Pathwise approximation of stochastic differential equations using coupling

A. M. Davie

May 28, 2015

1 Introduction

This paper describes two methods for the pathwise (or 'strong') approximation to solutions of stochastic differential equations driven by Brownian motion. The standard approach, as described in [16], uses a stochastic Taylor expansion at each time step to approximate the solution to the required order. The difficulty with this method is that, if the driving Brownian motion has dimension > 1, then to get order greater than $\frac{1}{2}$ one needs iterated stochastic integrals, which are hard to generate. An efficient method for generating double integrals, which suffice for order 1 approximation, is given in [8] for dimension two, but it has not been extended to higher dimensions. A method for approximating double integrals in any dimension using Fourier expansion is developed in [17, 23, 28] but it involves a significant computational cost. There is still a need for more efficient methods of order 1 or higher, and the methods described in this paper are an attempt to meet that need.

Our first method, which assumes nondegeneracy of the diffusion matrix, uses the Taylor expansion but generates an approximation to the expansion as a whole rather than attempting to generate the individual terms. It treats the expansion as a perturbation of the (dominant) linear term. We show that by replacing the iterated integrals by random variables with the same moments (up to a certain order) conditional on the linear term, we get a random vector which is a good approximation in distribution to the original Taylor expansion. The we use a technique from optimal transport theory to find a coupling which gives a good approximation in mean square. In principle the method can give approximations of arbitrarily high order, but the nondegeneracy condition is rather restrictive.

The second method, which does not assume a nondegeneracy conditions, uses a similar approach based on perturbation and coupling to improve the previous results using the Fourier method for double integrals. This gives a method for order 1 approximation with a small computational cost.

For some purposes it is desirable to simulate a solution (for the same Brownian path) simultaneously using two different step sizes (typically h and $h/2$). In our framework this means finding an effectively implementable coupling between the random variables generated for the two step sizes. For the first method, assuming nondegeneracy, for order 1 we are able to do this exactly in two dimensions and approximately (in the sense that the coupling is approximate) in higher dimensions.

Because of the way our methods use coupling, our error bounds can be regarded as bounds in Vaserstein metrics (see section 5 for definitions). We discuss the extent to which such bounds are adequate substitutes for standard 'strong' bounds.

The paper is organised as follows: sections 2 and 3 give background material on the Taylor expansion method and on two-level approximation. Section 4 motivates our approach by considering a simple approximation scheme which gives order 1 under the nondegeneracy

condition. Section 5 gives background material on coupling and the associated Vaserstein metrics. The following two sections give the main results on the first method, and section 9 the results on the Fourier method. There follow two sections on some extensions and examples, and a final discussion section, considering in general terms the interpretation of Vaserstein-type approximation as strong approximation and the suitability of our methods to various types of application.

We mention some previous work on numerical approximation of SDEs using coupling. Kanagawa [15] describes a variant of Euler using non-normal random variables, and the final chapter of volume 2 of Rachev and Ruschendorff [21] gives a number of extensions of this idea using the well-known 'KMT theorem' [19]. For the usual Euler method, Alfonso, Jourdain and Kohatsu-higa [1, 2] give an order $h^{\frac{2}{3}-\epsilon}$ bound for pathwise approximation for one-dimensional SDEs, and an order $h\sqrt{-\log h}$ bound for fixed-time approximation for vector SDEs. Fournier [6] applies a Vaserstein bound for the central limit theorem due to Rio [22] to numerical approximation of SDEs driven by Lévy processes. [5] uses weak approximation bounds and the Strassen-Dudley theorem to obtain bounds in Vaserstein metric; the results indicate that this approach does not achieve orders higher than the standard $\frac{1}{2}$ for Euler.

Using a different approach to ours, Cruzeiro, Malliavin and Thalmeier [3] obtain an order-one method for non-degenerate SDEs essentially equivalent to the order one method we describe in Sections 4 and 7.

We also note that that coupling methods have been used in existence and convergence proofs, for example Gyöngy and Krylov [18] uses coupling (in the shape of Skorohod's theorem, applied to sequences of discrete approximations) to prove existence results for strong solutions.

I am grateful to Xiling Zhang, Lukasz Szpruch, Mike Giles, Yazid Alhojilan and Yousef Alnafisah for corrections and/or suggested improvements to earlier versions of this paper.

2 Pathwise approximation of SDEs using stochastic Taylor expansions.

Here we give a brief review, with particular reference to order $\frac{1}{2}$ and 1 schemes.

Consider an Itô SDE

$$dx_i(t) = a_i(t, x(t))dt + \sum_{k=1}^{d} b_{ik}(t, x(t))dW_k(t), \qquad x_i(0) = x_i^{(0)}, \qquad i = 1, \cdots, q \tag{1}$$

on an interval $[0, T]$, for a q-dimensional vector $x(t)$, with a d-dimensional driving Brownian path $W(t)$. If the coefficients $b_{ik}(t, x)$ satisfy a global Lipschitz condition

$$|a_i(t, x) - a_i(t, y)| \leq C|x - y|, \qquad |b_{ik}(t, x) - b_{ik}(t, y)| \leq C|x - y| \tag{2}$$

for all $x, y \in \mathbb{R}^q$, $t \in [0, T]$ and all i, k, where C is a constant, and if a_i and b_i are continuous in t for each x, then (1) has a unique solution $x(t)$ which is a process adapted to the filtration induced by the Brownian motion. This solution satisfies satisfies $\mathbb{E}|x(t)|^p < \infty$ for each $p \in [1, \infty)$ and $t \in [0, T]$.

The standard approach to the strong or pathwise approximation of the solution of (1), as described for example in [16], is to divide $[0, T]$ into a finite number N of subintervals, which we shall usually assume to be of equal length $h = T/N$, and to approximate the equation on each subinterval using a stochastic Taylor expansion. Such expansions are described in

detail in chapter 5 of [16]. The simplest such approximation, using only the linear term in the expansion, gives the Euler (also known as Euler-Maruyama) scheme

$$\tilde{x}_i^{(j+1)} = \tilde{x}_i^{(j)} + a_i(t_j, \tilde{x}^{(j)})h + \sum_{k=1}^{d} b_{ik}(t_j \tilde{x}^{(j)})\Delta W_k^{(j)} \tag{3}$$

while adding the quadratic terms gives the Milstein scheme

$$\tilde{x}_i^{(j+1)} = \tilde{x}_i^{(j)} + a_i(t_j, \tilde{x}^{(j)})h + \sum_{k=1}^{d} b_{ik}(t_j, \tilde{x}^{(j)})\Delta W_k^{(j)} + \sum_{k,l=1}^{d} \rho_{ikl}(t_j, \tilde{x}^{(j)})I_{kl}^{(j)} \tag{4}$$

where $\Delta W_k^{(j)} = W_k((j+1)h) - W_k(jh)$, $I_{kl}^{(j)} = \int_{jh}^{(j+1)h} \{W_k(t) - W_k(jh)\} dW_l(t)$ and $\rho_{ikl}(t,x) = \sum_{m=1}^{q} b_{mk}(t,x) \frac{\partial b_{il}}{\partial x_m}(t,x)$.

Assuming (2) the Euler scheme has order $\frac{1}{2}$, in the sense that

$$\mathbb{E}(\max_{j=1}^{N} |\tilde{x}^{(j)} - x(jh)|^2) = O(h) \tag{5}$$

and under a stronger smoothness condition on the b_{ik} the Milstein scheme has order 1, i.e. $\mathbb{E}(\max_{j=1}^{N} |\tilde{x}^{(j)} - x(jh)|^2) = O(h^2)$ (see Kloeden and Platen [2], Section 10.3). In fact stronger versions of these bounds hold, for example if (2) holds then

$$\mathbb{E}(\max_{j} |\tilde{x}^{(j)} - x(jh)|^p) = O(h^{p/2}) \tag{6}$$

holds for every $p \in [1, \infty)$ for the Euler scheme. We note that from (5) (or (6)) one can deduce results on almost sure convergence. For example if $\tilde{x}^{(r,j)}$ for $j = 1, \cdots 2^r$ denotes the Euler approximation with stepsize $h^{(r)} = 2^{-r}T$, then from (5) it follows easily that a.s. $\tilde{x}^{(r,2^r)} - x(T) = O(2^{(1-\epsilon)r/2})$ for any $\epsilon > 0$.

The Euler scheme is straightforward to implement, as the only random variables one has to generate are the normally-distributed $\Delta W_k^{(j)}$, but for Milstein one has also to generate the 'area integrals' $A_{kl}^{(j)}$ which is non-trivial if $d \geq 2$. Order $\frac{1}{2}$ is the best one can do in general when the only random variables generated are the $\Delta W_k^{(j)}$. An efficient method for generating area integrals when $d = 2$ is given in [8], but it seems very hard to extend it to $d > 2$.

There is however one special class of equations for which Milstein can be implemented using only the $\Delta W_k^{(j)}$. This comes from the observation that $I_{kl}^{(j)} + I_{lk}^{(j)} = 2B_{kl}^{(j)}$ where $B_{kl}^{(j)} = \frac{1}{2}\Delta W_k^{(j)} \Delta W_l^{(j)}$ if $k \neq l$ and $B_{kk}^{(j)} = \frac{1}{2}\{(\Delta W_k^{(j)})^2 - h\}$, which is the same as $A_{kk}^{(j)}$. It follows that if the condition

$$\rho_{ikl}(t,x) = \rho_{ilk}(t,x) \tag{7}$$

for all $x \in \mathbb{R}^q$, $t \in [0,T]$ and all i, k, l is satisfied then the Milstein scheme (4) reduces to

$$\tilde{x}_i^{(j+1)} = \tilde{x}_i^{(j)} + a_i(t_j, \tilde{x}^{(j)})h + \sum b_{ik}(t_j, \tilde{x}^{(j)})\Delta W_k^{(j)} + \sum \rho_{ikl}(t_j, \tilde{x}^{(j)})B_{kl}^{(j)} \tag{8}$$

which does not require generation of any random variables other than the $\Delta W_k^{(j)}$. The condition (7) is known as the *commutativity condition*.

We conclude that when (7) holds then (8), being equivalent to (4) in this case, is of order 1 (assuming the required smoothness of the b_{ik}). In particular this is true when $d = 1$. But for $d > 1$, scheme (8) in general only has order $\frac{1}{2}$. We shall see, however, that a modified form of (8) gives order 1 for a large class of SDEs. But first we describe sequential approximation using Lévy's construction.

3 Two-level approximation

In order to use Euler or (8) to simulate approximate solutions to (1) one has to generate the increments $\Delta W_k^{(j)}$, which are independent $N(0,h)$ random variables. If one wishes to simulate a sequence of approximations converging to the solution, one can use Lévy's construction of the Brownian motion, as follows. For $r \in \mathbb{N}$ let $h^{(r)} = 2^{-r}T$ and let $\Delta W_k^{(r,j)} = W_k((j+1)h^{(r)}) - W_k(jh^{(r)})$ be the corresponding increments. Then we have

$$\Delta W_k^{(r,j)} = \Delta W_k^{(r+1,2j)} + \Delta W_k^{(r+1,2j+1)} \tag{9}$$

In Levy's method one starts by generating the $\Delta W_k^{(r,j)}$ for a given r, then for each j,k generates $\Delta W_k^{(r+1,2j)}$ and $\Delta W_k^{(r+1,2j+1)}$, being independent $N(0,h^{(r+1)})$, conditional on (9) holding for the already generated $\Delta W_k^{(r,j)}$. This can then be repeated for $r+2, r+3, \cdots$. Then we get a sequence of sets of Brownian increments $(\Delta W_k^{(n,j)})$ for $n = r, r+1, \cdots$, and in the limit we get a Brownian path $W(t) = (W_1(t), \cdots W_d(t))$ such that $W_k(j2^{-n}T) = \sum_{i=0}^{j-1} \Delta W_k^{(n,i)}$ for each $n \geq r$ and $0 \leq j \leq 2^n$. For each n we obtain an approximate solution $x^{(n,j)}$ using (8), and as $n \to \infty$ this will converge to the solution of (1) for the Brownian path $W(t)$. As previously mentioned, the convergence will, if $d > 1$, usually be only of order $\frac{1}{2}$.

We will see in section 8 that a modified interpretation of (8), described in section (4), leads to a relation between between the random variables generated at successive levels which is different to (9), and which leads to convergence of order 1 under a non-degeneracy condition.

The essential point in such constructions is the joint generation of the required random variables at two consecutive levels h and $h/2$. In the Lévy construction one can either generate ΔW at level h and then conditionally generate the increments on the two subintervals as described above, or else start by generating the increments at level $h/2$ and adding them to get the level h increment. Then by several iterations of either process one can generate approximations at several levels for the same Brownian path, either starting at the coarsest level or at the finest. We remark that, while for the Brownian increments it is straightforward to use either method, for some purposes it may be easier to start at the finest level. This applies, for example, to the method described in [8] for the double integrals in two dimensions; for this there is no problem in generating the integrals (together with the increments) at level $h/2$ and combining them to get the integrals at level h, but the reverse process is very much harder.

We now describe two other applications of such two-level simulation, which we will consider later as potential applications of our coupling methods.

Empirical estimation of the error of a numerical method.

Suppose we have a numerical method which assigns an estimate x_h for $x(T)$, where x is the solution of an SDE (1) for a given stepsize h. In general we do not know $x(T)$ so we cannot estimate the mean error $\mathbb{E}|x(T) - x_h|$ directly but we can estimate $\mathbb{E}|x_h - x_{h/2}|$ by making R independent simulations $(x_h^{(r)}, x_{h/2}^{(r)})$ for $r = 1, \cdots, R$, each simulation being for one Brownian path, and using $\frac{1}{R}\sum_{r=1}^{R} |x_h^{(i)} - x_{h/2}^{(i)}|$ as an estimator. To justify this procedure, as a means of testing order of convergence and as an estimator of the constant, suppose that we have a bound $\mathbb{E}|x(T) - x_h| \leq Ch^\gamma$ for the actual error, then using the triangle inequality we get $\mathbb{E}|x_h - x_{h/2}| \leq C(1+2^{-\gamma})h^\gamma$. In the other direction, if a bound $\mathbb{E}|x_h - x_{h/2}| \leq C_1 h^\gamma$ holds then we similarly get $\mathbb{E}|x_h - x(T)| \leq \sum_{h=0}^{\infty} C_1(2^{-k}h)^\gamma = \frac{C_1 h^\gamma}{1-2^{-\gamma}}$. So the order of convergence can be estimated by estimating $\mathbb{E}|x_h - x_{h/2}|$ for a range of values of h, and the constant can be estimated up to a known multiplicative error.

Multilevel method for weak approximation of SDEs.

The second application is to the estimation of $\mathbb{E}f(x(T))$, where x is the solution of (1) as before, and f is a given Lipschitz function. The simplest approach to this problem is to use the average $\frac{1}{N}\sum_{i=1}^{N} f(x_h^{(i)})$ using N independent simulations of a method such as Euler with stepsize h, for a suitable choice of h. If the method has weak order γ (i.e. $|\mathbb{E}(f(x_h) - f(x(T)))| = O(h^\gamma)$) then to attain an error not greater than a given $\mathcal{E}$ one needs $h \sim \mathcal{E}^{1/\gamma}$, for the sampling error one needs $N \sim \mathcal{E}^{-2}$. So the total computational load is of order $\mathcal{E}^{-2-1/\gamma}$. For the Euler method the weak order is 1 so the load is of order $\mathcal{E}^{-3}$.

The multilevel approach of Giles [10, 11] is a modification which reduces this load. Let $h^{(r)} = 2^{-r}T$ and define $\mu_r = \mathbb{E}f(x_{h^{(r)}})$; then as above we wish to estimate μ_n for a suitable large n. Instead of doing this directly we write $\mu_n = \mu_0 + \sum_{k=1}^{n}(\mu_k - \mu_{k-1})$, and estimate μ_0 and each $\mu_k - \mu_{k-1}$ independently. To estimate $\mu_k - \mu_{k-1}$ we simulate $x_{h^{(k)}}$ and $x_{h^{(k-1)}}$ using the same Brownian path, so that (if we are using Euler) the variance of $x_{h^{(k)}} - x_{h^{(k-1)}}$ is of order 2^{-k}. This reduces the number of simulations required, and one can calculate (see [10]) that the computational load needed for error $\leq \mathcal{E}$ is $O(\mathcal{E}^{-2}(\log \mathcal{E})^2)$. If we use instead a method with strong order greater than $\frac{1}{2}$, such as Milstein, then the load is $O(\mathcal{E}^{-2})$, as shown in [11], provided the method can be implemented so that each two-level simulation has load $O(h^{-1})$. The removal of the $(\log \mathcal{E})^2$ represents a relatively small improvement but is worthwhile if it can be done efficiently. The order $\frac{3}{4}$ method described in [20] achieves this in principle, in that the two-level simulation can be done with $O(h^{-1})$ operations, but its detailed implementation is quite involved, especially in higher dimensions. The antithetic multilevel method of [12] (which in fact makes use of the scheme (8)) achieves $O(\mathcal{E}^{-2})$, under a stronger regularity condition on f than Lipschitz.

Consequently there is room for an efficient and easily implemented method with strong order greater than $\frac{1}{2}$, which would achieve a load of $O(\mathcal{E}^{-2})$ for Lipschitz f.

4 A modified approach using (8)

Here we describe a modified version of (8) which gives order 1 under a nondegeneracy condition on the b_{ik}. In terms of the actual calculations this version is in fact identical to (8), the difference is in the interpretation of the normal random variables generated, as we explain below.

In the standard approach to using schemes such as Milstein, for fixed j one generates separately the random variables $\Delta W_k^{(j)}$ and $I_{kl}^{(j)}$ and combines them to obtain the RHS of (4). The approach we consider here is rather to generate an approximation to the combination we need, i.e. $Y_i := \sum b_{ik}(\tilde{x}^{(j)})\Delta W_k^{(j)} + \sum \rho_{ikl}(\tilde{x}^{(j)})I_{kl}^{(j)}$ directly. We consider a scheme

$$\tilde{x}_i^{(j+1)} = \tilde{x}_i^{(j)} + \sum b_{ik}(\tilde{x}^{(j)})X_k^{(j)} + \frac{1}{2}\sum \rho_{ikl}(\tilde{x}^{(j)})(X_k^{(j)}X_l^{(j)} - h\delta_{kl}) \tag{10}$$

where the $X_k^{(j)}$ are independent with $N(0,h)$ distribution. This is the same as (8) with $\Delta W_k^{(j)}$ replaced by $X_k^{(j)}$. But we do not assume $X_k^{(j)} = \Delta W_k^{(j)}$; rather, we want $Z_i := \sum b_{ik}(\tilde{x}^{(j)})X_k^{(j)} + \frac{1}{2}\sum \rho_{ikl}(\tilde{x}^{(j)})(X_k^{(j)}X_l^{(j)} - h\delta_{kl})$ to be a good approximation to Y_i. To be more explicit, for a fixed k we seek a joint distribution of the sets of random variables $(\Delta W_k(j), I_{kl}^{(j)})$ and $(X_k^{(j)})$, having the required marginal distributions and so that $\mathbb{E}(Y_i - Z_i)^2 = O(h^3)$. If we can do this for each k then we will obtain, on a suitable probability space, a solution $x(t)$ of (1) and an approximation $(\tilde{x}_i^{(j)})$ obtained using (10) with $\mathbb{E}(x(jh) - \tilde{x}^{(j)})^2 = O(h^2)$, i.e. an order 1 approximation.

