The Annals of Applied Probability
2012, Vol. 22, No. 3, 1008–1045
DOI: 10.1214/11-AAP774

A NEW EXTRAPOLATION METHOD FOR WEAK APPROXIMATION SCHEMES WITH APPLICATIONS

BY KOJIRO OSHIMA, JOSEF TEICHMANN[1,2] AND DEJAN VELUŠČEK[1,2,3]

Mizuho Corporate Bank, Ltd., ETH Zürich and ETH Zürich

Fujiwara's method can be considered as an extrapolation method of order 6 of the Ninomiya–Victoir weak approximation scheme for the numerical approximation of solution processes of SDEs. We present an extension of Fujiwara's method for arbitrarily high orders, which embeds the original Fujiwara method as the order 6 case. The approach can be considered as a variant of Richardson extrapolation, which allows one to reach high orders with few extrapolation steps. The most important contribution of our approach is that we only need m extrpolation steps in order to achieve order of approximation $2m$, which is half the number of steps in comparison to classical approaches.

1. Introduction. Consider the following framework: let $(\Omega, \mathscr{F}, P)$ be a probability space, and let $(B_t^1, \dots, B_t^d)_{t \in \mathbf{R}_+}$ be a d-dimensional standard Brownian motion. Define $B_t^0 := t$ and $B_t := (B_t^0, B_t^1, \dots, B_t^d)$. We consider stochastic differential equations driven by the Brownian motion $(B_t)_{t \in \mathbf{R}_+}$

$$X(t, x) = x + \sum_{i=0}^{d} \int_0^t V_i(X(s, x)) \circ dB_s^i, \tag{1}$$

where the initial values x lies in $\mathbf{R}^N$, the vector fields are bounded, C^∞-bounded, $V_i \in C_b^\infty(\mathbf{R}^N; \mathbf{R}^N)$, and $\circ$ denotes the Stratonovich integral.

Throughout the article we shall use the abbreviation

$$D_i := \frac{\partial}{\partial x^i}.$$

Take $m, p \in \mathbf{N}$. We denote by W_p^m the Sobolev space defined as the closure of C_0^∞ functions $\varphi: \mathbf{R}^N \to \mathbf{R}$ in the norm

$$\|\varphi\|_{m,p} := \left(\sum_{|\gamma| \le m} \int_{\mathbf{R}^N} |D^\gamma \varphi(x)|^p \, dx \right)^{1/p},$$

Received November 2009; revised February 2011.

[1]Supported in part by Vienna WWTF project "Mathematik und Kreditrisiken."

[2]Supported by means of the START prize Project Y 328 and the ETH Foundation.

[3]Supported by AMaMeF exchange Grant No. 2080 and the Sciex Project 10.072.

MSC2010 subject classifications. Primary 65H35; secondary 65C30.

Key words and phrases. Weak approximation schemes, high order methods, cubature methods, extrapolation, Ninomiya–Victoir scheme, Fujiwara extrapolation method, Richardson extrapolation.

where $D^\gamma := D_1^{\gamma_1} \cdots D_N^{\gamma_N}$ for multi-indeces $\gamma = (\gamma_1, \dots, \gamma_N)$ of length $|\gamma| := \gamma_1 + \cdots + \gamma_N$.

We are interested in weak approximations of the solution processes of the SDE (1), that is, numerical methods for approximation of $E[f(X(t,x))]$ for some $f \in C_b^\infty(\mathbf{R}^N)$ (see [10]). By introducing some additional integrability conditions on the function f, the smoothness conditions of f and of the vector fields can be relaxed:

ASSUMPTION 1. Fix an integer $l \geq 1$, and denote $V_j = \sum_{k=1}^N c_{j,k} \frac{\partial}{\partial x^k}$. Let $f \in W_p^m$, and let the coefficients $c_{j,k}$ of the vector fields be bounded Borel functions satisfying the conditions that the partial derivatives

$$D^\rho c_{j,k}, \qquad j = 0, \dots, d, k = 1, \dots, N,$$

exist and are bounded in norm for all multi-indeces ρ, satisfying $|\rho| \leq l$.

It is well known that under such assumptions $E[f(X(t,x))]$ as a function of t has values in W_p^m and uniquely solves the corresponding Kolmogorov backward equation; see, for example, [7].

In practice, for example, in real-life problems in mathematical finance, it is often the case that we cannot write down an explicit formula for $E[f(X(t,x))]$, or that the dimension of the problem becomes very high, or that the underlying stochastic process lives on a (fairly complicated) manifold rather than on $\mathbf{R}^N$. Hence, we are forced to use numerical methods respecting the geometry of the problem involving (quasi) Monte Carlo integration. The widely used (higher) Itô–Taylor methods do not respect the geometry of problems and have a high complexity of implementation, both leading to low numerical orders of convergence. The pioneering approaches from Kusuoka (see [11, 12] and [13]) and of Lyons and Victoir (see [15]) addressed exactly the previously mentioned problems, leading to easier implementation and methods which respect the geometry. Recently some more concrete implementations of this abstract approach emerged (e.g., [16, 17], to mention just a few), one of the most successful being the so called Ninomiya–Victoir weak approximation scheme (see [17]), which is comparable to the well-known Strang scheme of numerical analysis.

We associate for later use the following simpler stochastic differential equations:

$$X^{(i)}(t,x) = x + \int_0^t V_i\bigl(X^{(i)}(s,x)\bigr) \circ dB_s^i. \tag{2}$$

Let $\{P_t\}_{t \in \mathbf{R}_+}$ and $\{P_t^{(i)}\}_{t \in \mathbf{R}_+}$ be the associated heat semigroups on $C_b^\infty(\mathbf{R}^N)$ such that $P_t f(x) := E[f(X(t,x))]$ for $t \geq 0$, and $P_t^{(i)} f(x) := E[f(X^{(i)}(t,x))]$ for

$t \geq 0$. Notice here that the equation associated to the index 0 is a pure drift equation, the semigroup therefore a transport semigroup. Let $\theta \in \mathbf{N}$ and denote furthermore by

$$\mathcal{A} := V_0 + \frac{1}{2}\sum_{i=1}^{d} V_i^2,$$

$$\overrightarrow{Q}_t^{[\theta]} := \left(P_{t/\theta}^{(0)} \circ \cdots \circ P_{t/\theta}^{(d)}\right)^{\theta},$$

$$\overleftarrow{Q}_t^{[\theta]} := \left(P_{t/\theta}^{(d)} \circ \cdots \circ P_{t/\theta}^{(0)}\right)^{\theta},$$

$$Q_t^{[\theta]} := \tfrac{1}{2}\left(\overrightarrow{Q}_t^{[\theta]} + \overleftarrow{Q}_t^{[\theta]}\right),$$

the generator of the diffusion process (1), two ordered products of (semi-)flows with generators V_0 and V_i^2 and the average of the two ordered products $Q^{[\theta]}$. Then we have the well-known short time asymptotics, formulated in the language of k-norms (see Definition 7)

$$\left|P_t g(x) - Q_t^{[\theta]} g(x)\right| \leq C t^3 \|g\|_{6\theta(d+1)}$$

as $t \to 0$, leading—by iteration—to (a variant of) the Ninomiya–Victoir scheme. Indeed, when we define n-fold iteration of the operator $Q_{T/n}^{[\theta]}$

$$Q_{T,n}^{[\theta]} = Q_{T/n}^{[\theta]} \circ \cdots \circ Q_{T/n}^{[\theta]},$$

we obtain a scheme of weak approximation order $r = 2$, that is,

$$\left|P_T g(x) - Q_{T,n}^{[\theta]} g(x)\right| \leq \frac{C}{n^2} \|g\|_{6\theta(d+1)}.$$

Let us now define formally a weak approximation method of P_T for some fixed, finite $T \in \mathbf{R}_+$ of order r.

DEFINITION 1. A family of linear operators $\{Q_{T,n}\}_{n \in \mathbf{N}}$ on $C_b^\infty(\mathbf{R}^N)$, continuous with respect to the supremum norm topology, is called a *weak approximation method of order r* if there exists $C > 0$ and some number $k \geq 0$ such that

$$|P_T f(x) - Q_{T,n} f(x)| \leq \frac{C}{n^r} \|f\|_k \tag{3}$$

for all $x \in \mathbf{R}^N$ and for all $f \in C_b^\infty(\mathbf{R}^N)$.

Notice that the operator $Q_{T,n}$ is only supposed to be linear and continuous with respect to the supremum norm topology on the set of C_b^∞-function, but not necessarily of sub-Markovian type. This means in particular that classical (Richardson) extrapolations belong to this class of approximations.

REMARK 1. If, in addition, the operators $Q_{T,n}$ are of sub-Markovian type, the weak approximation method $\{Q_{T,n}\}_{n\in\mathbf{N}}$ of order r will be called a *weak approximation scheme of order* r.

Denote $h = \frac{T}{n}$. Having a weak approximation scheme $Q_{T,n} = (Q_h)^n$ for some operator Q_h related to some short time asymptotics h^{k+1}, the classical Richardson extrapolation method requires that one knows the expansion of the error term $\mathrm{Err}(T,h) = P_T f(x) - Q_{T,n} f(x)$ as a function of h

$$\begin{aligned}\mathrm{Err}(T,h) = {} & e_1(T)h^r + e_2(T)h^{k+1} + \cdots + e_m(T)h^{r+m} \\ & + O(h^{r+m+1})\end{aligned} \tag{4}$$

up to a certain degree m. By using a linear combination of $Q_{T,n}$ with different time refinements, that is, different multiples θh of the subdivision intervals of $[0,T]$, we can eliminate the error terms $e_1(T)h^r$ up to $e_m(T)h^{r+m}$ and thus reduce the discretization error of the method. In the case of smooth payoff function f, the error expansion of the type (4) is known to exist for the Euler–Maruyama scheme and for the Milshtein scheme up to an arbitrary order (see [18] and the impressive work [7]). Since the repeated use of the Richardson extrapolation, for example, with Euler–Maruyama schemes, exhibits numerical instabilities, it is preferable to start with higher order schemes and to perform as little as possible extrapolation steps. For the Ninomiya–Victoir scheme the error expansion has been proven only up to the first term $e_1(T)h$ so far (see [14] for the proof for bounded Borel payoff functions f). For references and several important remarks on the Richardson extrapolation of a splitting scheme of Euler type see in particular [7].

The main idea of our work is to combine several different approximation schemes of global order $r = 2$, which approximate the same equation and which have similar error expansions: a linear combination of those *different* schemes makes certain orders of the expansion disappear. We attribute this idea to Fujiwara [5]. Fujiwara constructs a weak approximation method of order six for smooth functions $C_b^\infty(\mathbf{R}^N)$, which consists of a linear combination of several previously described Ninomiya–Victoir schemes. Instead of a linear combination of approximation schemes with different time refinements θh, Fujiwara uses linear combination of $Q_{T,n}^{[\theta]}$ for different values of θ, that is, the extrapolation is performed by time refinement of the basic building blocks of the Ninomiya–Victoir scheme rather than of the whole scheme. The fact that Ninomiya–Victoir scheme is a particular implementation of the Kusuoka–Lyons–Victoir approach enables us to use techniques of free Lie algebra and of generalized power series to prove the annihilation of terms. In this paper, we define two weak approximation methods, namely the *generalized Fujiwara method of order* $r = 2m$, $m \in \mathbf{N}$, including the extrapolation method in [5] and the *extrapolated symmetrized splitting method of order* $r = 2m$. We finally obtain the following Theorem 6, which proves the convergence order of the generalized Fujiwara method and whose proof can be found in Section 4.

Let $Q_t^{[\theta]}$ be as above and let $\theta_1, \theta_2, \dots, \theta_m \in \mathbf{N}$ be pairwise distinct. Furthermore let $V = [V_{i,j}]_{i=1,\dots,m, j=1,\dots,m}$, $V_{i,j} = 1/\theta_i^{2(j-1)}$ be a Vandermonde matrix of dimension m. Denote by $f = [f_i]_{i=1,\dots,m} = [1, 0, \dots, 0] \cdot V^{-1}$, then

$$Q_{T,n} := \sum_{i=1}^{m} f_{\theta_i} \big(Q_{T/n}^{[\theta_i]}\big)^n$$

for $n \geq 0$ is a weak approximation method of order $2m$, where a choice of k is given by

$$k = 2(2m+1)(d+1) \max\{\theta_i | i = 1, \dots, m\}$$

that means

$$|P_T g(x) - Q_{T,n} g(x)| \leq \frac{C}{n^{2m}} \|g\|_k$$

for test functions $g \in C_b^\infty(\mathbb{R}^N)$.

On the other hand, Theorem 5 provides the order of convergence of the extrapolated symmetrized splitting method:

Let m, $Q_t^{[\theta]}$, $\theta_1, \dots, \theta_m$, V, f, n and k be as above, then

$$Q_{T,n}^s := \sum_{i=1}^{m} f_{\theta_i} Q_T^{[n\theta_i]}$$

is a weak approximation method of order $2m$; that means

$$|P_T g(x) - Q_{T,n}^s g(x)| \leq \frac{C}{n^{2m}} \|g\|_k$$

for test functions $g \in C_b^\infty(\mathbb{R}^N)$.

The results also hold—by results from [7]—with respect to Sobolev norms. Notice that in comparison to classical methods, we do only need half of the extrapolation steps to achieve order $2m$.

The use of weak approximation methods is prone to two kind of errors. The discretization error comes from the method/scheme itself and measures how well the scheme approximates the given operator P_T. On the other hand, the integration error is a result of the (quasi) Monte Carlo integration. By combining the schemes using a linear combination to reduce the discretization error, the integration error can gain an additional factor, which equals at most the sum of absolute values of the coefficients of the linear combination of the combined approximation schemes. Usually, for example, Richardson extrapolation or generalized Fujiwara method or of symmetrized Euler schemes, this sum gets significantly larger with every new summand in the linear combination, which finally can lead to numerical instabilities.

