\sum_{l=1}^d \left| \frac{\partial^2 y_l}{\partial x_k \partial x_j}(s,x) \right| \right) ds \\
&\leq \int_0^t \lambda_2 e^{2\lambda_2 s} + \lambda_2 \sum_{l=1}^d \left| \frac{\partial^2 y_l}{\partial x_k \partial x_j}(s,x) \right| ds \\
&= \tfrac{1}{2}(e^{2\lambda_2 t} - 1) + \int_0^t \lambda_2 \sum_{l=1}^d \left| \frac{\partial^2 y_l}{\partial x_k \partial x_j}(s,x) \right| ds.
\end{aligned} \tag{2.9}$$

This and Gronwall's inequality yield for all $j, k \in \{1, \ldots, d\}$, all $t \in [0, \infty)$, and all $x \in [0,1]^d$ that

$$\sum_{i=1}^d \left| \frac{\partial^2 y_i}{\partial x_k \partial x_j}(t,x) \right| \leq \tfrac{1}{2}(e^{2\lambda_2 t} - 1) e^{\lambda_2 t}. \tag{2.10}$$

It follows from the chain rule, (2.5), $\lambda_1 \leq \lambda_2$, and from (2.10) for all $j, k \in \{1, \ldots, d\}$, all $t \in [0, \infty)$, and all $x \in [0,1]^d$ that

$$\begin{aligned}
&\left| \frac{\partial^2 (f \circ y)}{\partial x_k \partial x_j}(t,x) \right| \\
&\leq \left| \sum_{i,l=1}^d \frac{\partial^2 f}{\partial y_l \partial y_i}(y(t,x)) \frac{\partial y_l}{\partial x_k}(t,x) \frac{\partial y_i}{\partial x_j}(t,x) \right| + \left| \sum_{i=1}^d \frac{\partial f}{\partial y_i}(y(t,x)) \frac{\partial^2 y_i}{\partial x_k \partial x_j}(t,x) \right| \\
&\leq \|f\|_{C^2} \left(\sum_{l=1}^d \left| \frac{\partial y_l}{\partial x_k}(t,x) \right| \right) \left(\sum_{i=1}^d \left| \frac{\partial y_i}{\partial x_j}(t,x) \right| \right) + \|f\|_{C^2} \sum_{i=1}^d \left| \frac{\partial^2 y_i}{\partial x_k \partial x_j}(t,x) \right|
\end{aligned}$$

$$\le \left(e^{2\lambda_2 t} + \tfrac{1}{2}(e^{2\lambda_2 t} - 1)e^{\lambda_2 t}\right)\|f\|_{C^2}$$
$$\le \left(e^{2\lambda_2 t} + (e^{2\lambda_2 t} - 1)e^{2\lambda_2 t}\right)\|f\|_{C^2} = e^{4\lambda_2 t}\|f\|_{C^2}.$$

Together with the case $m = 1$ and $\lambda_1 \le \lambda_2$, this shows for all $t \in [0, \infty)$ that

$$\|T_t^1 f\|_{C^2} = \max\left\{\|T_t^1 f\|_{C^1}, \max_{j,k\in\{1,\ldots,d\}} \sup_{x\in[0,1]^d}\left|\frac{\partial^2(f\circ y)}{\partial x_k \partial x_j}(t,x)\right|\right\} \le e^{4\lambda_2 t}\|f\|_{C^2}. \tag{2.11}$$

Case $m = 3$**:** The proof of the case $m = 3$ is analogous to the case $m = 2$ and therefore omitted. This finishes the proof of Lemma 2.1. □

3. Diffusion part

The goal of this section is to prove (1.2) for $m \in \{0, 1, 2, 3\}$ under suitable assumptions in the case where the drift coefficient is zero; see Lemma 3.8 below. For that, we first look at the one-dimensional case in Subsection 3.1 below, and then we lift this result to the multidimensional case in Subsection 3.2 below.

3.1. *One-dimensional case.* The following lemma on smoothness preservation of the semigroup is well-known if, for $m \in \{0, 1, 2, 3\}$, the norm $\|\cdot\|_{C^m}$ is replaced by the equivalent norm $\varphi \mapsto \sum_{k=0}^m \|\frac{d^k\varphi}{dx^k}\|_\infty$; see Dorea (1976). The proof of the new upper bound of the operator norm of the semigroup with respect to $\|\cdot\|_{C^m}$ for $m \in \{0, 1, 2, 3\}$ is a straightforward adaptation of the proofs in Dorea (1976).

Lemma 3.1 (Smoothness preservation of one-dimensional diffusive part)**.** *Let $a \in C^3([0,1],\mathbb{R})$ satisfy that $a(0) = 0 = a(1)$ and for all $x \in (0,1)$ that $a(x) > 0$, let $A\colon C^2([0,1],\mathbb{R}) \to C([0,1],\mathbb{R})$ satisfy for all $\varphi \in C^2([0,1],\mathbb{R})$ and all $x \in [0,1]$ that*

$$(A\varphi)(x) = \frac{1}{2}a(x)\frac{d^2\varphi}{dx^2}(x), \tag{3.1}$$

for all $m \in \mathbb{N}_0$ we define $\mathcal{D}_m(A) := C^2([0,1],\mathbb{R}) \cap C^m([0,1],\mathbb{R}) \cap A^{-1}C^m([0,1],\mathbb{R})$, we define $\nu_0 := 0$, $\nu_1 := 0$, $\nu_2 := \frac{1}{2}\|\frac{d^2a}{dx^2}\|_\infty$, and $\nu_3 := \|\frac{d^3a}{dx^3}\|_\infty + \frac{3}{2}\|\frac{d^2a}{dx^2}\|_\infty$, and we denote by $\{S_t\colon t \in [0,\infty)\}$ the strongly continuous contraction semigroup on $C([0,1],\mathbb{R})$ generated by $(A, \mathcal{D}_0(A))$; see Theorem 1 *on* p. 38 *in* Mandl (1968)*. Then it holds for all $m \in \{0, 1, 2, 3\}$ that*

(i) *it holds for all $t \in [0,\infty)$ that $S_t\colon C^m([0,1],\mathbb{R}) \to C^m([0,1],\mathbb{R})$,*
(ii) *$\{S_t\colon t \in [0,\infty)\}$ defines a strongly continuous semigroup on $C^m([0,1],\mathbb{R})$ with generator $(A, \mathcal{D}_m(A))$, and*
(iii) *it holds for all $t \in [0,\infty)$ and all $\varphi \in C^m([0,1],\mathbb{R})$ that*

$$\|S_t\varphi\|_{C^m} \le e^{\nu_m t}\|\varphi\|_{C^m}. \tag{3.2}$$

Proof: For every $m \in \{0, 1, 2, 3\}$ Theorem 1 and Remark 1 in Ethier (1978) and the Main Theorem in Dorea (1976) yield for all $t \in [0,\infty)$ that $S_t\colon C^m([0,1],\mathbb{R}) \to C^m([0,1],\mathbb{R})$ and that $\{S_s\colon s \in [0,\infty)\}$ restricted to $C^m([0,1],\mathbb{R})$ defines a strongly continuous semigroup with generator $(A, \mathcal{D}_m(A))$. This proves (i) and (ii).