It turns out that this is possible if the matrix $(b_{ik}(x))$ is of rank q for each x, with bounded right inverse, and also smooth enough. One can then show that the random vectors Y and Z have smooth densities on $\mathbb{R}^d$ and by estimating the difference between the densities one

can obtain the required joint density. So in effect we can get order 1 convergence using (8), provided the rank condition holds. We prove this in section 7, as a special case of a more general result including higher order methods. For the case $q = d = 2$ an alternative proof is outlined in section 8, using couplings between approximations using (8) at different levels.

We remark that scheme (10) is essentially equivalent to the scheme described in [3].

5 Coupling and Vaserstein metrics

The question raised in section 4, of seeking a joint distribution of two random vectors with given marginal distributions is an example of the *coupling* of probability distributions. In the most general context, if we have two probability spaces $(\mathcal{X}, \mathcal{F}, \mathbb{P})$ and $(\mathcal{Y}, \mathcal{G}, \mathbb{Q})$ then a coupling between $\mathbb{P}$ and $\mathbb{Q}$ is a measure on $\mathcal{X} \times \mathcal{Y}$ which has $\mathbb{P}$ and $\mathbb{Q}$ as its marginal distributions. The problem considered in section 2 was effectively one of finding a coupling between the distributions of the $\mathbb{R}^d$-valued random vectors Y and Z, in other words a joint distribution consistent with the given marginals, with a good bound for $\mathbb{E}|Y-Z|^2$. In general the problem of finding a coupling which *minimises* $\mathbb{E}|Y - Z|^2$ is an example of an *optimal transport* problem. In our context we are interested in couplings with good bounds rather then in necessarily finding the minimum, but the size of the minimum of $\mathbb{E}|Y - Z|^2$ is a relevant quantity.

The general optimal transport problem is to minimise $\mathbb{E}c(Y, Z)$ where c is a non-negative 'cost function' on $\mathcal{X} \times \mathcal{Y}$. The case $c(X, Y) = |X - Y|^2$ for $\mathcal{X} = \mathcal{Y} = \mathbb{R}^n$ as considered above is the best understood one; the case $c(X, Y) = |X - Y|^p$ for general p is also well-studied. In this connection we write $\mathcal{T}_p(\mathbb{P}_1, \mathbb{P}_2)$ for the infimum of $\mathbb{E}|X - Y|^p$, taken over all joint distributions with the given marginals. We also write $\mathbb{W}_p(\mathbb{P}_1, \mathbb{P}_2)$ for $(\mathcal{T}_p(\mathbb{P}_1, \mathbb{P}_2))^{1/p}$. For $p \geq 1$ one can then show that $\mathbb{W}_p$ is a metric on the set of all probability measures $\mathbb{P}$ on $\mathbb{R}^n$ having finite pth moment (i.e. satisfying $\int_{\mathbb{R}^n} |x|^p d\mathbb{P}(x) < \infty$). $\mathbb{W}_p$ is known as the p-Vaserstein metric.

It is often convenient to abuse notation somewhat and write $\mathcal{T}_p(X, Y)$ for $\mathcal{T}_p(\mathbb{P}_1, \mathbb{P}_2)$ when X and Y have distributions $\mathbb{P}_1$ and $\mathbb{P}_2$ respectively, and similarly for $\mathbb{W}_p$. Note that if X and Y are random vectors on the same probability space it is not in general the case that $\mathcal{T}_p(X, Y) = \mathbb{E}|X - Y|^p$ - this will only be the case when the coupling given by the joint distribution of X and Y minimises $\mathbb{E}|X - Y|^p$.

When $n = 1$ we may also write $\mathcal{T}_p(F, G)$ where F and G are the corresponding distribution functions.

In the applications of coupling to the simulation of SDEs, typically at each step we are trying to simulate a random variable X and instead simulate Y which is an approximation to X (and easier to simulate). If we are simulating the SDE using a single stepsize h, we only requires the existence of a coupling between X and Y with a good bound for $\mathbb{E}|X-Y|^p$, without having to construct the coupling - in other words, we require to bound $\mathcal{T}_p(X, Y)$. But for multilevel simulation, where we are simulating approximations at two levels for the same driving path, we need an explicit coupling between the Y variables at the different levels which gives good error bounds.

For such purposes it is useful to be able to estimate $\mathcal{T}_p$, and we now mention two relevant classical results from optimal transport theory. We will not explicitly use them but they are useful for general orientation. The first concerns the one dimensional case:

Proposition 1. Let F and G be distribution functions on $\mathbb{R}$, and suppose F is continuous. Let $p \geq 1$. Then the infimum of $\mathbb{E}|X - Y|^p$, subject to $F_X = F$ and $F_Y = G$, is given by taking $Y = G^{-1}(F(X))$.

It follows then that, if F is absolutely continuous, $\mathcal{T}_p(F, G) = \int_{\mathbb{R}} |G^{-1}(F(x)) - x|^p f(x)dx$

where $f = F'$ is the corresponding density function. We can also write $\mathcal{T}_p(F,G) = \int_0^1 |F^{-1}(t) - G^{-1}(t)|^p dt$; this form is valid without any continuity assumption on F (or G).

Proposition 2. Let $\mathbb{P}_1$ and $\mathbb{P}_2$ be probability measures on $\mathbb{R}^n$, and suppose that $\mathbb{P}_1$ is absolutely continuous w.r.t. Lebesgue measure on $\mathbb{R}^n$. Then (1) there is a unique convex function ϕ on $\mathbb{R}^n$ such that $\nabla\phi(\mathbb{P}_1) = \mathbb{P}_2$, and (2) $\mathcal{T}_2(\mathbb{P}_1, \mathbb{P}_2) = \int_{\mathbb{R}^n} |x - \nabla\phi(x)|^2 d\mathbb{P}_1(x)$.

We note that a mapping $\psi : \mathbb{R}^n \to \mathbb{R}^n$ is of the form $\psi = \nabla\phi$, where ϕ is convex, if and only if the Jacobian matrix $D\psi$ is symmetric and positive semi-definite everywhere. So the coupling which minimises $\mathbb{E}|X - Y|^2$ is given by $Y = \psi(X)$ where ψ is the unique mapping with the above property such that $\psi(X)$ has distribution $\mathbb{P}_2$.

In general it is quite hard to find the ϕ (or corresponding $\psi = \nabla\phi$) whose existence is asserted in proposition 2.

We also note the elementary result (see e.g. proposition 7.10 in [26]) that

$$\mathbb{W}_p(\mu,\nu) \le 2^{(p-1)/p} \left\{ \int |x|^p d|\mu - \nu|(x) \right\}^{1/p} \tag{11}$$

for any two probability measures μ, ν on $\mathbb{R}^n$ and for any $p \ge 1$.

This is quite a good bound if $p = 1$ but is less good for $p > 1$; we shall however use it (with $p = 2$) for bounding some small remainder terms. A coupling which gives (11) is described at the end of this section.

Much information on coupling and optimal transport (far more than is needed for our problem) can be found in the books by Rachev and Ruschendorff [21] and by Villani [26, 27]. The second volume of [21] contains numerous applications to probability, including numerical methods for SDE.

An effective coupling for (11).

For future reference we note here an explicit procedure for generating coupled random variables which achieves the bound (11), in the case where μ and ν have densities f and g respectively, w.r.t. Lebesgue measure on $\mathbb{R}^n$. It is based on the proof of proposition 7.10 of [26], and is intended for the case where $\alpha := \int_{\mathbb{R}^n} |f(x) - g(x)| dx$ is small.

We first generate X with density f and (independently) U uniform on $[0,1]$. If $Uf(X) \le g(X)$ we set $Y = X$; otherwise we generate Y independently with density $2\alpha^{-1}(g - f)_+$. Then Y has density g and

$$\mathbb{E}|X-Y|^p = 2\alpha^{-1} \int_{\mathbb{R}^{2n}} |x-y|^p (f(x)-g(x))_+ (g(y)-f(y))_+ dxdy \le 2^{p-1} \int_{\mathbb{R}^n} |x|^p |f(x)-g(x)| dx$$

using the facts that $|x-y|^p \le 2^{p-1}(|x|^p + |y|^p)$, that $\int (f(x)-g(x)_+)dx = \int (g(x)-f(x))_+ dx = \alpha/2$ and that $|a-b| = (a-b)_+ + (b-a)_+$ for $a, b \in \mathbb{R}$. Note that (11) then follows.

The above procedure can be implemented provided one has effective algorithms for calculating $f(x)$ and $g(x)$, and for generating random variables with these densities. The procedure also requires generation of a variable with density $2\alpha^{-1}(g - f)_+$ and one way of doing this is a rejection method: generate Y with density g and accept it with probability $\max(0, 1 - \frac{f(x)}{g(x)})$; if it is not accepted, generate a new Y and repeat until accepted. This will require $2\alpha^{-1}$ attempts on average, but as this case occurs with probability $\alpha/2$ the computational load is still $O(1)$ even if α is small (as is typically the case).

It may happen that one of the densities (say f) may be difficult to calculate exactly, but a convergent sequence (f_n) of approximations can be calculated iteratively. Then, for example to test whether $Uf(X) \le g(X)$, we first modify the sequence (f_n) to give two sequences (f_n^+) and (f_n^-) converging to f, with $f_n^+ \ge f$ and $f_n^- \le f$ (for example by setting $f_n^\pm(x) = f_n(x) \pm \epsilon_n(x)$ where $\epsilon_n(x)$ is a calculable upper bound for $|f_n(x) - f(x)|$ such that $\epsilon_n(x) \to 0$ as $n \to \infty$). Then if $Uf_1^+(X) \le g(X)$ then we know that $Uf(X) \le g(X)$.

Otherwise we test whether $Uf_1^-(X) > g(X)$; if so then $Uf(X) > g(X)$, otherwise we test with f_2^+ and so on. If f_1 is a good approximation and the convergence is reasonably fast then the first test will usually suffice and the expected number of tests will be ≤ 2.

An application of this method is outlined at the end of section 8.

Note on spelling. Vaserstein's original paper [25] is in Russian; we have used the transliteration 'Vaserstein' from the Cyrillic alphabet as that is the one used by Vaserstein himself in his English-languauge publications. The alternative spelling 'Wasserstein' is also widely used in the literature.

6 Perturbation method

To illustrate the idea behind this method, consider the following situation: suppose we wish to simulate a random variable $U = X + \epsilon Y$, where X and Y are independent, X has a smooth density, and ϵ is small. Suppose also that X has a distribution which is easy to generate (e.g. normal) but Y is hard to generate. Then generating U by generating X and Y will be hard. An alternative is as follows: suppose Z is another random variable (independent of X), which is easy to generate and has the same moments up to order $m-1$ as Y (i.e. $\mathbb{E}(Z^k) = \mathbb{E}(Y^k)$ for $k = 1, \cdots, m-1$). The idea is that then $V = X + \epsilon Z$ can be used as an approximation to U, with error of order ϵ^m.

To see why this might be true, writing f_X for the density of X etc, we have $f_U(x) = \mathbb{E}f_X(x - \epsilon Y) = f(x) + \sum_{k=1}^{m-1} \frac{(-\epsilon)^k}{k!} f^{(k)}(x)\mathbb{E}(Y^k) + O(\epsilon^m)$. Because Z has the same moments we get the same expression for $f_V(x)$, so $f_U(x) - f_V(x) = O(\epsilon^m)$. One might expect to get from this a Vaserstein distance estimate of the same order $\mathbb{W}_2(U, V) = O(\epsilon^m)$ which gives a coupling with $\mathbb{E}(U - V)^2 = O(\epsilon^{2m})$.

This reasoning is not rigorous and is only an outline; one needs some restrictions on the distributions for it to work. Also we are mainly interested in an extension of this approach to vector random variables. We now outline also the idea of the method we shall use to deduce a bound for the Vaserstein distance from a bound for $f_U - f_V$; we do this in the vector case.

So we suppose U and V are $\mathbb{R}^d$-valued random variables, with density functions g and h respectively. Suppose u is an $\mathbb{R}^d$-valued function on $\mathbb{R}^d$ such that $\nabla.u(x) = h(x) - g(x)$ for $x \in \mathbb{R}^d$. For $t \in [0, 1]$ write $g_t = g + t(h - g)$ and $\zeta_t(x) = g_t(x)^{-1}u(x)$. For fixed x define $\rho_t(x)$ as the solution of the differential equation system $\frac{d}{dt}\rho_t(x) = \zeta_t(\rho_t(x))$ with $\rho_0(x) = x$. Provided everything is well-defined (in other words $g_t(x)$ does not vanish and the solution of the ODE does not explode for $t \in [0, 1]$), a standard calculation shows that $\rho_t(U)$ has density g_t and in particular $\rho_1(U)$ has density $g_1 = h$ and so a coupling between the two distributions is given by $V = \rho_1(U)$ and hence $\mathbb{W}_2(g, h) \leq (\int_{\mathbb{R}^d} |x - \rho_1(x)|^2 dx)^{1/2}$. The idea is that if $|h - g|$ is small then we can choose u to be small, so that ζ_t and hence $|x - \rho_1(x)|$ will be small. In practice there are several details to check, and in particular there are potential problems if g gets close to 0. We apply this method in two different situations in sections 7 and 9.

7 Higher-order approximations in non-degenerate case

Now we consider the use of coupling to generate high-order approximations to the stochastic Taylor expansions used to approximate SDE solutions, in the case where the diffusion matrix is nondegenerate. We first briefly review the construction of Itô-Taylor approximations, as described in detail in chapter 5 and section 10.6 of [16].

Following the notation of chapter 5 of [16] we let $\mathcal{M}$ be the set of all multi-indices $\alpha = (j_1, \cdots, j_l)$ of length $l = l(\alpha)$ with $0 \leq j_k \leq d$. We define the iterated integral

$I_{\alpha,s,t} = \int_s^t \int_s^{t_l} \cdots \int_s^{t_2} dW_{j_1}(t_1) \cdots dW_{j_l}(t_l)$. Note that in this notation, $\Delta W_k^{(j)}$ as defined in section 2 is $I_{k,jh,(j+1)h}$ and $I_{kl}^{(j)}$ is $I_{kl,jh,(j+1)h}$.

For $m \geq 2$ in $\mathbb{N}$ we define

$$A_m = \left\{ \alpha \in \mathcal{M} : l(\alpha) \geq 2 \text{ and either } l(\alpha) + n(\alpha) \leq m \text{ or } l(\alpha) = n(\alpha) = \frac{m+1}{2} \right\}$$

where $n(\alpha)$ is the number of zero indices in α.

We also have the Itô diffusion operators defined by

$$L^0 = \frac{\partial}{\partial t} + \sum_{k=1}^{q} a_k \frac{\partial}{\partial x_k} + \frac{1}{2} \sum_{k,l=1}^{q} \sum_{j=1}^{d} b_{kj} b_{lj} \frac{\partial^2}{\partial x_k \partial x_l}$$

and $L^j = \sum_{k=1}^{q} b_{kj} \frac{\partial}{\partial x_k}$ for $j = 1, \cdots, d$.

Then the order $\gamma = \frac{m}{2}$ Taylor approximation to (1) is given by

$$\tilde{x}_i^{(j+1)} = \tilde{x}_i^{(j)} + a_i(jh, \tilde{x}^{(j)})h + \sum_{k=1}^{d} b_{ik}(jh, \tilde{x}^{(j)}) \Delta W_k^{(j)} + \sum_{\alpha \in A_m} f_{\alpha,i}(jh, \tilde{x}^{(j)}) I_{\alpha,t_j,t_{j+1}} \qquad (12)$$

where the R^q-valued functions $f_\alpha(t,x)$ are defined recursively by $f_v(t,x) = x$ and $f_{j\alpha} = L^j f_\alpha$ for $j = 0, \cdots, d$, where v is the multi-index of zero length, and $f_{\alpha,i}$ denotes the ith component of f_α. For $\gamma = \frac{1}{2}$ we have $A_{1/2} = \{v, (0), (1), \cdots, (d)\}$ which gives the Euler scheme, and for $\gamma = 1$ we get also all the multi-indices (jk) for $j, k = 1, \cdots, d$, giving Milstein.

Our objective is to approximate the RHS of (12) by a random variable which is easier to generate. We do this using the perturbation and coupling approach of section 6, but in modified form as we require to apply the perturbation argument with the random variables X and Y of section 6 not being independent. We first prove three lemmas which will be needed.

Let Σ be a positive definite real $q \times q$ matrix and let f be the density function on $\mathbb{R}^q$ of the $N(0, \Sigma)$ normal distribution. Let $\mathcal{P}$ denote the set of polynomials in d variables with real coefficients and let the projection operator P on $\mathcal{P}$ be defined by $(Pp)(x) = P(x) - \overline{p}$ where $\overline{p} = \int_{\mathbb{R}^q} p(x) f(x) dx$. Then $\overline{Pp} = 0$. We have the following:

Lemma 1. *Let $p \in \mathcal{P}$. Then we can find a vector polynomial $\psi \in \mathcal{P}^d$ such that $\nabla.(f\psi) = fPp$.*

Proof. By an orthogonal linear change of coordinates we may suppose that Σ is diagonal with entries $\sigma_1^2, \cdots, \sigma_d^2$.

We prove the lemma by induction on the degree n of p. If $n = 0$ then $Pp = 0$ so $\psi = 0$ will do. Then we suppose that $n > 0$ and the lemma is true for polynomials of degree $< n$. By linearity we can assume $p(x) = x_1^{k_1} \cdots x_d^{k_q}$ where $\sum_{j=1}^{q} k_j = n > 0$. Then define $\lambda \in \mathcal{P}^q$ by $\lambda = -(\sum_{j=1}^{q} \sigma_j^{-2} k_j)^{-1} \nabla p$; then $\nabla.(f\lambda) = (p + r)f$ where r has degree $< n$. By the inductive hypothesis we can find $\alpha \in \mathcal{P}^d$ such that $\nabla.(f\alpha) = fPr$ and then $\psi = \lambda - \alpha$ gives $\nabla.(f\psi) = fPp$ as required. □

We remark that the proof shows that ψ can be chosen to be the gradient of a polynomial (and indeed there is a unique choice of ψ of this form).

The next lemma summarises the required coupling argument.

Lemma 2. *Let $n \leq N$ and R be positive integers, and for $j = 1, \cdots, N$ let $p_j, r_j \in \mathcal{P}$, all having degree $\leq R$, and such that $p_j = r_j$ for $j \leq n$. Let $\eta > 0$ with $\eta R \leq n$ and let $K > 0$.*

Then we can find $C > 0$ such that, if $\epsilon > 0$ and we write $\mu_0 = pf\chi_B dx$ and $\nu_0 = rf\chi_B dx$ where $p = 1 + \sum_{j=1}^{N} \epsilon^j p_j$, $r = 1 + \sum_{j=1}^{N} \epsilon^j r_j$ and $B = \{x \in \mathbb{R}^q : |x| \leq \epsilon^{-\eta}\}$, and if μ and ν are probability measures on $\mathbb{R}^q$ with $\int_{\mathbb{R}^q} (1 + |x|^2) d|\mu - \mu_0|(x) < K\epsilon^{2n+2}$ and $\int_{\mathbb{R}^q} (1 + |x|^2) d|\nu - \nu_0|(x) < K\epsilon^{2n+2}$, then $\mathbb{W}_2(\mu, \nu) < C\epsilon^{n+1}$.

Proof. By Lemma 1 we can find $\psi_j \in \mathcal{P}^q$ for $j = 1, \cdots, N$ such that $\nabla.(f\psi_j) = fP(p_j - r_j)$. Let $\psi = \sum_{j=1}^N \epsilon^j \psi_j$ and then $\nabla.(f\psi) = f\theta$ where $\theta = P(p-r)$. We write $p_t = p - t\theta$.

