Weak Error Analysis and Small-Noise Optimization of Stochastic Linear Multistep Methods (SLMMs) for Itô and Lévy SDEs

Anay Aggarwal

Under the direction of

Joonsoo Lee
Massachusetts Institute of Technology

Research Science Institute
July 29, 2025

Abstract

We study the weak convergence behavior of stochastic linear multistep methods (SLMMs) for solving stochastic differential equations (SDEs). While SLMMs have been previously investigated in terms of strong error, their weak convergence properties remain largely unexplored. We develop a general framework for computing the full weak error expansion of arbitrary SLMMs using the Kolmogorov backward equation and linear recurrence techniques. This framework enables a systematic analysis of both existing and newly proposed multistep methods. In particular, we construct SLMMs that achieve improved weak accuracy in small noise regimes. We extend our analysis to Lévy-driven SDEs and derive precise weak error bounds for a class of multistep schemes that interpolate drift and diffusion terms. Numerical simulations confirm the theoretical predictions and illustrate regimes in which multistep methods significantly outperform standard one-step approaches in weak approximation accuracy.

Summary

Many real-world systems, from stock prices to weather patterns, are influenced by randomness, which makes their behavior impossible to predict exactly. To understand and forecast these systems, we use computer simulations based on mathematical models called *stochastic differential equations*. These simulations work by taking small steps forward in time, using known information to estimate what happens next. One-step methods only use the current state of the system to make predictions about the next state. This research focuses on more advanced "multistep methods," which also take into account past behavior. We developed a new mathematical framework to analyze how accurate these multistep methods are, especially in cases where randomness plays a small but important role. Using this framework, we proved that several key methods are not only reliable, but can also outperform one-step approaches. These results give scientists more powerful tools to model uncertainty and make better predictions in complex, unpredictable environments.

1 Introduction

Stochastic differential equations (SDEs) are fundamental tools for modeling systems influenced by randomness. They arise naturally in physics, biology, quantitative finance, and other domains where uncertainty plays a critical role. A typical one-dimensional Itô SDE is of the form

$$dX_t = f(X_t, t)\,dt + g(X_t, t)\,dW_t,$$

where f and g are the drift and diffusion coefficients, respectively, and W_t is a standard Brownian motion. Since closed-form solutions are rare, such equations are typically studied through numerical approximation.

Classical one-step methods for numerical approximation, such as Euler–Maruyama and Milstein, are widely used and well understood. Their strong and weak convergence properties are documented extensively; see, for example, [1, Chapter 5] and [2, Chapter 14]. However, one-step methods rely only on information at the current timestep, potentially limiting their accuracy and stability.

By contrast, multistep methods incorporate information from several previous timesteps and are central to the numerical solution of deterministic ordinary differential equations (ODEs) due to their improved efficiency and accuracy. The stochastic analogues—stochastic linear multistep methods (SLMMs)—offer a promising yet underexplored alternative in the stochastic setting. While SLMMs can improve drift approximation without requiring higher-order stochastic integrals, challenges with stability and convergence have limited their widespread adoption. Prior work has primarily focused on strong convergence for specific two-step SLMMs under restrictive assumptions [3], with comparatively little attention given to weak convergence.

Weak approximation is of particular importance in applications where expectations of functionals of the solution are of interest, such as option pricing in finance or moment estimation in stochastic models of population dynamics [4]. This motivates a systematic study of the weak behavior of SLMMs.

Moreover, many real-world systems experience not just continuous fluctuations but also sudden jumps, which are better captured by SDEs driven by Lévy processes (processes like the Wiener process but with discontinuous jumps). The analysis of weak convergence for such jump-driven systems is more delicate. Existing results on Lévy-driven SDEs focus largely on one-step schemes; see [5], [6], and [7] for surveys and algorithmic developments, and [8], [9] for results on strong approximation.

This paper develops a general framework to analyze the weak error of SLMMs, both in the Brownian and Lévy-driven settings. The approach is based on using the Kolmogorov backward equation to turn quantities based on SDEs into quantities based on elliptic partial differential equations (PDEs) and yields a systematic algorithm for computing the full weak error expansion of any SLMM. We apply this framework to existing two-step methods, derive new multistep schemes optimized for small noise regimes, and extend the analysis to Lévy-driven systems. Numerical simulations confirm the theoretical results and demonstrate regimes in which multistep schemes significantly outperform one-step methods in terms of weak error.