It remains to check that (3.2) can be established with our choice of the norm on $C^m([0,1],\mathbb{R})$. For every $m \in \{0, 1, 2\}$ Theorem k in Dorea (1976) with $k = m$ yields for all $\lambda > \nu_m$ and all $\varphi \in C^m([0,1],\mathbb{R})$ that $J_\lambda\varphi := (\lambda - A)^{-1}\varphi \in \mathcal{D}_m(A)$ exists and its proof shows that

$$\left\|\tfrac{d^m J_\lambda\varphi}{dx^m}\right\|_\infty \le \tfrac{1}{\lambda-\nu_m}\left\|\tfrac{d^m\varphi}{dx^m}\right\|_\infty. \tag{3.3}$$

Fix $m \in \{0,1,2\}$ for the rest of this paragraph. Consider $G := A - \nu_m$ with domain $\mathcal{D}(G) = \mathcal{D}_m(A)$. Since $C^\infty([0,1],\mathbb{R}) \subseteq \mathcal{D}(G)$, it follows that $\mathcal{D}(G)$ is dense in $C^m([0,1],\mathbb{R})$ w.r.t. $\|\cdot\|_{C^m}$. Equation (3.3) implies for all $\lambda,\lambda' > 0$ with $\lambda = \lambda' + \nu_m$ and all $\varphi \in C^m([0,1],\mathbb{R})$ that $(\lambda' - G)^{-1}\varphi = J_\lambda\varphi \in \mathcal{D}(G)$ and

$$\begin{aligned} \|(\lambda' - G)^{-1}\varphi\|_{C^m} = \|J_\lambda\varphi\|_{C^m} &= \max_{k\in\{0,\dots,m\}} \left\|\tfrac{d^k J_\lambda\varphi}{dx^k}\right\|_\infty \\ &\le \max_{k\in\{0,\dots,m\}} \tfrac{1}{\lambda-\nu_k}\left\|\tfrac{d^k\varphi}{dx^k}\right\|_\infty \\ &\le \tfrac{1}{\lambda-\nu_m}\|\varphi\|_{C^m} = \tfrac{1}{\lambda'}\|\varphi\|_{C^m}. \end{aligned} \tag{3.4}$$

Thus $\mathcal{D}(G)$ is dense in $C^m([0,1],\mathbb{R})$, G is dissipative, and $\mathcal{R}(1-G) = C^m([0,1],\mathbb{R})$. Consequently, the Hille-Yosida theorem (see, e.g. Theorem 1.2.6 in Ethier and Kurtz, 1986) yields that G generates a unique strongly continuous contraction semigroup $\{P_t\colon t\in[0,\infty)\}$ on $C^m([0,1],\mathbb{R})$. This implies that $\{e^{\nu_m t}P_t\colon t\in[0,\infty)\}$ is a strongly continuous semigroup on $C^m([0,1],\mathbb{R})$ with infinitesimal generator $\nu_m + G = A$. It follows that $\{S_t\colon t\in[0,\infty)\}$ restricted to $C^m([0,1],\mathbb{R})$ is given by $\{e^{\nu_m t}P_t\colon t\in[0,\infty)\}$ and that it holds for all $t\in[0,\infty)$ and all $\varphi\in C^m([0,1],\mathbb{R})$ that

$$\|S_t\varphi\|_{C^m} = e^{\nu_m t}\|P_t\varphi\|_{C^m} \le e^{\nu_m t}\|\varphi\|_{C^m}. \tag{3.5}$$

Since $m \in \{0,1,2\}$ was arbitrary, (3.2) is shown for all $m\in\{0,1,2\}$.

To prove (iii), it remains to treat the case $m = 3$. Define $\tilde\nu_3 := \nu_3 - \frac{1}{2}\|\frac{d^3a}{dx^3}\|_\infty$. Theorem 3 in Dorea (1976) yields for all $\lambda > \tilde\nu_3$ and all $\varphi\in C^3([0,1],\mathbb{R})$ that $J_\lambda\varphi := (\lambda - A)^{-1}\varphi \in \mathcal{D}_3(A)$ exists and its proof shows that

$$\left\|\tfrac{d^3 J_\lambda\varphi}{dx^3}\right\|_\infty \le \tfrac{1}{\lambda-\tilde\nu_3}\Big(\left\|\tfrac{d^3\varphi}{dx^3}\right\|_\infty + \tfrac{1}{2}\left\|\tfrac{d^3a}{dx^3}\right\|_\infty\left\|\tfrac{d^2J_\lambda\varphi}{dx^2}\right\|_\infty\Big). \tag{3.6}$$

This, (3.3), and the inequality $\nu_0 \le \nu_1 \le \nu_2 \le \tilde\nu_3$ yield for all $\lambda > \tilde\nu_3$ and all $\varphi\in C^3([0,1],\mathbb{R})$ that

$$\|J_\lambda\varphi\|_{C^3} \le \tfrac{1}{\lambda-\tilde\nu_3}\Big(\|\varphi\|_{C^3} + \tfrac{1}{2}\left\|\tfrac{d^3a}{dx^3}\right\|_\infty\|J_\lambda\varphi\|_{C^3}\Big). \tag{3.7}$$

If $\lambda > \nu_3$, then $\lambda > \tilde\nu_3$ and $1 - \frac{1}{2}\|\frac{d^3a}{dx^3}\|_\infty(\lambda-\tilde\nu_3)^{-1} = \frac{\lambda-\nu_3}{\lambda-\tilde\nu_3} > 0$, rearranging (3.7) therefore yields for all $\lambda > \nu_3$ and all $\varphi\in C^3([0,1],\mathbb{R})$ that

$$\|J_\lambda\varphi\|_{C^3} \le \tfrac{\lambda-\tilde\nu_3}{\lambda-\nu_3}\tfrac{1}{\lambda-\tilde\nu_3}\|\varphi\|_{C^3} = \tfrac{1}{\lambda-\nu_3}\|\varphi\|_{C^3}. \tag{3.8}$$

The remaining part of the proof of (iii) follows from an application of the Hille-Yosida theorem as in the previous paragraph. This finishes the proof of Lemma 3.1. □

3.2. *Multidimensional case.* Throughout this subsection, we use the definitions and the notation introduced in the following Setting 3.2.