We let $B_+ = \{x \in \mathbb{R}^q : |x| \leq \epsilon^{-\eta} + 1\}$ and $B_- = \{x \in \mathbb{R}^q : |x| \leq \epsilon^{-\eta} - 1\}$. Then if ϵ is small enough we have $|\frac{\psi}{p_t}| \leq C_1\epsilon^{n+1}|x|^R \leq 1$ on B_+ and hence, if we define $\phi_t(x)$ to be the solution of $\frac{d\phi_t}{dt} = \frac{\psi(\phi_t)}{p_t(\phi_t)}$ with initial condition $\phi_0 = x$, we have that, for $x \in B$, $|\phi_t(x) - x| \leq C_1\epsilon^{n+1}|x|^R \leq 1$ and so $\phi_t(x)$ remains in B_+ for $0 \leq t \leq 1$. It follows that $B_- \subseteq \phi_1(B) \subseteq B_+$. Then $\phi_1(\mu_0)$ is a measure on $\phi_1(B)$ with density $(p-\theta)f$. We see also that $\int |\phi_1(x) - x|^2 d\mu_0(x) < C_2\epsilon^{2n+2}$ so that $\mathbb{W}(\phi_1(\mu_0), \mu_0) < C_2^{1/2}\epsilon^{n+1}$. And, noting that $p - r - \theta$ is a constant bounded by $2K\epsilon^{2n+1}$, and also noting the exponential bound for f on $B_+ \backslash B_-$, we have $\int(1 + |x|^2)d|\phi_1(\mu_0) - \nu_0|(x) < C_1\epsilon^{2n+2}$, so that using (11) we have $\mathbb{W}_2(\mu_0, \nu_0) < C_3\epsilon^{n+1}$. Finally the required result follows from this and the hypotheses on $\mu - \mu_0$ and $\nu - \nu_0$, using (11) again. □

In order to apply lemma 2, we need to use the (modified) perturbation procedure of section 6 to obtain approximations of the form pf, with polynomial p, for the relevant densities. This requires some polynomial calculations, given in the next lemma.

Lemma 3. *Let $Q : \mathbb{R} \times \mathbb{R}^q \times \mathbb{R}^r \to \mathbb{R}^q$ be a polynomial function (i.e. each component is a polynomial in the $1 + q + r$ variables) such that $Q(0, z, 0) = 0$ for all $z \in \mathbb{R}^q$ and let $n \in \mathbb{N}$. The one can find polynomial functions $T : \mathbb{R} \times \mathbb{R}^q \times \mathbb{R}^r \to \mathbb{R}^q$ and $S : \mathbb{R} \times \mathbb{R}^q \times \mathbb{R}^q \times \mathbb{R}^r \to \mathbb{R}^q$ such that $T(0, z, 0) = 0$ for all $z \in \mathbb{R}^q$ and, if $x \in \mathbb{R}^q$, $v \in \mathbb{R}^r$ and $z = x + Q(\epsilon, x, \epsilon v)$ then $x = z + T(\epsilon, z, \epsilon v) + \epsilon^n S(\epsilon, x, z, v)$.*

Proof. First note that we can write $Q(\epsilon, x, \epsilon v) = \epsilon R(\epsilon, x, v)$ where R is a polynomial. Next define polynomials $Q_k : \mathbb{R} \times \mathbb{R}^q \times \mathbb{R}^q \times \mathbb{R}^r \to \mathbb{R}^q$ and $P_k : \mathbb{R} \times \mathbb{R}^q \times \mathbb{R}^r \to \mathbb{R}^q$ for $k = 1, 2, \cdots$ recursively by $Q_{k+1}(\epsilon, x, z, u) = Q_k(\epsilon, z - Q(\epsilon, x, u), z, u) - Q_k(\epsilon, z, z, u)$ and $P_{k+1}(\epsilon, z, u) = P_k(\epsilon, z, u) - Q_k(\epsilon, z, z, u)$, starting with $Q_1(\epsilon, x, z, u) = Q(\epsilon, x, u)$ and $P_1(\epsilon, z, u) = u$. The following properties are easily shown by induction on k: $P_k(0, x, 0) = 0$ for all x; $Q_k(\epsilon, x, z, \epsilon v) = \epsilon^k R_k(\epsilon, x, z, v)$ where R_k is a polynomial; if $z = x + Q(\epsilon, x, \epsilon v)$ then $P_k(\epsilon, z, \epsilon v) - Q_k(\epsilon, x, z, \epsilon v) = x$. The lemma then follows by taking $T(\epsilon, z, u) = P_n(\epsilon, z, u)$ and $S(\epsilon, x, z, v) = -R_n(\epsilon, x, z, v)$. □

We now return to (12). For notational simplicity we consider the first step, from 0 to h, of (12), and write I_α for $I_{\alpha,0,h}$. We can write this step as $x_j^{(1)} = x_j^{(0)} + a_j(0, x^{(0)})h + Y_j$ where

$$Y_i = \sum_{k=1}^d b_{ik} W_k(h) + \sum_{\alpha \in A_m} c_{\alpha,i} I_\alpha$$

for $i = 1, \cdots, q$, defining an $\mathbb{R}^q$-valued random variable $Y = (Y_1, \cdots, Y_q)$. Here $b_{ik} = b_{ik}(0, x^{(0)})$ and $c_{\alpha,i} = f_{\alpha,i}(0, x^{(0)})$ are real constants.

To isolate the dependence on h, and also to separate the random variables into parts dependent and independent of $W(h)$, we write $W_j(t) = h^{1/2}B_j(t/h) + th^{-1/2}V_j$ where $B_1, \cdots, B_d$ are independent standard Brownian bridges on $(0, 1)$ and $V_j = h^{-1/2}W_j(h)$ are independent $N(0, 1)$ (and are independent of the B_j). We also write $B_0(t) = t$ and $K_\alpha = \int_0^1 \int_0^{t_l} \cdots \int_0^{t_2} dB_{j_1}(t_1) \cdots dB_{j_l}(t_l)$.

Then for $\alpha = (j_1, \cdots, j_l)$ we can write I_α as

$$I_\alpha = h^{(l(\alpha)+n(\alpha))/2} \sum_{\beta=(i_1,\cdots,i_l)} K_\beta \prod_{k:i_k<j_k} V_{j_k} \tag{13}$$

where the sum is over all $\beta = (i_1, \cdots, i_l)$ such that for each $k \in \{1, \cdots, l\}$ we have either $i_k = j_k$ or $i_k = 0 < j_k$. (For example, $I_{120} = h^2(K_{120} + K_{020}V_1 + K_{100}V_2 + K_{000}V_1V_2)$). Setting

$\epsilon = h^{1/2}$, and defining $\mathcal{M}_m = \{\alpha \in \mathcal{M} : 2 \le l(\alpha) \le m\}$, we can then write

$$Y_i = \epsilon \sum_{k=1}^{d} b_{ik} V_k + Q_i(\epsilon, V, (\epsilon^{l(\alpha)} K_\alpha)_{\alpha \in \mathcal{M}_m})$$

where Q_i is a polynomial, each monomial of which has overall order at least 2 in ϵ (i.e. there are no monomials of the form V^β or ϵV^β).

The idea is to generate Y approximately by replacing the K_α by other random variables L_α which are easier to generate. Suppose we have random variables L_α, defined for $\alpha \in \mathcal{M}^*$ such that $2 \le l(\alpha) \le m$. Then we can define

$$\tilde{Y}_i = \epsilon \sum_{k=1}^{d} b_{ik} \tilde{V}_k + Q_i(\epsilon, V, (\epsilon^{l(\alpha)} L_\alpha)_{\alpha \in \mathcal{M}_m})$$

where the $\tilde{V}_k$ are independent $N(0,1)$ and independent of the L^α (but not necessarily equal to the V_k) The main result is:

Theorem 4. *Assume the matrix (b_{ik}) has rank q. Suppose the random variables L_α have all moments finite and that $\mathbb{E}(K_{\alpha_1} \cdots K_{\alpha_r}) = \mathbb{E}(L_{\alpha_1} \cdots L_{\alpha_r})$ whenever $\alpha_1, \cdots, \alpha_r \in \mathcal{M}_m$ satisfy $\sum_{k=1}^{r} (l(\alpha_k) - 1) \le m - 1$. Then $\mathbb{W}_2(Y, \tilde{Y}) \le C\epsilon^{m+1}$ where the constant C depends only on d, m, upper bounds for the constants b_{ik}, $c_{i,\alpha}$ and a right inverse of the matrix (b_{ik}), and moment bounds for the L_α.*

Proof. We write $r = |\mathcal{M}_m|$ and introduce the notation $X_i = \sum_{k=1}^{d} b_{ik} V_k$ and $Z_i = \epsilon^{-1} Y_i$; then we have $Z_i = X_i + \hat{Q}_i(\epsilon, X, (\epsilon^{l(\alpha)-1} K_\alpha)_{\alpha \in \mathcal{M}_m})$ where $\hat{Q}_i$ is a polynomial on $\mathbb{R} \times \mathbb{R}^q \times \mathbb{R}^r$ whose coefficients are polynomials in $(b_{ik}), (c_{i,\alpha})$, and the polynomial $\hat{Q}_i$ contains no monomials of the form X^β. Replacing K_α by L_α we can define $\tilde{Z}_i = X_i + \hat{Q}_i(\epsilon, X, (\epsilon^{l(\alpha)-1} L_\alpha)_{\alpha \in \mathcal{M}_m})$ and then $\tilde{Y}_i = \epsilon \tilde{Z}_i$. Note that X has a nondegenerate normal distribution.

Now we apply lemma 3 to $\hat{Q} = (\hat{Q}_1, \cdots, \hat{Q}_q)$, with $r = |\mathcal{M}_m|$ as above, $v = (\epsilon^{l(\alpha)-2} K_\alpha)_{\alpha \in \mathcal{M}_m}$ and $n = 2m + 1$, and obtain polynomial functions T and S such that if $z = x + \hat{Q}(\epsilon, x, \epsilon v)$ then $x = z + T(\epsilon, z, \epsilon v) + \epsilon^n S(\epsilon, x, z, v)$. Let N be the maximum of the degrees of $\hat{Q}$, S and T, and let $\eta = \frac{1}{2N}$ and $R = \epsilon^{-\eta}$.

Next we fix a smooth function $\chi : \mathbb{R}^q \to \mathbb{R}^q$ such that $\chi(x) = x$ for $|x| \le R+2$ and $R+2 \le \chi(x) \le 2R$ for $|x| > R + 2$. Then set $X^* = \chi(X)$, $K^*_\alpha = \chi(K_\alpha)$, $v^* = (\epsilon^{l(\alpha)-2} K^*_\alpha)_{\alpha \in \mathcal{M}_m}$, $\tilde{v}^* = (\epsilon^{l(\alpha)-2} L^*_\alpha)_{\alpha \in \mathcal{M}_m}$, $L^*_\alpha = \chi(L_\alpha)$, $Z_* = X + \hat{Q}(\epsilon, X^*, \epsilon v^*)$ and $\tilde{Z}_* = X + \hat{Q}(\epsilon, X^*, \epsilon \tilde{v}^*)$. Then

$$\mathbb{E}|K^*_\alpha - K_\alpha|^2 = O(\epsilon^M), \quad \mathbb{E}|L^*_\alpha - L_\alpha|^2 = O(\epsilon^M), \quad \mathbb{E}|Z^* - Z|^2 = O(\epsilon^M), \quad \mathbb{E}|\tilde{Z}^* - \tilde{Z}|^2 = O(\epsilon^M) \tag{14}$$

for any given $M \in \mathbb{N}$.

Now if we fix values of (K_α) then we can write $Z^* = \phi(X^*)$ where

$$\phi_i(x) = x_i + \hat{Q}_i(\epsilon, x, \epsilon v^*)$$

Since $N \ge \deg(\hat{Q})$, we have for any $x \in B_{R+1}$ that $|\hat{Q}(\epsilon, x, \epsilon v^*)| = O(\epsilon^{1/2})$ and $|D\hat{Q}(\epsilon, x, \epsilon v^*)| = O(\epsilon^{1/2})$. Then for ϵ small enough, $|\hat{Q}(\epsilon, x, \epsilon v^*)| \le 1$ and $\|DQ(\epsilon, x, \epsilon v^*)\| \le \frac{1}{2}$, where DQ is the matrix of partialderivatives w.r.t. x. Then ϕ is one-one on B_{R+1} and $B_R \subseteq \phi(B_{R+1}) \subseteq \phi(B_{R+2})$.

Let $f^c_{Z^*}$ denote the conditional density of Z^* given (K_α). Then for $z \in \phi(B_{R+1})$ we have $f^c_{Z^*}(z) = f_X(\phi^{-1}(z)) |J_\phi(\phi^{-1}(z))|^{-1}$ where J_ϕ denotes the Jacobian determinant of ϕ. Write the Taylor expansion of $f_X(z+h)$ about z as $f_X(z)\{1 + \sum_{|k|=1}^{\infty} \theta_k(z) h^k\}$ where θ_k is a polynomial of degree $|k|$, the sum being over multi-indices $k = (k_1, \cdots, k_q)$. Let $t(z,h) =$

$\sum_{|k|=1}^{2m-1}\theta_k(z)h^k$ and note that, since all derivatives of f_X are bounded on $\mathbb{R}^q$, we have $f(z+h)=f_X(z)\{1+t(z,h)\}+O(\epsilon^{2m})$. Then if $x\in B_{R+1}$ and $z=\phi(x)$, we have $x=z+T(z)+O(\epsilon^{2m})$ and then

$$f_X(x)=f_X(z+T(z))+O(\epsilon^{2m})=f_X(z)\{1+t(z,T(z))\}+O(\epsilon^{2m})$$

We also have that $J_\phi(u)=\det(I+D_u\hat{Q}(\epsilon,u,\epsilon v^*))=1+\psi(\epsilon,u,\epsilon v^*)$, where ψ is a polynomial on $\mathbb{R}\times\mathbb{R}^q\times\mathbb{R}^r$, with degree $\le N^q$, and then $|\psi(\epsilon,u,\epsilon v^*)|=O(\epsilon^{1/2})$ for $u\in B_{R+1}$. So for $z\in\phi(B_{R+1})$ we have

$$\begin{aligned}J_\phi(\phi^{-1}(z))^{-1}&=1+\sum_{j=1}^{4m}(-1)^j\psi(\epsilon,\phi^{-1}(z),\epsilon v^*)^j+O(\epsilon^{2m})\\&=1+\sum_{j=1}^{4m}(-1)^j\psi(\epsilon,z+T(z),\epsilon v^*)^j+O(\epsilon^{2m})\end{aligned}$$

So

$$\begin{aligned}f^c_{Z^*}(z)&=f_X(z)\{1+t(z,T(z))\}\left(1+\sum_{j=1}^{4m}(-1)^j\psi(\epsilon,z+T(z),\epsilon v^*)^j\right)+O(\epsilon^{2m})\\&=f_X(z)(1+\sum_{j=1}^{S}\epsilon^jU_j)+O(\epsilon^{2m})\end{aligned}$$

for some integer S, where U_j is a polynomial in z and in those K^*_α with $l(\alpha)-1\le j$, and for any monomial $z^kK^*_{\alpha_1}\cdots K^*_{\alpha_r}$ we have $\sum_{j=1}^r(l(\alpha_j)-1)\le j$. Finally taking expectation w.r.t. the K^*_α we obtain for the unconditional density $f_{Z^*}(z)=f_X(z)(1+\sum_{j=1}^{2m-1}\epsilon^j\mathbb{E}U_j)+O(\epsilon^{2m})$. Replacing K^*_α by L^*_α we obtain $f_{\tilde{Z}^*}(z)=f_X(z)(1+\sum_{j=1}^{2m-1}\epsilon^j\mathbb{E}\tilde{U}_j)+O(\epsilon^{2m})$ where $\tilde{U}_j$ is obtained from U_j by replacing K_α by L_α for each α.

The hypothesis on the moments, together with (14), implies that $\mathbb{E}U_j-\mathbb{E}\tilde{U}_j=O(\epsilon^{2m})$ for $j\le 2m-1$ and then Lemma 2 says that $\mathbb{W}_2(Z^*,\tilde{Z}^*)=O(\epsilon^m)$. Again from (14) we have $\mathbb{W}_2(Z^*,Z)=O(\epsilon^m)$ and $\mathbb{W}_2(\tilde{Z}^*,\tilde{Z})=O(\epsilon^m)$ and the result follows.

□

Application of theorem.

We now consider the application of theorem 4 to the generation of approximate solutions of (1) of given order $\gamma=\frac{m}{2}$. We suppose the coefficients a_k,b_{ik} are sufficiently regular (uniform bounds for the coefficients and their derivatives up to order two will certainly suffice) and that the matrix (b_{ik}) has rank q everywhere with uniformly bounded right inverse. We modify (12) by replacing $I_{\alpha,jh,(j+1)h}$ with $\tilde{I}_{\alpha,jj,(j+1)h}$ obtained as above by expressing the I_α in terms for integrals K_β of the Brownian bridge using (13) and then replacing K_β by suitable L_β. The modified scheme is then

$$\hat{x}^{(j+1)}=\hat{x}^{(j)}+a(jh,\hat{x}^{(j)})h+B(jh,\hat{x}^{(j)})\tilde{V}^{(j)}+\sum_{\alpha\in A_m}f_\alpha(t_j,\hat{x}^{(j)})\tilde{I}_{\alpha,t_j,t_{j+1}}\tag{15}$$

Theorem 4 tells us that, with a suitable coupling (which will depend on $x^{(j)}$), we have

$$\mathbb{E}\left|B(jh,\hat{x}^{(j)})\tilde{V}^{(j)}+\sum_{\alpha\in A_m}f_\alpha(jh,\hat{x}^{(j)})\tilde{I}_{\alpha,jh,(j+1)h}-B(jh,\hat{x}^{(j)})V^{(j)}\sum_{\alpha\in A_m}f_\alpha(jh,\hat{x}^{(j)})I_{\alpha,jh,(j+1)h}\right|^2\le Ch^{m+1}$$

Then the proof of Theorem 10.6.3 of [16] shows that

$$\mathbb{E}(\max_{j=1}^{N} |\hat{x}^{(j)} - x(jh)|^2) \leq C_1 h^m \tag{16}$$

in other words the method is of order $\gamma = \frac{m}{2}$.

We note that, in the case $m = 2$, this result can also be deduced from the main result of [3].

Moment calculations.

In order to generate suitable L_β we need to calculate the relevant expectations of products of K_β. We start with two lemmas which help with such calculations. The first is a simple observation that certain K_β are actually deterministic. Here we write $I_\alpha = I_{\alpha,0,1}$ (i.e. we take $h = 1$).

Lemma 5. *Let $\beta = (jj\cdots j)$ with length $l \geq 2$. Then (i) if $j = 0$ then $K_\beta = \frac{1}{l!}$, and (ii) if $j > 0$ then $K_\beta = 0$ if l is odd, while $K_\beta = \frac{(-1)^r}{2^r r!}$ if $l = 2r$.*

Proof. (i) is a simple calculation. For (ii) we use equation (5.2.21) of [16] which expresses I_β as a Hermite polynomial in $W_j(1)$. Also from $W_j(t) = B_j(t) + tW_j(1)$ we an write $I_\beta = \sum_{r=0}^{l} \psi_r W_j(1)^r$ where each ψ_r is a linear combination of some K_α, with $\psi_0 = K_\beta$. Since $W_j(1)$ is independent of all K_α we can compare coefficients of the two expansions of I_β as polynomials in $W_j(1)$, and conclude that K_β is the constant term in the Hermite polynomial from (5.2.21) of [16] (taking $t = 1$), and this term is 0 if l is odd and $\frac{(-1)^r}{2^r r!}$ if $l = 2r$. □

Next, certain moments are automatically zero:

Lemma 6. *If $\beta_1, \cdots, \beta_s \in \mathcal{M}$ and if some $j \geq 1$ occurs an odd number of times in the concatenated multi-index $\beta_1 \cdots \beta_s$ then $\mathbb{E}(K_{\beta_1} \cdots K_{\beta_s}) = 0$.*

Proof. This follows from the fact that the law of the process $B_j(t)$ is the same as that of $-B_j(t)$, and that replacing B_j by $-B_j$ in K_{β_i} multiplies it by a factor of $(-1)^l$ where l is the number of occurrences of j in β_i. □

For example, $\mathbb{E}(K_{202}K_{121}) = 0$ since 2 occurs 3 times in (202121).