Here the most important contribution of this work gets relevant: according to Theorems 5 or 6 we need to combine only m Ninomiya–Victoir or Euler schemes in order to achieve the numerical order $2m$, $m \in \mathbf{N}$, as discretization error. This leads to a tremendous acceleration and stabilization of approximation methods.

The remainder of the article is organized as follows: in Section 2 we introduce all algebraic prerequisites, in Section 3 we show the main algebraic result of this article, which is then applied in Section 4 to prove the existence of generalized Fujiwara method. In Section 5 we provide an implementation result for the generalized Fujiwara method, where the results can be compared to [16].

2. Algebraic prerequisites and their relation to weak approximation. Let A be the set of elements $a_0, \dots, a_d$. We call A an alphabet and $a_0, \dots, a_d$ letters. A word in alphabet A is a finite sequence of letters. Let 1 be an empty word and A^* a set of words including 1. If we impose a total ordering on A, then A^* together with word concatenation and lexicographic ordering becomes an ordered unital semigroup. Let $\mathbf{R}\langle A\rangle$ be a set of noncommutative polynomials on A^* over $\mathbf{R}$, that is, a set of $\mathbf{R}$-linear combinations of elements of A^*, and let $\mathbf{R}\langle\langle A\rangle\rangle$ be a set of noncommutative series of elements of A^* with coefficients in $\mathbf{R}$, that is, a set of functions $f\colon A^* \to \mathbf{R}$ with well-ordered support. Using componentwise addition and multiplication, which is induced by word concatenation, makes $\mathbf{R}\langle\langle A^*\rangle\rangle$ an $\mathbf{R}$-algebra; see [4] for more details. The degree of a monomial is a number of letters contained in the monomial and the degree of a noncommutative polynomial and a noncommutative series are the maximum degree of monomials contained in them. Let $\mathbf{R}\langle A\rangle_m$ and $\mathbf{R}\langle A\rangle_{\leq m}$ be the set of homogeneous polynomials of the degree m and the set of polynomials of the degree less or equal to m, respectively. Define $\mathbf{R}\langle\langle A\rangle\rangle_m$ and $\mathbf{R}\langle\langle A\rangle\rangle_{\leq m}$ in the same manner. Since every $u \in \mathbf{R}\langle\langle A\rangle\rangle$ has a well-ordered support, we can define $\mathbf{R}\langle\langle A\rangle\rangle_{>m} = \{u \in \mathbf{R}\langle\langle A\rangle\rangle \,|\, \deg(\inf(\operatorname{supp}(u))) > m\}$ and $\mathbf{R}\langle\langle A\rangle\rangle_{\geq m} = \{u \in \mathbf{R}\langle\langle A\rangle\rangle \,|\, \deg(\inf(\operatorname{supp}(u))) \geq m\}$, and it is easy to see that $\mathbf{R}\langle\langle A\rangle\rangle_{>m}$ and $\mathbf{R}\langle\langle A\rangle\rangle_{\geq m}$ are double-sided ideals in algebra $\mathbf{R}\langle\langle A\rangle\rangle$. Let j_m and $j_{\leq m}$ be the natural surjective maps from $\mathbf{R}\langle\langle A\rangle\rangle$ onto $\mathbf{R}\langle\langle A\rangle\rangle_m$ and $\mathbf{R}\langle\langle A\rangle\rangle_{\leq m}$, respectively.

Since every subset of A^* has a least element regarding lexicographical ordering, we have $\mathbf{R}\langle\langle A\rangle\rangle = \mathbf{R}^{A^*}$. The set A^* is countable, therefore taking metric topology in $\mathbf{R}$ makes $\mathbf{R}^{A^*}$ with induced product topology into a Polish space. Hence, we can consider its Borel σ-algebra $\mathcal{B}(\mathbf{R}\langle\langle A\rangle\rangle)$, $\mathbf{R}\langle\langle A\rangle\rangle$-valued random variables and expectations, and other notions as usual.

For $u \in \mathbf{R}\langle\langle A\rangle\rangle$ we define the exponential map

$$\exp(u) := \sum_{n\geq 0} \frac{u^n}{n!},$$

and for $u \in \mathbf{R}\langle\langle A\rangle\rangle$ with vanishing constant term, we define the logarithm

$$\log(1+u) := \sum_{n\geq 1} \frac{(-1)^{n-1}}{n} u^n.$$

It is easy to check that

$$\log(\exp(u)) = u, \tag{5}$$

$$\exp(\log(u)) = u \tag{6}$$

on the respective domains. For $\theta \in \mathbf{N}$ define

$$p := \exp\left(\sum_{i=0}^{d} a_i\right),$$

$$\overrightarrow{q}^{[\theta]} := \left(\exp\left(\frac{1}{\theta}a_0\right)\cdots\exp\left(\frac{1}{\theta}a_d\right)\right)^{\theta},$$

$$\overleftarrow{q}^{[\theta]} := \left(\exp\left(\frac{1}{\theta}a_d\right)\cdots\exp\left(\frac{1}{\theta}a_0\right)\right)^{\theta},$$

$$q^{[\theta]} := \tfrac{1}{2}(\overrightarrow{q}^{[\theta]} + \overleftarrow{q}^{[\theta]}).$$

Let us make the substitution, which is the heart of the transfer from algebra to numerical schemes, namely $a_0 \mapsto V_0, a_1 \mapsto V_1^2/2, \dots, a_d \mapsto V_d^2/2$, formally correct. First let B be another alphabet with elements $v_0, v_1, \dots, v_d$ and set $B^*, \mathbf{R}\langle B\rangle$, etc. in the same manner. For all $t \in \mathbf{R}_+$ define an algebra homomorphism $\Psi_t : \mathbf{R}\langle\langle A\rangle\rangle \longrightarrow \mathbf{R}\langle\langle B\rangle\rangle$ by setting

$$\Psi_t(a_0) := t v_0, \tag{7}$$

$$\Psi_t(a_i) := t v_i^2/2 \tag{8}$$

for all $i \in \{1, \dots, d\}$.

Recall that the vector fields under consideration are C^∞-bounded. Define next the algebra homomorphism Φ from the algebra $\mathbf{R}\langle B\rangle$ to the vector space of smooth differential operators (of finite order) on $C_b^\infty(\mathbf{R}^N)$ by setting

$$\Phi(v_i) = V_i. \tag{9}$$

The algebra of noncommutative words plays a major role in the analysis of weak approximation schemes due to the following well-known asymptotic expansion theorem.

THEOREM 1. *For all function $f \in C_b^\infty(\mathbf{R}^N)$, $x \in \mathbf{R}^N$ and $n \in \mathbf{N}$,*

$$P_t f(x) = \sum_{k=0}^{n} \frac{t^k}{k!}\mathcal{A}^k f(x) + \mathcal{O}(t^{n+1}) = \Phi(\Psi_t(j_{\le n} p)) f(x) + \mathcal{O}(t^{n+1}) \tag{10}$$

as $t \to 0$, holds true.

PROOF. By the Feynman–Kac formula (see [9]), $u(t,x) := P_t f(x)$ satisfies the PDE

$$\frac{d}{dt}u(t,x) = \mathcal{A}u(t,x) \tag{11}$$

with initial condition $u(0,x) = f(x)$. Again by [9], we have

$$\mathcal{A} P_t f = P_t \mathcal{A} f \tag{12}$$

for the Feller semigroup P. Hence $P_t f(x)$ is smooth with respect to $t \geq 0$. Thus we can use Taylor's formula around $t = 0$ and obtain the result. $\square$

Hence we can, for example, express the generator $\mathcal{A}$ of the diffusion process (1) by

$$\Phi\big(\Psi_1(a_0 + \cdots + a_d)\big) = \mathcal{A};$$

in particular we obtain the following crucial asymptotic formulas:

$$\Phi\Psi_t(j_{\leq n}(\exp(a_i)))f = P_t^{(i)} f + \mathcal{O}(t^{n+1})$$

as $t \to 0$ and $i = 0, \ldots, d$ again due to Theorem 1.

To be more precise on the goal of our paper, Theorem 1 allows us to approximate p by linear combinations of $(q^{[\theta]})^n$ up to a certain degree $2m-1$ within the algebra $\mathbf{R}\langle\langle A \rangle\rangle$ such that the remainder term is $\mathcal{O}(\frac{1}{n^{2m}})$. In this case $P_t f(x)$ can be approximated by linear combinations of

$$\Phi\big(\Psi_t\big(j_{\leq 2m}\big(q^{[\theta]}\big)^n\big)\big)f(x)$$

in a weak sense of order $2m$.

Since V_i are bounded C^∞-bounded vector fields, it is sufficient to consider global transport flows, whence the following definition.

DEFINITION 2. *The flow of the bounded C^∞-bounded vector field V* is a smooth map $\mathrm{Fl}^V : \mathbf{R} \times \mathbf{R}^N \to \mathbf{R}^N$ with the apparent flow properties:

(1) $\mathrm{Fl}_t^V(\mathrm{Fl}_s^V(x)) = \mathrm{Fl}_{s+t}^V(x)$, $x \in \mathbf{R}^N$;
(2) $\frac{d}{dt}\mathrm{Fl}_t^V(x) = V(\mathrm{Fl}_t^V(x))$, $x \in \mathbf{R}^N$;
(3) $\mathrm{Fl}_0^V(x) = x$ for any $x \in \mathbf{R}^N$.

Let $\{\exp(tV_i)\}_{t \in \mathbf{R}_+}$ be the transport group on $C_b^\infty(\mathbf{R}^N)$ generated by V_i, that is, for all $f \in C_b^\infty(\mathbf{R}^N)$, $\exp(tV_i)$ satisfies

$$\frac{d}{dt}\exp(tV_i)f = V_i \exp(tV_i) f, \tag{13}$$

which means in turn that $\exp(tV_i)f = f \circ \mathrm{Fl}_t^{V_i}$.

Analogously and motivated by Theorem 1 we denote

$$\begin{aligned} P^{(i)} f &= \exp\left(\frac{t}{2}V_i^2\right) f, \\ P_t f &= \exp\left(tV_0 + \frac{t}{2}\sum_{i=1}^d V_i^2\right) f \end{aligned} \tag{14}$$

for $f \in C_b^\infty(\mathbf{R}^N)$.

Due to Stratonovich calculus the processes $X^{(i)}(t,x)$ from (2), for $i = 1, \dots, d$, are given by the evaluation of the flow of the vector field V_i at times B_t^i, that is, $X^{(i)}(t,x) = \mathrm{Fl}^{V_i}_{B_t^i}(x)$.

The previous apparent consequences of Itô calculus can be applied to the algebraic setting itself by considering certain (stochastic) differential equations with state spaces $\mathbf{R}\langle\langle B\rangle\rangle_{\leq m}$. Let us define the linear vector field $W_i(x) := x v_i$, for $i = 0, \dots, d$. Its flow is apparently given by

$$\mathrm{Fl}_t^{W_i}(x) = x \exp(t v_i). \tag{15}$$

Then $Y^{(i)}(t,x) = \mathrm{Fl}^{W_i}_{B_t^i}(x)$, for $x \in \mathbf{R}\langle\langle B\rangle\rangle_{\leq m}$, holds true, where

$$dY^{(i)}(t,x) = W_i\big(Y^{(i)}(t,x)\big) \circ dB_t^i.$$

Taking expectations we are led to solutions of the PDE with generator $\frac{1}{2} W_i^2$. We solve the Kolmogorov backward PDE with a linear inital value functional f, its solution is

$$x \mapsto f\left(x \exp\left(t \frac{v_i^2}{2}\right)\right). \tag{16}$$

On the other hand the solution of the Kolmogorov equation is the expectation of $f(Y^{(i)}(t,x))$, which in turn equals $f(\mathrm{Fl}^{W_i}_{B_t^i}(x))$; hence we obtain

$$E[f(x \exp(B_t^i v_i))] = f\left(x \exp\left(t \frac{v_i^2}{2}\right)\right) \tag{17}$$

for $x \in \mathbf{R}\langle\langle B\rangle\rangle_{\leq m}$ and any linear functional f and for any m. Therefore the equation

$$E[\exp(B_t^i v_i)] = \exp\left(t \frac{v_i^2}{2}\right) \tag{18}$$

holds in the universal enveloping algebra $\mathbf{R}\langle\langle B\rangle\rangle$. This closes the circle to the previous general equations: namely, if we apply Φ to the last equation and evaluate at f, we obtain precisely the equation

$$E[f(\mathrm{Fl}_t^{V_i}(x))] = P_t^{(i)} f(x). \tag{19}$$

On the algebraic level we can also consider the differential operator $W_0 + \frac{1}{2} \times \sum_{i=1}^d W_i^2$, and hence derive the noncommutative generalization

$$E[Y(t,x)] = \exp\left(t\left(v_0 + \frac{1}{2}\sum_{i=1}^d v_i^2\right)\right), \tag{20}$$

where Y solves the stochastic differential equation

$$dY(t,x) = \sum_{i=0}^{d} W_i(Y(t,x)) \circ dB_t^i$$

in any $\mathbf{R}\langle\langle B\rangle\rangle_{\leq m}$ for any m.

3. How to approximate *p* by *q*? Let us repeat the basic abbreviations: for $\theta \in \mathbf{N}$ we have defined

$$\begin{aligned}
p &:= \exp\left(\sum_{i=0}^{d} a_i\right);\\
\overrightarrow{q}^{[\theta]} &:= \left(\exp\left(\frac{1}{\theta}a_0\right)\cdots\exp\left(\frac{1}{\theta}a_d\right)\right)^{\theta};\\
\overleftarrow{q}^{[\theta]} &:= \left(\exp\left(\frac{1}{\theta}a_d\right)\cdots\exp\left(\frac{1}{\theta}a_0\right)\right)^{\theta};\\
q^{[\theta]} &:= \tfrac{1}{2}\left(\overrightarrow{q}^{[\theta]} + \overleftarrow{q}^{[\theta]}\right).
\end{aligned}$$

LEMMA 1 ([5], Lemma 2.1). *We have*

$$\log \overleftarrow{q}^{[1]} = \sum_{i=1}^{\infty} (-1)^{i+1} j_i\left(\log \overrightarrow{q}^{[1]}\right). \tag{21}$$

We give an alternative proof of Lemma 1.