Setting 3.2 (Diffusion coefficients). Let $d\in\mathbb{N}$, let $(\Omega,\mathcal{F},\mathbb{P},(\mathbb{F}_t)_{t\in[0,\infty)})$ be a stochastic basis, let $W = (W(1),\dots,W(d))\colon [0,\infty)\times\Omega\to\mathbb{R}^d$ be a standard $(\mathbb{F}_t)_{t\in[0,\infty)}$-Brownian motion with continuous sample paths, let $a_1,\dots,a_d \in C^3([0,1],\mathbb{R})$ satisfy for all $i\in\{1,\dots,d\}$ and all $x\in(0,1)$ that $a_i(0) = 0 = a_i(1)$ and $a_i(x) > 0$, and we define $\mu_0 := 0$, $\mu_1 := 0$, $\mu_2 := \max_{i\in\{1,\dots,d\}} \frac{1}{2}\|\frac{d^2a_i}{dx^2}\|_\infty$, and $\mu_3 := \max_{i\in\{1,\dots,d\}}(\|\frac{d^3a_i}{dx^3}\|_\infty + \frac{3}{2}\|\frac{d^2a_i}{dx^2}\|_\infty)$.

Theorem 3.2 in Shiga and Shimizu (1980) implies that there exist $(\mathbb{F}_t)_{t\in[0,\infty)}$-adapted processes $Y^x = (Y^x(1),\dots,Y^x(d))\colon [0,\infty)\times\Omega\to[0,1]^d$, $x\in[0,1]^d$, with continuous sample paths satisfying for all $i \in \{1,\ldots,d\}$, all $t \in [0,\infty)$, and all $x = (x_1,\ldots,x_d) \in [0,1]^d$ that $\mathbb{P}$-a.s.

$$Y_t^x(i) = x_i + \int_0^t \sqrt{a_i(Y_s^x(i))}\, dW_s(i). \tag{3.9}$$

We denote by $\{T_t^2 \colon t \in [0,\infty)\}$ the associated strongly continuous contraction semigroup on $C([0,1]^d,\mathbb{R})$, which satisfies for all $t \in [0,\infty)$, all $f \in C([0,1]^d,\mathbb{R})$, and all $x \in [0,1]^d$ that $(T_t^2 f)(x) = \mathbb{E}[f(Y_t^x)]$; see Remark 3.2 in Shiga and Shimizu (1980). For every $i \in \{1,\ldots,d\}$ we denote by $\{S_t^i \colon t \in [0,\infty)\}$ the strongly continuous contraction semigroup on $C([0,1],\mathbb{R})$ associated with $Y^\cdot(i)$, which satisfies for all $t \in [0,\infty)$, all $\varphi \in C([0,1],\mathbb{R})$, and all $x \in [0,1]$ that $(S_t^i\varphi)(x) = \mathbb{E}[\varphi(Y_t^x(i))]$, and by

$$[0,\infty) \times [0,1] \times \mathcal{B}(\mathbb{R}) \ni (t,x,A) \mapsto p_t^i(x,A) \in [0,1] \tag{3.10}$$

the corresponding transition kernel.

Note that $Y^\cdot(i)$, $i \in \{1,\ldots,d\}$, are independent diffusion processes with generators $A_i \colon C^2([0,1],\mathbb{R}) \to C([0,1],\mathbb{R})$, $i \in \{1,\ldots,d\}$, satisfying for all $i \in \{1,\ldots,d\}$, all $\varphi \in C^2([0,1],\mathbb{R})$, and all $x \in [0,1]$ that

$$(A_i\varphi)(x) = \frac{1}{2} a_i(x) \frac{d^2\varphi}{dx^2}(x), \tag{3.11}$$

so that the result of Subsection 3.1 applies. Moreover, it holds for all $i \in \{1,\ldots,d\}$, all $t \in [0,\infty)$, all $\varphi \in C([0,1],\mathbb{R})$, and all $x \in [0,1]$ that

$$(S_t^i\varphi)(x) = \int p_t^i(x,dy)\varphi(y) \tag{3.12}$$

and it holds for all $t \in [0,\infty)$, all $f \in C([0,1]^d,\mathbb{R})$, and all $x = (x_1,\ldots,x_d) \in [0,1]^d$ that

$$(T_t^2 f)(x) = \int \bigotimes_{k=1}^d p_t^k(x_k, dy_k) f(y). \tag{3.13}$$

The aim of this subsection is to show for all $m \in \{0,1,2,3\}$ that it holds for all $t \in [0,\infty)$ that $T_t^2 \colon C^m([0,1]^d,\mathbb{R}) \to C^m([0,1]^d,\mathbb{R})$ and for all $t \in [0,\infty)$ and all $f \in C^m([0,1]^d,\mathbb{R})$ that $\|T_t^2 f\|_{C^m} \le e^{\mu_m t}\|f\|_{C^m}$; see Lemma 3.8 below.

Lemma 3.3 (Continuity property)**.** *Assume* Setting 3.2*, let* $t \in [0,\infty)$*, let* $f \in C([0,1]^d,\mathbb{R})$*, and let* $I \subseteq \{1,\ldots,d\}$*. Then the function*

$$[0,1]^d \ni x \mapsto \int \bigotimes_{k\in\{1,\ldots,d\}\setminus I} p_t^k(x_k,dy_k) f\big((x_i \mathbb{1}_{i\in I} + y_i \mathbb{1}_{i\notin I})_{i\in\{1,\ldots,d\}}\big) \tag{3.14}$$

is continuous.

Proof: Throughout this proof, we denote by $f_I \colon [0,1]^d \times [0,1]^d \to \mathbb{R}$ the function satisfying for all $x,y \in [0,1]^d$ that $f_I(x,y) = f((x_i\mathbb{1}_{i\in I} + y_i\mathbb{1}_{i\notin I})_{i\in\{1,\ldots,d\}})$. Let $\{x^n \colon n \in \mathbb{N}\} \subseteq [0,1]^d$ be a convergent sequence with $\lim_{n\to\infty} x^n = x \in [0,1]^d$.

Then it holds for all $n \in \mathbb{N}$ that

$$\begin{aligned}
&\left| \int \bigotimes_{k\in\{1,\ldots,d\}\setminus I} p_t^k(x_k^n, dy_k) f_I(x^n, y) - \int \bigotimes_{k\in\{1,\ldots,d\}\setminus I} p_t^k(x_k, dy_k) f_I(x,y) \right| \\
&\quad \le \left| \int \bigotimes_{k\in\{1,\ldots,d\}\setminus I} p_t^k(x_k^n, dy_k) \big(f_I(x^n, y) - f_I(x,y) \big) \right| \\
&\qquad + \left| \int \bigotimes_{k\in\{1,\ldots,d\}\setminus I} p_t^k(x_k^n, dy_k) f_I(x,y) - \int \bigotimes_{k\in\{1,\ldots,d\}\setminus I} p_t^k(x_k, dy_k) f_I(x,y) \right| \\
&\quad \le \sup_{y\in[0,1]^d} \big| f_I(x^n, y) - f_I(x,y) \big| \\
&\qquad + \left| \int \bigotimes_{k\in\{1,\ldots,d\}\setminus I} p_t^k(x_k^n, dy_k) f_I(x,y) - \int \bigotimes_{k\in\{1,\ldots,d\}\setminus I} p_t^k(x_k, dy_k) f_I(x,y) \right|.
\end{aligned} \tag{3.15}$$