A systematic approach to calculating $\mathbb{E}(K_{\beta_1} \cdots K_{\beta_s})$ in general is as follows. First, $K_{\beta_1} \cdots K_{\beta_s}$ can be expressed as a linear combination of terms of the form $I_{\alpha_1} \cdots I_{\alpha_r}$ (using $B_j(t) = W_j(t) - tW_j(1)$). Then using the shuffle algebra structure of the algebra of Itô integrals [7], such terms can in turn be expressed as linear combinations of single integrals I_α. The expectation of such integrals is easy to find, since $\mathbb{E}I_\alpha = 0$ unless $\alpha = (0 \cdots 0)$. This approach can be tedious as a large number of terms may be produced, and quicker methods can often be used, as we now illustrate with the following lemma which, together with lemmas 5 and 6, gives all the moments needed for the cases $m = 2$ and $m = 3$.

Lemma 7. *(i) if $0 \leq k < l$ then $K_{lk} = -K_{kl}$ and $\mathbb{E}(K_{kl}^2) = \frac{1}{12}$.*
(ii) if $k > 0$ then $\mathbb{E}K_{0kk} = \mathbb{E}K_{k0k} = \mathbb{E}K_{kk0} = -\frac{1}{6}$.

Proof. (i) $K_{lk} = -K_{kl}$ follows using integration by parts. For the other part, first suppose $k = 0$. Then $\mathbb{E}(I_{0l}^2) = \mathbb{E}(\int_0^1 t dW_l)^2 = \int_0^1 t^2 dt = \frac{1}{3}$. Also, using $K_{00} = \frac{1}{2}$ from Lemma 5, we have $I_{0l} = \frac{1}{2} I_l + K_{0l}$. We square both sides and take expectations; since all K_α are independent of all I_i, we see that the cross-terms vanish and so $\frac{1}{3} = \mathbb{E}(I_{0l}^2) = \frac{1}{4} + \mathbb{E}(K_{0l}^2)$ so $\mathbb{E}(K_{0l}^2) = \frac{1}{12}$. Next, if $k > 0$ then similarly we have $\mathbb{E}(I_{kl}^2) = \mathbb{E}(\int_0^1 W_k(t) dW_l(t))^2 =$

$\mathbb{E}(\int_0^1 W_k(t)^2 dt) = \int_0^1 t dt = \frac{1}{2}$. And $I_{kl} = \frac{1}{2} I_k I_l + K_{0l} I_k + K_{k0} I_l + K_{kl}$. Squaring and taking expectations again, and noting that by the first part we have $K_{k0}^2 = K_{0k}^2$, we obtain

$$\frac{1}{2} = \mathbb{E}(I_{kl}^2) = \frac{1}{4} + \mathbb{E}(K_{0k}^2) + \mathbb{E}(K_{0l}^2) + \mathbb{E}(K_{kl}^2) = \frac{1}{4} + \frac{1}{12} + \frac{1}{12} + \mathbb{E}(K_{kl}^2)$$

so $\mathbb{E}(K_{kl}^2) = \frac{1}{12}$.
(ii) Using $K_{000} = \frac{1}{6}$ we have $0 = I_{0kk} = \frac{1}{6} I_k^2 + K_{00k} I_k + K_{0k0} I_k + K_{0kk}$ and taking expectations we find $\mathbb{E} K_{0kk} = -\frac{1}{6}$. Similarly for the other two. □

Construction of schemes for $m = 2$ and $m = 3$.

We now give details for these two cases.

For $m = 2$ the only α needed in (15) are $\alpha = (kl)$ with $k, l > 0$. Then the K_β needed for (13) are K_{kl} with $k, l \geq 0$, and the only moments required are $\mathbb{E} K_{kl}$ and by lemmas 5 and 6 the only non-zero moments are the deterministic ones $K_{00} = \frac{1}{2}$ and $K_{kk} = -\frac{1}{2}$ for $k > 0$. So we can take $L_{00} = \frac{1}{2}$, $L_{kk} = -\frac{1}{2}$ for $k > 0$ and all other $L_\beta = 0$. Then (13) gives $I_{kl} = h(K_{00} V_k V_l + K_{k0} V_l + K_{0l} V_k + K_{kl})$ for $k, l \geq 1$ and hence we take $\tilde{I}_{kl} = \frac{1}{2}(V_k V_l - \delta_{kl})$. Finally $f_{kl,i} = L^k b_{il} = \rho_{ikl}$ and we obtain scheme (10) which has order 1 under the conditions of the theorem.

For $m = 3$ the only α needed in (15) are all α of length 2 and α of length 3 with no zero indices. The β needed for (13) are all those with length 2 or 3, and the relevant moments are $\mathbb{E} K_\beta$ for all such β, and all $\mathbb{E}(K_\alpha K_\beta)$ where α and β have length two. From lemmas 7 and 6 we see that the K_{kl} for $0 \leq k < l$ have mean 0, variance $\frac{1}{12}$, and are uncorrelated; also $K_{lk} = -K_{kl}$ for $0 \leq k < l$. Together with $K_{00} = \frac{1}{2}$ and $K_{kk} = -\frac{1}{2}$ from lemma 5 these facts cover all K_β of length two.

Also from lemmas 6, 5 and 7 we see that the only β of length 3 with non-zero $\mathbb{E} K_\beta$ are $\mathbb{E} K_{000} = \frac{1}{6}$ and $\mathbb{E} K_{0kk} = \mathbb{E} K_{k0k} = \mathbb{E} K_{0kk} = -\frac{1}{6}$ for $k > 0$.

Hence we should generate L_{kl} for $0 \leq k < l$ to be uncorrelated with mean 0 and variance $\frac{1}{12}$, then set $L_{lk} = -L_{kl}$, and $L_{00} = \frac{1}{2}$ and $L_{kk} = -\frac{1}{2}$ for $k > 0$. And we set $L_{000} = \frac{1}{6}$, $L_{0kk} = L_{k0k} = L_{kk0} = -\frac{1}{6}$ for $k > 0$, and all other L_β of length 3 to 0. We have then $\tilde{I}_{00} = \frac{1}{2} h^2$, $\tilde{I}_{0k} = h^{3/2}(\frac{1}{2} V_k + L_{0k})$, $\tilde{I}_{k0} = h^{3/2}(\frac{1}{2} V_k - L_{0k})$, $\tilde{I}_{kl} = h(\frac{1}{2} V_k V_l + L_{0l} V_k - L_{0k} V_l + L_{kl})$ and $\tilde{I}_{nkl} = \frac{1}{6} h^{3/2}(V_n V_k V_l - \delta_{kl} V_n - \delta_{nl} V_k - \delta_{nk} V_l)$ for $k, l, n > 0$.

Then after calculating the relevant f_α, considering for simplicity the case where $a_i = 0$ and b_{ik} is independent of t, and writing $c_{ln} = \sum_{k=1}^d b_{lk} b_{nk}$, we obtain the order $\frac{3}{2}$ scheme which is summarised below:

$$\begin{aligned}
\hat{x}_i^{(j+1)} =& \hat{x}_i^{(j)} + h^{1/2} \sum_{k=1}^d b_{ik}(\hat{x}^{(j)}) V_k^{(j)} + h \sum_{k,l=1}^d \rho_{ikl}(\hat{x}^{(j)}) \left(\frac{1}{2} V_k^{(j)} V_l^{(j)} + L_{0l}^{(j)} V_k^{(j)} - L_{0k}^{(j)} V_l^{(j)} + L_{kl}^{(j)} \right) \\
&+ \frac{1}{2} h^{3/2} \sum_{k=1}^d \left\{ \left(\frac{1}{2} V_k^{(j)} + L_{0,k}^{(j)} \right) \sum_{l,n=1}^q c_{ln}(\hat{x}^{(j)}) \frac{\partial^2 b_{ik}}{\partial x_l \partial x_n}(\hat{x}^{(j)}) \right\} \\
&+ \frac{1}{6} h^{3/2} \sum_{k,l,n=1}^d \left\{ \left(V_n^{(j)} V_k^{(j)} V_l^{(j)} - \delta_{kl} V_n^{(j)} - \delta_{kn} V_l^{(j)} - \delta_{ln} V_k^{(j)} \right) \sum_{p=1}^q b_{pk}(\hat{x}^{(j)}) \frac{\partial \rho_{iln}}{\partial x_p}(\hat{x}^{(j)}) \right\}
\end{aligned}$$

where for each j the random variables $V_k^{(j)}$, $1 \leq k \leq d$ and $L_{kl}^{(j)}$, $0 \leq k < l \leq d$ are generated independently so that the $V_k^{(j)}$ are each $N(0,1)$ and the $L_{kl}^{(j)}$ have zero mean and variance $\frac{1}{12}$; we then set $L_{lk}^{(j)} = -L_{kl}^{(j)}$ for $k < l$, and also $L_{kk}^{(j)} = -\frac{1}{2}$ for $k > 0$.

A final remark for this section: we have concentrated on L^2 bounds for convergence, as they are the easiest to handle. However it is not hard to modify the proof of theorem 4 to give $\mathbb{W}_p$ bounds for $p > 2$, and then the proofs of the approximation bounds can be adapted using standard methods to obtain L^p bounds so that for example in our application of theorem 4 we could obtain $\mathbb{E}(\max_{j=1}^N |\hat{x}^{(j)} - x(jh)|^p) \leq C(p) h^{pm/2}$.

8 Two-level approximation using (10)

In section 7 we considered a fixed value of h. Now we describe an approach to the scheme (10) using a modification of the the Levy construction described in section 3. We still use a coupling, this time between the random variables $X_k^{(j)}$ in (10) at consecutive levels. For simplicity we assume that $q = d$, $a_i = 0$ and that b_{ik} depends only on x. We also assume again that the matrix (b_{ik}) has rank d, i.e. it is invertible.

The modification is to replace (9) by a different requirement. For stepsize $h^{(r)}$ we will have Nd independent random variables $X_k^{(r,j)}$, each with $N(0,h)$ distribution, and we apply (8) with $X_k^{(r,j)}$ replacing $\Delta W_k^{(r,j)}$, obtaining our approximate solution $\tilde{x}_i^{(r,j)}$ by the recurrence relation

$$\tilde{x}_i^{(r,j+1)} = \tilde{x}_i^{(r,j)} + \sum b_{ik}(\tilde{x}^{(r,j)})X_k^{(r,j)} + \sum \rho_{ikl}(\tilde{x}^{(r,j)})B_{kl}^{(r,j)} \tag{17}$$

where now $B_{kl}^{(r,j)} = \frac{1}{2}X_k^{(r,j)}X_l^{(r,j)}$ if $k \neq l$ and $B_{kk}^{(r,j)} = \frac{1}{2}\{(X_k^{(r,j)})^2 - h^{(r)}\}$. The idea is that, conditional on the $X_k^{(r,j)}$ for a fixed j, we generate $X_k^{(r+1,2j)}$ and $X_k^{(r+1,2j+1)}$ so that they are independent of each other and each $N(0,h^{(r+1)})$, but do so in such a way that $\mathbb{E}|\tilde{x}_i^{r+1,2j} - \tilde{x}_i^{r,j}|^2 \leq C(h^{(r)})^2$ for some constant C independent of r.

The analogue of (9) will no longer hold exactly, but it will hold approximately, with the result that, if $m \in \mathbb{N}$ and $0 \leq j \leq 2^m$, then $\sum_{i=0}^{j2^n-1} X_k^{(m+n,i)}$ converges to a limit as $n \to \infty$, and that limit is $W_k(j2^{-m}T)$ for a limiting Brownian path W. Then we will find that the 'approximate solutions' $\tilde{x}^{(r)}$ will converge with order 1 to the solution of (1) for this path W.

We now indicate how to make the transition from r to $r+1$. To achieve order 1 convergence we need a global error $O(h)$ hence a local error $O(h^{3/2})$. In the following discussion, when we say a random variable X is $O(h^\alpha)$, we mean $\mathbb{E}(X^2) = O(h^{2\alpha})$. The main point is to achieve the local error $O(h^{3/2})$, and to describe how this is done we look at the first step, i.e. we compare $\tilde{x}_k^{(r,1)}$ with $\tilde{x}_k^{(r+1,2)}$. We have

$$\tilde{x}_i^{(r,1)} = x_i^{(0)} + \sum_{k=1}^d b_{ik}(x^{(0)})X_k^{(r,0)} + \frac{1}{2}\sum_{k,l=1}^d \rho_{ikl}(x^{(0)})(X_k^{(r,0)}X_l^{(r,0)} - h^{(r)}\delta_{kl}) \tag{18}$$

and also

$$\tilde{x}_i^{(r+1,1)} = x_i^{(0)} + \sum_{k=1}^d b_{ik}(x^{(0)})X_k^{(r+1,0)} + \frac{1}{2}\sum_{k,l=1}^d \rho_{ikl}(x^{(0)})(X_k^{(r+1,0)}X_l^{(r+1,0)} - h^{(r+1)}\delta_{kl}), \tag{19}$$

$$\tilde{x}_i^{(r+1,2)} = \tilde{x}_i^{(r+1,1)} + \sum_{k=1}^d b_{ik}(\tilde{x}^{(r+1,1)})X_k^{(r+1,1)} + \frac{1}{2}\sum_{k,l=1}^d \rho_{ikl}(\tilde{x}^{(r+1,1)})(X_k^{(r+1,1)}X_l^{(r+1,1)} - h^{(r+1)}\delta_{kl}) \tag{20}$$

Now $b_{ik}(\tilde{x}^{(r+1,1)}) = b_{ik}(x^{(0)}) + \sum_{l,n=1}^d \frac{\partial b_{ik}}{\partial x^l}(x^{(0)})b_{ln}(x^{(0)})X_n^{(r+1,0)} + O(h)$ and $\rho_{ikl}(\tilde{x}^{(r+1,1)}) = \rho_{ikl}(x^{(0)}) + O(h^{1/2})$. Using these relations in (20) and combining it with (19) one finds that

$$\begin{aligned}\tilde{x}_i^{(r+1,2)} =& x_i^{(0)} + \sum_{k=1}^d b_{ik}(x^{(0)})\left(X_k^{(r+1,0)} + X_k^{(r+1,1)}\right) + \sum_{k,l=1}^d \rho_{ikl}(x^{(0)})X_k^{(r+1,1)}X_l^{(r+1,0)} \\ &+ \frac{1}{2}\sum_{k,l=1}^d \rho_{ikl}(x^{(0)})\left(X_k^{(r+1,0)}X_l^{(r+1,0)} + X_k^{(r+1,1)}X_l^{(r+1,1)} - h^{(r)}\delta_{kl}\right) + O((h^{(r)})^{3/2})\end{aligned} \tag{21}$$

Now let (c_{ij}) be the matrix inverse of $(b_{ik}(x^{(0)}))$, so that $\sum_j c_{ij}b_{jk}(x^{(0)}) = \delta_{ik}$. Then from (18) and (21) we find that in order to obtain $\tilde{x}^{(r,1)} - \tilde{x}^{(r+1,2)} = O((h^{(r)})^{3/2})$ we require

$$X_i^{(r,0)} = X_i^{(r+1,0)} + X_i^{(r+1,1)} + \sum_{k,l=1}^{d} \tau_{ikl}\left(X_k^{(r+1,1)}X_l^{(r+1,0)} - X_l^{(r+1,1)}X_k^{(r+1,0)}\right) + O((h^{(r)})^{3/2}) \quad (22)$$

where $\tau_{ikl} = \frac{1}{2}\sum_j c_{ij}\rho_{jkl}$.

It is convenient to reformulate (22) by a scaling. We fix r and write $\epsilon = (h^{(r)})^{1/2}$, $X_i^{(r,0)} = \epsilon V_i$, $X_i^{(r+1,0)} = \epsilon Y_i$ and $X_i^{(r+1,1)} = \epsilon Z_i$. Then $V_1, \cdots, V_d$ are independent and $N(0,1)$, while $(Y_1, \cdots, Y_d, Z_1, \cdots, Z_d)$ are independent and $N(0, \frac{1}{2})$. We seek a coupling between (V_i) and (Y_i, Z_i) so that $V_i = Y_i + Z_i + \epsilon\sum_{k,l=1}^{d}\tau_{ikl}(Z_kY_l - Z_lY_k) + O(\epsilon^2)$. A slight simplification is obtained by writing $U_i = Y_i + Z_i$ and $U_i^* = Y_i - Z_i$, so that the U_i and U_i^* should all be independent and $N(0,1)$, and we require

$$V_i = U_i + \frac{\epsilon}{2}\sum_{k,l=1}^{d}\tau_{ikl}(U_l^*U_k - U_k^*U_l) + O(\epsilon^2) \quad (23)$$

To summarise: we require a coupling between $(V_1, \cdots, V_d)$ and $(U_1, \cdots, U_d, U_1^*, \cdots U_d^*)$, so that all these random variables are $N(0,1)$, $(V_1, \cdots, V_d)$ are mutually independent, $(U_1, \cdots, U_d, U_1^*, \cdots U_d^*)$ are also mutually independent, and (23) holds. The existence of such a coupling follows from Theorem 4 but for simultaneous simulation of approximations at levels r and $r+1$ we need a coupling that can be implemented efficiently. We now describe such a coupling for the case $d = 2$, and then give a method for generating approximate couplings for general d.

Exact coupling in two-dimensional case.

When $d = 2$, (23) reduces to $V_i = U_i + \epsilon a_i(U_2^*U_1 - U_1^*U_2) + O(\epsilon^2)$ where $a_i = \tau_{i12} - \tau_{i21}$. Then we can write $\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = R_\theta \begin{pmatrix} a \\ 0 \end{pmatrix}$ where R_θ is a rotation matrix and $a = (a_1^2 + a_2^2)^{1/2}$. Writing $V = R_\theta V'$, $U = R_\theta U'$, and $U^* = R_\theta \tilde{U}$ our required condition becomes

$$V_1' = U_1' + \epsilon a(\tilde{U}_2U_1' - \tilde{U}_1U_2') + O(\epsilon^2), \qquad V_2' = U_2' + O(\epsilon^2) \quad (24)$$

Now consider the following: suppose U and α are independent random variables, U being $N(0,1)$ and α taking the values ± 1 each with probability $\frac{1}{2}$, and let b and c be fixed real numbers. If $|b| \le \frac{1}{2}$ and $|c| \le \frac{1}{2}$ then we define $\mathcal{Y} = U + \alpha(bU + c)$ and $V = \Phi^{-1}(F(\mathcal{Y}))$ where $F(y) = \frac{1}{2}\{\Phi(\frac{y-c}{1+b}) + \Phi(\frac{y+c}{1-b})\}$ is the c.d.f. of $\mathcal{Y}$ (here Φ is the c.d.f. of $N(0,1)$); then V is $N(0,1)$. Otherwise we generate V independently to be $N(0,1)$.