We review background on weak approximation and SLMMs in Section 2. In Section 3, we introduce the general algorithm for computing weak error expansions. We apply the theory to specific schemes, including Adams–Bashforth variants, and derive their weak order behavior in Section 4, showing that, in the small-noise case, SLMMs provide asymptotic improvement over their one-step analogs. In Section 5, we extend the framework to Lévy-driven SDEs. Finally, in Section 6, we present numerical results that validate the theory and highlight practical implications. Details of straightforward proofs are omitted.

2 Preliminaries

In order to approximate solutions to an SDE, we define a discrete temporal grid on $I = [0, T]$, with constant stepsize h, of the form $0 = t_0 < t_1 < t_2 < \cdots < t_N = T$. We define h to be T/N so that $t_\ell = \ell \cdot h$ for $\ell = 0, 1, \ldots, N$. We make discrete approximations $X_{t_0}, X_{t_1}, \ldots, X_{t_N}$. For brevity, we denote $X_i := X_{t_i}$ for indices i. A k-step SLMM takes the form (for $\ell \geq k$)

$$\sum_{j=0}^{k} \alpha_j X_{\ell-j} = h \sum_{j=0}^{k} \beta_j f(X_{\ell-j}, t_{\ell-j}) + \sum_{j=1}^{k} \Gamma_j(X_{\ell-j}, t_{\ell-j}) I^{t_{\ell-j}, t_{\ell-j+1}},$$

where the α_j and β_j are constants. We set $\alpha_0 = 1$. Every diffusion term $\Gamma_j(x,t)\, I^{t,t+h}$ is a finite sum of terms each containing an appropriate function q of x and t multiplied by a multiple Wiener integral over $[t, t+h]$, i.e., it takes the general form

$$\Gamma_j(x,t)\, I^{t,t+h} = q_j(x,t)\, I_r^{t,t+h} + q_j(x,t)\, I_{r_1,r_2}^{t,t+h} + \ldots$$

A general multiple Wiener integral is given by

$$I_{r_1,r_2,\ldots,r_j}^{t,t+h}(y) = \int_t^{t+h} \int_t^{s_1} \cdots \int_t^{s_{j-1}} y\left(X_{s_j}, s_j\right)\, dW_{r_1}(s_j) \cdots dW_{r_j}(s_1),$$

where $r_i \in \{0, 1, \ldots, m\}$ and $dW_0(s) = ds$. If $y \equiv 1$ we omit y and write $I_{r_1,r_2,\ldots,r_j}^{t,t+h}$. Note that the integral $I_r^{t,t+h}$ is simply the increment $W_r(t+h) - W_r(t)$ of the scalar Wiener process W_r. The term $I^{t,t+h}$ denotes the collection of multiple Wiener integrals associated with the interval $[t, t+h]$.

We introduce two common SLMMs defined in [3] that will be studied later in the paper.

Example 2.1. The two-step Milstein-Adams-Bashforth method (MAB2) is given by the relation

$$X_\ell - X_{\ell-1} = h\left(\frac{3}{2}f(X_{\ell-1}, t_{\ell-1}) - \frac{1}{2}f(X_{\ell-2}, t_{\ell-2})\right) + g(X_{\ell-1}, t_{\ell-1})\, I_1^{t_{\ell-1},t_\ell} \\ + (g\partial_x g)(X(t_{\ell-1}), t_{\ell-1})\, I_{1,1}^{t_{\ell-1},t_\ell}.$$

Example 2.2. Two-step Maruyama schemes take the general form

$$\sum_{j=0}^{2} \alpha_j X_{\ell-j} = h\sum_{j=0}^{2} \beta_j f(X_{\ell-j}, t_{\ell-j}) + \sum_{j=1}^{2} \gamma_j g(X_{\ell-j}, t_{\ell-j})\, I_1^{t_{\ell-j},t_{\ell-j+1}}$$

for coefficients $\alpha_i, \beta_i, \gamma_i$.

Let us denote the mean square norm of a random variable as $||Z||_{L_2} = (\mathbb{E}[|Z|^2])^{1/2}$. A scheme is mean square convergent with order γ if

$$\max_{\ell=1,2,\ldots,N} ||X(t_\ell) - X_\ell||_{L_2} \leq Ch^\gamma$$

for some constant C independent of our partition of I, where $X(t_\ell)$ denotes the true solution to the SDE at $t = t_\ell$ and X_ℓ is our approximation.

Buckwar and Winkler show in [3] that two-step Maruyama schemes are mean-square convergent with weak order $\frac{1}{2}$, assuming certain consistency coefficients on $\alpha_i, \beta_i, \gamma_i$ (and that even higher accuracy can be obtained when the noise term is small). This is done by analyzing the local error accumulated at each step, then computing global error based on local errors.