By uniform continuity of f on $[0,1]^d$, the first summand on the right-hand side converges to zero as $n \to \infty$. For fixed $x \in [0,1]^d$, the function $[0,1]^d \ni y \mapsto f_I(x,y)$ is continuous, which implies the continuity of $[0,1]^d \ni z \mapsto \int \bigotimes_{k\in\{1,\ldots,d\}\setminus I} p_t^k(z_k, dy_k) f_I(x,y)$. Therefore, the second summand on the right-hand side converges to zero as $n \to \infty$. This finishes the proof of Lemma 3.3. $\square$

Lemma 3.4 (Continuity of pure derivatives)**.** *Assume* Setting 3.2*, let $m \in \{0,1,2,3\}$, let $t \in [0,\infty)$, and let $f \in C^m([0,1]^d, \mathbb{R})$. Then it holds for every $i \in \{1,\ldots,d\}$ that the partial derivative*

$$[0,1]^d \ni x \mapsto \frac{\partial^m}{\partial x_i^m} \int p_t^i(x_i, dy_i) f(x_1, \ldots, x_{i-1}, y_i, x_{i+1}, \ldots, x_d) \tag{3.16}$$

exists and is continuous.

Proof: It suffices to prove the claim for $i = 1$. For fixed $x \in [0,1]^d$, the function $[0,1] \ni y \mapsto f(y, \hat{x}_1)$ is in $C^m([0,1],\mathbb{R})$, so Lemma 3.1 implies that the function $[0,1] \ni z \mapsto \int p_t^1(z, dy_1) f(y_1, \hat{x}_1)$ is in $C^m([0,1],\mathbb{R})$. This shows the existence of the partial derivative (3.16). It remains to show continuity on $[0,1]^d$. For that, let $\{x^n : n \in \mathbb{N}\} \subseteq [0,1]^d$ be a convergent sequence with $\lim_{n\to\infty} x^n = x \in [0,1]^d$. Lemma 3.1 implies for all $n \in \mathbb{N}$ that

$$\begin{aligned}
&\left| \frac{\partial^m}{\partial (x_1^n)^m} \int p_t^1(x_1^n, dy_1) \left(f(y_1, \widehat{x^n}_1) - f(y_1, \hat{x}_1) \right) \right| \\
&\qquad \le e^{\mu_m t} \max_{k\in\{0,\ldots,m\}} \sup_{z\in[0,1]} \left| \frac{\partial^k f}{\partial z^k}(z, \widehat{x^n}_1) - \frac{\partial^k f}{\partial z^k}(z, \hat{x}_1) \right|.
\end{aligned} \tag{3.17}$$

Since $f \in C^m([0,1]^d, \mathbb{R})$, it follows for all $k \in \{0,\ldots,m\}$ that $[0,1]^d \ni x \mapsto \frac{\partial^k f}{\partial x_1^k}(x)$ is uniformly continuous. Therefore, the right-hand side of (3.17) converges to zero as $n \to \infty$. It holds for all $n \in \mathbb{N}$ that

$$\begin{aligned}
&\left| \frac{\partial^m}{\partial (x_1^n)^m} \int p_t^1(x_1^n, dy_1) f(y_1, \widehat{x^n}_1) - \frac{\partial^m}{\partial x_1^m} \int p_t^1(x_1, dy_1) f(y_1, \hat{x}_1) \right| \\
&\quad \le \left| \frac{\partial^m}{\partial (x_1^n)^m} \int p_t^1(x_1^n, dy_1) \big(f(y_1, \widehat{x^n}_1) - f(y_1, \hat{x}_1) \big) \right| \\
&\qquad + \left| \frac{\partial^m}{\partial (x_1^n)^m} \int p_t^1(x_1^n, dy_1) f(y_1, \hat{x}_1) - \frac{\partial^m}{\partial x_1^m} \int p_t^1(x_1, dy_1) f(y_1, \hat{x}_1) \right|.
\end{aligned} \tag{3.18}$$

The first summand on the right-hand side of (3.18) converges to zero as $n \to \infty$ by (3.17). We have shown above that $[0,1] \ni z \mapsto \int p_t^1(z, dy_1) f(y_1, \hat{x}_1)$ is in $C^m([0,1], \mathbb{R})$, so also the second summand on the right-hand side of (3.18) converges to zero as $n \to \infty$. This finishes the proof of Lemma 3.4. □

Lemma 3.5 (Continuity of pure derivatives, continued)**.** *Assume* Setting 3.2*, let $m \in \{0,1,2,3\}$, let $t \in [0,\infty)$, and let $f \in C^m([0,1]^d, \mathbb{R})$. Then it holds for every $i \in \{1, \ldots, d\}$ that the partial derivative*

$$[0,1]^d \ni x \mapsto \frac{\partial^m}{\partial x_i^m} \int \bigotimes_{k=1}^d p_t^k(x_k, dy_k) f(y) \tag{3.19}$$

exists and is continuous.

Proof: It suffices to show the claim for $i = 1$. By Fubini's theorem, it holds for all $x \in [0,1]^d$ that

$$\int \bigotimes_{k=1}^d p_t^k(x_k, dy_k) f(y) = \int p_t^1(x_1, dy_1) \int \bigotimes_{k=2}^d p_t^k(x_k, dy_k) f(y). \tag{3.20}$$

For fixed $x \in [0,1]^d$, the fact that $f \in C^m([0,1]^d, \mathbb{R})$ and the dominated convergence theorem imply that the function $[0,1] \ni z \mapsto \int \bigotimes_{k=2}^d p_t^k(x_k, dy_k) f(z, \hat{y}_1)$ is in $C^m([0,1], \mathbb{R})$. Therefore, (3.20) and Lemma 3.1 prove the existence of the partial derivative (3.19). Moreover, Fubini's theorem, the fact that $f \in C^m([0,1]^d, \mathbb{R})$, Lemma 3.1, and the dominated convergence theorem imply for all $x \in [0,1]^d$ that

$$\begin{aligned}
\frac{\partial^m}{\partial x_1^m} \int \bigotimes_{k=1}^d p_t^k(x_k, dy_k) f(y) &= \frac{\partial^m}{\partial x_1^m} \int \bigotimes_{k=2}^d p_t^k(x_k, dy_k) \int p_t^1(x_1, dy_1) f(y) \\
&= \int \bigotimes_{k=2}^d p_t^k(x_k, dy_k) \frac{\partial^m}{\partial x_1^m} \int p_t^1(x_1, dy_1) f(y).
\end{aligned} \tag{3.21}$$

Consequently, Lemma 3.4 and Lemma 3.3 imply the continuity of (3.19). This completes the proof of Lemma 3.5. □

Lemma 3.6 (Continuity of mixed second derivatives)**.** *Assume* Setting 3.2 *and let $t \in [0,\infty)$ and $f \in C^2([0,1]^d, \mathbb{R})$. Then it holds for every $i, j \in \{1, \ldots, d\}$ that the partial derivative*

$$[0,1]^d \ni x \mapsto \frac{\partial^2}{\partial x_i \partial x_j} \int \bigotimes_{k=1}^d p_t^k(x_k, dy_k) f(y) \tag{3.22}$$

exists and is continuous.