Claim: $\mathbb{E}(V - \mathcal{Y})^2 \le K(b^2 + c^2)^2$ where K is a constant independent of b and c.

Proof of claim (outline). First note that $\mathbb{E}(V - \mathcal{Y})^2 \le 2\mathbb{E}(V^2 + \mathcal{Y}^2) = 2(2 + b^2 + c^2)$ so it suffices to prove the claim for $|b| \le \frac{1}{2}$, $|c| \le \frac{1}{2}$. Using the expression for F, we find that for $|y| \le (b^2 + c^2)^{-1/4}$ we have $|F(y) - \Phi(y)| \le K_1(b^2 + c^2)(1 + y^2)\phi(y)$ and so $|y - \Phi^{-1}(F(y))| \le K_2(b^2 + c^2)(1 + y^2)$, where $\phi = \Phi'$ is the $N(0,1)$ density function, and then the claim follows since $\mathbb{E}(V - \mathcal{Y})^2 = \mathbb{E}(\mathcal{Y} - \Phi^{-1}(F(\mathcal{Y})))^2$ and the contribution from $|\mathcal{Y}| > (b^2 + c^2)^{-1/4}$ is negligible if $b^2 + c^2$ is small.

We return to (24), and recall that we require to generate the six random variables $V_1', V_2', U_1', U_2', \tilde{U}_1, \tilde{U}_2$ so that each is $N(0,1)$ and that V_1', V_2' are independent and that $U_1', U_2', \tilde{U}_1, \tilde{U}_2$ are also mutually independent. We also require these two sets of random variables are coupled so that (24) holds. We describe two ways of doing this. The first, which is slightly simpler, generates all six variables together, and is suitable for two-level simulation.

For the second we generate V_1' and V_2' first (or assume they are given) and then generate the other 4 conditionally on V_1', V_2'. This is more useful if we wish to simulate several levels sequentially, going from level r to level $r+1$.

For the first method we start by generating independent $N(0,1)$ variables U_1', U_2', Q, R and α taking value ± 1 with probability $\frac{1}{2}$ each. Then set $V_2' = U_2'$, $\tilde{U}_1 = \alpha Q$ and $\tilde{U}_2 = \alpha R$. We also define $\mathcal{Y} = U_1' + \alpha(bU_1' + c)$ and $V_1' = \Phi^{-1}(F(\mathcal{Y}))$ as above with $b = \epsilon a R$ and $c = -\epsilon a Q U_2'$. This gives $\mathcal{Y} = U_1' + \epsilon a(\tilde{U}_2 U_1' - \tilde{U}_1 U_2')$. Also, conditional on Q, R, U_2' we see that V_1' is $N(0,1)$, so V_1' is independent of V_2' and all six variables have $N(0,1)$ distribution. Finally from the 'claim' above we have

$$\mathbb{E}(V_1' - \mathcal{Y})^2 \le K a^4 \epsilon^4 \mathbb{E}(R^2 + Q^2 (U_2')^2)^2 = 14 K a^4 \epsilon^4$$

and (24) follows.

The second method is essentially the same as the first, but done conditionally on V_1', V_2'. First we note that, if $\mathcal{Y} = U + \alpha(bU + c)$ as above, then conditional on $\mathcal{Y} = y$, we have that the pair (α, U) takes the values $(1, \frac{y-c}{1+b})$ with probability p and $(-1, \frac{y+c}{1-b})$ with probability $1-p$, where $p = \frac{(1-b)\phi(\frac{y-c}{1+b})}{(1-b)\phi(\frac{y-c}{1+b}) + (1+b)\phi(\frac{y+c}{1-b})}$ (here $\phi = \Phi'$ is the standard normal density).

Next, given V_1' and V_2', we generate independent $N(0,1)$ variables Q and R as before, then apply the procedure in the previous paragraph with $\mathcal{Y} = F^{-1}(\Phi(V_1))$, $b = \epsilon a R$ and $c = -\epsilon a Q V_2'$ to obtain U_1 $(= U)$ and α. Finally we set $U_2' = V_2'$, $\tilde{U}_1 = \alpha Q$ and $\tilde{U}_2 = \alpha R$. The six random variables so generated have the same joint distribution as given by the first method, so (24) still holds.

We remark that, as pointed out by M. B. Giles, for the implementation of either method one can avoid the use of α, generating $\tilde{U}_1$ and $\tilde{U}_2$ as independent $N(0,1)$ directly and using them in place of Q and R in the definition of b and c, the point being that F is unaltered if one replaces (b, c) by $(-b, -c)$.

Using either of the above methods we can achieve (22) and hence $\tilde{x}_k^{(r,1)} - \tilde{x}_k^{(r+1,2)} = O((h^{(r)})^{3/2})$ in the $d = 2$ case, provided the matrix function b_{ik} has a bounded inverse. We apply the same argument to each interval of length $h^{(r)}$. There is a complication in that the initial data for the two levels are no longer the same - when we go from $\tilde{x}^{(r,j)}$ and $\tilde{x}^{(r+1,2j)}$ to $\tilde{x}^{(r,j+1)}$ and $\tilde{x}^{(r+1,2j+2)}$, we follow the same procedure as above, taking b_{ik} and ρ_{ikl} evaluated at $\tilde{x}^{(r,j)}$. There is an extra error term arising from the fact that $\tilde{x}^{(r,j)} \ne \tilde{x}^{(r+1,2j)}$ in general. We can control this using the standard procedure for getting global bounds from local bounds, and we obtain $\tilde{x}_k^{(r,j)} - \tilde{x}_k^{(r+1,2j)} = O(h^{(r)})$. Then we get a sequence of approximations which converge with order 1 as $r \to \infty$.

One also finds from (22) that $X^{(r,0)} = X^{(r+1,0)} + X^{(r+1,1)} + O(h^{(r)})$, and similarly $X^{(r,j)} = X^{(r,2j)} + X^{(r,2j+1)}$, and then we can deduce that the 'approximate Brownian path' defined by the $X_k^{(r,j)}$ converges with order $\frac{1}{2}$ to a limiting Brownian path $W(t)$, in the sense that $\mathbb{E}|\sum_{l=1}^{j} X^{(r,l)} - W(jh^{(r)})|^2 \le C 2^{-r}$. Then the approximate solutions $\tilde{x}_i^{(r,j)}$ will converge (with order 1) to the solution $x(t)$ of (1) for the path W, in the sense that

$$\mathbb{E}(\max_{j=1}^{2^r} |\tilde{x}^{(r,j} - x(jh^{(r)})|^2) = O(2^{-2r}) \tag{25}$$

We remark that (25) implies that for any γ with $0 < \gamma < 1$, the approximate solution almost surely converges uniformly to $x(t)$, in the sense that $\max_{j=1}^{2^r} |\tilde{x}^{(r,j} - x(jh^{(r)})| = O(2^{-\gamma r})$ a.s.

Approximate coupling for general d

It is not so obvious how to extend the above exact coupling method to general d. Here we describe a construction which satisfies (23) with U, U^* having the required distribution but V being only approximately $N(0,1)$. We are somewhat less precise here with error bounds but the estimates can easily be made rigorous. We start with a lemma.

Lemma 8. *Let $U = (U_1, \cdots, U_d)$ be a random vector with $N(0, I)$ distribution and let A be a fixed $d \times d$ matrix. Let $Y = U + \epsilon AU$. Then the density function of Y satisfies*

$$f_Y(y) = (2\pi)^{-d/2} e^{-|y|^2/2} \{1 + \epsilon(y^t Ay - \mathrm{tr}A) + \epsilon^2 \Omega\} + O(\epsilon^3) \tag{26}$$

where $\Omega = -(\mathrm{tr}A) y^t Ay - y^t A^2 y - \frac{1}{2}|Ay|^2 + \frac{1}{2}(y^t Ay)^2 + \frac{1}{2}(\mathrm{tr}A)^2 + \frac{1}{2}\mathrm{tr}(A^2)$.

Proof. Writing $y = (I + \epsilon A)u$, we have

$$f_Y(y) = \det(I + \epsilon A)^{-1} f_U(u) = (2\pi)^{-d/2} \det(I + \epsilon A)^{-1} e^{-|u|^2/2}$$

Now $\mathrm{tr}\log(I + \epsilon A) = \epsilon \mathrm{tr}A - \frac{\epsilon^2}{2}\mathrm{tr}(A^2) + O(\epsilon^3)$ so

$$\det(I + \epsilon A)^{-1} = \exp\left(-\epsilon \mathrm{tr}A + \frac{\epsilon^2}{2}\mathrm{tr}(A^2)\right) + O(\epsilon^3) = 1 - \epsilon \mathrm{tr}A + \frac{\epsilon^2}{2}\{(\mathrm{tr}A)^2 + \mathrm{tr}(A^2)\} + O(\epsilon^3) \tag{27}$$

And a simple calculation gives $|u|^2 = |y|^2 - 2\epsilon y^t Ay + \epsilon^2(2y^t A^2 y + |Ay|^2) + O(\epsilon^3)$ so that

$$e^{-|u|^2/2} = e^{-|y|^2/2}\left\{1 + \epsilon y^t Ay - \epsilon^2\left(y^t A^2 y + \frac{1}{2}|Ay|^2\right) + \frac{\epsilon^2}{2}(y^t Ay)^2\right\} + O(\epsilon^3) \tag{28}$$

Combining (27) and (28) gives the result. □

We now apply this to (23). Suppose U and U^* are as in (23) and define $Y_i = U_i + \frac{\epsilon}{2}\Sigma_{k,l=1}^d \tau_{jkl}(U_l^* U_k - U_k^* U_l)$ which we can rewrite as $Y_i = U_i + \epsilon \Sigma_{k,l=1}^d \sigma_{ikl} U_k U_l^*$ where $\sigma_{ikl} = \frac{1}{2}(\tau_{ikl} - \tau_{ilk})$. We wish to find V such that $V - U = O(\epsilon^2)$ and V is close to $N(0, I)$. To do this we first apply lemma 8 to approximate the density function of Y.

We write $Y = U + \epsilon AU$ where the matrix $A = (a_{ik})$ is given by $a_{ik} = \sum_{l=1}^d \sigma_{ikl} U_l^*$. Then by lemma 8 the density of Y, conditional on U^*, is given by (26). Substituting for A in (26) and taking expectation w.r.t. U^*, we find after some calculation that the unconditional density of Y satisfies

$$(2\pi)^{-d/2} e^{-|y|^2/2}\left\{1 + \epsilon^2\left(\mathcal{K} - \sum_{i,k=1}^d \theta_{ik} y_i y_k + \sum_{i,j,k,m=1}^d \Psi_{ijkm} y_i y_j y_k y_m\right)\right\} + O(\epsilon^4) \tag{29}$$

where $\mathcal{K} = \frac{1}{2}\sum_{i,k,l=1}^d (\sigma_{ikl}\sigma_{kil} + \sigma_{iil}\sigma_{kkl})$, $\theta_{ik} = \sum_{l,m=1}^d (\sigma_{iml}\sigma_{mkl} + \sigma_{ikl}\sigma_{mml} + \frac{1}{2}\sigma_{mil}\sigma_{mkl})$ and $\Psi_{ijkm} = \frac{1}{2}\sum_{l=1}^d \sigma_{ikl}\sigma_{jml}$. Note that the density is invariant under $\epsilon \to -\epsilon$, so there is no ϵ^3 term.

We now try to modify Y by adding a correction term to make the distribution closer to $N(0, I)$. We consider an $\mathbb{R}^d$-valued random variable V given by

$$V_i = U_i + \epsilon \Sigma_{k,l=1}^d \sigma_{ikl} U_k U_l^* + \epsilon^2 p_i(U) \tag{30}$$

where p is an $\mathbb{R}^d$-valued polynomial on $\mathbb{R}^d$, to be chosen. We find that the density function f_V satisfies

$$f_V(y) = (2\pi)^{-d/2} e^{-|y|^2/2}\{1 + \epsilon^2(\mathcal{K} - \theta(y) + \Psi(y) + y.p(y) - \nabla.p(y))\} + O(\epsilon^3)$$

where $\theta(y) = \sum_{i,k=1}^d \theta_{ik} y_i y_k$ and $\Psi(y) = \sum_{i,j,k,m=1}^d \Psi_{ijkm} y_i y_j y_k y_m$. According to lemma 1, we can choose the polynomial p so that $\nabla.p(y) - y.p(y) = \Psi(y) - \theta(y) + \mu$ where μ is a constant; from the fact that f_V is a density we see that $\mu = \mathcal{K}$. If we follow the procedure described in the proof of lemma 1 we find in fact that $p = \frac{1}{8}\nabla(4\theta - \nabla^2\Psi - 2\Psi)$ has this property. We see that then f_V is an even function of ϵ, so there is no ϵ^3 term in its expansion, and the ϵ^2 term is 0, hence $f_V(y) = (2\pi)^{-d/2} e^{-|y|^2/2} + O(\epsilon^4)$. We also see that this V satisfies (23).

Writing $\phi(y) = (2\pi)^{-d/2}e^{-|y|^2/2}$, the standard normal density in $\mathbb{R}^d$, we can expand the density f_V further as $f_V(y) = \phi(y)(1+q(y)) + O(\epsilon^8)$ where q is a polynomial in ϵ and y with $q = O(\epsilon^4)$, and then lemma 2 shows that the $\mathbb{W}_2$ distance of V to $N(0, I)$ is $O(\epsilon^4)$. This means that we can find a random variable $\tilde{V}$ with $N(0, I)$ distribution, which is coupled to V so that $\mathbb{E}|V - \tilde{V}|^2 = O(\epsilon^8)$.

In the same way as with the exact coupling for $d = 2$ we can use the above construction of V, U, U^* to generate coupled approximate solutions $\tilde{x}^{(r,j)}$ and $\tilde{x}^{(r+1,j)}$ at levels r and $r+1$. Note that $\tilde{x}^{(r,j)}$ is not a true implementation of (17) since V does not have $N(0, I)$ distribution exactly. But if we replace V by $\tilde{V}$ we get a solution which we denote $\overline{x}^{(r,j)}$ which is a true implementation of (17), but we do not have a means of generating it jointly with the level $r+1$ solution, which is why we use $\tilde{x}^{(r,j)}$ as an approximation. In the same way as before we have the bound $\tilde{x}^{(r,j)} - \tilde{x}^{(r+1,2j)} = O(h)$, and from the above bound for $V - \tilde{V}$ we get the bound $\overline{x}^{(r,j)} - \tilde{x}^{(r,j)} = O(h^2)$. We see next how these methods can be applied.

Applications of two-level coupling.

The application of the exact two-dimensional method to empirical error estimation or to the multilevel method, as described in section 3, does not present any difficulty, provided the invertibility assumption on the matrix (b_{ik}) is satisfied. So we concentrate on the above method for general d, as the effect of its approximate nature requires attention.

We start with empirical estimation of the error of scheme (17). Following the description in section 3, but using the notation of the last but one paragraph, we require to estimate $\mathbb{E}|\overline{x}^{(r,N)} - \tilde{x}^{(r+1,2N)}|$, where N is the total number of steps at level r. But, as indicated in the discussion in that paragraph, what we are able to estimate empirically, using the approximate coupling, is $\mathbb{E}|\tilde{x}^{r,N} - \tilde{x}^{r+1,2N}|$. The difference between these two quantities is $O(h^2)$, by our results above. As we expect $\mathbb{E}|\overline{x}^{(r,N)} - \tilde{x}^{(r+1,2N)}|$ to be of order h, this difference should be negligible for small h, so we conclude that the approximate method is effective for empirical estimation.

For the multilevel method using (17) things are rather different. Using the notation of section 3, we wish to estimate $\mu_r - \mu_{r+1}$ using the approximate coupling between levels r and $r+1$. This means using $\mathbb{E}f(\tilde{x}^{(r,N)})$ to estimate μ_r, and this introduces an error as it is using only an approximate implementation of (17). Our analysis above of the approximate coupling shows that this error is $O((h^{(r)})^4)$. The problem is that for the multilevel method to be efficient the initial $h^{(0)}$ should not be too small, so that the errors for these initial intervals will be too big. So the approximate coupling does not seem to be useful for the multilevel method.

For this reason it would be desirable to find easily implemented exact couplings for general d. We now outline a possible method of doing this, using the method described at the end of section 5.

An approach to exact coupling for general d.

The idea is to start from the approximate coupling and use the coupling method described at the end of section 5 to bridge the gap to an exact coupling. We start with V as given by (30) and seek to apply this coupling method with $f = f_V$ and g being the standard normal density on $\mathbb{R}^d$. From the estimates for f_V (slightly tightened up) we can get $\int_{\mathbb{R}^d} |y|^2 |f_V(y) - g(y)| dy = O(\epsilon^4)$ and then provided our coupling achieves (11) we obtain $\mathbb{E}|V - V^*|^2 = O(\epsilon^4)$, where V^* is standard normal.

In order to implement the coupling procedure we need to calculate f_V, or at least have a sequence of calculable approximations converging to f_V. This may well be possible but we have not found a way of doing it. So instead we suggest a modification which we now describe.

As before we write $Y = U + \epsilon AU$ where A is given by $a_{ik} = \sum_{l=1}^d \sigma_{ikl} U_l^*$, and apply lemma 8. But we do not take expectation w.r.t. U^*. Instead, in a similar fashion to the

$d = 2$ construction, we write $U^* = \alpha R$ where α and R are independent, with α taking values ± 1 with probability $\frac{1}{2}$ each, and R being $\mathbb{R}^d$-valued with $N(0, I)$ distribution. Then we fix R and take expectation w.r.t. α. Then, conditional on R, we find that the density of Y satisfies (29), where now we have

$$\theta_{ik} = \sum_{l,m=1}^{d} (\sigma_{iml}\sigma_{mkn} + \sigma_{ikl}\sigma_{mmn} + \frac{1}{2}\sigma_{mil}\sigma_{mkn})R_l R_n,$$

$$\mathcal{K} = \frac{1}{2}\sum_{i,k,l,n=1}^{d} (\sigma_{ikl}\sigma_{kin} + \sigma_{iil}\sigma_{kkn})R_l R_n \quad \text{and} \quad \Psi_{ijkm} = \frac{1}{2}\sum_{l,n=1}^{d} \sigma_{ikl}\sigma_{jmn}R_l R_n$$

Note that R_l, R_n could be replaced by U_l^*, U_n^* in these expressions.

We then define V using (30), where the polynomial p is defined as before, but using the new definitions of θ and Ψ. Then p is an $\mathbb{R}^d$-valued cubic polynomial whose coefficients depend on R. Again we find that the density of V (conditional on R) satisfies $f_V(y) = (2\pi)^{-d/2}e^{-|y|^2/2} + O(\epsilon^{-4})$. Then we can apply the coupling method from section 5 provided we have a means of calculating the (conditional) density $f_V(y)$, or an iterative sequence of approximations, for given y. This can be approached using a similar method to that used previously for the case $d = 2$. One needs to solve the system of cubic equations $y_i = u_i + \epsilon\alpha\sum_{k,l=1}^{d}\sigma_{ikl}u_k R + \epsilon^2 p_i(u)$, $i = 1, \cdots, d$, for u given R, α, y. The natural iterative method $u_i^{(n+1)} = y_i - \epsilon\alpha\sum_{k,l=1}^{d}\sigma_{ikl}u_k^{(n)}R_l - \epsilon^2 p_i(u^{(n)})$ will converge fast enough provided ϵ is small and $|y|$ not too large. Some truncation will be needed to exclude large y and also to exclude any contribution from solutions to the equations other than the one given by the iteration.

We remark that, as in the two-dimensional case, one can avoid the generation of α and generate U^* directly.