The main quantity we study is the weak error

$$\mathrm{Err}(T, h) = |\mathbb{E}[\phi(X_T)] - \mathbb{E}[\phi(X_{t_N})]|$$

for a smooth function $\phi : \mathbb{R} \to \mathbb{R}$. The scheme has weak order γ if $\mathrm{Err}(T, h) \leq C_\phi h^\gamma$ for some constant C_ϕ only dependent on ϕ. We call a *weak error expansion* an equation of the form

$$\mathrm{Err}(T, h) = c_0 h^\gamma + c_1 h^{\gamma+1} + \cdots + c_r h^{\gamma+r} + O(h^{\gamma+r+1}),$$

where all the constants c_i are dependent on the function ϕ used. The weak error expansion of one-step schemes is understood. Talay and Tubaro [10] compute the weak error expansion of classical one-step schemes (such as Euler-Maruyama), and Iguchi and Yamada [11] compute the weak error expansion of a more involved scheme. Due to the Feynman-Kac formula [12] the main idea is to use the solution $u(t,x)$ to the Kolmogorov Backwards Partial Differential Equation (KB-PDE) [10]:

$$\begin{cases} \partial_t u + \mathcal{L}u = 0, \\ u(T,x) = \phi(x). \end{cases}$$

Here, $\mathcal{L}$ is the infinitesimal generator of the SDE. The solution to this PDE is

$$u(t,x) = \mathbb{E}[\phi(X_t)|X_0 = x].$$

Therefore,

$$\mathrm{Err}(T,h) = u(T, X_{t_n}) - u(0, X_0) = \sum_{i=1}^{n} u(t_i, X_{t_i}) - u(t_{i-1}, X_{t_{i-1}}).$$

In other words, to understand the error in approximation of our scheme, we simply need to understand the behavior of u under small perturbations dependent on the scheme. We understand the one-step case; we can bound $u(t_\ell + h, X_\ell + (X_{\ell+1} - X_\ell))$ using a two-variable Taylor expansion or other related methods (for example using integration and the Itô-Taylor expansion). The difficulty with multistep methods is that the differences $X_{\ell+1} - X_\ell$ are harder to track, because they may not be directly expressed using terms dependent on f and g. We overcome this difficulty using linear recurrence theory.

We require consistency conditions on the SDE approximation schemes used in order for the schemes to attain a certain weak error expansion. One of these conditions is the *Dahlquist root condition*, introduced by Dahlquist in [13] for Ordinary Differential Equations (ODEs). Buckwar and Winkler deduced in [3] that this condition must hold for multistep approximation schemes in order for them to attain a certain mean square convergence order. The SLMM (2.1) fulfills this condition if the roots of $p(\zeta) = \alpha_k \zeta^k + \alpha_{k-1}\zeta^{k-1} + \cdots + \alpha_0$ lie on or within the unit circle, and the roots that are on the unit circle are simple. This condition is useful because it ensures that a sequence a_n satisfying $a_n = \sum_{j=1}^{k} \alpha_j a_{n-j}$ does not grow to infinity.

We additionally study Lévy-driven stochastic differential equations. The general form of a Lévy-driven SDE is

$$dX_t = h(X_{t-})\, dZ_t,$$

where $h : \mathbb{R} \to \mathbb{R}$ is a measurable function and Z_t is a real-valued Lévy process, which may include both continuous and jump components. By the Lévy-Itô decomposition and the Lévy-Khintchine characterization [14], we can equivalently express such SDEs in the more explicit form:

$$dX_t = f(X_{t-})\, dt + g(X_{t-})\, dW_t + \int_{|z|<1} \gamma(X_{t-}, z)\, \tilde{J}(dt, dz) + \int_{|z|\geq 1} \delta(X_{t-}, z)\, J(dt, dz),$$

where:

- $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ are the drift and diffusion coefficients,
- $\tilde{J}(dt, dz) = J(dt, dz) - \nu(dz)\, dt$ is the compensated Poisson random measure for small jumps $|z| < 1$,
- $\gamma : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is the small-jump coefficient,
- $\delta : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is the large-jump coefficient, and
- $\nu(dz)$ is the Lévy measure, which governs the intensity and distribution of the jumps.

This representation separates the contributions of continuous and discontinuous components and enables precise analysis and numerical approximation.