Proof: The case where $i = j$ is treated by Lemma 3.5. It suffices to consider $i = 1$ and $j = 2$. The dominated convergence theorem implies for all $x \in [0,1]^d$ that

$$\frac{\partial}{\partial x_1} \int \bigotimes_{k=2}^{d} p_t^k(x_k, dy_k) f(x_1, \hat{y}_1) = \int \bigotimes_{k=2}^{d} p_t^k(x_k, dy_k) \frac{\partial f}{\partial x_1}(x_1, \hat{y}_1). \tag{3.23}$$

Using (3.23) and Fubini's theorem, it follows for all $x \in [0,1]^d$ that

$$\frac{\partial}{\partial x_1} \int \bigotimes_{k=2}^{d} p_t^k(x_k, dy_k) f(x_1, \hat{y}_1) = \int p_t^2(x_2, dy_2) \int \bigotimes_{k=3}^{d} p_t^k(x_k, dy_k) \frac{\partial f}{\partial x_1}(x_1, \hat{y}_1). \tag{3.24}$$

For fixed $x \in [0,1]^d$, the fact that $f \in C^2([0,1]^d, \mathbb{R})$ and the dominated convergence theorem imply that the function $[0,1] \ni z \mapsto \int \bigotimes_{k=3}^{d} p_t^k(x_k, dy_k) \frac{\partial f}{\partial x_1}(x_1, z, \hat{y}_{12})$ is in $C^1([0,1], \mathbb{R})$. Therefore, (3.24) and Lemma 3.1 imply the existence of the partial derivative $[0,1]^d \ni x \mapsto \frac{\partial^2}{\partial x_2 \partial x_1} \int \bigotimes_{k=2}^{d} p_t^k(x_k, dy_k) f(x_1, \hat{y}_1)$. Fubini's theorem, Lemma 3.1, and the dominated convergence theorem imply for all $x \in [0,1]^d$ that

$$\begin{aligned} &\frac{\partial^2}{\partial x_2 \partial x_1} \int \bigotimes_{k=2}^{d} p_t^k(x_k, dy_k) f(x_1, \hat{y}_1) \\ &= \int \bigotimes_{k=3}^{d} p_t^k(x_k, dy_k) \frac{\partial}{\partial x_2} \int p_t^2(x_2, dy_2) \frac{\partial f}{\partial x_1}(x_1, \hat{y}_1). \end{aligned} \tag{3.25}$$

Lemma 3.4 and Lemma 3.3 show that (3.25) is continuous as a function of $x \in [0,1]^d$. Consequently, Schwarz's theorem (see, e.g. Theorem 9.41 in Rudin (1976)) implies that the partial derivative $[0,1]^d \ni x \mapsto \frac{\partial^2}{\partial x_1 \partial x_2} \int \bigotimes_{k=2}^{d} p_t^k(x_k, dy_k) f(x_1, \hat{y}_1)$ exists and satisfies for all $x \in [0,1]^d$ that

$$\frac{\partial^2}{\partial x_1 \partial x_2} \int \bigotimes_{k=2}^{d} p_t^k(x_k, dy_k) f(x_1, \hat{y}_1) = \frac{\partial^2}{\partial x_2 \partial x_1} \int \bigotimes_{k=2}^{d} p_t^k(x_k, dy_k) f(x_1, \hat{y}_1). \tag{3.26}$$

In particular, for fixed $x \in [0,1]^d$, the function $z \mapsto \frac{\partial}{\partial x_2} \int \bigotimes_{k=2}^{d} p_t^k(x_k, dy_k) f(z, \hat{y}_1)$ is in $C^1([0,1], \mathbb{R})$. From this and Lemma 3.1, it follows that the partial derivative (3.22) exists. Fubini's theorem, Lemma 3.1, and the dominated convergence theorem further show for all $x \in [0,1]^d$ that

$$\begin{aligned} &\frac{\partial^2}{\partial x_1 \partial x_2} \int \bigotimes_{k=1}^{d} p_t^k(x_k, dy_k) f(y) \\ &= \int \bigotimes_{k=3}^{d} p_t^k(x_k, dy_k) \frac{\partial}{\partial x_1} \int p_t^1(x_1, dy_1) \frac{\partial}{\partial x_2} \int p_t^2(x_2, dy_2) f(y). \end{aligned} \tag{3.27}$$

Then Lemma 3.4 and Lemma 3.3 imply that (3.27) is continuous as a function of $x \in [0,1]^d$. This concludes the proof of Lemma 3.6. □

The proof of the following Lemma 3.7 is analogous to the proofs of Lemma 3.5 and Lemma 3.6 above and therefore omitted here.

Lemma 3.7 (Continuity of mixed third derivatives)**.** *Assume* Setting 3.2 *and let* $t \in [0, \infty)$ *and* $f \in C^3([0,1]^d, \mathbb{R})$*. Then it holds for every* $i, j, l \in \{1, \ldots, d\}$ *that the* *partial derivative*

$$[0,1]^d \ni x \mapsto \frac{\partial^3}{\partial x_i \partial x_j \partial x_l} \int \bigotimes_{k=1}^{d} p_t^k(x_k, dy_k) f(y) \tag{3.28}$$

exists and is continuous.

Lemma 3.8 (C^m-estimate for multidimensional diffusive part)**.** *Assume* Setting 3.2*, let* $m \in \{0,1,2,3\}$*, let* $t \in [0,\infty)$*, and let* $f \in C^m([0,1]^d, \mathbb{R})$*. Then it holds that* $T_t^2 f \in C^m([0,1]^d, \mathbb{R})$ *and*

$$\|T_t^2 f\|_{C^m} \le e^{\mu_m t} \|f\|_{C^m}. \tag{3.29}$$

Proof: Existence and continuity of the partial derivatives follow from Lemma 3.5, Lemma 3.6, and Lemma 3.7. It follows from Lemma 3.1 and from the dominated convergence theorem for all $n \in \mathbb{N}_0$ with $n \le m$ and all $x \in [0,1]^d$ that

$$\begin{aligned} \left| \frac{\partial^n (T_t^2 f)}{\partial x_1^n}(x) \right| &= \left| \frac{\partial^n}{\partial x_1^n} \int p_t^1(x_1, dy_1) \int \bigotimes_{i=2}^{d} p_t^i(x_i, dy_i) f(y) \right| \\ &\le e^{\mu_n t} \max_{k \in \{0,\ldots,n\}} \sup_{z \in [0,1]} \left| \frac{\partial^k}{\partial z^k} \int \bigotimes_{i=2}^{d} p_t^i(x_i, dy_i) f(z, \hat{y}_1) \right| \\ &= e^{\mu_n t} \max_{k \in \{0,\ldots,n\}} \sup_{z \in [0,1]} \left| \int \bigotimes_{i=2}^{d} p_t^i(x_i, dy_i) \frac{\partial^k f}{\partial z^k}(z, \hat{y}_1) \right| \\ &\le e^{\mu_n t} \max_{k \in \{0,\ldots,n\}} \left\| \frac{\partial^k f}{\partial x_1^k} \right\|_\infty . \end{aligned} \tag{3.30}$$