9 Coupling applied to the Fourier method

When the diffusion coefficient matrix (b_{ik}) has rank $< q$ the theory developed in sections 7 and 8 is not applicable. The Milstein method (4) can be used provided we have a means of generated the double integrals of the Brownian path. An approach to the latter is given by Kloeden, Platen and Wright [17], using a Fourier series expansion of the Brownian bridge process. In this section we describe how coupling can be used to improve this method. We start by outlining the method of [17], as described in section 5.8 of [16].

We consider the simulation of double integrals $A_{jk} = \int_0^1 W_j(t)dW_k(t)$ of a d-dimensional Brownian motion, jointly with the Brownian increment $W_j(1)$. As shown in section 5.8 of [16], we can write

$$A_{jk} = \frac{1}{2}(W_j(1)W_k(1) - \delta_{jk}) + \frac{1}{\pi\sqrt{2}}(z_k W_j(1) - z_j W_k(1)) + \frac{1}{2\pi}\lambda_{jk}$$

where $\lambda_{jk} = \sum_{r=1}^{\infty} r^{-1}(x_{jr}y_{kr} - x_{kr}y_{jr})$ and $z_j = \sum_{r=1}^{\infty} r^{-1}x_{jr}$, where x_{jr} and y_{jr} are independent $N(0,1)$ random variables, for $r \in \mathbb{N}$ and $j = 1, \cdots, d$, all independent of the $W_j(1)$. The idea is to approximate λ_{jk} and z_j by the truncated sums $\lambda_{jk}^{(n)} = \sum_{r=1}^{n} r^{-1}(x_{jr}y_{kr} - x_{kr}y_{jr})$ and $z_j^{(n)} = \sum_{r=1}^{n} r^{-1}x_{jr}$. We write $\tilde{\lambda}^{(n)} = \lambda - \lambda^{(n)}$ and $\tilde{z}^{(n)} = z - z^{(n)}$ and use the notation $v = (\lambda, z)$, $v^{(n)} = (\lambda^{(n)}, z^{(n)})$ and $\tilde{v}^{(n)} = (\tilde{\lambda}^{(n)}, \tilde{z}^{(n)})$.

We will show that coupling can be used to improve the accuracy attained in this way for a given n. As pointed out by Wiktorsson [28], if one uses the error estimate in [16], to get sufficient accuracy to obtain an $O(h)$ error using Milstein one needs n to be of order

h^{-1}, the same order as the number of steps, and then the overall computational load is of order h^{-2} which is the same as using Euler with stepsize of order h^2 which gives the same order of error. Wiktorsson shows that the addition of an extra term to the truncated series gives an improved order and a genuine improvement on Euler, so that one can obtain an average error bounded by $\mathcal{E}$ with load $O(\mathcal{E}^{-3/2})$ (as opposed to $O(\mathcal{E}^{-2})$ for Euler). Here we show that further improvements can be attained using coupling, and indeed that a method computationally equivalent to the Wiktorsson method can give error $\mathcal{E}$ for Milstein with load $O(\mathcal{E}^{-5/4})$, and potentially can give further improvements.

Our approach is to apply the perturbation method of section 7 with X being the truncated Fourier sum $(\lambda_{jk}^{(n)}), (z_j^{(n)})$ and ϵY being the sum of the remainder of the series (with $\epsilon = n^{-1/2}$). In one sense, namely that X and Y are independent, this application of the perturbation method is simpler than that of section 7, and essentially follows the description in section 6. On the other hand the distribution of X is much more complicated in the present case than the normal distribution of section 7, and this leads to some technical difficulties.

The main result is Theorem 15. Several technical lemmas, mainly concerning the properties of the density of $(\lambda_{jk}^{(n)}), (z_j^{(n)})$ are needed order to get the Vaserstein distance estimate. A discussion of the application of the result to the Milstein method is given at the end.

We start with two general lemmas. The first gives a standard type of interpolation result for derivative bounds; we give a proof for completeness.

Lemma 9. *Given m and d in $\mathbb{N}$, we can find $C > 0$ such that if B is an open ball of radius $r > 0$ in $\mathbb{R}^d$ and f has bounded continuous derivatives up to order m on B then*

$$|f|_k \leq C \max\{(|f|_0^{m-k}|f|_m^k)^{1/m}, |f|_0 r^{-k}\} \tag{31}$$

for $k = 0, 1, \cdots, m$, where $|f|_k = \sup_{|\beta|=k} \sup_{x\in B} |D^\beta f(x)|$.

Proof. We consider first the case $d = 1$, so that B is an interval. We start with the observation that if $|f^{(j)}| \geq K > 0$ on some interval in B of length l, then $|f^{(j-1)}| \geq \frac{KL}{4}$ on a subinterval of length at least $\frac{l}{4}$, and repeating we find that $|f| \geq 4^{-j(j+1)/2} K l^j$ on some interval in B, so that $|f|_0 \geq 4^{-j(j+1)/2} K l^j$.

Now write $P = |f|_0$ and let $k \in \{1, 2, \cdots, m\}$. Let $x \in B$ and let $\alpha = |f^{(k-1)}(x)|$. Then $|f^{(k-1)}| \geq \alpha/2$ on an interval in B of length at least $\min(\frac{\alpha}{2|f|_k}, r/2)$. Then the observation above gives $P \geq C_1 \alpha \min\{(\alpha/|f|_k)^{k-1}, r^{k-1}\}$. From this it follows that $|f^{(k-1)}(x)| = \alpha \leq C_1^{-1} \max\{(P|f|_k^{k-1})^{1/k}, P r^{1-k}\}$. We can write this inequality as $J_{k-1} \leq C_1^{-1} J_k^{(k-1)/k}$ where $J_k = \max\{P^{-1}|f|_k, r^{-k}\}$. It follows then that $J_k \leq C J_m^{(m-k)/m}$ for $k = 1, 2, \cdots, m$, which implies (31) in this case.

We can now apply this to prove the case of general d by expressing (mixed) derivatives as linear combinations of derivatives along lines. To see this, let $k \in \{1, 2, \cdots, m\}$ and note that we can fine $v_1, \cdots, v_N$ in the unit sphere of $\mathbb{R}^d$ centred at 0, such that a homogeneous polynomial of degree k on $\mathbb{R}^d$ is determined by its values at $v_1, \cdots, v_N$; here $N = \binom{k+d-1}{k}$, the dimension of the space of homogeneous polynomials for degree k. This means that any derivative D^β with $|\beta| = k$ is a linear combination of kth order line derivatives in the directions v_j. Also note that there is $a > 0$ such that for any $j \in \{1, \cdots, N\}$ and any $x = (x_1, 0, \cdots, 0) \in \mathbb{R}^d$ with $-1 < x_1 \leq 0$, the line segment $x + t v_j,\ 0 \leq t \leq a$ is contained in the unit ball. Scaling by r and translating to our ball B, we then see that for any $x \in B$ there is an orthogonal matrix R such that for each j the line segment $x + tRv_j,\ 0 \leq t \leq ra$ is contained in B. We can then bound $D^\beta f(x)$ by applying the $d = 1$ result to the restrictions of f to these line segments. □

Lemma 10. *Let $n \geq k$ be positive integers. Then there is a constant $C > 0$ such that, if $\psi : \mathbb{R}^n \to \mathbb{R}^k$ is a C^2 map with $\|D^2\psi\| \leq L$ everywhere, if $x_0 \in \mathbb{R}^n$ and B is an $n \times k$*

matrix such that $D\psi(x_0)B = I_k$, then for any $r < r_0 = 1/(2KL)$ the image under ψ of Lebesgue measure on $B(x_0, r)$ has density at least $C(r/K)^{n-k}\|D\psi(x_0)\|^{-k}$ at x_0, where $K = \max(1, \|B\|)$.

Proof. We may suppose $x_0 = 0$ and $\psi(0) = 0$. Let $A = D\psi(0)$ and let $E \subseteq \mathbb{R}^n$ be the null space of A. Define $F : \mathbb{R}^n \to \mathbb{R}^k \times E$ by $F(x) = (\psi(x), P_E x)$ where P_E is the orthogonal projection on E. Write $M = DF(0)$ so that $Mx = (Ax, P_E x)$.

Now if $x \in \mathbb{R}^n$ then $Ax = ABAx$ so $x - BAx \in E$ and then $|(I - P_E)x| = |(I - P_E)BAx| \leq \|B\||Ax|$ and it follows that $|x| \leq K|Mx|$ so $\|M^{-1}\| \leq K$.

Next define $G(x) = M^{-1}F(x)$. Then $DG(0) = I$ and if $|x| < r_0$ then $\|DG(x) - I\| \leq K\|DF(x) - M\| \leq KL|x| < \frac{1}{2}$. So if $0 < r < r_0$ then G maps $B(0, r)$ diffeomorphically onto a domain containing $B(0, r/2)$, and then F maps $B(0, r)$ diffeomorphically onto a domain containing $B(0, \frac{r}{2K})$. Now if μ denotes Lebesgue measure on $B(0, r)$ in $\mathbb{R}^n$, and $x \in B(0, r)$, then the density of $F(\mu)$ at $F(x)$ is $|\det(DF(x))|^{-1} \geq C_1|\det(M)|^{-1} \geq C_1\|D\psi(0)\|^{-k}$ and the result follows by projecting onto the first component. □

Now we return to the truncated random variable $v^{(n)}$ defined at the start of this section. We let f_n be the density function of $v^{(n)}$; we can treat f_n as a function on $\mathbb{R}^{d(d-1)/2} \times \mathbb{R}^d = \mathbb{R}^{d(d+1)/2}$. Then we have the following:

Lemma 11. *(1) Let $N \in \mathbb{N}$ and suppose $n > N + d(d+1)/2$. Then f_n has continuous bounded derivatives up to order N on $\mathbb{R}^{d(d+1)/2}$, with bounds independent of n.*
(2) For given d there are positive constants C_1 and C_2 such that $\mathbb{E}e^{C_1|v^{(p)}|} \leq C_2$ and $\mathbb{E}e^{C_1 p^{1/2}|\bar{v}^{(p)}|} \leq C_2$.

Proof. (1) Let $\hat{f}_n$ denote the Fourier transform of f_n, which is the same as the characteristic function of $(\lambda^{(n)}, z^{(n)})$. We have $\hat{f}_n(\alpha, w) = \mathbb{E}\exp\{i(\sum_{j<k}\alpha_{jk}\lambda_{jk}^{(n)} + \sum_j w_j x_j^{(n)})\} = \mathbb{E}\exp\{i\sum_{r=1^n}\frac{1}{r}(\sum_{j<k}\alpha_{jk}(x_{jr}y_{kr} - x_{kr}y_{jr}) + \sum_j w_j x_{jr})\}$. Now let D be the $d \times d$ skew-symmetric matrix by defining $\alpha_{kj} = -\alpha_{jk}$ for $j < k$ and $\alpha_{jj} = 0$, and let $A = D\hat{D}$ $(= -D^2)$. Then a standard type of calculation gives that

$$\hat{f}_n(\alpha, w) = \exp\{-\frac{1}{2}\sum_{r=1}^{n} r^{-2}w^t(I + r^{-2}A)^{-1}w\}\prod_{r=1}^{n}\det(I + r^{-2}A)^{-1/2} \tag{32}$$

As A is positive semi-definite we see that $0 \leq \hat{f}_n(\alpha, w) \leq \hat{f}_{n_0}(\alpha, w)$ where $n_0 = 1 + N + d(d+1)/2$.

Note that $\mathrm{tr}(A) = 2|\alpha|^2$. So $A \leq 2|\alpha|^2 I$ and hence

$$\sum_{r=1}^{n} r^{-2}w^t(I + r^{-2}A)^{-1}w \geq w^t(I + A)^{-1}w \geq \frac{|w|^2}{1 + 2|\alpha|^2}$$

And $\det(I + r^{-2}A) \geq 1 + r^{-2}\mathrm{tr}(A) = 1 + 2r^{-2}|\alpha|^2$. So $\hat{f}_{n_0}(\alpha, w) \leq C_1(1 + |\alpha|)^{-n_0}\exp(-\frac{|w|^2}{2(1+2|\alpha|^2)}) \leq C_2(|\alpha| + |w|)^{-n_0}$. As $n_0 > N + d(d+1)/2$ it follows that $(1 + |\alpha| + |w|)^N \hat{f}_n(\alpha, w) \in L^1(\mathbb{R}^{d(d+1)/2})$, with L^1 norm bounded independently of n, and the result follows.

(2) Replacing (α, w) by $(i\alpha, iw)$ in (32) we get the moment generating function of $v^{(p)}$:

$$\mathbb{E}\exp\left(\sum_{j<k}\lambda_{jk}^{(p)} + \sum_j w_j z_j^{(p)}\right) = \exp\left\{\frac{1}{2}\sum_{r=1}^{p} r^{-2}w^t(I - r^{-2}A)^{-1}w\right\}\prod_{r=1}^{n}\det(I - r^{-2}A)^{-1/2}$$

provided $|\alpha| < 1$, and similarly we have

$$\mathbb{E}\exp\left(\sum_{j<k}\tilde{\lambda}_{jk}^{(p)} + \sum_j w_j \tilde{z}_j^{(p)}\right) = \exp\left\{\frac{1}{2}\sum_{r=p+1}^{\infty} r^{-2}w^t(I - r^{-2}A)^{-1}w\right\}\prod_{r=p+1}^{\infty}\det(I - r^{-2}A)^{-1/2}$$

provided $|\alpha| < p$. The results then follow by straightforward estimates. □

We use ψ_n to denote the map from $(x,y) \in \mathbb{R}^{dn} \times \mathbb{R}^{dn}$ to $(\lambda^{(n)}, z^{(n)}) \in \mathbb{R}^{d(d+1)/2}$ as defined above. It will be convenient to express ψ_n in terms of an inner product on $\mathbb{R}^n$ defied by $\langle \xi, \eta \rangle = \sum_{r=1}^n r^{-1} \xi_r \eta_r$. We write $x_j = (x_{j1}, \cdots, x_{jn}) \in \mathbb{R}^n$ for $j = 1, \cdots, d$, and y_j similarly. Also let $\rho = (1, 1, \cdots, 1) \in \mathbb{R}^n$. Then $\psi_n(x,y) = (\lambda^{(n)}, z^{(n)})$ where $\lambda_{jk}^{(n)} = \langle x_j, y_k \rangle - \langle x_k, y_j \rangle$ and $z_j = \langle x_j, \rho \rangle$. We use $|x|$, $|\lambda|$ etc for the usual Euclidean norm on the relevant space.

Now ψ_n maps onto $\mathbb{R}^{d(d+1)/2}$ provided $n \geq d+1$. To see this, let $v = (\lambda, z) \in \mathbb{R}^{d(d+1)/2}$ and define x and y in $\mathbb{R}^{dn}$ by $x_{jr} = r|\lambda|^{1/2}\delta_{jr}$ for $r \neq d+1$, $x_{j,d+1} = (d+1)(z_j - |\lambda|^{1/2})$, while $y_{kr} = |\lambda|^{-1/2}\lambda_{rk}$ for $r < k$ and 0 for $r \geq k$. Then, writing $u = (x,y)$, we have $\psi_n(u) = v$ and $|u| \leq K(|\lambda|^{1/2} + |z|)$ where K is a constant depending only on d.

We can express the last conclusion in an equivalent way by introducing some notation: let $\mathcal{M}(v) = |\lambda| + |z|^2$ and $\mathcal{M}_n(v) = \inf\{|u|^2 : \psi_n(u) = v\}$. Then $K_1 \mathcal{M}(v) \leq \mathcal{M}_n(v) \leq K_2 \mathcal{M}(v)$ where K_1 and K_2 are positive constants depending only on d.

Lemma 12. *Suppose $n > d(d+1)/2$. Then there is a constant C such that*

$$f_n(v) \leq C(1 + \mathcal{M}_n(v))^{nd} e^{-\mathcal{M}_n(v)/2}$$

for all $v \in \mathbb{R}^{d(d+1)/2}$.

Proof. Given $v \in \mathbb{R}^{d(d+1)/2}$, fix $u^* \in \mathbb{R}^{2nd}$ so that $\psi_n(u^*) = v$ and $|u^*|^2 = \mathcal{M}_n(v)$. Let $K = 2\sqrt{1 + \mathcal{M}_n(v)}$. Then if $u' \in \mathbb{R}^{2nd}$ is such that $|u^* + Ku'| \geq |u^*|$ we claim that

$$|u^* + Ku'|^2 \geq |u^*|^2 + |u'|^2 - 2 \tag{33}$$

This is immediate if $|u'| \leq \sqrt{2}$, while if $|u'| > \sqrt{2}$ then $K|u'| \geq \sqrt{2}|u'|(1 + |u^*|) > |u'| + 2|u^*|$ so that $|u^* + Ku'| > |u'| + |u^*|$ and (33) follows.

From (33) it follows that $\phi(u^* + Ku') \leq e^{1-|u^*|^2/2}\phi(u')$ where ϕ denotes the $N(0,I)$ density on $\mathbb{R}^{2nd}$. Hence ϕ, restricted to $\{x : |x| \geq |u^*|\}$, is $\leq e^{1-|u^*|^2/2} K^{2nd}$ times the density of $u^* + KU'$, where U' is an $N(0,I)$ random variable. This implies that $f_n(v) \leq e^{1-|u^*|^2/2} K^{2nd} \rho(v)$ where ρ is the density function of the $\mathbb{R}^{d(d+1)/2}$-valued random variable $\psi_n(u^* + KU')$.

A calculation similar to that used in the proof of part (1) of Lemma 11 shows that the Fourier transform $\hat{\rho}$ satisfies

$$|\hat{\rho}(\alpha, w)| = \exp\left\{ -\frac{K^2}{2} \sum_{r=1}^n \frac{1}{r} (\beta_r^t (I + r^{-2}K^4 A)^{-1} \beta_r + \gamma_r^t (I + r^{-2}K^4 A)^{-1} \gamma_r) \right\} \prod_{r=1}^n \det(I + r^{-2}K^4 A)^{-1/2}$$

where A is defined as in the proof of Lemma 11, β_r is the vector in $\mathbb{R}^d$ with components $\beta_{rj} = \sum_{k=1}^d \alpha_{jk} y_{kr}^* + w_j$ and γ_r is defined similarly with $\gamma_{rj} = \sum_{k=1}^d \alpha_{jk} x_{kr}^*$.

Again using similar arguments to those used at the end of the proof of part (1) of Lemma 11, we deduce that the L^1 norm of $\hat{\rho}$, and hence also $\rho(0)$, is bounded by $Ce^{-|u^*|^2/2}$ and the result follows. □

Lemma 13. *Suppose $n > 3d$ and that v and v' in $\mathbb{R}^{d(d+1)/2}$ satisfy $|v - v'| \leq (1 + \mathcal{M}_n(v))^{-1}$. Then*

$$f_n(v) \geq C(1 + \mathcal{M}_n(v'))^{-n} e^{-\mathcal{M}_n(v')/2}$$

where the constant C depends only on d and n.