DEFINITION 3. Let $\mathfrak{g}$ be a Lie algebra. For $X, Y \in \mathfrak{g}$ define $c_1(X,Y) = X + Y$ and $c_n(X,Y)$ by the following recursion formula:

$$\begin{aligned}
&(n+1)c_{n+1}(X,Y)\\
&\quad = \frac{1}{2}[X - Y, c_n(X,Y)]\\
&\qquad + \sum_{p\geq 1, 2p\leq n} K_{2p} \sum_{\substack{k_1,\ldots,k_{2p}\geq 0,\\ k_1+\cdots+k_{2p}=n}} \left[c_{k_1}(X,Y), [\ldots, [c_{k_{2p}}(X,Y), X+Y]\ldots]\right],
\end{aligned}$$

where the coefficients $K_{2p} \in \mathbf{R}$ are defined in [22], (2.15.9).

Recall that the expression $c_k(X,Y)$, $k \in \mathbf{N}$, is exactly the homogeneous part of degree k of the Hausdorff series of the $\log(\exp(X)\exp(Y))$ which follows from the famous Baker–Campbell–Hausdorff formula. For more details about $c_n(X,Y)$ see [22], Section 2.15.

LEMMA 2.

$$c_n(X,Y) = (-1)^{n+1} c_n(Y,X). \tag{22}$$

PROOF. For $n=1$ the assertion is clear.

Suppose we have $c_m(X,Y) = (-1)^{m+1} c_m(Y,X)$ for all $m \le n$. By recursion we obtain

$$\begin{aligned}
&(n+1)c_{n+1}(Y,X) \\
&\quad = \frac{1}{2}[Y-X, c_n(Y,X)] \\
&\qquad + \sum_{p\ge 1, 2p\le n} K_{2p} \sum_{\substack{k_1,\dots,k_{2p}\ge 0,\\ k_1+\dots+k_{2p}=n}} \bigl[c_{k_1}(Y,X), [\dots, [c_{k_{2p}}(Y,X), X+Y]\dots]\bigr].
\end{aligned}$$

Using the induction hypothesis and bilinearity of Lie brackets, the above equation transforms into

$$\begin{aligned}
&(n+1)c_{n+1}(Y,X) \\
&\quad = \frac{1}{2}(-1)^{n+2}[X-Y, c_n(X,Y)] \\
&\qquad + \sum_{p\ge 1, 2p\le n} K_{2p} \sum_{\substack{k_1,\dots,k_{2p}\ge 0,\\ k_1+\dots+k_{2p}=n}} (-1)^{k_1+\dots+k_{2p}+2p} \\
&\qquad\qquad \times \bigl[c_{k_1}(X,Y), [\dots, [c_{k_{2p}}(X,Y), X+Y]\dots]\bigr] \\
&\quad = (-1)^{n+2}\Bigl(\frac{1}{2}[X-Y, c_n(X,Y)] \\
&\qquad\qquad + \sum_{p\ge 1, 2p\le n} K_{2p} \sum_{\substack{k_1,\dots,k_{2p}\ge 0,\\ k_1+\dots+k_{2p}=n}} \bigl[c_{k_1}(X,Y), \\
&\qquad\qquad\qquad [\dots, [c_{k_{2p}}(X,Y), X+Y]\dots]\bigr]\Bigr) \\
&\quad = (-1)^{n+2}(n+1)c_{n+1}(X,Y),
\end{aligned}$$

which proves the assertion. $\square$

DEFINITION 4. As before, assume that the alphabet A consists of letters $a_0,\dots,a_d$ for some $d \in \mathbf{N}_0$. Let $k \in \mathbf{N}_0$, $k \le d$ and let

$$\overrightarrow{q}^{[\theta]}(k) := \biggl(\exp\Bigl(\frac{1}{\theta}a_0\Bigr)\cdots\exp\Bigl(\frac{1}{\theta}a_k\Bigr)\biggr)^{\theta},$$

$$\overleftarrow{q}^{[\theta]}(k) := \left(\exp\left(\frac{1}{\theta}a_k\right)\cdots\exp\left(\frac{1}{\theta}a_0\right)\right)^{\theta}.$$

Clearly $\overrightarrow{q}^{[\theta]} = \overrightarrow{q}^{[\theta]}(d)$. Finally, let $\tau_{i,k}$ denote $j_i(\log(\overrightarrow{q}^{[1]}(k)))$.

PROOF OF LEMMA 1. The case $d = 0$ is trivial. Next we consider the case $d = 1$. Using the Baker–Campbell–Hausdorff formula to expand $\tau_{l,1}$ and $j_l(\log(\overleftarrow{q}^{[1]}(1)))$ and applying (22) proves the formula (21).

By applying the Baker–Campbell–Hausdorff formula to the definition of $\tau_{l,d}$, we get

$$\begin{aligned}\tau_{l,d} &= j_l\big(\log(\overrightarrow{q}^{[1]}(d))\big)\\ &= j_l\big(\log\big(\exp\big(\log(\overrightarrow{q}^{[1]}(d-1))\big)\exp(a_d)\big)\big)\\ &= j_l\left(\sum_{k=1}^{l} c_k\left(\sum_{j=1}^{l}\tau_{j,d-1}, a_d\right)\right).\end{aligned}$$

Suppose that for all $n \in \mathbf{N}$, $n < d$ we have

$$\log\big(\overleftarrow{q}^{[1]}(n)\big) = \sum_{i=1}^{\infty}(-1)^{i+1}\tau_{i,n}.$$

Using Lemma 2, the induction hypothesis and the BCH formula on

$$j_l\big(\log(\overleftarrow{q}^{[1]}(d))\big)$$

gives us

$$\begin{aligned}j_l\big(\log(\overleftarrow{q}^{[1]}(d))\big) &= j_l\big(\log\big(\exp(a_d)\exp\big(\log\big(\overleftarrow{q}^{[1]}(d-1)\big)\big)\big)\big)\\ &= j_l\left(\sum_{k=1}^{l} c_k\big(a_d, \log(\overleftarrow{q}^{[1]}(d-1))\big)\right)\\ &= j_l\left(\sum_{k=1}^{l} c_k\left(a_d, \sum_{j=1}^{l}(-1)^{j+1}\tau_{j,d-1}\right)\right)\\ &= j_l\left(\sum_{k=1}^{l}(-1)^{k+1}c_k\left(\sum_{j=1}^{l}(-1)^{j+1}\tau_{j,d-1}, a_d\right)\right).\end{aligned}$$

Thus, it is sufficient to show that for all $k \in \{1, \ldots, l\}$ and $l \in \mathbf{N}$ we have

$$j_l\left(c_k\left(\sum_{j=1}^{l}(-1)^{j+1}\tau_{j,d-1}, a_d\right)\right) = (-1)^{k+l} j_l\left(c_k\left(\sum_{j=1}^{l}\tau_{j,d-1}, a_d\right)\right). \tag{23}$$

Note that the equality in (23) holds trivially for $k > l$.

Since $\tau_{j,d-1}$ is a homogeneous polynomial of degree j, the assertion is clear for $k = 1$ and all $l \in \mathbf{N}$. It is easy to see that for $l' < l$ we have

$$
\begin{aligned}
& j_{l'}\left(c_m\left(\sum_{j=1}^{l}(-1)^{j+1}\tau_{j,d-1}, a_d\right)\right) \\
& \quad = j_{l'}\left(c_m\left(\sum_{j=1}^{l'}(-1)^{j+1}\tau_{j,d-1}, a_d\right)\right).
\end{aligned}
\tag{24}
$$

Let now

$$
j_l\left(c_m\left(\sum_{j=1}^{l}(-1)^{j+1}\tau_{j,d-1}, a_d\right)\right) = (-1)^{k+l} j_l\left(c_m\left(\sum_{j=1}^{l}\tau_{j,d-1}, a_d\right)\right)
$$

for all $m \in \{1, \dots, k\}$ and $l \in \mathbf{N}$; then we have

$$
\begin{aligned}
& j_l\left(c_{k+1}\left(\sum_{j=1}^{l}(-1)^{j+1}\tau_{j,d-1}, a_d\right)\right) \\
& \quad = \frac{1}{k+1} \\
& \qquad \times j_l\Biggl(\frac{1}{2}\left[\sum_{j=1}^{l}(-1)^{j+1}\tau_{j,d-1} - a_d, c_k\left(\sum_{j=1}^{l}(-1)^{j+1}\tau_{j,d-1}, a_d\right)\right] \\
& \qquad\quad + \sum_{\substack{p\geq 1\\ 2p\leq k}} K_{2p} \\
& \qquad\qquad \times \sum_{\substack{k_1,\dots,k_{2p}>0\\ k_1+\cdots+k_{2p}=k}} \Biggl[c_{k_1}\left(\sum_{j=1}^{l}(-1)^{j+1}\tau_{j,d-1}, a_d\right), \\
& \qquad\qquad\qquad \Biggl[c_{k_2}\left(\sum_{j=1}^{l}(-1)^{j+1}\tau_{j,d-1}, a_d\right), \dots, \\
& \qquad\qquad\qquad\quad \Biggl[c_{k_{2p}}\left(\sum_{j=1}^{l}(-1)^{j+1}\tau_{j,d-1}, a_d\right), \\
& \qquad\qquad\qquad\qquad \sum_{j=1}^{l}(-1)^{j+1}\tau_{j,d-1} + a_d\Biggr]\dots\Biggr]\Biggr]\Biggr)
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{1}{k+1} \\
&\quad \times \Biggl(\frac{1}{2}\sum_{j'=1}^{l-1}\Biggl[(-1)^{j'+1}\tau_{j',d-1},\, j_{l-j'}\Biggl(c_k\Biggl(\sum_{j=1}^{l}(-1)^{j+1}\tau_{j,d-1},a_d\Biggr)\Biggr)\Biggr] \\
&\qquad + \Biggl[-a_d,\, j_{l-1}\Biggl(c_k\Biggl(\sum_{j=1}^{l}(-1)^{j+1}\tau_{j,d-1},a_d\Biggr)\Biggr)\Biggr] \\
&\qquad + \sum_{\substack{p\ge 1\\ 2p\le k}} K_{2p} \sum_{\substack{k_1,\dots,k_{2p}>0\\ k_1+\cdots+k_{2p}=k}} \\
&\qquad\ \times \sum_{\substack{m_1,\dots,m_{2p+1}>0\\ m_1+\cdots+m_{2p+1}=l}} \Biggl[j_{m_1}\Biggl(c_{k_1}\Biggl(\sum_{j=1}^{l}(-1)^{j+1}\tau_{j,d-1},a_d\Biggr)\Biggr), \\
&\qquad\qquad\qquad \Biggl[j_{m_2}\Biggl(c_{k_2}\Biggl(\sum_{j=1}^{l}(-1)^{j+1}\tau_{j,d-1},a_d\Biggr)\Biggr),\dots, \\
&\qquad\qquad\qquad\quad \Biggl[j_{m_{2p}}\Biggl(c_{k_{2p}}\Biggl(\sum_{j=1}^{l}(-1)^{j+1}\tau_{j,d-1},a_d\Biggr)\Biggr), \\
&\qquad\qquad\qquad\qquad (-1)^{m_{2p+1}+1}\tau_{m_{2p+1},d-1} + j_{m_{2p+1}}(a_d)\Biggr]\dots\Biggr]\Biggr]\Biggr).
\end{aligned}
$$

Using (24), the induction hypothesis and the bilinearity of Lie brackets the above expression transforms into

$$
\begin{aligned}
&j_l\Biggl(c_{k+1}\Biggl(\sum_{j=1}^{l}(-1)^{j+1}\tau_{j,d-1},a_d\Biggr)\Biggr) \\
&\quad = \frac{1}{k+1} \\
&\qquad \times \Biggl(\frac{1}{2}\sum_{j'=1}^{l-1}(-1)^{j'+1+l-j'+k}\Biggl[\tau_{j',d-1},\, j_{l-j'}\Biggl(c_k\Biggl(\sum_{j=1}^{l}\tau_{j,d-1},a_d\Biggr)\Biggr)\Biggr] \\
&\qquad\quad + (-1)^{l+k-1}\Biggl[-a_d,\, j_{l-1}\Biggl(c_k\Biggl(\sum_{j=1}^{l}\tau_{j,d-1},a_d\Biggr)\Biggr)\Biggr] \\
&\qquad\quad + \sum_{\substack{p\ge 1\\ 2p\le k}} K_{2p} \sum_{\substack{k_1,\dots,k_{2p}>0\\ k_1+\cdots+k_{2p}=k}}
\end{aligned}
$$

$$\times \sum_{\substack{m_1,\dots,m_{2p+1}>0\\ m_1+\cdots+m_{2p+1}=l}} (-1)^{m_1+\cdots+m_{2p+1}+1+k_1+\cdots+k_{2p}}$$

$$\times \Bigg[j_{m_1}\Bigg(c_{k_1}\Bigg(\sum_{j=1}^{l}\tau_{j,d-1}, a_d\Bigg)\Bigg),$$

$$\Bigg[j_{m_2}\Bigg(c_{k_2}\Bigg(\sum_{j=1}^{l}\tau_{j,d-1}, a_d\Bigg)\Bigg), \dots,$$

$$\Bigg[j_{m_{2p}}\Bigg(c_{k_{2p}}\Bigg(\sum_{j=1}^{l}\tau_{j,d-1}, a_d\Bigg)\Bigg),$$

$$\tau_{m_{2p+1},d-1} + j_{m_{2p+1}}(a_d)\Bigg]\dots\Bigg]\Bigg]\Bigg).$$

Thus,

$$j_l\Bigg(c_{k+1}\Bigg(\sum_{j=1}^{l}(-1)^{j+1}\tau_{j,d-1}, a_d\Bigg)\Bigg)$$

$$= (-1)^{k+l+1}\frac{1}{k+1}$$

$$\times j_l\Bigg(\frac{1}{2}\Bigg[\sum_{j=1}^{l}\tau_{j,d-1} - a_d, c_k\Bigg(\sum_{j=1}^{l}\tau_{j,d-1}, a_d\Bigg)\Bigg]$$

$$+ \sum_{\substack{p\ge 1\\ 2p\le k}} K_{2p}$$

$$\times \sum_{\substack{k_1,\dots,k_{2p}>0\\ k_1+\cdots+k_{2p}=k}} \Bigg[c_{k_1}\Bigg(\sum_{j=1}^{l}\tau_{j,d-1}, a_d\Bigg),$$

$$\Bigg[c_{k_2}\Bigg(\sum_{j=1}^{l}\tau_{j,d-1}, a_d\Bigg), \dots,$$

$$\Bigg[c_{k_{2p}}\Bigg(\sum_{j=1}^{l}\tau_{j,d-1}, a_d\Bigg), \sum_{j=1}^{l}\tau_{j,d-1} + a_d\Bigg]\dots\Bigg]\Bigg]\Bigg)$$

$$= (-1)^{k+l+1} j_l\Bigg(c_{k+1}\Bigg(\sum_{j=1}^{l}\tau_{j,d-1}, a_d\Bigg)\Bigg),$$

which is the desired result. $\square$

PROPOSITION 1 ([5], Proposition 2.2). *There exists $c_i \in \mathbf{R}\langle\langle A\rangle\rangle_{\geq 2i+1}$ such that for all $\theta \in \mathbf{N}$,*

$$q^{[\theta]} = p + \sum_{i=1}^{\infty} \frac{c_i}{\theta^{2i}}$$

holds true.