If $m \ge 2$, then Lemma 3.1 and the dominated convergence theorem show for all $x \in [0,1]^d$ that

$$\begin{aligned} \left| \frac{\partial^2 (T_t^2 f)}{\partial x_1 \partial x_2}(x) \right| &= \left| \frac{\partial^2}{\partial x_1 \partial x_2} \int p_t^1(x_1, dy_1) \int \bigotimes_{i=2}^{d} p_t^i(x_i, dy_i) f(y) \right| \\ &\le \max_{k \in \{0,1\}} \sup_{z_1 \in [0,1]} \left| \frac{\partial^k}{\partial z_1^k} \frac{\partial}{\partial x_2} \int \bigotimes_{i=2}^{d} p_t^i(x_i, dy_i) f(z_1, \hat{y}_1) \right| \\ &= \max_{k \in \{0,1\}} \sup_{z_1 \in [0,1]} \left| \frac{\partial}{\partial x_2} \int p_t^2(x_2, dy_2) \int \bigotimes_{i=3}^{d} p_t^i(x_i, dy_i) \frac{\partial^k f}{\partial z_1^k}(z_1, \hat{y}_1) \right| \\ &\le \max_{k,l \in \{0,1\}} \sup_{z_1, z_2 \in [0,1]} \left| \frac{\partial^l}{\partial z_2^l} \int \bigotimes_{k=3}^{d} p_t^k(x_k, dy_k) \frac{\partial^k f}{\partial z_1^k}(z_1, z_2, \hat{y}_{12}) \right| \\ &\le \max_{k,l \in \{0,1\}} \left\| \frac{\partial^{k+l} f}{\partial x_1^k \partial x_2^l} \right\|_\infty . \end{aligned} \tag{3.31}$$

Similarly, if $m = 3$, it follows for all $x \in [0,1]^d$ that

$$\left| \frac{\partial^3 (T_t^2 f)}{\partial x_1 \partial x_2^2}(x) \right| \le e^{\mu_2 t} \max_{k \in \{0,1\}, l \in \{0,1,2\}} \left\| \frac{\partial^{k+l} f}{\partial x_1^k \partial x_2^l} \right\|_\infty \tag{3.32}$$

and

$$\left|\frac{\partial^3(T_t^2 f)}{\partial x_1 \partial x_2 \partial x_3}(x)\right| \leq \max_{k,l,n\in\{0,1\}} \left\|\frac{\partial^{k+l+n} f}{\partial x_1^k \partial x_2^l \partial x_3^n}\right\|_\infty. \tag{3.33}$$

All of the above estimates also hold for the partial derivatives in the remaining coordinate directions. Combining all of these estimates shows (3.29). This completes the proof of Lemma 3.8. □

4. Main result: Spatial derivatives of semigroups

The following main result, Theorem 4.1 establishes upper bounds for the C^m-norms, $m \in \{0,1,2\}$, of the semigroup corresponding to the SDE (4.1). We note that $\lambda_1, \lambda_2, \lambda_3, \mu_2, \mu_3$, which in principle depend on the dimension, are in certain situations bounded in the dimension; see, e.g., Example 4.2.

Theorem 4.1 (C^m-estimate for semigroups of square-root diffusions)**.** *Let $d \in \mathbb{N}$, let $(\Omega, \mathcal{F}, \mathbb{P}, (\mathbb{F}_t)_{t\in[0,\infty)})$ be a stochastic basis, let $W = (W(1),\dots,W(d))\colon [0,\infty) \times \Omega \to \mathbb{R}^d$ be a standard $(\mathbb{F}_t)_{t\in[0,\infty)}$-Brownian motion with continuous sample paths, let $a_1,\dots,a_d \in C^3([0,1],\mathbb{R})$ satisfy for all $i \in \{1,\dots,d\}$ and all $x \in (0,1)$ that $a_i(0) = 0 = a_i(1)$ and $a_i(x) > 0$, let $b_1,\dots,b_d \in C^3([0,1]^d,\mathbb{R})$ satisfy for all $i \in \{1,\dots,d\}$ and all $x = (x_1,\dots,x_d) \in [0,1]^d$ with $x_i \in \{0,1\}$ that $(-1)^{x_i} b_i(x) \geq 0$, for every $m \in \{1,2,3\}$ we define $\lambda_m := \max_{\alpha\in\mathbb{N}_0^d, 0<|\alpha|\leq m} \sum_{i=1}^d \|\partial^\alpha b_i\|_\infty$, and we define $\lambda_0 := 0$, $\mu_0 := 0$, $\mu_1 := 0$, $\mu_2 := \max_{i\in\{1,\dots,d\}} \frac{1}{2}\|\frac{d^2 a_i}{dx^2}\|_\infty$, and $\mu_3 := \max_{i\in\{1,\dots,d\}}(\|\frac{d^3 a_i}{dx^3}\|_\infty + \frac{3}{2}\|\frac{d^2 a_i}{dx^2}\|_\infty)$. Then*

(i) *there exist $(\mathbb{F}_t)_{t\in[0,\infty)}$-adapted processes $X^x = (X^x(1),\dots,X^x(d))\colon [0,\infty)\times \Omega \to [0,1]^d$, $x \in [0,1]^d$, with continuous sample paths satisfying for all $i \in \{1,\dots,d\}$, all $t \in [0,\infty)$, and all $x = (x_1,\dots,x_d) \in [0,1]^d$ that $\mathbb{P}$-a.s.*

$$X_t^x(i) = x_i + \int_0^t b_i(X_s^x)\,ds + \int_0^t \sqrt{a_i(X_s^x(i))}\,dW_s(i) \tag{4.1}$$

and

(ii) *it holds for all $m \in \{0,1,2\}$, all $t \in [0,\infty)$, and all $f \in C^m([0,1]^d,\mathbb{R})$ that the function $[0,1]^d \ni x \mapsto \mathbb{E}[f(X_t^x)] \in \mathbb{R}$ is an element of $C^m([0,1]^d,\mathbb{R})$ and satisfies*

$$\|[0,1]^d \ni x \mapsto \mathbb{E}[f(X_t^x)] \in \mathbb{R}\|_{C^m} \leq e^{(m^2\lambda_m+\mu_m)t}\|f\|_{C^m}. \tag{4.2}$$

Proof: Theorem 3.2 in Shiga and Shimizu (1980) implies (i).