Proof. We choose $u = (x,y) \in \mathbb{R}^{2nd}$ such that $|u|^2 = \mathcal{M}_n(v')$ and $\psi_n(u) = v'$. Define $K = 1 + |u|$ and $h = \langle \rho, \rho \rangle^{-1} = (\sum_{r=1}^n r^{-2})^{-1}$. As $n > 3d$ we can find $\tau_1, \cdots, \tau_d \in \mathbb{R}^n$ which are mutually orthogonal (w.r.t. the inner product $\langle\ ,\ \rangle$), and orthogonal to $x_1, \cdots, x_d, y_1, \cdots, y_d$ and ρ, and also so that $\langle \tau_j, \tau_j \rangle = 1$ for each j. Then we can define a linear map $T : \mathbb{R}^{d(d+1)/2} \to \mathbb{R}^{2nd}$ by $T(\mu, \zeta) = (\tilde{x}, \tilde{y})$ where $\tilde{x}_j = h\zeta_j \rho \in \mathbb{R}^n$ and $\tilde{y}_k = \sum_{j=1}^{k-1} (\mu_{jk} - h(\zeta_j q_k - \zeta_k q_j))\tau_j \in \mathbb{R}^n$,

where $q_j = \langle y_j, \rho\rangle$. Note that $\|T\| \le C_1 K$. We define also $\hat{T}(\mu,\zeta) = u^* + T(\mu,\zeta)$, where $u^* = u + (\tau, 0)$, and note that then $\psi_n(\hat{T}(\mu,\zeta)) = (\lambda+\mu, z+\zeta)$.

In particular, if we define $u_0 = \hat{T}(v - v')$, then $\psi_n(u_0) = v$. Then $\|u - u_0\| = \|(\tau,0) + T(v-v')\| \le C_1 K$. Also $D\psi_n(u_0)T$ is the identity on $\mathbb{R}^{d(d+1)/2}$, $\|D\psi_n(u_0)\| \le C_2 K$, and the second derivative matrix of ψ_n is constant. By Lemma 10 the image under ψ_n of Lebesgue measure on a ball $B(u_0, C_3K^{-1})$ has density at least $C_4 K^{-n}$ at v. Also the density of the $N(0,I)$ distribution is at least $C_5 e^{-\mathcal{M}_n(v')/2}$ on this ball and the result follows. □

Lemma 14. *Given $d \ge 2$ there exist positive δ and C such that whenever $v, v' \in \mathbb{R}^{d(d+1)/2}$ with $|v - v'| \le \delta(1+\mathcal{M}(v))^{-1}$, we have*

$$f_p(v) \ge C(1+\mathcal{M}(v'))^{-n(d+1)} f_p(v')$$

for all $p \ge n = 1 + \max(3d, d(d+1)/2)$

Proof. We use the notation $\mathcal{N}(v) = 1 + \mathcal{M}(v)$. Then we note that, by Lemmas 12 and 13, we have that $f_n(v) \ge C_1 \mathcal{N}(v)^{-n(d+1)} f_n(v')$ whenever $|v - v'| \le \mathcal{N}(v)^{-1}$. We have to extend this to f_p with a uniform constant. We note also that, from Lemma 12, and 13 with $v' = v$, we can find a constant K such that whenever v_1 and v_2 satisfy $\mathcal{N}(v_1) \ge K\mathcal{N}(v_2)$ we have $f_n(v_1) \le f_n(v_2)$.

We now assume $p > n$, and express f_p as a convolution involving f_n. Write $v^{n,p}$ for the random variable $v^{(p)} - v^{(n)}$, noting that then $v^{(n,p)}$ is independent of $v^{(n)}$. We also note that $\mathbb{E}|v^{(n,p)}|^2 = d^2 \sum_{r=n+1}^{p} r^{-2} < d^2\pi^2/3$ and so $\mathbb{P}(v^{(n,p)} > 3d) \le \frac{\pi^2}{27} < \frac{1}{2}$. Then $f_p(v) = \mathbb{E} f_n(v - v^{(n,p)})$. We write $X = v - v^{(n,p)}$ and $X' = v' - v^{(n,p)}$, noting that then $X - X' = v - v'$.

Let U be the random variable which is 1 when $\mathcal{N}(X') \le 10d^2 K \mathcal{N}(v')$ and 0 otherwise. Then $\mathbb{E}\{f_n(X')U\} \ge \frac{1}{2}\mathbb{E} f_n(X')$. Now let $\delta = (10d^2K)^{-1}$. Then $|v - v'| \le \delta \mathcal{N}(v')^{-1}$ implies $|X - X'| \le \mathcal{N}(X')^{-1}$ and so

$$\begin{aligned} f_p(v) &= \mathbb{E} f_n(X) \ge \mathbb{E}\{f_n(X)\chi_A\} \ge C_1 \mathbb{E}\{\mathcal{N}(X')^{-n(d+1)} f_n(X')\} \\ &\ge C_2 \mathcal{N}(v')^{-n(d+1)} \mathbb{E}\{f_n(X')\chi_A\} \ge \frac{1}{2} C_2 \mathcal{N}(v')^{-n(d+1)} \mathbb{E}\{f_n(X') \\ &= \frac{1}{2} C_2 \mathcal{N}(v')^{-n(d+1)} f_p(v') \end{aligned}$$

as required. □

Theorem 15. *Suppose $p \in \mathbb{N}$ and $\overline{v}$ is a $\mathbb{R}^{d(d+1)/2}$-valued random variable, independent of $v^{(p)}$, such that for some positive a, b we have $\mathbb{E}\{\exp(ap^{1/2}|\overline{v}|)\} \le b$, and such that $\overline{v}$ has the same moments up to order $m-1$ as $\tilde{v}^{(p)}$).*

Then $\mathbb{W}_2(v, v^{(p)} + \overline{v}) \le Cp^{-m/2}$, where C is a constant depending only on d, m, a, b.

Proof. First note that, by Theorem 11, if p is large enough then f_p has uniformly bounded derivatives up to order $2m-1$.

We use the notation B_r for the ball $|x| \le r$ in $\mathbb{R}^{d(d+1)/2}$. From part (2) of lemma 11, we can choose $A > 0$ (independent of p), so that, if we let $R = A \log p$ and $r = p^{-1/2}R$, then $\mathbb{E}\{|v^{(p)}|^2 \chi_{|v^{(p)}|>R}\} = O(p^{-m})$, and $\mathbb{E}\{|\tilde{v}^{(p)}|^2 \chi_{|\tilde{v}^{(p)}|>r}\} = O(p^{-m})$ with the same for $\overline{v}$. Then we can find a random variable X such that $X = v^{(p)}$ if $v^{(p)} \in B_{2R}$, and $X \in B_{3R} \backslash B_{2R}$ otherwise. Then we have $\mathbb{E}|X - v^{(p)}|^2 \le C_1 p^{-m}$. We also let Y and Z be random variables which are in B_{2r} always and $\mathbb{E}|Y - \tilde{v}^{(p)}|^2 \le C_2 p^{-m}$ and $\mathbb{E}|Z - \overline{v}|^2 \le C_2 p^{-m}$. Then $\mathbb{W}_2(v, X+Y) = O(p^{-m/2})$ and $\mathbb{W}_2(v^{(p)} + \overline{v}, X + Z) = O(p^{-m/2})$. Note also that then $\mathbb{E}(Y^\beta - Z^\beta) = O(p^{-m})$ for $|\beta| < m$.

We let $g(x)$ be the density of $X+Y$ and $h(x)$ the density of $X+Z$. Then $g(x) = \mathbb{E} f_p(x - Y)$ for $x \in B_{2R}$. Now for $y \in B_{2r}$ we have

$$f_p(x-y) = \sum_{0 \le |\beta| \le 2m} (-1)^{|\beta|} (D^\beta f_p)(x) y^\beta + O(p^{-m})$$

and hence $g(x) = \sum_{0\le|\beta|\le 2m}(-1)^{|\beta|}(D^\beta f_p)(x)\mathbb{E}(Y^\beta) + O(p^{-m})$. Similarly $h(x) = \sum_{0\le|\beta|\le 2m}(-1)^{|\beta|}(D^\beta f_p)(x)\mathbb{E}(Z^\beta) + O(p^{-m})$. Then we have

$$h(x) - g(x) = \sum_{m\le|\beta|\le 2m}(-1)^{|\beta|}(D^\beta f_p)(x)\mathbb{E}(Z^\beta - Y^\beta) + O(p^{-m})$$

For a multi-index $\beta = (\beta_1, \cdots, \beta_{d(d+1)/2}$ with $|\beta| > 0$ we write $\beta' = (\beta_1, \cdots, \beta_j - 1, \cdots, \beta_{d(d+1)/2})$ where $j = j(\beta)$ is the first j such that $\beta_j > 0$. Then if we write

$$u(x) = \sum_{m\le|\beta|\le 2m}(-1)^{|\beta|}(D^{\beta'} f_p)(x)\mathbb{E}(Z^\beta - Y^\beta)e_{j(\beta)}$$

where e_j is the unit vector in the jth coordinate direction, we have $h(x) - g(x) = \nabla.u(x) + O(p^{-m})$.

For $x \in B$ let $\Omega(x)$ denote the set of $y \in \mathbb{R}^{d(d+1)/2}$ such that $|y - x| \le \delta(1 + \mathcal{M}(v))^{-1}$ where δ is as in lemma 14. Now, if $\eta > 0$ is given, then combining lemmas 11, 9 and 14 we find that if p is large enough then $|D^\beta f_p(y)| \le C_1 f_p(x)^{1-\eta}$. Then if η was chosen small enough we can deduce that $|f(y) - g(y)|$, $|u(y)|$ and $|\nabla.u(y)|$ can all be bounded for $y \in \Omega$ by $C_2 p^{-1/4} f(x)$, and also that $|u(y)| \le C_3 f(x)^{2/3} p^{-m/2}$. Then we can define, for $0 \le t \le 1$ and $y \in \Omega(x)$, $\zeta_t(y) = u(u)(g(y) + t\nabla.u(y))^{-1}$, and we have

$$|\zeta_t(y)| \le C_4 \min(p^{-1/4} f_p(x), p^{-m/2} f(x)^{-1/3})$$

Now, given $x \in B$ we let $\rho_t(x)$ be the solution of the ODE $\frac{d}{dt}\rho_t(x) = \zeta_t(\rho_t(x))$ with $\rho_0(x) = x$. From the above bounds we see that for $0 \le t \le 1$ the solution stays in $\Omega(x)$ and $|\rho_1(x) - x| \le C_4 p^{-m/2} f(x)^{-1/3}$. We see that $\rho_1(X + Y)$ has density

$$g(x) + \nabla.u(x) = h(x) + O(\epsilon^{2m})$$

on $\rho_t(B)$, and so $\mathbb{W}_2(\rho_1(X+Y), X+Z) = O(p^{-m/2})$. And $\mathbb{E}|\rho_1(X+Y) - (X+Y)|^2 = O(p^{-m})$ so $\mathbb{W}_2(\rho_1(X + Y), X + Z) = O(p^{-m/2})$. The result then follows from the triangle inequality for the $\mathbb{W}_2$ metric. □

Application to the Milstein method.

We consider the application of theorem 15 to the approximate solution of a vector SDE driven by a d-dimensional Brownian path using the Milstein method with step-size h. For each step, we first generate the Brownian increments, and then for the double integrals A_{jk} we need to generate approximations X_{jk} with $\mathbb{E}|X_{jk} - A_{jk}|^2 = O(h)$ in order to get $O(h)$ for the Milstein approximation. By theorem 15, we can do this using $v^{(p)}$ and $\overline{v}$ satisfying the hypotheses of the theorem provided $p^{-m} \le h$, i.e. $p \ge h^{-1/m}$. For fixed m the computational load in generating $\overline{v}$ should be independent of p, while that for generating $v^{(p)}$ is proportional to P. So for fixed m the load is $O(p) = O(h^{-1/m})$ for each step and $O(h^{-1-1/m})$ for the whole approximation on a fixed interval $[0, T]$.

Regarding the construction of $\overline{v}$, we first note that $\tilde{v}^{(p)}$ has a distribution symmetric about the origin, in the sense that $-\tilde{v}^{(p)}$ has the same distribution, and hence all moments of odd order are 0. If we choose $\overline{v}$ to have the same property, then we have only to match even-order moments, and then it makes sense to assume m is even. We now give explicit constructions in the cases $m = 2, 4, 6$.

The simplest case is $m = 2$ for which, since the components of $\tilde{v}^{(p)}$ all have mean 0, we can simply take $\overline{v} = 0$. Then we need $p \sim h^{-1/2}$ and a total computational load $\sim h^{-3/2}$ to get the order 1 accuracy. But this is using $v^{(p)}$ as the approximation for v, just as in [16], and the error estimate in [16] leads to $p \sim h^{-1}$ as pointed out in [28]. The reason for the difference is that our interpretation of the random vector $v^{(p)}$ we generate is different from

that in [16] - in [16] it is the truncation of the Fourier series for v, but here it is a different $\mathbb{R}^{d(d+1)/2}$-valued random variable X, with exactly the same distribution, but giving a better approximation to v, i.e. $\mathbb{E}|X - v|^2 = O(p^{-2})$ as opposed to $\mathbb{E}|v^{(p)} - v|^2 = O(p^{-1})$ for the interpretation in [16].

The case $m = 4$ is also straightforward. The random variables $\tilde{v}_{jk}$ are mutually uncorrelated, in fact if $j < k$ and $i < l$ then $\mathbb{E}(\tilde{\lambda}_{jk}\tilde{\lambda}_{il}) = \begin{cases} \eta_p & \text{if } j = i \text{ and } k = l \\ 0 & \text{otherwise} \end{cases}$, where $\eta_p = \sum_{r=p+1}^{\infty} r^{-2}$, and similarly the z_j each have variance η_p and they are uncorrelated with each other and all the λ_{jk}. So if P_{jk} for $1 \le j < k \le d$ and Q_j for $1 \le j \le d$ are i.i.d. random variables with even distribution, variance 1 and exponential decay then $\overline{v} = \eta_p^{1/2}(P, Q)$ will do. One possibility is to choose the P_{jk} and Q_j to be $N(0, 1)$, which gives a method computationally equivalent to that described by Wiktorsson [28]. But with our approach we get a computational load $O(h^{-5/4})$, compared with $O(h^{-3/2})$ from Wiktorsson's estimate - again the difference arises from the coupling interpretation.

We can also handle the case $m = 6$ without much trouble. For this we need to match the 4th order moments as well. These moments are more complicated, as there are several cases, but we give an argument which avoids the need to calculate the moments explicitly. It will be more convenient to work with the cumulants, so we try to generate $\overline{v}$ having the same 2nd and 4th order cumulants as $\tilde{v}^{(p)}$. We start by generating P_{jk} and Q_j to be independent $N(0, 1)$ as in the last paragraph, and set $v' = (P, Q)$. Next we generate x_j and y_j to be independent $N(0, 1)$ for $j = 1, \cdots, d$, and write $v^* = (\mu, x)$ where $\mu_{jk} = x_j y_k - x_k y_j$. Now since the cumulants of a sum of independent random variables are sums of the cumulants, we see that the 4th order cumulants of $\tilde{v}^{(p)}$ are the corresponding cumulants of v^* multiplied by $\nu_p := \sum_{r=p+1}^{\infty} r^{-4}$. Then we try $\overline{v} = \alpha v' + \beta v^*$ where α and β are constants. Using the fact the v', being Gaussian, has zero 4th order cumulants, we find that to match all 2nd and 4th order cumulants we require $\alpha^2 + \beta^2 = \eta_p$ and $\beta^4 = \nu_p$. So we take $\beta = \nu_p^{1/4}$ and $\alpha = (\eta_p - \nu_p^{1/2})^{1/2}$ to get $\overline{v}$ with the required property, and then we get a computational load $O(h^{-7/6})$ for an order 1 approximation using Milstein.

Extension to longer iterated integrals.

It is natural to try to extend the above results to iterated integrals of length > 2, for application to approximation schemes of order > 1. However for length 3 there are already significant difficulties, because the independence of the perturbation $\tilde{v}^{(p)}$ from the truncated sum $v^{(p)}$, which we used in the above analysis for length 2, no longer holds for length 3. Indeed, the expansion of an iterated integral I_{jkl} contains sums like $\sum_{r,s=1}^{\infty} \frac{1}{rs} x_{jr} x_{k,s+r} y_{ls}$ (as shown in section 5.8 of [16]). Then the truncated expansion $v^{(p)}$ would contain the sum over just those pairs (r, s) such that $r, s, r + s$ are all $\le p$. The sum over the remaining would be contained in $\tilde{v}^{(p)}$, which would therefore contain terms involving some x_{jr} and l_{ls} for $r, s \le p$.

This would considerably complicate the proof of an analogue of theorem 15. But it would also mean that the moments to be matched would have to be conditional on $v^{(p)}$, which in practice probably means conditional on all the x_{jr}, y_{jr} for $r \le p$. This would complicate the calculation of the moments and the matching process, so it is not clear to what extent a scheme based integrals of length 3 or more would be computationally efficient. Further investigation of this question is needed.

10 Relaxation of assumptions on coefficients

In the above discussion we have assumed a global Lipschitz condition for the b_{lk}, and for the results of sections 7 and 8 a bounded matrix inverse. These are quite restrictive conditions.

In the standard theory one can replace the global Lipschitz condition by a local one,

which is much less restrictive. Then the solution $x(t)$ can 'explode' in finite time and one cannot in general expect L^p error bounds of the type discussed above. The usual approach to such situations is to define a sequence of stopping times τ_n by $\tau_n = \inf\{t : x(t) \geq n\}$. Then a.s. $\tau_n \to \tau \in [0, \infty]$ where τ is a stopping time - the 'explosion time' - and one can show that a.s. $|x(t)| \to \infty$ as $t \to \tau$ from below. See the 'Theoretical Complement' to section 7.3 of [4] for a discussion of this.

Then one can introduce sequences of modified coefficients $a_n(t,x), b_n(t,x)$ such that $a_n(t,x) = a(t,x)$ for $|x| \leq n+1$ and the same for b_n, and such that a_n and b_n satisfy global Lipschitz conditions. This allows one to apply the known a.s. convergence results, such as (5) for Euler, to the modified equations and deduce that, with probability 1, for any $\epsilon > 0$ the Euler approximations with stepsize $2^{-r}T$ will converge uniformly with order $\frac{1}{2} - \epsilon$ on any interval $[0, t_1]$ such that $t_1 < \tau$.

One would expect similar results for scheme (10) with the coupling approach - assuming the b_{ik} are C^2, if we construct a sequence of approximations as described in section 8, then one can deduce in the same way from the bound (25) that a.s. the approximation converges with order $1 - \epsilon$ on any interval $[0, t_1]$ on which the solution is bounded and avoids any x where the matrix $(b_{lk}(x))$ is not invertible. Rather more generally, if we consider a single-level approximation using a scheme (15) with order $\frac{m}{2}$, and assume sufficient smoothness of the coefficients (but no global bounds), then one can deduce from the bound (16) a bound 'in probability' of the following form: given $\gamma < \frac{m}{2}$ and given $T > 0$ and $\epsilon > 0$, one can find $h_0 > 0$ such that if $0 < h < h_0$ then, with probability $> 1 - \epsilon$, the approximation $\hat{x}$ given by (15) with stepsize h satisfies $|\hat{x}^{(j)} - x(jh)| < h^{-\gamma}$ for all j such that $0 \leq jh \leq \min(T, \tau - \epsilon, \tau' - \epsilon)$, where τ is the explosion time and τ' is the first time the matrix $(b_{ik}(t, x(t))$ has rank $< q$.

One can go further and consider what happens when the solution passes through points where the matrix has rank $< q$; one possible approach is to use a variable stepsize approach, taking smaller steps when close to such points. In such a way one might hope still to get higher order for equations where the matrix $(b_{ik}(t,x))$ is invertible except on a subvariety of codimension 1 (which when $q \leq d$ is the generic situation). See [9] for some background on variable stepsize methods for SDE.

Some other relevant references for this section are [13, 14, 24].

11 Local couplings and examples

It is a feature of schemes of the form (15), as well as those described in section 9, that the couplings are between increments for a single step, so that the discrete processes $(x^{(j)})$ and $(x(jh))$ are adapted to the same filtration. Such couplings we call *local*. Local couplings may in some respects be easier to handle, but we give examples below to indicate that non-local couplings can give better approximations.