Let us recall the proof of the Proposition 1 from [5]:

PROOF OF PROPOSITION 1. Let $t \in \mathbf{R}$ and let us denote

$$\begin{aligned}
\overrightarrow{q}_t^{[\theta]} &:= \left(\exp\left(\frac{t}{\theta}a_0\right)\cdots\exp\left(\frac{t}{\theta}a_d\right)\right)^{\theta},\\
\overleftarrow{q}_t^{[\theta]} &:= \left(\exp\left(\frac{t}{\theta}a_d\right)\cdots\exp\left(\frac{t}{\theta}a_0\right)\right)^{\theta},\\
q_t^{[\theta]} &:= \tfrac{1}{2}(\overrightarrow{q}_t^{[\theta]} + \overleftarrow{q}_t^{[\theta]}).
\end{aligned}$$

It is clear that $\overrightarrow{q}_1^{[\theta]} = \overrightarrow{q}^{[\theta]}$, $\overleftarrow{q}_1^{[\theta]} = \overleftarrow{q}^{[\theta]}$ and $q_1^{[\theta]} = q^{[\theta]}$. Obviously,

$$\log(\overrightarrow{q}^{[\theta]}) = \log((\overrightarrow{q}_{1/\theta}^{[1]})^{\theta}).$$

By definition of $\tau_{i,d}$, the former expression amounts to

$$\log(\overrightarrow{q}^{[\theta]}) = \sum_{i=1}^{\infty} \frac{1}{\theta^{i-1}}\tau_{i,d}.$$

Analogously, by using the Lemma 1, we get

$$\log(\overleftarrow{q}^{[\theta]}) = \sum_{i=1}^{\infty} \frac{1}{(-\theta)^{i-1}}\tau_{i,d}.$$

Hence,

$$q^{[\theta]} = \frac{1}{2}\left(\sum_{k=0}^{\infty}\frac{1}{k!}\left(\sum_{i=1}^{\infty}\frac{1}{\theta^{i-1}}\tau_{i,d}\right)^{k} + \sum_{k=0}^{\infty}\frac{1}{k!}\left(\sum_{i=1}^{\infty}\frac{1}{(-\theta)^{i-1}}\tau_{i,d}\right)^{k}\right).$$

Since $\tau_{1,d} = \sum_{i=0}^{d} a_i$, this proves the assertion. $\square$

COROLLARY 1. *Let q be a linear combination of $q^{[\theta]}$ for some $\theta \in \mathbf{N}$. If there exists $n \in \mathbf{N}$ such that $j_{\leq 2n-1}(q) = j_{\leq 2n-1}(p)$, then $j_{\leq 2n}(q) = j_{\leq 2n}(p)$.*

PROOF. For all $\theta \in \mathbf{N}$, $j_{\leq 2}(q^{[\theta]}) = j_{\leq 2}(p)$ holds. Hence, the case $n = 1$ is clear. Suppose $n \geq 2$ and $j_{\leq 2n-1}(q) = j_{\leq 2n-1}(p)$. Since $q = \sum_{j=1}^{k} \alpha_j q^{[\theta_j]}$ for some $\theta_i \in \mathbf{N}$, $\theta_i \neq \theta_j$ for $i \neq j$ and since $j_{\leq 2}(q^{[\theta]}) = j_{\leq 2}(p)$ for all $\theta \in \mathbf{N}$, it follows $\sum_{j=1}^{k} \alpha_j = 1$. According to Proposition 1,

$$q = p + \sum_{i=1}^{\infty} c_i \left(\sum_{j=1}^{k} \alpha_j \frac{1}{\theta_j^{2i}} \right)$$

for some $c_i \in \mathbf{R}\langle\langle A \rangle\rangle_{\geq 2i+1}$. Since $j_{\leq 2n-1}(q) = j_{\leq 2n-1}(p)$, we have

$$\sum_{j=1}^{k} \alpha_j \frac{1}{\theta_j^{2i}} = 0$$

for all $i = 1, \dots, n-1$. Then

$$q - p = \sum_{i=n}^{\infty} c_i \left(\sum_{j=1}^{k} \alpha_j \frac{1}{\theta_j^{2i}} \right) \qquad \text{where } c_n \in \mathbf{R}\langle\langle A \rangle\rangle_{\geq 2n+1},$$

which proves the corollary. $\square$

Set

$$V := \begin{bmatrix} 1 & \cdots & 1 \\ 1/\theta_1^2 & \cdots & 1/\theta_m^2 \\ \vdots & \ddots & \vdots \\ 1/\theta_1^{2(m-1)} & \cdots & 1/\theta_m^{2(m-1)} \end{bmatrix}.$$

COROLLARY 2.

$$j_{\leq 2m} \left(\left(V^{-1} \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \right)^T \begin{bmatrix} q^{[\theta_1]} - p \\ \vdots \\ q^{[\theta_m]} - p \end{bmatrix} \right) = 0 \tag{25}$$

holds true.

COROLLARY 3. *For all $l \in \{1, \dots, m-1\}$,*

$$\left(V^{-1} \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \right)^T \begin{bmatrix} \frac{1}{\theta_1^{2l}} \\ \vdots \\ \frac{1}{\theta_m^{2l}} \end{bmatrix} = 0. \tag{26}$$

DEFINITION 5 (Abstract extrapolated symmetrized splitting scheme). A family of series

$$\left\{ q_n^s := \sum_{i=1}^{m} f_{\theta_i} q^{[n\theta_i]} \right\}_{n \in \mathbf{N}} \tag{27}$$

is called an *abstract extrapolated symmetrized splitting scheme* of (symbolic) order $2m$ if

$$\begin{aligned} f &= [\, f_{\theta_1} \quad \cdots \quad f_{\theta_m} \,]^T \\ &= V^{-1} \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \end{aligned}$$

holds.

REMARK 2. Apparently $q_n^s - p = \sum_{i \geq 2m} \frac{d_i}{n^i}$ for some $d_i \in \mathbf{R}\langle\langle A \rangle\rangle$, giving rise to an extrapolation method.

4. Abstract generalized Fujiwara scheme and its property. The following section deals with the (abstract) generalized Fujiwara scheme and its properties. Notice that the abstract generalized Fujiwara scheme seems more complicated than the abstract extrapolated symmetrized splitting scheme.

DEFINITION 6 (Abstract generalized Fujiwara scheme). A family of series

$$\left\{ q_n := \sum_{i=1}^{m} f_{\theta_i} \left(q^{[\theta_i]} \right)^n \right\}_{n \in \mathbf{N}} \tag{28}$$

is called an *abstract generalized Fujiwara scheme* of (symbolic) order $2m$ if

$$\begin{aligned} f &= [\, f_{\theta_1} \quad \cdots \quad f_{\theta_m} \,]^T \\ &= V^{-1} \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \end{aligned}$$

holds.

A straightforward calculation involving induction gives the following connection concerning the powers of series in $\mathbf{R}\langle\langle A \rangle\rangle$. Notice that we split the product $q^n - p^n$ into telescoping summands, where one, two up to m terms of the form $q - p$ appear.

PROPOSITION 2. *For $p, q \in \mathbf{R}\langle\langle A\rangle\rangle$ and for all $1 \le m \le n$*

$$
\begin{aligned}
q^n - p^n &= \sum_{l=1}^{m-1}\Biggl(\sum_{k_l=l-1}^{n-1}\sum_{k_{l-1}=l-2}^{k_l-1}\cdots\sum_{k_2=1}^{k_3-1}\sum_{k_1=0}^{k_2-1} p^{k_1}(q-p)p^{(k_2-k_1-1)}(q-p)\times\cdots \\
&\qquad\qquad \times p^{(k_l-k_{l-1}-1)}(q-p)p^{n-k_l-1}\Biggr) \qquad (29)\\
&\quad + \sum_{k_m=m-1}^{n-1}\sum_{k_{m-1}=m-2}^{k_m-1}\cdots\sum_{k_2=1}^{k_3-1}\sum_{k_1=0}^{k_2-1} q^{k_1}(q-p)p^{k_2-k_1-1}(q-p)\times\cdots \\
&\qquad\qquad \times p^{k_m-k_{m-1}-1}(q-p)p^{n-k_m-1}
\end{aligned}
$$

holds true.

PROOF. Fix $n \in \mathbf{N}$. For $m = 1$ we get the well-known identity

$$q^n - p^n = \sum_{k=0}^{n-1} q^k(q-p)p^{n-k-1}.$$

For $m = 2$ the assertion follows using straightforward calculation, and for $m = 3$ the decomposition can be found in [5]. Let $M \in \mathbf{N}$, $M \le n$, and suppose the decomposition is valid for all $m \le M - 1$. Let us start from the RHS and observe the second set of summands for $m = M$,

$$
\begin{aligned}
&\sum_{k_M=M-1}^{n-1}\sum_{k_{M-1}=M-2}^{k_M-1}\cdots\sum_{k_2=1}^{k_3-1}\sum_{k_1=0}^{k_2-1} q^{k_1}(q-p)p^{k_2-k_1-1}(q-p)\times\cdots \\
&\qquad\qquad \times p^{k_M-k_{M-1}-1}(q-p)p^{n-k_M-1}. \qquad (30)
\end{aligned}
$$

By multiplying the first two terms on the left, expression (30) decomposes into two parts,

$$
\begin{aligned}
&\sum_{k_M=M-1}^{n-1}\sum_{k_{M-1}=M-2}^{k_M-1}\cdots\sum_{k_2=1}^{k_3-1}\sum_{k_1=0}^{k_2-1} q^{k_1+1}p^{k_2-k_1-1}(q-p)\times\cdots \\
&\qquad\qquad \times p^{k_M-k_{M-1}-1}(q-p)p^{n-k_M-1}
\end{aligned}
$$

and

$$
\begin{aligned}
&-\sum_{k_M=M-1}^{n-1}\sum_{k_{M-1}=M-2}^{k_M-1}\cdots\sum_{k_2=1}^{k_3-1}\sum_{k_1=0}^{k_2-1} q^{k_1}p^{k_2-k_1}(q-p)\times\cdots \\
&\qquad\qquad \times p^{k_M-k_{M-1}-1}(q-p)p^{n-k_M-1}.
\end{aligned}
$$

By canceling the terms of the form

$$q^{k_1+1}p^{k_2-k_1-1}(q-p)\cdots p^{k_M-k_{M-1}-1}(q-p)p^{n-k_M-1}$$

with the terms

$$-q^{k_1+1}p^{k_2-(k_1+1)}(q-p)\cdots p^{k_M-k_{M-1}-1}(q-p)p^{n-k_M-1}$$

in the last expression, we get the following remaining terms:

$$\begin{aligned}
&-\sum_{k_M=M-1}^{n-1}\sum_{k_{M-1}=M-2}^{k_M-1}\cdots\sum_{k_2=1}^{k_3-1}p^{k_2}(q-p)p^{k_3-k_2-1}(q-p)\times\cdots\\
&\qquad\times p^{k_M-k_{M-1}-1}(q-p)p^{n-k_M-1}\\
&+\sum_{k_M=M-1}^{n-1}\sum_{k_{M-1}=M-2}^{k_M-1}\cdots\sum_{k_2=1}^{k_3-1}q^{k_2}(q-p)p^{k_3-k_2-1}(q-p)\times\cdots\\
&\qquad\times p^{k_M-k_{M-1}-1}(q-p)p^{n-k_M-1}.
\end{aligned}$$

Reindexing the above expression gives

$$\begin{aligned}
&-\sum_{k_{M-1}=M-1}^{n-1}\sum_{k_{M-2}=M-2}^{k_M-1}\cdots\sum_{k_1=1}^{k_2-1}p^{k_1}(q-p)p^{k_2-k_1-1}(q-p)\times\cdots\\
&\qquad\times p^{k_{M-1}-k_{M-2}-1}(q-p)p^{n-k_{M-1}-1}
\end{aligned}\tag{31}$$

and

$$\begin{aligned}
&+\sum_{k_{M-1}=M-1}^{n-1}\sum_{k_{M-2}=M-2}^{k_{M-1}-1}\cdots\sum_{k_1=1}^{k_2-1}q^{k_1}(q-p)p^{k_2-k_1-1}(q-p)\times\cdots\\
&\qquad\times p^{k_{M-1}-k_{M-2}-1}(q-p)p^{n-k_{M-1}-1}.
\end{aligned}\tag{32}$$