We denote by $\{T_t\colon t \in [0,\infty)\}$ the family of operators on $C([0,1]^d,\mathbb{R})$ that satisfy for all $t \in [0,\infty)$, all $f \in C([0,1]^d,\mathbb{R})$, and all $x \in [0,1]^d$ that $(T_t f)(x) = \mathbb{E}[f(X_t^x)]$. Then $\{T_t\colon t \in [0,\infty)\}$ is the strongly continuous contraction semigroup on $C([0,1]^d,\mathbb{R})$ associated with the diffusion process $X^\cdot$; see Remark 3.2 in Shiga and Shimizu (1980). Let $G\colon C^2([0,1]^d,\mathbb{R}) \to C([0,1]^d,\mathbb{R})$ satisfy for all $f \in C^2([0,1]^d,\mathbb{R})$ and all $x = (x_1,\dots,x_d) \in [0,1]^d$ that

$$(Gf)(x) = \sum_{i=1}^d b_i(x)\frac{\partial f}{\partial x_i}(x) + \frac{1}{2}\sum_{i=1}^d a_i(x_i)\frac{\partial^2 f}{\partial x_i^2}(x). \tag{4.3}$$

Then the generator of $\{T_t\colon t \in [0,\infty)\}$ is given by the closure of G (see, e.g., Remark 3.2 in Shiga and Shimizu, 1980), so $C^2([0,1]^d,\mathbb{R})$ is a core (cf., e.g., Section I.3 in Ethier and Kurtz, 1986) for G. Let $\{T_t^1\colon t \in [0,\infty)\}$ be as in Lemma 2.1, let

$\{T_t^2 \colon t \in [0,\infty)\}$ be as in Setting 3.2, and let $G_1, G_2 \colon C^2([0,1]^d,\mathbb{R}) \to C([0,1]^d,\mathbb{R})$ satisfy for all $f \in C^2([0,1]^d,\mathbb{R})$ and all $x = (x_1,\dots,x_d) \in [0,1]^d$ that

$$(G_1 f)(x) = \sum_{i=1}^{d} b_i(x) \frac{\partial f}{\partial x_i}(x) \tag{4.4}$$

and

$$(G_2 f)(x) = \frac{1}{2} \sum_{i=1}^{d} a_i(x_i) \frac{\partial^2 f}{\partial x_i^2}(x). \tag{4.5}$$

Then the closures of G_1 and G_2 are the generators of the strongly continuous contraction semigroups on $C([0,1]^d,\mathbb{R})$ given by $\{T_t^1 \colon t \in [0,\infty)\}$ and $\{T_t^2 \colon t \in [0,\infty)\}$, respectively. Hence, it holds that $C^2([0,1]^d,\mathbb{R})$ is a core for G, that $C^2([0,1]^d,\mathbb{R})$ is a subset of the domains of both G_1 and G_2, and that $G = G_1 + G_2$ on $C^2([0,1]^d,\mathbb{R})$. Therefore, it follows from Trotter's product formula (see, e.g., Corollary I.6.7 in Ethier and Kurtz, 1986) that the semigroup $\{T_t \colon t \in [0,\infty)\}$ satisfies for all $t \in [0,\infty)$ and all $f \in C([0,1]^d,\mathbb{R})$ that

$$\lim_{n\to\infty} \| T_t f - (T^1_{t/n} T^2_{t/n})^n f \|_\infty = 0. \tag{4.6}$$

By induction, it follows from Lemma 2.1 and Lemma 3.8 for all $n \in \mathbb{N}$, all $m \in \{0,1,2,3\}$, all $t \in [0,\infty)$, and all $f \in C^m([0,1]^d,\mathbb{R})$ that $(T^1_{t/n} T^2_{t/n})^n f \in C^m([0,1]^d,\mathbb{R})$ and

$$\| (T^1_{t/n} T^2_{t/n})^n f \|_{C^m} \leq e^{\left((m^2 + 4\cdot \mathbb{1}_{\{3\}}(m))\lambda_m + \mu_m\right)t} \|f\|_{C^m}. \tag{4.7}$$

Equation (4.7) shows for all $m \in \{0,1,2\}$, all $t \in [0,\infty)$, and all $f \in C^{m+1}([0,1]^d,\mathbb{R})$ that the sequence $\{(T^1_{t/n} T^2_{t/n})^n f \colon n \in \mathbb{N}\}$ is bounded in $C^{m+1}([0,1]^d,\mathbb{R})$. Therefore, the Arzelà-Ascoli theorem guarantees for all $m \in \{0,1,2\}$, all $t \in [0,\infty)$, and all $f \in C^{m+1}([0,1]^d,\mathbb{R})$ that every subsequence of $\{(T^1_{t/n} T^2_{t/n})^n f \colon n \in \mathbb{N}\}$ has a convergent subsequence in $C^m([0,1]^d,\mathbb{R})$, whose limit is given by $T_t f$ due to (4.6). This and (4.7) imply for all $m \in \{0,1,2\}$, all $t \in [0,\infty)$, and all $f \in C^{m+1}([0,1]^d,\mathbb{R})$ that $T_t f \in C^m([0,1]^d,\mathbb{R})$ and

$$\|T_t f\|_{C^m} \leq e^{(m^2\lambda_m + \mu_m)t} \|f\|_{C^m}. \tag{4.8}$$

For the rest of the proof, fix $m \in \{0,1,2\}$, fix $t \in [0,\infty)$, and fix $f \in C^m([0,1]^d,\mathbb{R})$. Since $C^{m+1}([0,1]^d,\mathbb{R})$ is dense in $C^m([0,1]^d,\mathbb{R})$, we find a sequence $\{f_k \colon k \in \mathbb{N}\} \subseteq C^{m+1}([0,1]^d,\mathbb{R})$ with the property that $\lim_{k\to\infty} \|f - f_k\|_{C^m} = 0$. By the previous step, it holds for all $k \in \mathbb{N}$ that $T_t f_k \in C^m([0,1]^d,\mathbb{R})$ and for all $k, l \in \mathbb{N}$ that

$$\|T_t f_k - T_t f_l\|_{C^m} = \|T_t(f_k - f_l)\|_{C^m} \leq e^{(m^2\lambda_m + \mu_m)t} \|f_k - f_l\|_{C^m}, \tag{4.9}$$

which shows that $\{T_t f_k \colon k \in \mathbb{N}\}$ is a Cauchy sequence in $C^m([0,1]^d,\mathbb{R})$. By completeness, it follows that $\{T_t f_k \colon k \in \mathbb{N}\}$ converges in $C^m([0,1]^d,\mathbb{R})$. Moreover, since T_t is a contraction on $C([0,1]^d,\mathbb{R})$, it holds for all $k \in \mathbb{N}$ that

$$\|T_t f - T_t f_k\|_\infty = \|T_t(f - f_k)\|_\infty \leq \|f - f_k\|_\infty. \tag{4.10}$$