We first give an example showing that in general the Euler scheme (3) cannot give order better than $\frac{1}{2}$, with any type of coupling.

Example 1. Consider the system with $d = 1$, $q = 2$ given by

$$dx_1 = x_2 dW, \quad dx_2 = -x_1 dW; \qquad x_1(0) = 1,\ x_2(0) = 0 \tag{34}$$

Note that for any solution of this system, S defined by $S(t) = x(t)^2 + y(t)^2$ satisfies the (deterministic) equation $dS = S dt$ with $S(0) = 1$ so we must have $S(t) = e^t$. Now consider an Euler approximation to (34) on the interval $0 \leq t \leq 1$ with N steps of size $h = N^{-1}$. The approximate solution is given by $x_1^{(n+1)} = x_1^{(n)} + x_2^{(n)} \Delta W^{(n)}$, $x_2^{(n+1)} = x_2^{(n)} - x_1^{(n)} \Delta W^{(n)}$. Then writing $S^{(n)} = (x_1^{(n)})^2 + (x_n^{(n)})^2$ we see that $S^{(n+1)} = S^{(n)}(1 + (\Delta W^{(n)})^2)$ and so $S^{(N)} = \prod_{n=0}^{N-1}(1 + (\Delta W^{(n)})^2)$. Also $S(1) = e$ and one shows easily that $|e - S^{(N)}|$ is typically

of order $h^{1/2}$. As $S(1)$ is independent of the Brownian path W, this shows that the error in the Euler approximation will not be less than $O(h^{1/2})$ for any coupling.

We remark that since $d = 1$, approximation using (8) will give error $O(h)$ for this system, even in the usual strong sense.

Our second example shows that in one dimension, with a local coupling Euler cannot give order better than $\frac{1}{2}$, although as shown in [1] higher order can be attained with a non-local coupling. We first prove a lemma:

Lemma 16. *Let Y be a $N(0,1)$ random variable and define $Z_\epsilon = Y + \epsilon(Y^2 - 1)$. There is a constant $C > 0$ such that, if U is a random variable with distribution symmetric about 0, then $\mathbb{W}_2(Z_\epsilon, U) \geq c\epsilon$ for all sufficiently small positive ϵ.*

Proof. We find that the density function f_ϵ of Z_ϵ satisfies

$$f_\epsilon(z) = (2\pi)^{-1/2} e^{-z^2/2}(1 + \epsilon z(z^2 - 3)) + O(\epsilon^{-2})$$

Now fix a smooth nonnegative function ϕ on $\mathbb{R}$, supported on $[1, \frac{3}{2}]$, such that $\int_1^{3/2} \phi(x)dx = 1$. Define ψ on $\mathbb{R}$ by $\psi(x) = \phi(x) - \phi(-x)$. Then we see that $\mathbb{E}\psi(Z_\epsilon) = \int \psi(x) f_\epsilon(x)dx \geq C_1\epsilon + O(\epsilon^2)$. Also $\mathbb{E}\psi(U) = 0$ and $|\mathbb{E}\psi(U) - \mathbb{E}\psi(Z_\epsilon)| \leq C_2 \mathbb{W}_2(Z_\epsilon, U)$ and the result follows. □

Example 2. Consider the one-dimensional equation $dx = xdW$ on [0,1] with $x(0) = 1$. We show that the Euler method using a local coupling cannot approximate $x(1)$ to better than order $h^{1/2}$ in mean-square sense.

To show this, first note that the solution is given explicitly by $x(t) = e^{W(t) - \frac{t}{2}}$. Now suppose we have an Euler approximation with step-size $h = \frac{1}{N}$ given by $x^{(j+1)} = x^{(j)} + x^{(j)}V_j$, where $V_0, \cdots V_{N-1}$ are independent $N(0,h)$ random variables, locally coupled to the increments $\Delta W_j = W((j+1)h) - W_{jh}$, so that V_j is $\mathcal{F}_{(j+1)h}$ measurable and $\mathbb{E}(V_j|\mathcal{F}_{jh}) = 0$. Let e_j denote the error $e_j = x^{(j)} - x(jh)$. Then the sequences $(x^{(j)})$ and $x(jh)$ are martingales, hence so is (e_j).

Using the fact that $x((j+1)h) = x(jh)e^{\Delta W_j - \frac{h}{2}}$, we can then write $e_{j+1} - e_j = -\rho_j - z_j$ where $\rho_j = x(jh)\{\Delta W_j + \frac{1}{2}(\Delta W_j^2 - h)\} - x^{(j)}V_j$ and $z_j = x(jh)\{e^{\Delta W_j - \frac{h}{2}} - 1 - \Delta W_j - \frac{1}{2}(\Delta W_j^2 - h)\}$. Applying lemma 16 with $Y = h^{-1/2}\Delta W_j$ and $\epsilon = \frac{1}{2}h^{1/2}$ we find that $\mathbb{E}(\rho_j^2) \geq \frac{C}{2}h^2\mathbb{E}(x(jh)^2) = \frac{C}{2}h^2 e^{jh} \geq \frac{C}{2}h^2$. Also $\mathbb{E}(z_j)^2 = O(h^3)$ so for h small enough we have $\mathbb{E}(e_{j+1} - e_j)^2 \geq \frac{C}{4}h^2$.

Finally, the martingale property of (e_j) gives $\mathbb{E}(e_N)^2 = \sum_{j=0}^{N-1} \mathbb{E}(e_{j+1} - e_j)^2 \geq \frac{Ch}{4}$, in other words the mean square error in the approximation $x^{(N)}$ of $x(1)$ is at least a constant times $h^{1/2}$.

Finally a similar example for the scheme (8) in higher dimensions.

Example 3. Consider the system with $d = 2$, $q = 3$ given by

$$dx_1 = dW_1, \quad dx_2 = dW_2, \quad dx_3 = x_1 dW_2; \qquad x_i(0) = 0, \quad i = 1, 2, 3 \tag{35}$$

This system of equations has solution $x_1(t) = W_1(t)$, $x_2(t) = W_2(t)$, $x_3(t) = \int_0^t W_1(\tau)dW_2(\tau)$. We consider approximation using (8) on the interval [0,1] with stepsize $h = N^{-1}$. We will show that with a *local* coupling it is not possible to improve on $O(h^{1/2})$ error. It seems likely however that one could construct a non-local coupling which gives higher order in general.

We write $W_i^{(s)} = W_i(sh)$ and $\Delta W_i^{(s)} = W_i^{(s+1)} - W_i^{(s)}$, and note first that from the above expression for the solution we can write

$$x_3(1) = \frac{1}{2}\sum_{s=0}^{N-1}(W_1^{(s+1)} + W_1^{(s)})\Delta W_2^{(s)} + Y \tag{36}$$

where $Y = \sum_{s=0}^{N-1}\{\int_{sh}^{(s+1)h}(W_1(t) - W_1^{(s)})dW_2(t) - \frac{1}{2}\Delta W_1^{(s)}\Delta W_2^{(s)}\}$.

Now suppose $U_i^{(j)}$ for $i = 1, 2$ and $j = 1, \cdots, N$ are random variables such that the increments $\Delta U_i^{(j)}$ are mutually independent, and that the $U_i^{(j)}$ are locally coupled to the $W_i^{(j)}$. Then we define $U_3^{(j)} = \frac{1}{2}\sum_{s=0}^{j-1}(U_1^{(s+1)} + U_1^{(s)})\Delta U_2^{(s)}$, so that $U_i^{(j)}$ is the approximation to$x_i(jh)$ using (8). We also define the errors $Z_i^{(j)} = x_i(jh) - U_i^{(j)}$ for $i = 1, 2, 3$ and note that for $i = 1, 2$ we have $Z_i^{(j)} = W_i^{(j)} - U_i^{(j)}$ so that $(Z_1^{(j)})$ and $(Z_2^{(j)})$ are martingales.

Now after some calculation we find that $Z_3^N = Y + \frac{1}{2}(V - R) + Z_2^{(N)}U_1^{(N)}$ where $V = \sum_{s=0}^{N-1}(Z_1^{(s+1)} + Z_1^{(s)})\Delta W_2^{(s)}$ and $R = \sum_{s=0}^{N-1}(Z_2^{(s+1)} + Z_2^{(s)})\Delta U_2^{(s)}$. Now we have $V = 2P + Q$ where $P = \sum_{s=0}^{N-1} Z_1^{(s)}\Delta_2 W^{(s)}$ and $Q = \sum_{s=0}^{N-1}\Delta Z_1^{(s)}\Delta W_2^{(s)}$. Then using the martingale property we have

$$\mathbb{E}(P^2) = \sum_{s=0}^{N-1}\mathbb{E}(Z_1^{(s)})^2\mathbb{E}(\Delta W_2^{(s)})^2 = h\sum_{s=0}^{N-1}\mathbb{E}(Z_1^{(s)})^2 \le \mathbb{E}(Z_1^{(N)})^2$$

and also by Cauchy-Schwarz,

$$(\mathbb{E}|Q|)^2 \le \left\{\sum_{s=0}^{N-1}\mathbb{E}(Z_1^{(s+1)} - Z_1^{(s)})^2\right\}\left\{\sum_{s=0}^{N-1}\mathbb{E}(W_2^{(s+1)} - W_2^{(s)})^2\right\} = \mathbb{E}(Z_2^{(N)})^2$$

So $\mathbb{E}|V| \le 2\{\mathbb{E}((Z_1^{(N)})^2 + (Z_2^{(N)})^2)\}^{1/2}$, with the same bound obtained in the same way for R, and then we conclude that $\mathbb{E}|Y| \le 3\{\sum_{i=1}^3\mathbb{E}(Z_i^{(N)})^2\}^{1/2}$. On the other hand there is a constant $C > 0$ (independent of N) such that $\mathbb{E}|Y| \ge Ch^{1/2}$ and hence $\{\sum_{i=1}^3\mathbb{E}(Z_i^{(N)})^2\}^{1/2} \ge Ch^{1/2}/3$, showing that it is not possible to improve on $O(h^{1/2})$ with a local coupling.

A natural question is whether better approximations can be achieved using non-local couplings for the system in example 3, or indeed for general SDEs. It seems likely that this is the case, and that the 'KMT method' of [19], which is a powerful tool for finding couplings, would be useful for this purpose (examples of the application of this method to non-local couplings for numerical approximation of SDE can be found in the final chapter of volume 2 of [21]).

12 General discussion

We return to the SDE (1) on $[0, T]$ and consider the general question of assessing the accuracy of a given approximation scheme in the context of our 'coupling' approach. As before we consider a uniform subdivision of $[0, T]$ into N steps of size $h = T/N$, and suppose that our scheme generates an approximation $(x_i^{(j)})$, $j = 1, 2, \cdots, N$, $i = 1, 2, \cdots, q$. We look for a coupling between the distributions of the $\mathbb{R}^{qN}$-valued random vectors $(x_i^{(j)})$ and $(x_i(jh))$ which minimises the 'distance' between these two random vectors, using a suitable measure of distance between two $\mathbb{R}^{qN}$-valued random vectors.

One measure of distance, which is commonly used for discrete approximations to SDE, is defined as follows: given $\mathbb{R}^{qN}$-valued random vectors $x = (x^{(j)})_{j=1}^N$ and $y = (y^{(j)})_{j=1}^N$ on the same probability space, where $x^{(j)}$ and $y^{(j)}$ are $\mathbb{R}_q$-valued, we define

$$d_{2,\infty}(x, y) = \{\mathbb{E}(\max_{j=1}^N |x^{(j)} - y^{(j)}|^2)\}^{1/2} \tag{37}$$

and then for example the bound (6) for the error of the Euler approximation can be written $d_{2,\infty}((\tilde{x}^{(j)}), (x(jh))) = O(h^{1/2})$.

Then for two $\mathbb{R}^{qN}$-valued random variables (in general on different probability spaces) we can define a Vaserstein-type distance

$$\mathbb{W}_{2,\infty}(x,y) = \inf d_{2,\infty}(x,y)$$

the infimum being between all couplings between (the distributions of) x and y. The main results of this paper can be regarded as bounds for $\mathbb{W}_{2,\infty}((\tilde{x}^{(j)}),(x(jh)))$ for various schemes for generating approximations $(\tilde{x}^{(j)})$.

One can argue that a bound in $\mathbb{W}_{2,\infty}$ for an approximation scheme is in fact equivalent to a standard strong approximation bound, as follows: suppose we have a bound $\mathbb{W}_{2,\infty}((\tilde{x}^{(j)}),(x(jh))) = O(h^\gamma)$. This means that there is an $\mathbb{R}^{qN}$-valued random variable $Z = (Z^{(j)})$, on the same probability space as the solution $x(t)$, having the same distribution as $(\tilde{x}^{(j)})$, and such that $d_{2,\infty}(Z,(x(jh))) = O(h^\gamma)$. When we implement the scheme and generate $(\tilde{x}^{(j)})$, we can regard the process as generating Z, which approximates the solution in strong sense.

This means that for simulating individual solution trajectories approximation in $\mathbb{W}$ sense is equivalent to strong approximation. We now consider to what extent this holds for some other applications of strong approximation.

One such application is to the simulation of the stochastic flow defined by an SDE, in which one simulates a number R of trajectories of the SDE, with the same driving Brownian path, but with R different starting points (some examples can be found in section 13.1 of [16]). One can think of this as a single simulation of an SDE for Rq variables, driven by d-dimensional path. For such an application the method of section 7 will fail, as the SDE in Rq variables will usually be degenerate if $R > 1$, even if the original SDE in q variables is nondegenerate. On the other hand the Fourier method of section 9 will work because we can generate the approximate double integrals at each step, and use the same ones for each trajectory.

The application of two-level approximation to the multilevel weak approximation method has been discussed in section 8. It is not clear how two-level approximation could readily be achieved using our version of the Fourier method (this limitation also applies to the other versions of the Fourier method in the literature). It would be highly desirable to find higher-order methods which are valid for degenerate diffusions and for which two-level approximation is possible.

One can similarly define $\mathbb{W}_{p,\infty}$ metrics; we have worked mainly with $p = 2$ as it is the easiest to work with. In practice $p = 1$ bounds would usually be sufficient but one needs L^2 bounds for the local errors to get efficient bounds for global errors. It is also not hard to get bounds for $p > 2$.

We have concentrated on approximation over the whole path but sometimes one may be interested only in simulating the final value $x(T)$, and if X is an approximate simulation of $x(T)$ we can consider Vaserstein distances $\mathbb{W}_p(X, x(T))$. One can of course deduce bounds for such distances from whole-path bounds, but with some methods (especially using non-local couplings) it may be possible to get better bounds for the final time than for the whole path. In this connection we mention a connection with weak approximation bounds - it follows from the Kantorovich-Rubinstein theorem (see, e.g., Theorem 1.14 of [26]) that

$$\mathbb{W}_1(X, x(T)) = \sup |\mathbb{E}f(X) - \mathbb{E}f(x(T))|$$

where the supremum is over all functions $f : \mathbb{R}^q \to \mathbb{R}$ satisfying $|f(u) - f(v)| \le |u - v|$ for all $u, v \in \mathbb{R}^q$.

References

[1] A. Alfonsi, B. Jourdain and A. Kohatsu-Higa, Pathwise optimal transport bounds between a one-dimensional diffusion and its Euler scheme, *Ann. Appl. Probab* **24** (2014), 1049-1080.

[2] A. Alfonsi, B. Jourdain and A. Kohatsu-Higa, Optimal transport bounds between the time-marginals of a multidimensional diffusion and its Euler scheme, *arXiv*:1405.7007.

[3] A. B. Cruzeiro, P. Malliavin and A. Thalmaier, Geometrization of Monte-Carlo numerical analysis of an elliptic operator: strong approximation, C. R. Math. Acad. Sci. Paris **338** (2004), 481-486.

[4] R. N. Bhattacharya and E. C. Waymire, *Stochastic Processes with Applications*, Wiley 1990.

[5] B. Charbonneau, Y. Svyrydov and P. Tupper, Weak convergence in the Prokhorov metric of methods for stochastic differential equations, *IMA J. Numer. Anal.* **30** (2010), 579-594.

[6] N. Fournier, Simulation and approximation of Lévy-driven SDEs, *ESIAM Probab. Stat.* **15** (2011), 233-248.

[7] J. G. Gaines, The algebra of iterated stochastic integrals, *Stochastics Stochastics Reports* **49** (1994), 169-179.

[8] J. G. Gaines and T. J. Lyons, Random generation of stochastic area integrals, *SIAM J. Appl. Math.* **54** (1994), 1132-1146.

[9] J. G. Gaines and T. J. Lyons, Variable step size control in the numerical solution of stochastic differential equations, *SIAM J. Appl. Math.* **57** (1997), 1455-1484

[10] M. B. Giles, Multilevel Monte Carlo path simulation, *Operations Research* **56** (2008), 607-617.

[11] M. B. Giles, Improved multilevel Monte Carlo convergence using the Milstein scheme, *Monte Carlo and quasi-Monte Carlo methods 2006*, 343-358, Springer, Berlin 2008.

[12] M. B. Giles and L. Szpruch, Antithetic multilevel Monte Carlo estimation for multi-dimensional SDEs without Lévy area simulation, arXiv:1202.6283.

[13] D. J. Higham, X. Mao and A. M. Stuart, Strong convergence of Euler-type methods for nonlinear stochastic differential equations, *SIAM J. Numer. Anal.* **40** (2002), 1041-1063.

[14] D. J. Higham, X. Mao and C. Yuan, Almost sure and moment exponential stability in the numerical simulation of stochastic differential equations, *SIAM J. Numer. Anal.* **45** (2007), 592-609.

[15] S. Kanagawa, The rate of convergence for approximate solutions of SDEs, *Tokyo J. Math* **12** (1989), 33-48.

[16] P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations*, Springer-Verlag 1995.

[17] P. E. Kloeden, E. Platen and I. Wright, The approximation of multiple stochastic integrals, *J. Stoch. Anal. Appl.* **10** (1992), 431-441.

[18] I. Gyöngy and N. Krylov, Existence of strong solutions for Itô's stochastic equations via approximations, *Probab. Theory Related Fields* **105** (1996), 143-158.

[19] J. Komlós, P. Major and G. Tusnády, An approximation of partial sums of independent RV's and the sample DF. I, *Z. Wahr. und Wer. Gebiete* **32** (1975), 111-131.

[20] Q. Li, *Numerical Approximation for SDE*, PhD thesis, University of Edinburgh 2007.

[21] S. T. Rachev and L. Ruschendorff, *Mass Transportation Problems*, Volume 1, Theory; Volume 2, Applications. Springer-Verlag 1998.

[22] E. Rio, Upper bounds for minimal distances in the central limit theorem, *Ann. Inst. Henri Poincaré Probab. Stat.* **45** (2009), 802-817.

[23] T. Rydén and M. Wiktorsson, On the simulation of iterated Itô integrals, *Stochastic Processes Appl.* **91** (2001), 151-168.

[24] S. Sabanis, A note on tamed Euler approximations, *Electron. Commun. Probab.* **18** (2013), no. 47.

[25] L. N. Vaserstein, Markov processes over denumerable products of spaces describing large system of automata (Russian), *Problemy Peredaci Informacii* **5** (1969), 64-72.

[26] C. Villani, *Topics in Optimal Transportation*, AMS 2003.

[27] C. Villani, *Optimal Transport Old and New*, Springer-Verlag 2009.

[28] M. Wiktorsson, Joint characteristic function and simultaneous simulation of iterated Itô integrals for multiple independent Brownian motions, *Ann. Appl. Probab.* **11** (2001), 470-487.