Let us closely observe the summand at $l=M-1$ in the original sum (29) for $m=M$, that is,

$$\begin{aligned}
&\sum_{k_{M-1}=M-2}^{n-1}\sum_{k_{M-2}=M-3}^{k_{M-1}-1}\cdots\sum_{k_2=1}^{k_3-1}\sum_{k_1=0}^{k_2-1}p^{k_1}(q-p)p^{(k_2-k_1-1)}(q-p)\times\cdots\\
&\qquad\times p^{(k_{M-1}-k_{M-2}-1)}(q-p)p^{n-k_{M-1}-1}.
\end{aligned}\tag{33}$$

Summing up (33) with (31) gives us all the summands of (33) with the exponent $k_1=0$, that is,

$$\begin{aligned}
&\sum_{k_{M-1}=M-2}^{n-1}\sum_{k_{M-2}=M-3}^{k_{M-1}-1}\cdots\sum_{k_2=1}^{k_3-1}(q-p)p^{(k_2-k_1-1)}(q-p)\times\cdots\\
&\qquad\times p^{(k_{M-1}-k_{M-2}-1)}(q-p)p^{n-k_{M-1}-1}.
\end{aligned}$$

Combining this with (32) produces

$$\sum_{k_{M-1}=M-2}^{n-1}\sum_{k_{M-2}=M-3}^{k_{M-1}-1}\cdots\sum_{k_2=1}^{k_3-1}\sum_{k_1=0}^{k_2-1} q^{k_1}(q-p)p^{k_2-k_1-1}(q-p)\times\cdots \\ \times p^{k_{M-1}-k_{M-2}-1}(q-p)p^{n-k_{M-1}-1}.$$

Adding the above sum to the remaining summands in (29) for $l=1,2,\dots,M-2$ results in

$$\sum_{l=1}^{M-2}\left(\sum_{k_l=l-1}^{n-1}\sum_{k_{l-1}=l-2}^{k_l-1}\cdots\sum_{k_2=1}^{k_3-1}\sum_{k_1=0}^{k_2-1} p^{k_1}(q-p)p^{(k_2-k_1-1)}(q-p)\times\cdots \right.\\ \left.\times p^{(k_l-k_{l-1}-1)}(q-p)p^{n-k_l-1}\right) \\ +\sum_{k_{M-1}=M-2}^{n-1}\sum_{k_{M-2}=M-3}^{k_{M-1}-1}\cdots\sum_{k_2=1}^{k_3-1}\sum_{k_1=0}^{k_2-1} q^{k_1}(q-p)p^{k_2-k_1-1}(q-p)\times\cdots \\ \times p^{k_{M-1}-k_{M-2}-1}(q-p)p^{n-k_{M-1}-1},$$

which equals to q^n-p^n by the induction hypothesis. □

LEMMA 3. *For* $z_1,z_2\in\mathbf{R}\langle\langle A\rangle\rangle$, *if* $j_{\le l}(z_1)=0$ *and* $j_{\le m}(z_2)=0$, *then* $j_{\le l+m+1}(z_1z_2)=0$.

PROOF. By the assumption, monomials with the lowest degree contained in z_1 and z_2 are of the degree $l+1$ and $m+1$. Then, monomial with the lowest degree contained in z_1z_2 has the degree $l+m+2$. Hence $j_{\le l+m+1}(z_1z_2)=0$. □

COROLLARY 4. *For* $z_1,z_2,z_3\in\mathbf{R}\langle\langle A\rangle\rangle$, *if* $j_{\le l}(z_1)=j_{\le l}(z_2)$ *and* $j_{\le m}(z_3)=0$, *then* $j_{\le l+m+1}(z_1z_3)=j_{\le l+m+1}(z_2z_3)$.

COROLLARY 5. *For* $z\in\mathbf{R}\langle\langle A\rangle\rangle$, *if* $j_{\le l}(z)=0$, *then* $j_{\le ml+m-1}(z^m)=0$.

THEOREM 2. *If a series*

$$q_n := \sum_{i=1}^{m} f_{\theta_i}\big(q^{[\theta_i]}\big)^n \tag{34}$$

is an abstract Fujiwara scheme of order $2m$, *then for all* $l\in\{2,\dots,m-1\}$,

$$j_{\le 2m+l-1}\left(\sum_{i=1}^{m} f_{\theta_i}\big(q^{[\theta_i]}-p\big)^l\right)=0 \tag{35}$$

holds true.

PROOF. Fix $l \in \{2, \dots, m-1\}$. By Proposition 1,

$$\sum_{i=1}^{m} f_{\theta_i}\big(q^{[\theta_i]} - p\big)^l = \sum_{i=1}^{m} f_{\theta_i} \sum_{i_1,\dots,i_l=1}^{\infty} \frac{c_{i_1}\cdots c_{i_l}}{\theta_i^{2(i_1+\cdots+i_l)}} \tag{36}$$

holds. It is easy to see that

$$c_{i_1}\cdots c_{i_l} \in \mathbf{R}\langle\langle A\rangle\rangle_{\geq 2(i_1+\cdots+i_l)+l}.$$

Hence, we have

$$\begin{aligned}
&j_{\leq 2m+l-1}\left(\sum_{i_1,\dots,i_l=1}^{\infty} \frac{c_{i_1}\cdots c_{i_l}}{\theta_i^{2(i_1+\cdots+i_l)}}\right)\\
&\quad = j_{\leq 2m+l-1}\left(\sum_{\substack{i_1,\dots,i_l\geq 1\\ i_1+\cdots+i_l\leq m-1}} \frac{c_{i_1}\cdots c_{i_l}}{\theta_i^{2(i_1+\cdots+i_l)}}\right)\\
&\quad = j_{\leq 2m+l-1}\left(\sum_{k=l}^{m-1}\sum_{\substack{i_1,\dots,i_l\geq 1\\ i_1+\cdots+i_l=k}} \frac{c_{i_1}\cdots c_{i_l}}{\theta_i^{2k}}\right).
\end{aligned}$$

Thus, we have

$$\begin{aligned}
&j_{\leq 2m+l-1}\left(\sum_{i=1}^{m} f_{\theta_i}\big(q^{[\theta_i]} - p\big)^l\right)\\
&\quad = j_{\leq 2m+l-1}\left(\sum_{k=l}^{m-1}\sum_{\substack{i_1,\dots,i_l\geq 1\\ i_1+\cdots+i_l=k}} \sum_{i=1}^{m} f_{\theta_i}\frac{1}{\theta_i^{2k}} c_{i_1}\cdots c_{i_l}\right)\\
&\quad = j_{\leq 2m+l-1}\left(\sum_{k=l}^{m-1}\sum_{\substack{i_1,\dots,i_l\geq 1\\ i_1+\cdots+i_l=k}} f^T \begin{bmatrix} \frac{1}{\theta_1^{2k}}\\ \vdots \\ \frac{1}{\theta_m^{2k}}\end{bmatrix} c_{i_1}\cdots c_{i_l}\right)\\
&\quad = 0
\end{aligned}$$

by Corollary 3. □

THEOREM 3. *If $\{q_n\}_{n\in\mathbf{N}}$ is an abstract Fujiwara scheme of order m, then*

$$\Psi_{1/n}q_n - \Psi_1 p = \sum_{l=1}^{m}\sum_{k_l=l-1}^{n-1}\sum_{k_{l-1}=l-2}^{k_l-1}\cdots\sum_{k_2=1}^{k_3-1}\sum_{k_1=0}^{k_2-1} \Psi_{1/n} a_{l,(k_l,\dots,k_1)},$$

where $j_{\leq 2m+l-1}(a_{l,(k_l,\dots,k_1)}) = 0$ *for* $l = 1, \dots, m$.

PROOF. The case $m=1$ is trivial.

Let $m \geq 2$, and let $\{q_n := \sum_{i=1}^m f_{\theta_i}(q^{[\theta_i]})^n\}_{n\in\mathbf{N}}$ be an abstract Fujiwara scheme of order m. Note that $\Psi_1 p = (\Psi_{1/n} p)^n$. Then by Proposition 2, we have

$$
\begin{aligned}
&\Psi_{1/n} q_n - \Psi_1 p \\
&\quad = \sum_{i=1}^m f_{\theta_i} \sum_{k=0}^{n-1} (\Psi_{1/n} p)^k \big(\Psi_{1/n} q^{[\theta_i]} - \Psi_{1/n} p\big)(\Psi_{1/n} p)^{n-k-1} \\
&\qquad + \sum_{i=1}^m f_{\theta_i} \sum_{l=2}^{m-1} \sum_{k_l=l-1}^{n-1} \sum_{k_{l-1}=l-2}^{k_l-1} \cdots \sum_{k_2=1}^{k_3-1} \sum_{k_1=0}^{k_2-1} (\Psi_{1/n} p)^{k_1} \big(\Psi_{1/n} q^{[\theta_i]} - \Psi_{1/n} p\big) \\
&\qquad\qquad \times (\Psi_{1/n} p)^{k_2-k_1-1} \times \cdots \\
&\qquad\qquad \times \big(\Psi_{1/n} q^{[\theta_i]} - \Psi_{1/n} p\big) \\
&\qquad\qquad \times (\Psi_{1/n} p)^{k_l-k_{l-1}-1} \\
&\qquad\qquad \times \big(\Psi_{1/n} q^{[\theta_i]} - \Psi_{1/n} p\big) \\
&\qquad\qquad \times (\Psi_{1/n} p)^{n-k_l-1} \\
&\qquad + \sum_{i=1}^m f_{\theta_i} \sum_{k_m=m-1}^{n-1} \sum_{k_{m-1}=m-2}^{k_m-1} \cdots \sum_{k_2=1}^{k_3-1} \sum_{k_1=0}^{k_2-1} \big(\Psi_{1/n} q^{[\theta_i]}\big)^{k_1} \big(\Psi_{1/n} q^{[\theta_i]} - \Psi_{1/n} p\big) \\
&\qquad\qquad \times (\Psi_{1/n} p)^{k_2-k_1-1} \times \cdots \\
&\qquad\qquad \times \big(\Psi_{1/n} q^{[\theta_i]} - \Psi_{1/n} p\big) \\
&\qquad\qquad \times (\Psi_{1/n} p)^{k_m-k_{m-1}-1} \\
&\qquad\qquad \times \big(\Psi_{1/n} q^{[\theta_i]} - \Psi_{1/n} p\big) \\
&\qquad\qquad \times (\Psi_{1/n} p)^{n-k_m-1}.
\end{aligned}
$$

Set

$$
a_{1,(k_1)} = \sum_{i=1}^m f_{\theta_i} p^{k_1} \big(q^{[\theta_i]} - p\big) p^{n-k_1-1}.
$$

For $l \in \{2, \ldots, m-1\}$ set

$$
\begin{aligned}
a_{l,(k_l,\ldots,k_1)} = \sum_{i=1}^m f_{\theta_i} p^{k_1} \big(q^{[\theta_i]} - p\big) p^{k_2-k_1-1} \big(q^{[\theta_i]} - p\big) \times \cdots \\
\times p^{k_l-k_{l-1}-1} \big(q^{[\theta_i]} - p\big) p^{n-k_l-1},
\end{aligned}
$$

and for $l = m$ define

$$a_{m,(k_m,\dots,k_1)} = \sum_{i=1}^{m} f_{\theta_i} \big(q^{[\theta_i]}\big)^{k_1} \big(q^{[\theta_i]} - p\big) p^{k_2-k_1-1} \big(q^{[\theta_i]} - p\big) \times \cdots$$
$$\times\, p^{k_m-k_{m-1}-1} \big(q^{[\theta_i]} - p\big) p^{n-k_m-1}.$$

In particular the summand $a_{1,(k_1)}$ can be written as

$$a_{1,(k_1)} = p^{k_1} \sum_{i=1}^{m} f_{\theta_i} \big(q^{[\theta_i]} - p\big) p^{n-k_1-1}. \tag{37}$$

Let $c_i \in \mathbf{R}\langle\langle A\rangle\rangle_{\geq 2i+1}$, $i \in \mathbf{N}$, be as in Proposition 1. By Theorem 2,

$$j_{\leq 2m}\left(\sum_{i=1}^{m} f_{\theta_i} \big(q^{[\theta_i]} - p\big)\right) = 0;$$

thus $j_{\leq 2m}(a_{1,(k_1)}) = 0$. Also it holds that

$$j_{2m+1}\left(\sum_{i=1}^{m} f_{\theta_i} \big(q^{[\theta_i]} - p\big)\right) = j_{2m+1}(c_m)[1/\theta_1^{2m}, \dots, 1/\theta_m^{2m}] f.$$

By Corollary 2, for all $\theta \in \mathbf{N}$, $j_{\leq 2}(q^{[\theta]} - p) = 0$ holds. Thus, by Corollaries 4, 5 and Proposition 1,

$$\begin{aligned} &j_{\leq 3m-1}\Big(\big(q^{[\theta_i]}\big)^{k_1} \big(q^{[\theta_i]} - p\big) p^{k_2-k_1-1} \big(q^{[\theta_i]} - p\big) \times \cdots \\ &\qquad\qquad \times\, p^{k_m-k_{m-1}-1} \big(q^{[\theta_i]} - p\big) p^{n-k_m-1}\Big) = 0 \end{aligned} \tag{38}$$

holds, hence $j_{\leq 3m-1}(a_{m,(k_m,\dots,k_1)}) = 0$. Moreover,

$$\begin{aligned} &j_{3m}\Big(\big(q^{[\theta_i]}\big)^{k_1} \big(q^{[\theta_i]} - p\big) p^{k_2-k_1-1} \big(q^{[\theta_i]} - p\big) \cdots p^{k_m-k_{m-1}-1} \big(q^{[\theta_i]} - p\big) p^{n-k_m-1}\Big) \\ &\quad = (j_3(c_1))^m \frac{1}{\theta_i^{2m}}. \end{aligned}$$

Let $p_1, \dots, p_{l+1} \in \mathbf{R}\langle\langle A\rangle\rangle$ with property $j_0(p_i) = 1$ for $i \in \{1, \dots, l+1\}$. By using similar arguments as in the proof of Theorem 2, we get

$$\begin{aligned} &j_{\leq 2m+l-1}\left(\sum_{i=1}^{m} f_{\theta_i} p_1 \big(q^{[\theta_i]} - p\big) p_2 \big(q^{[\theta_i]} - p\big) \cdots p_l \big(q^{[\theta_i]} - p\big) p_{l+1}\right) \\ &\quad = j_{\leq 2m+l-1}\left(\sum_{k=l}^{m-1} \sum_{\substack{i_1,\dots,i_l \geq 1 \\ i_1+\cdots+i_l = k}} f^T \begin{bmatrix} \frac{1}{\theta_1^{2k}} \\ \vdots \\ \frac{1}{\theta_m^{2k}} \end{bmatrix} p_1 c_{i_1} p_2 \cdots p_l c_{i_l} p_{l+1}\right) \\ &\quad = 0 \end{aligned}$$

and

$$j_{2m+l}\left(\sum_{i=1}^{m} f_{\theta_i} p_1(q^{[\theta_i]}-p)p_2(q^{[\theta_i]}-p)\cdots p_l(q^{[\theta_i]}-p)p_{(l+1)}\right)$$
$$=\sum_{\substack{i_1,\dots,i_l\geq 1\\ i_1+\dots+i_l=m}} f^T \begin{bmatrix} \frac{1}{\theta_1^{2m}} \\ \vdots \\ \frac{1}{\theta_m^{2m}} \end{bmatrix} j_{2m+l}(c_{i_1}\cdots c_{i_l})$$

for all $l \in \{2,\dots,m-1\}$. We conclude, that $j_{\leq 2m+l-1}(a_{l,(k_l,\dots,k_1)}) = 0$. $\square$

5. From algebra to analysis. Now we are able—by means of our homomorphisms Ψ and Φ—to transfer the algebraic results into the realm of weak approximation methods.