This identifies the limit point of $\{T_t f_k \colon k \in \mathbb{N}\} \subseteq C^m([0,1]^d,\mathbb{R})$ and shows that $T_t f \in C^m([0,1]^d,\mathbb{R})$ and that $\lim_{k\to\infty} \|T_t f - T_t f_k\|_{C^m} = 0$. Then it follows from (4.8) that

$$\|T_t f\|_{C^m} = \lim_{k\to\infty} \|T_t f_k\|_{C^m} \le \lim_{k\to\infty} e^{(m^2\lambda_m+\mu_m)t}\|f_k\|_{C^m} = e^{(m^2\lambda_m+\mu_m)t}\|f\|_{C^m}. \tag{4.11}$$

Since $m \in \{0,1,2\}$, $t \in [0,\infty)$, and $f \in C^m([0,1]^d, \mathbb{R})$ were arbitrary, this proves (ii) and completes the proof of Theorem 4.1. □

The following example applies Theorem 4.1 to a system of interacting diffusions in Hutzenthaler et al. (2015); Hutzenthaler and Pieper (2020) (with $a = 2$, $\kappa = \alpha = \beta = 1$, $(\mu_D)_{D\in\mathbb{N}\cup\{\infty\}} \equiv 0$).

Example 4.2. Let $d \in \mathbb{N}$ and for every $i \in \{1,\dots,d\}$ let $b_i = ([0,1]^d \ni (x_1,\dots,x_d) \mapsto \frac{1}{d}\sum_{j=1}^d \frac{2-x_i}{2-x_j}(x_j - x_i) - x_i(1-x_i) \in \mathbb{R})$ and $a_i = ([0,1] \ni x \mapsto (2-x)x(1-x) \in \mathbb{R})$. Then

$$\begin{aligned}\lambda_1 &= \max_{j\in\{1,\dots,d\}} \sum_{i=1}^d \sup_{x\in[0,1]^d} \left|\tfrac{\partial}{\partial x_j} b_i(x)\right| \\ &= \max_{j\in\{1,\dots,d\}} \sum_{i=1}^d \sup_{x\in[0,1]^d} \left|\tfrac{1}{d}\mathbb{1}_{i\neq j}\left(\tfrac{2-x_i}{(2-x_j)^2}(x_j - x_i) + \tfrac{2-x_i}{2-x_j}\right) + \mathbb{1}_{i=j}(2x_i - 1)\right| \\ &\le \sum_{i=1}^d (\tfrac{4}{d} + \mathbb{1}_{i=j}) = 5,\end{aligned} \tag{4.12}$$

$\lambda_2 \le 10$, $\lambda_3 \le 15$, $\mu_2 = 3$, and $\mu_3 = 15$.

Acknowledgment. We thank an anonymous referee for many helpful comments. This paper has been partially supported by the DFG Priority Program “Probabilistic Structures in Evolution” (SPP 1590), grant HU 1889/3-2.

References

Cox, J. T., Fleischmann, K., and Greven, A. Comparison of interacting diffusions and an application to their ergodic theory. *Probab. Theory Related Fields*, **105** (4), 513–528 (1996). MR1402655.

Dawson, D. and Greven, A. Multiple time scale analysis of interacting diffusions. *Probab. Theory Related Fields*, **95** (4), 467–508 (1993). MR1217447.

Dorea, C. C. Y. Differentiability preserving properties of a class of semigroups. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, **36** (1), 13–26 (1976). MR415790.

Epstein, C. L. and Pop, C. A. Transition probabilities for degenerate diffusions arising in population genetics. *Probab. Theory Related Fields*, **173** (1-2), 537–603 (2019). MR3916114.

Ethier, S. N. A class of degenerate diffusion processes occurring in population genetics. *Comm. Pure Appl. Math.*, **29** (5), 483–493 (1976). MR428488.

Ethier, S. N. Differentiability-preserving properties of Markov semigroups associated with one-dimensional diffusions. *Z. Wahrsch. Verw. Gebiete*, **45** (3), 225–238 (1978). MR510027.

Ethier, S. N. and Kurtz, T. G. *Markov processes. Characterization and convergence*. Wiley Series in Probability and Mathematical Statistics: Probability and

Mathematical Statistics. John Wiley & Sons, Inc., New York (1986). ISBN 0-471-08186-8. MR838085.

Gikhman, I. I. and Skorokhod, A. V. *Introduction to the theory of random processes.* Translated from the Russian by Scripta Technica, Inc. W. B. Saunders Co., Philadelphia, Pa.-London-Toronto, Ont. (1969). MR0247660.

Hairer, M., Hutzenthaler, M., and Jentzen, A. Loss of regularity for Kolmogorov equations. *Ann. Probab.*, **43** (2), 468–527 (2015). MR3305998.

Hartman, P. *Ordinary differential equations.* Classics in Applied Mathematics. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA (2002). ISBN 0-89871-510-5. MR1929104.

Hutzenthaler, M. Interacting diffusions and trees of excursions: convergence and comparison. *Electron. J. Probab.*, **17**, no. 71, 49 (2012). MR2968678.

Hutzenthaler, M., Jordan, F., and Metzler, D. Costly defense traits in structured populations. *ArXiv Mathematics e-prints* (2015). arXiv: 1505.02154v3.

Hutzenthaler, M. and Pieper, D. Propagation of chaos and the many-demes limit for weakly interacting diffusions in the sparse regime. *Ann. Appl. Probab.*, **30** (5), 2311–2354 (2020). MR4149530.

Hutzenthaler, M. and Wakolbinger, A. Ergodic behavior of locally regulated branching populations. *Ann. Appl. Probab.*, **17** (2), 474–501 (2007). MR2308333.

Mandl, P. *Analytical treatment of one-dimensional Markov processes.* Die Grundlehren der mathematischen Wissenschaften, Band 151. Academia Publishing House of the Czechoslovak Academy of Sciences, Prague; Springer-Verlag New York Inc., New York (1968). MR0247667.

Peng, S. G. Probabilistic interpretation for systems of quasilinear parabolic partial differential equations. *Stochastics Stochastics Rep.*, **37** (1-2), 61–74 (1991). MR1149116.

Rudin, W. *Principles of mathematical analysis.* McGraw-Hill Book Co., New York-Auckland-Düsseldorf, third edition (1976). MR0385023.

Shiga, T. and Shimizu, A. Infinite-dimensional stochastic differential equations and their applications. *J. Math. Kyoto Univ.*, **20** (3), 395–416 (1980). MR591802.

Talay, D. and Tubaro, L. Expansion of the global error for numerical schemes solving stochastic differential equations. *Stochastic Anal. Appl.*, **8** (4), 483–509 (1990). MR1091544.