Recall the conditions from Assumption 1 if we want to switch from the bounded, C^∞-bounded case to Sobolev spaces W_p^m for some $m, p \in \mathbf{N}$:

Fix an integer $l \geq 1$, and denote $V_j = \sum_{k=1}^{N} c_{j,k}\frac{\partial}{\partial x^k}$. Let $f \in W_p^m$ and let the coefficients $c_{j,k}$ of the vector fields be bounded Borel functions satisfying the conditions that the partial derivatives

$$D^\rho c_{j,k}, \qquad j=0,\dots,d, k=1,\dots,N,$$

exist and by magnitude are bounded for all multi-indeces ρ, satisfying $|\rho| \leq l$.

First, let us recall some immediate consequences of the results from [6–8]. Note that for any $\theta \in \mathbf{N}$ and $f \in C_b^\infty(\mathbf{R}^N)$ the expressions $\overrightarrow{Q}_t^{[\theta]} f$ and $\overleftarrow{Q}_t^{[\theta]} f$ define a splitting-up approximation of the $P_t f$ of order 1. The following theorem is an immediate consequence of [7], Theorem 2.2, and its proof (see also [8], Theorem 4.3), even though our formulation here is slightly more general than Gyöngy and Krylov's formulation in [7].

THEOREM 4. *Let f, N, T and d be as before. Let $m, k \in \mathbf{N}$ and set $h = T/n$. Let Assumption 1 hold with $l \geq 4+m+4k$. Then for all $n \geq 1$ and $x \in \mathbf{R}^N$ the following representations hold true*:

$$\begin{aligned}\overrightarrow{Q}_T^{[n\theta]} f(x) &= \left(\overrightarrow{Q}_{T/n}^{[\theta]}\right)^n f(x)\\ &= P_T f(x) + \left(\frac{h}{\theta}\right) u_1^{(1)}(T,x) + \left(\frac{h}{\theta}\right)^2 u_1^{(2)}(T,x) + \cdots \\ &\quad + \left(\frac{h}{\theta}\right)^k u_1^{(k)}(T,x) + R_{1,n}^{(k)}(T,x),\end{aligned}$$

$$\begin{aligned}\overleftarrow{Q}_T^{[n\theta]} f(x) &= \big(\overleftarrow{Q}_{T/n}^{[\theta]}\big)^n f(x)\\ &= P_T f(x) + \left(\frac{h}{\theta}\right) u_2^{(1)}(T,x) + \left(\frac{h}{\theta}\right)^2 u_2^{(2)}(T,x) + \cdots\\ &\quad + \left(\frac{h}{\theta}\right)^k u_2^{(k)}(T,x) + R_{2,n}^{(k)}(T,x)\end{aligned}$$

and

$$\begin{aligned}\big(Q_{T/n}^{[\theta]}\big)^n f(x) &= P_T f(x) + h u_3^{(1)}(T,x,\theta) + h^2 u_3^{(2)}(T,x,\theta) + \cdots\\ &\quad + h^k u_3^{(k)}(T,x,\theta) + R_{3,n}^{(k)}(T,x,\theta),\end{aligned}$$

where the functions $u_i^{(1)}, \dots, u_i^{(k)}$ and $R_{i,n}^{(k)}$, $i = 1, 2, 3$, are W_p^m-valued and weakly continuous in T. Furthermore, $u_i^{(j)}$, $i = 1, 2, 3$, $j = 1, \dots, k$, are independent of n, and

$$\|R_{i,n}^{(k)}(T)\|_{m,p} \le C h^{k+1}$$

for all n, where C depends only on k, N, d, f, T, m and p.

REMARK 3. By the important results of [7], Theorem 2.2, and [8], Theorem 4.3, the terms $u_i^{(j)}$ in the error expansion from Theorem 4 are calculated by applying certain polynomials of operators to the specified functions. The form of the polynomials is universal, that is, depends only on the number of summands into which the generator $\mathcal{A}$ of the semigroup P_t is split.

DEFINITION 7. Let $m \in \mathbf{N}$ and let $g \in C_b^\infty(\mathbf{R}^N)$. Define

$$\|g\|_m := \sup_{i \le m} \|\nabla^i g\|_\infty.$$

REMARK 4. The function $\|\cdot\|_m$ is a norm on $C_b^\infty(\mathbf{R}^N)$.

REMARK 5. The assertions of Theorem 4 also hold with respect to the uniform norms $\|\cdot\|_m$.

We are now able to formulate results on weak approximation methods of order $2m$ by means of our algebraic preparations. Recall therefore the definitions

$$\begin{aligned}\overrightarrow{Q}_t^{[\theta]} &:= \big(P_{t/\theta}^{(0)} \circ \cdots \circ P_{t/\theta}^{(d)}\big)^\theta,\\ \overleftarrow{Q}_t^{[\theta]} &:= \big(P_{t/\theta}^{(d)} \circ \cdots \circ P_{t/\theta}^{(0)}\big)^\theta,\\ Q_t^{[\theta]} &:= \tfrac{1}{2}\big(\overrightarrow{Q}_t^{[\theta]} + \overleftarrow{Q}_t^{[\theta]}\big)\end{aligned}$$

of the building blocks of the Ninomiya–Victoir schemes.

THEOREM 5. *Let* $\{q_n^s\}_{n\in\mathbf{N}}$ *be a abstract extrapolated symmetrized splitting scheme of order* $2m$; *then*

$$Q_{T,n}^s := \sum_{i=1}^m f_{\theta_i}(Q_T^{[n\theta_i]})$$

for $n \geq 0$ *is a weak approximation method of order* $2m$, *where a choice of* k *is given by*

$$k = 2(2m+1)(d+1)\max\{\theta_i | i = 1, \dots, m\},$$

that means

$$|P_T g(x) - Q_{T,n} g(x)| \leq \frac{C}{n^{2m}} \|g\|_k$$

for test functions $g \in C_b^\infty(\mathbb{R}^N)$.

PROOF. According to the Theorem 4, the coefficients $u_i^{(j)}$ in the error expansion are constructed by applying certain polynomials of operators to specified functions. The form of the polynomials is universal; hence we can replace in a first step the variables a_l by bounded linear operators: plugging the bounded linear operators instead of the variables a_l into the algebraic expressions shows that the algebraic argument is in fact an analytic one, since the respective series will converge absolutely and reproduce the expression of Theorem 4. However, due to Proposition 1 we know that for the abstract symmetrized extrapolation scheme, all odd power terms in the error expansion vanish. Hence we can conclude that all terms up to order $2m-1$ vanish in error expansion $P_T - Q_{T,n}$, since—on appropriate smooth functions—the bounded linear approximations converge to the respective terms in Theorem 4.

Apparently each term in $Q_t^{[\theta]}$ increases the number of derivatives necessary to do the estimation by $2(2m+1)$, which leads to the formula for k. □

THEOREM 6. *Let* $\{q_n\}_{n\in\mathbf{N}}$ *be an abstract generalized Fujiwara scheme of order* $2m$; *then*

$$Q_{T,n} := \sum_{i=1}^m f_{\theta_i}(Q_{T/n}^{[\theta_i]})^n$$

for $n \geq 0$ *is a weak approximation method of order* $2m$, *where a choice of* k *is given by*

$$k = 2(2m+1)(d+1)\max\{\theta_i | i = 1, \dots, m\};$$

that means

$$|P_T g(x) - Q_{T,n} g(x)| \leq \frac{C}{n^{2m}} \|g\|_k$$

for test functions $g \in C_b^\infty(\mathbb{R}^N)$.

PROOF. Theorem 4 provides us with the error expansion. As in the proof of Theorem 5 it suffices to consider only bounded linear operators on some Banach space. Plugging these bounded linear operators in the algebraic expressions instead of the variables a_l shows that the algebraic argument is in fact the analytic one, since the respective series will converge absolutely. These polynomials of operators have been calculated in Theorem 3 for the generalized Fujiwara scheme. Thus, in the case of bounded linear operators, the first $2m-1$ terms vanish, and we obtain, therefore, an order of approximation $2m$. Using the universality of the polynomials which define the error terms, the first $2m-1$ terms vanish for bounded linear approximations of the vector fields V_i, and therefore for the limit which exists due to Theorem 4 on appropriate smooth functions. Apparently each term in $Q_t^{[\theta]}$ increases the necessary number of derivatives to do the estimation by $2(2m+1)$, which leads to the formula for k. $\square$

REMARK 6. The approximation method $Q^s_{T,n}$ of order $2m$ from the Theorem 5 will be called *the extrapolated symmetrized splitting method of order* $2m$. The approximation method $Q_{T,n}$ of order $2m$ from the Theorem 6 will be called *the generalized Fujiwara method of order* $2m$.

EXAMPLE 1. The generalized Fujiwara method in the case $m=1$ apparently corresponds to a version of the original Ninomiya–Victoir scheme.

EXAMPLE 2. The generalized Fujiwara case $m=2$ corresponds to a weak approximation method already presented in [5]. One can choose $\theta_1=1$ and $\theta_2=2$ and $f_{\theta_1}=-\frac{1}{3}$ and $f_{\theta_2}=\frac{4}{3}$.

EXAMPLE 3. The generalized Fujiwara in the case $m=3$ corresponds to Fujiwara's originally presented weak approximation method, which in our language reads as follows. Notice that we do not need the full strength of our previous proof, which is built on Theorem 2.

For all mutually different numbers $\theta_1,\theta_2,\theta_3\in\mathbf{N}$, we can construct an abstract Fujiwara scheme q of order 6 with the form

$$q=f_{\theta_1}(q^{[\theta_1]})^n+f_{\theta_2}(q^{[\theta_2]})^n+f_{\theta_3}(q^{[\theta_3]})^n.$$

For the proof, which is presented for convenience here, we assume without loss of generality that $\theta_1<\theta_2<\theta_3$. We have

$$f=\begin{bmatrix} f_{\theta_1}\\ f_{\theta_2}\\ f_{\theta_3}\end{bmatrix}=\begin{bmatrix} \dfrac{\theta_1^4}{(\theta_2^2-\theta_1^2)(\theta_3^2-\theta_1^2)}\\ \dfrac{-\theta_2^4}{(\theta_3^2-\theta_2^2)(\theta_2^2-\theta_1^2)}\\ \dfrac{\theta_3^4}{(\theta_3^2-\theta_1^2)(\theta_3^2-\theta_2^2)}\end{bmatrix}.$$

By Corollary 2, we have

$$j_{\leq 4}(q^{[\theta_2]} - p) = j_{\leq 4}\left(\frac{\theta_1^2}{\theta_2^2}(q^{[\theta_1]} - p)\right),$$

$$j_{\leq 4}(q^{[\theta_3]} - p) = j_{\leq 4}\left(\frac{\theta_1^2}{\theta_3^2}(q^{[\theta_1]} - p)\right).$$

Then, by Corollary 4, we have

$$j_{\leq 7}\big((q^{[\theta_2]} - p)^2\big) = j_{\leq 7}\left(\frac{\theta_1^4}{\theta_2^4}(q^{[\theta_1]} - p)^2\right),$$

$$j_{\leq 7}\big((q^{[\theta_3]} - p)^2\big) = j_{\leq 7}\left(\frac{\theta_1^4}{\theta_3^4}(q^{[\theta_1]} - p)^2\right).$$

Thus,

$$\begin{aligned}
j_{\leq 7}&\left(\sum_{i=1}^{3} f_{\theta_i}(q^{[\theta_i]} - p)^2\right) \\
&= j_{\leq 7}\left(\left(f_{\theta_1} + f_{\theta_2}\frac{\theta_1^4}{\theta_2^4} + f_{\theta_3}\frac{\theta_1^4}{\theta_3^4}\right)(q^{[\theta_1]} - p)^2\right) \\
&= \left(\frac{\theta_1^4}{(\theta_2^2 - \theta_1^2)(\theta_3^2 - \theta_1^2)} - \frac{\theta_1^4}{(\theta_3^2 - \theta_2^2)(\theta_2^2 - \theta_1^2)} + \frac{\theta_1^4}{(\theta_3^2 - \theta_1^2)(\theta_3^2 - \theta_2^2)}\right) \\
&\quad \times j_{\leq 7}\big((q^{[\theta_1]} - p)^2\big) \\
&= 0.
\end{aligned}$$

6. Implementation of a generalized Fujiwara method of order $2m$. A weak approximation method of order six was first introduced by Fujiwara [5]. In previous sections we theoretically constructed weak approximation methods of order $2m$ for arbitrary $m \in \mathbf{N}$. Since the extrapolated symmetrized splitting method of order $2m$ is just a linear combination of various splitting schemes, its implementation and its computational costs can be easily obtained. Its error analysis will be performed elsewhere.

In this section we show how to construct a practical generalized Fujiwara approximation method with approximating flow of vector fields V_i, which drive the SDE (1), by some suitable integration schemes. The usual choice for the integration schemes are Runge–Kutta methods. In our concrete example from mathematical finance we will use a seventh-order nine-stage explicit Runge–Kutta method with a very good stability, given by Tanaka et al. (see [19–21]). Higher order

Runge–Kutta methods often lose stability with respect to rounding error, truncated error and piling error. In addition, these affect decrease order of approximating error. Since in a concrete application of the algorithm, for example, in mathematical finance, some of the ODEs can be very close to being stiff, the stability of the Runge–Kutta algorithm is of high importance. We show a relation between the convergence order of a weak approximation method and an mth order Runge–Kutta method. In addition we construct a concrete algorithm of a generalized Fujiwara method of order $2m$ and analyze its computational cost and its approximating error. At the end we present a concrete numerical experiment.

The results of this section can be compared to those from [16].

6.1. *Runge–Kutta method.* For $V \in C_b^\infty(\mathbf{R}^N, \mathbf{R}^N)$, the map

$$\exp : C_b^\infty(\mathbf{R}^N, \mathbf{R}^N) \times \mathbf{R}_+ \times \mathbf{R}^N \to \mathbf{R}^N$$

represents the flow driven by the vector field V starting at x_0, that is, the solution of the ordinary differential equation

$$\begin{aligned} \frac{d}{dt}x(t) &= V(x(t)), \\ x(0) &= x_0. \end{aligned} \tag{39}$$

In case there are no ambiguities, we shall omit the x_0 from the notation.

DEFINITION 8 (s stage explicit Runge–Kutta method of order m for autonomous systems). A s stage explicit Runge–Kutta method of order m for autonomous systems is determined by a lower triangular matrix $A = [a_{ij}]_{i,j=1}^s$ and a row $b = [b_1 \cdots b_s]$ such that the following hold:

- Let $h \in \mathbf{R}$, $t_0 \in \mathbf{R}$, and let $t_n = t_{n-1} + h$ for all $n \in \mathbf{N}$. Given the vector x_{n-1} as an approximation to $x(t_{n-1})$, where x satisfies the equation (39), the approximation x_n to $x(t_n)$ is computed by evaluating, for $i = 1, 2, \dots, s$,

  $$F_i = V(X_i),$$

  where $X_1, X_2, \dots, X_s$ are given by

  $$X_i = x_{n-1} + h \sum_{j<i} a_{ij} F_j$$

  and then evaluating

  $$y_n = y_{n-1} + h \sum_{j=1}^{s} b_j F_j.$$

- The Taylor expansion of x_n as a function of h around 0 should coincide with the Taylor expansion of $x(t_n) = x(t_{n-1} + h)$ up to (including) the term at the power h^{m+1}.

  From this point onwards, let $R_m(t, V)(x_0)$ denote the approximation for the solution of the system (39) at time t, defined by a Runge–Kutta method of order m.

REMARK 7. Usually Runge–Kutta methods are studied for general nonautonomous systems. In these cases the method is uniquely identified by a triplet A, b and c, where A and b are as above, and $c = [c_1 \cdots c_s]^T$ is a suitable column vector.

See Butcher [2] and [3] for more details about the theory of the Runge–Kutta method.

The next theorem shows that we need at least a 12th order Runge–Kutta method for the generalized Fujiwara method of order 6.

THEOREM 7. *For all $f \in C_b^\infty(\mathbf{R}^N \to \mathbf{R}^N)$, $t \in \mathbf{R}_+$ and $x \in \mathbf{R}^N$, there exists $C_i > 0$ such that*

$$|f(\exp(tV_0)) - f(R_m(t, V_0)(x))| \le C_0 t^{m+1},$$
$$\left|E\left[f\left(\exp(\sqrt{t}ZV_i)\right) - f\left(R_{2m}\left(\sqrt{t}Z, V_i\right)(x)\right)\right]\right| \le C_i t^{m+1},$$

where $i \in \{1, \dots, d\}$ and $Z \sim \mathcal{N}(0, 1)$.

PROOF. The first inequality follows from the definition of an mth order Runge–Kutta method and Taylor's theorem. Set $i \in \{1, \dots, d\}$. By the definition of the Runge–Kutta method and Taylor's theorem again, we have

$$\begin{aligned} &f\left(\exp(\sqrt{t}ZV_i)\right) - f\left(R_{2m}\left(\sqrt{t}ZV_i\right)(x)\right) \\ &\quad = \frac{t^{m+1/2}Z^{2m+1}}{2(m+1)!} V_i^{2m+1} f(x) + O(t^{m+1}). \end{aligned}$$

Note that for all $k \in \mathbf{N}$, $E[Z^{2k+1}] = 0$ holds. Thus the conclusion is true. □

The next theorem shows that if we do not urge to have the computational cost of the algorithm to be linear with respect to the number of subdivision points n, a 4th order Runge–Kutta method is enough for a generalized Fujiwara method of order six.

THEOREM 8. *For $k, n \in \mathbf{N}$, for all $f \in C_b^\infty(\mathbf{R}^N)$, for all $i \in \{1, \dots, d\}$ and for all $x \in \mathbf{R}^N$, there exists $C_i > 0$ such that*

$$\left|E\left[f\left(\exp\left(\frac{Z}{\sqrt{n}}V_i\right)\right) - f\left(R_m\left(\frac{Z}{n^k\sqrt{n}}, V_i\right)^{n^k}(x)\right)\right]\right| \le \frac{C_i}{n^{km+k+m/2+1}}$$

holds where $Z \sim \mathcal{N}(0, 1)$.

PROOF.

$$
\begin{aligned}
&\left|E\left[f\left(\exp\left(\frac{Z}{\sqrt{n}}V_i\right)(x)\right)-f\left(R_m\left(\frac{Z}{n^k\sqrt{n}}1,V_i\right)^{n^k}(x)\right)\right]\right| \\
&\quad=\left|E\left[\left(\exp\frac{Z}{n^k\sqrt{n}}V_i\right)^{n^k}f(x)-R_m\left(\frac{Z}{n^k\sqrt{n}},V_i\right)^{n^k}f(x)\right]\right| \\
&\quad=\left|E\left[\sum_{l=0}^{n^k-1}\left(\exp\frac{Z}{n^k\sqrt{n}}V_i\right)^l\left(\left(\exp\frac{Z}{n^k\sqrt{n}}V_i\right)-R_m\left(\frac{Z}{n^k\sqrt{n}},V_i\right)\right)\right.\right. \\
&\qquad\qquad\qquad\qquad\qquad\left.\left.\times R_m\left(\frac{Z}{n^l\sqrt{n}},V_i\right)^{n^k-l-1}f(x)\right]\right| \\
&\quad\le\frac{C_i}{n^{2(k+1/2)(m/2+1)}}n^k \\
&\quad\le\frac{C_i}{n^{(km+k+m/2+1)}}. \qquad \square
\end{aligned}
$$

6.2. *Recipe for a generalized Fujiwara method of order* $2m$. In the following subsection we will provide the pseudocode for implementation of the generalized Fujiwara method of order $2m$ with fixed coefficients $\theta_1,\theta_2,\dots,\theta_m$. Let $f=[f_1\cdots f_m]^T$ be as in the Section 4, and let the function *solveDE*(V,x_0,t) return the solution of the ODE (39) at time t with initial condition $x(0)=x_0$.

REMARK 8. Usually in modern computers memory size is no longer an issue. From this perspective it seems sensible to generate all needed random variables in advance. Namely, the random variables for various θ_i's do not have to be independent; therefore we can reduce its number by reusing them, and there exist efficient algorithms which speed up the process of their generation if we do it in one batch instead of step by step as it is written in Algorithm 2.

6.3. *Computational cost.*

THEOREM 9. *Let $d,n,M,m,T,\theta_1,\dots,\theta_m$ be as above, such that T/n is sufficiently small. Furthermore, assume that each step of the method solveDE needs a operations, that is, additions, multiplications and function evaluations, that B operations are needed to generate a (pseudo or quasi) Bernoulli random variable and that Z operations are needed to generate a standard d-dimensional normally distributed (pseudo or quasi) random variable. Then the computational cost of Algorithm 1 is* $M(5m+n((d+1)a+Z+1)\sum_{k=1}^m\theta_k+nB+1)+2m$.

PROOF. Let us denote the computational cost of Algorithm 2 by C. A straightforward calculation shows that the computational cost of Algorithm 1 is equal to $M(C+1)+2m$.

Algorithm 1: Fujiwara

Data: function g, vector fields $V_0, V_1, \dots, V_d$, time T, initial condition x_0, number of partition points n, number of samples M

Result: approximation E of the expectation $\mathrm{E}[f(X_T)]$, where X_t is a process defined by the SDE (1)

```
Q ← 0 ∈ R^{1×m};
for o ← 1 to M do /* expectation (MonteCarlo or quasi Monte Carlo)      */
|  Q ← Q + samplePath(g, V_0, ..., V_d, T, x_0, n);
end
Q ← (1/M) Q;
/* approx. for E(g(X(T, x_0))) is the linear combination ∑_i f_i ∗ Q_i    */
E ← Q f;
return E
```

For fixed $j \in \{1, \dots, n\}$ in Algorithm 2 we have $\theta_k(d+1)a$ operations. Hence, for fixed $k \in \{1, \dots, m\}$ there are $5 + n\theta_k((d+1)a + Z + 1)$ operations. It follows that $C = 5m + n((d+1)a + Z + 1)\sum_{k=1}^{m}\theta_k + nB$. $\square$

REMARK 9. Rigorous use of Runge–Kutta algorithms for solving ODEs in the algorithm is only suitable for building a universal solver for SDEs of the type (1). In concrete practical applications it is to be expected that many of the ODEs of the type (39) have a nice enough explicit solution.

The error of the algorithm consists of discretization part, that is, the error due to numerical solution of ODEs and the error which comes from the scheme, and of the convergence error which comes from the Monte Carlo or quasi Monte Carlo simulation.

THEOREM 10. *For $n, M \in \mathbf{N}$ such that T/n is sufficiently small, the approximation error of Algorithm 1 is $O(1/n^{2m}) + O(1/\sqrt{M})$.*

REMARK 10. One should take great care when choosing a suitable subdivision of the interval, since the coefficient of the discretisation error directly depends on function f and vector fields V_i; thus, although bounded, the coefficient can get fairly large in some cases. Moreover, the convergence error of the Monte Carlo simulation is directly proportional to the square root of variance of $f(X(T, x))$. As in the case of discretisation error this should be taken into account, since, although constant, the variance can be large comparing to the size of error we would like to achieve.

6.4. *Numerical example.* For our numerical example we have chosen the genearlized Fujiwara method of order 6 with $\theta_1 = 1$, $\theta_2 = 2$ and $\theta_3 = 3$, that is, the method that first appeared in [5], and the generalized Fujiwara method of order 8 with the choice of parameters $\theta_1 = 1$, $\theta_2 = 2$, $\theta_3 = 3$ and $\theta_4 = 4$.

Algorithm 2: samplePath

Data: function g, vector fields $V_0, V_1, \dots, V_d$, time T, initial condition x_0, number of partition points n

Result: row vector $Q = [Q_s^{[\theta_1]} \cdots Q_s^{[\theta_m]}] \in \mathbf{R}^{1\times m}$ calculated for a random simulated path

$Q \leftarrow 0 \in \mathbf{R}^{1\times m}$;

generate independent Bernoulli(1/2) random variables $\Lambda = [\Lambda_1, \dots, \Lambda_n]$;

for $k \leftarrow 1$ **to** m **do**

  `/* generate all standard normal random variables that are needed */`

  $D \leftarrow T/(n\theta_k)[1 \cdots 1] \in \mathbf{R}^{1\times(\theta_k n)}$;

  $N \leftarrow \sqrt{T/(n\theta_k)}[N_1, \dots, N_{\theta_k n}]$, where $N_1, \dots, N_{\theta_k n}$ are i.i.d., with $N_1 \sim N(0, I_d)$;

  `/* first row serves for the component without Brownian motion */`

  $Z \leftarrow \begin{bmatrix} D \\ N \end{bmatrix}$;

  $X \leftarrow x_0$;

  for $j \leftarrow 1$ **to** n **do** `/* consequtively solve the ODEs */`

    if $\Lambda_j = 1$ **then** `/* solve appropriate ODE */`

      for $\Theta \leftarrow 1$ **to** θ_k **do** `/* repetition because of finer dissection */`

        for $i \leftarrow 0$ **to** d **do** `/* solving ODEs */`

          $X \leftarrow \mathit{solveDE}(V_i, X, Z_{i+1,(\Theta-1)*n+j})$;

        end

      end

    else

      for $\Theta \leftarrow 1$ **to** θ_k **do** `/* repetition because of finer dissection */`

        for $i \leftarrow 0$ **to** d **do** `/* solving ODEs */`

          $X \leftarrow \mathit{solveDE}(V_{d-i}, X, Z_{d-i+1,(\Theta-1)*n+j})$;

        end

      end

    end

  end

  $Q_s^{[\theta_k]} \leftarrow g(X)$

end

return Q

In order to compare the algorithm to the basic Ninomiya–Victoir scheme we consider an Asian call option written on an asset whose price process follows the Heston stochastic volatility model. Let X_1 be the price process of an asset following the Heston model:

$$
\begin{aligned}
X_1(t,x) &= x_1 + \int_0^t \mu X_1(s,x)\,ds + \int_0^t X_1(s,x)\sqrt{X_2(s,t)}\,dB^1(s), \\
X_2(t,x) &= x_2 + \int_0^t \alpha\big(\theta - X_2(s,x)\big)\,ds \\
&\quad + \int_0^t \beta\sqrt{X_2(s,t)}\big(\rho\,dB^1(s) + \sqrt{1-\rho^2}\,dB^2(s)\big),
\end{aligned} \tag{40}
$$

where $x = (x_1, x_2) \in (\mathbf{R}_{>0})^2$, $(B^1(t), B^2(t))$ is a two-dimensional standard Brownian motion, $-1 \le \rho \le 1$ and α, θ, μ are some positive coefficients satisfying $2\alpha\theta - \beta^2 > 0$. We are aware that our conditions on the vector fields' regularity are not satisfied at $x_2 = 0$; however, we want to provide the same example as in previous papers in the subject. On the other hand the condition $2\alpha\theta - \beta^2 > 0$ ensures that the volatility process can be considered as a perturbation of a square of a Brownian motions, which leads us back into our setting.

The payoff of the Asian call option on this asset with maturity T and strike K is $\max(X_3(T, x)/T - K, 0)$, where

$$X_3(t, s) = \int_0^t X_1(s, x)\, ds. \tag{41}$$

Hence, the price of this option becomes $D \times E[\max(X_3(T, x)/T - K, 0)]$ where D is an appropriate discount factor on which we do not focus in this experiment. As in [17] take $T = 1$, $K = 1.05$, $\mu = 0.05$, $\alpha = 2.0$, $\beta = 0.1$, $\theta = 0.09$, $\rho = 0$ and $x = (1.0, 0.09)$.

Up to the error of the magnitude 10^{-6} we have

$$E\big[\max\big(X_3(T, x)/T - K, 0\big)\big] = 6.0473534496 * 10^{-2}$$

obtained from [16]. Let $X(t, x) = (X_1(t, x), X_2(t, x), X_3(t, x))^T$. SDEs (40) and (41) can be transformed in the Stratonovich form since $X_2 \neq 0$.

$$X(t, x) = \sum_{i=0}^{2} \int_0^t V_i(X(s, x)) \circ dB^i(s),$$

where

$$\begin{aligned} V_0(y_1, y_2, y_3) &= \left(y_1\left(\mu - \frac{y_2}{2} - \frac{\rho\beta}{4}\right), \alpha(\theta - y_2) - \frac{\beta^2}{4}, y_1\right)^T, \\ V_1(y_1, y_2, y_3) &= \big(y_1\sqrt{y_2}, \rho\beta\sqrt{y_2}, 0\big)^T, \\ V_2(y_1, y_2, y_3) &= \big(0, \beta\sqrt{(1 - \rho^2)y_2}, 0\big)^T. \end{aligned} \tag{42}$$

Taking our choice of $\rho = 0$ into consideration we get exact solutions of ODEs of the type (39) driven by vector fields V_1 and V_2 (see [17] for more details):

$$\begin{aligned} \exp(tV_1)(x_1, x_2, x_3)^T &= \big(x_1 e^{t\sqrt{x_2}}, x_2, x_3\big), \\ \exp(tV_2)(x_1, x_2, x_3)^T &= \left(x_1, \left(\frac{\beta t}{2} + \sqrt{x_2}\right)^2, x_3\right). \end{aligned} \tag{43}$$

According to the proof of Theorem 7 we need a Runge–Kutta method of order of at least 6 to approximate the solution $\exp(tV_0)(x_1, x_2, x_3)^T$ for generalized Fujiwara method of order 6 and a Runge–Kutta method of order of at least 8 for a generalized Fujiwara method of order 8 if we want a linear algorithm. If we allow

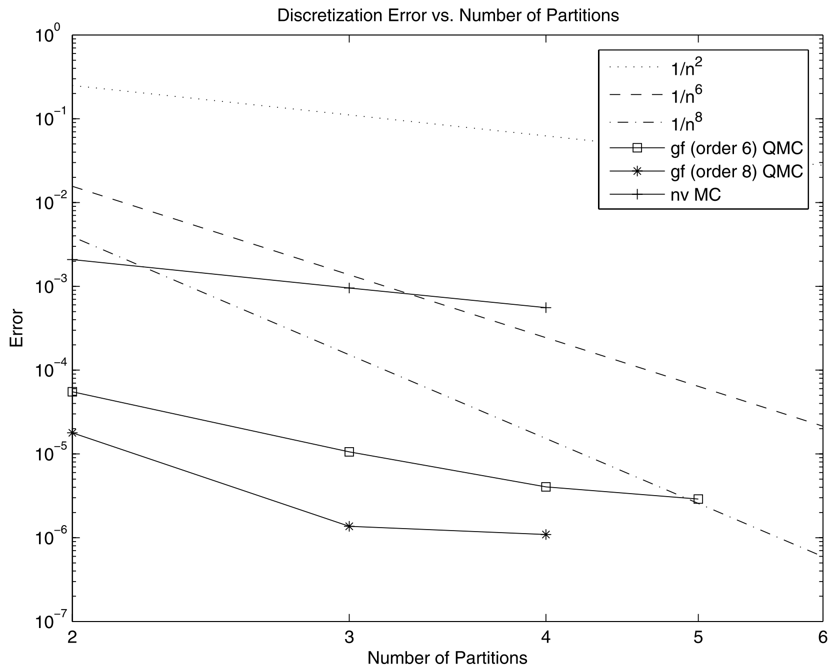

FIG. 1. *Error coming from discretization.*

quadratic computational cost for the the generalized Fujiwara method of order 8, it is sufficient to use a Runge–Kutta method of order 4. In our example we used a 9-stage 7th order Runge–Kutta method from [19].

The pseudorandom numbers in MC were generated by the Mersenne twister algorithm. The QMC was performed using a Sobol sequence, generated by the library SobolSeq51.dll provided by Broda (see [1]). Both MC and QMC integration were performed using 10^8 sample paths.

The use of exact solutions of ODEs driven by vector fields V_1 and V_2 reduces the computational cost of the algorithm by $2Mna\sum_{k=1}^{o}\theta_k$, where o designates the order of the generalized Fujiwara method divided by 2, M denotes the number of MC/QMC sample paths, n is the number of subdivision points and a is the number of operations required for solving ODEs driven by V_1 or V_2, if we compare it to the results of Theorem 9.

Method/*n*	**2**	**3**	**4**	**5**
NV	$0.20854 \cdot 10^{-2}$	$0.9554 \cdot 10^{-3}$	$0.5569 \cdot 10^{-3}$	
GF (order 6) MC	$0.615 \cdot 10^{-4}$	$0.365 \cdot 10^{-4}$	$0.352 \cdot 10^{-4}$	
GF (order 6) QMC	$0.552 \cdot 10^{-4}$	$0.108 \cdot 10^{-4}$	$0.403 \cdot 10^{-5}$	$0.290 \cdot 10^{-5}$
GF (order 8) MC	$0.454 \cdot 10^{-4}$	$0.369 \cdot 10^{-4}$	$0.551 \cdot 10^{-4}$	
GF (order 8) QMC	$0.178 \cdot 10^{-4}$	$0.137 \cdot 10^{-5}$	$0.109 \cdot 10^{-5}$	

The graph in Figure 1 clearly shows that the new extrapolation method reduces the order of the discretization error in comparison to the original Ninomiya–Victoir

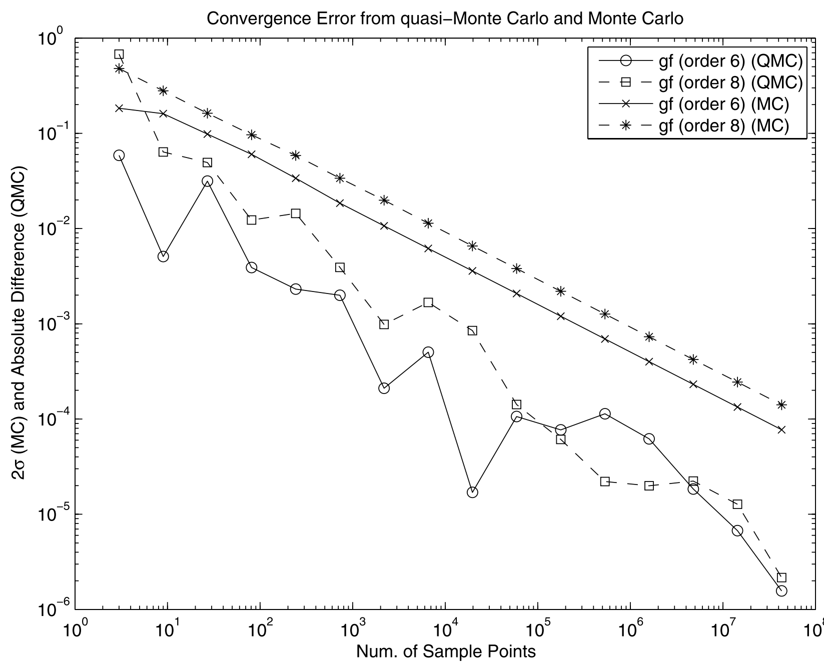

FIG. 2. *Error coming from integration.*

algorithm for several magnitudes (see Figures 1 and 2). In the QMC case the discretization error is soon (for small n) overshadowed by the integration error caused by QMC integration (see Figure 2), the order of the extrapolated algorithms can still be observed from the slope of curves in the graph in Figure 1.

Acknowledgments. A part of the research was done while the first and third authors were guests at the Research Unit of Financial and Actuarial Mathematics at the Vienna University of Technology.

The authors wish to thank Philipp Dörsek for valuable discussions and suggestions regarding the article.

REFERENCES

[1] BRITISH–RUSSIAN OFFSHORE DEVELOPMENT AGENCY (BRODA). Available at http://www.broda.co.uk/.

[2] BUTCHER, J. C. (1987). *The Numerical Analysis of Ordinary Differential Equations*: *Runge–Kutta and General Linear Methods*. Wiley, Chichester. MR0878564

[3] BUTCHER, J. C. (2003). *Numerical Methods for Ordinary Differential Equations*. Wiley, Chichester. MR1993957

[4] COHN, P. M. (1995). *Skew Fields*: *Theory of General Division Rings*. *Encyclopedia of Mathematics and Its Applications* **57**. Cambridge Univ. Press, Cambridge. MR1349108

[5] FUJIWARA, T. (2006). Sixth order methods of Kusuoka approximation. *Preprint Series*. Graduate School of Mathematical Sciences, Univ. Tokyo.

[6] GYÖNGY, I. and KRYLOV, N. (2003). On the splitting-up method and stochastic partial differential equations. *Ann. Probab.* **31** 564–591. MR1964941

[7] GYÖNGY, I. and KRYLOV, N. (2005). An accelerated splitting-up method for parabolic equations. *SIAM J. Math. Anal.* **37** 1070–1097 (electronic). MR2192288

[8] GYÖNGY, I. and KRYLOV, N. (2006). Expansion of solutions of parameterized equations and acceleration of numerical methods. *Illinois J. Math.* **50** 473–514 (electronic). MR2247837

[9] IKEDA, N. and WATANABE, S. (1989). *Stochastic Differential Equations and Diffusion Processes*, 2nd ed. *North-Holland Mathematical Library* **24**. North-Holland, Amsterdam. MR1011252

[10] KOHATSU-HIGA, A. (2001). Weak approximations. A Malliavin calculus approach. *Math. Comp.* **70** 135–172. MR1680895

[11] KUSUOKA, S. (2001). Approximation of expectation of diffusion process and mathematical finance. In *Taniguchi Conference on Mathematics Nara'98*. *Advanced Studies in Pure Mathematics* **31** 147–165. Math. Soc. Japan, Tokyo. MR1865091

[12] KUSUOKA, S. (2003). Malliavin calculus revisited. *J. Math. Sci. Univ. Tokyo* **10** 261–277. MR1987133

[13] KUSUOKA, S. (2004). Approximation of expectation of diffusion processes based on Lie algebra and Malliavin calculus. *Adv. Math. Econ.* **6** 69–83.

[14] KUSUOKA, S. (2009). Gaussian K-scheme. *Preprint Series*. Graduate School of Mathematical Sciences, Univ. Tokyo.

[15] LYONS, T. and VICTOIR, N. (2004). Cubature on Wiener space: Stochastic analysis with applications to mathematical finance. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **460** 169–198. MR2052260

[16] NINOMIYA, M. and NINOMIYA, S. (2009). A new higher-order weak approximation scheme for stochastic differential equations and the Runge–Kutta method. *Finance Stoch.* **13** 415–443. MR2519839

[17] NINOMIYA, S. and VICTOIR, N. (2008). Weak approximation of stochastic differential equations and application to derivative pricing. *Appl. Math. Finance* **15** 107–121. MR2409419

[18] TALAY, D. and TUBARO, L. (1990). Expansion of the global error for numerical schemes solving stochastic differential equations. *Stoch. Anal. Appl.* **8** 483–509. MR1091544

[19] TANAKA, M., KASAHARA, E., MURAMATSU, S. and YAMASHITA, S. (1992). On a solution of the order conditions for the nine-stage seventh-order explicit Runge–Kutta method (in Japanese). *Information Processing Society of Japan* **33** 1506–1511.

[20] TANAKA, M., MURAMATSU, S. and YAMASHITA, S. (1992). On the optimization of some nine-stage seventh-Order Runge–Kutta method (in Japanese). *Information Processing Society of Japan* **33** 1512–1526.

[21] TANAKA, M., YAMASHITA, S., KUBO, E. and NOZAKI, Y. (1993). On seventh-order nine-stage explicit Runge–Kutta methods with extended region of stability (in Japanese). *Information Processing Society of Japan* **34** 52–61.

[22] VARADARAJAN, V. S. (1984). *Lie Groups, Lie Algebras, and Their Representations. Graduate Texts in Mathematics* **102**. Springer, New York. Reprint of the 1974 edition. MR0746308

K. OSHIMA
MIZUHO CORPORATE BANK, LTD.
1-3-3, MARUNOUCHI
CHIYODA-KU
TOKYO
JAPAN
E-MAIL: kojiro.oshima@gmail.com

J. TEICHMANN
D. VELUŠČEK
D-MATH
ETH ZÜRICH
RÄMISTRASSE 101
8092 ZÜRICH
SWITZERLAND
E-MAIL: josef.teichmann@math.ethz.ch
dejan.veluscek@math.ethz.ch