One may define a k-step SLMM for Lévy-driven SDEs in the same way as was done for continuous SDEs. With the same grid on I, the general form is

$$\sum_{j=0}^{k} \alpha_j X_{\ell-j} = h \sum_{j=0}^{k} \beta_j f(X_{\ell-j}, t_{\ell-j}) + \sum_{j=1}^{k} \Gamma_j (X_{\ell-j}, t_{\ell-j}) I^{t_{\ell-j}, t_{\ell-j+1}} + \sum_{j=1}^{k} \Xi_j (X_{\ell-j}, t_{\ell-j}).$$

Here the α_i, β_i, and Γ_i are as before. The functions Ξ_j are any functions meant to approximate the jump terms (we do not specify a general form).

Example 2.3. The Euler-Maruyama scheme for Lévy Driven SDEs, a 1-step SLMM, is given by

$$\begin{aligned} X_n &= X_{n-1} + hf(X_n, t_n) + \Delta W_n g(X_n, t_n) \\ &\quad + \int_{|z|<1} \gamma(X_n, z) \tilde{J}((t_n, t_{n+1}], dz) + \int_{|z|\geq 1} \delta(X_n, z) J((t_n, t_{n+1}], dz). \end{aligned}$$

A common example of a nonlinear choice for Ξ_j is that made by the Asmussen-Rosiński approximation [15], in which the small jumps are approximated by Brownian motion.

The weak error expansion of Lévy-driven SDEs can be understood in the same sense as the continuous case. The KB-PDE is still well-defined, and one can again turn the question of weak error into a question about elliptic PDEs.

3 Algorithm for Weak Error Estimation

To obtain a weak error expansion for an SLMM, we must first estimate the local deviations $X_{i+1} - X_i$. Denote $\Delta x_\ell = X_\ell - X_{\ell-1}$. We say that $X \overset{\mathbb{E}}{=} Y$ for two random variables X, Y if

$\mathbb{E}[X] = \mathbb{E}[Y]$. We can then Taylor expand

$$u(t+h, x+\Delta x_\ell) - u(t,x) \overset{\mathbb{E}}{=} \sum_w \frac{\Delta w}{|w|!} \partial_w u$$
$$\overset{\mathbb{E}}{=} \sum_{a \geq 1} \sum_{b \geq 1} \frac{1}{a! b!} h^a (\Delta x_\ell)^b \partial_t^a \partial_x^b u. \tag{2.1}$$

Here, the summation in the first line is over all words w taken from the alphabet $\mathcal{A} = \{t, x\}$. The quantity Δw is the product of all Δc over characters c in w. One can define $\partial_w u$ in a similar fashion. Notice that the partial derivatives commute by smoothness assumptions on u. Therefore, understanding the deviation in $u(t,x)$ comes down to understanding moments $\mathbb{E}[\Delta x_\ell^b]$ of Δx_ℓ. We may compute these moments using linear recurrence theory, in particular, the following lemma.

Lemma 3.1. *Let k be a natural number, and let a_n and b_n be sequences for indices $n \geq -k+1$ such that, for some constants c_j, sequence a_n satisfies the k-step linear recurrence*

$$a_n = b_n + \sum_{j=1}^{k} c_j a_{n-j}.$$

Then the general solution is given by

$$a_n = d_n + \sum_{j=0}^{n} G_{n-j} b_j,$$

where

1. *Sequence G_n satisfies $G_n = \sum_{i=1}^{k} c_i G_{n-i}$ with $G_0 = 1$ and $G_j = 0$ for $-k+1 \leq j \leq -1$.*
2. *Sequence d_n satisfies $d_n = \sum_{j=1}^{k} c_j d_{n-j}$ with $d_0 = a_0$, $d_{-1} = a_{-1}$, ..., $d_{-k+1} = a_{-k+1}$.*

Proof. It can be routinely checked that the initial conditions are satisfied. Define the linear operator L such that $L(a_n) = a_n - \sum_{j=1}^{k} c_j a_{n-j}$. Then $L(a_n) = b_n$, $L(d_n) = 0$, and $L(G_n) = \delta_{n,0}$, where δ is the Kronecker delta function. We wish to show that $L(d + G * b)_n = 0$, where $*$ is the convolution operator defined as

$$(x * y)_n = \sum_{j=0}^{n} x_j y_{n-j}$$

for two sequences x_n and y_n. But $L(d + G * b)_n = Ld_n + L(G * b)_n = L(G * b)_n$, and

$$L(G * b)_n = \sum_{j=0}^{n} (LG)_{n-j} b_j = \sum_{j=0}^{n} \delta_{n-j,0} b_j = b_n,$$

as desired. ■

In order to obtain a recurrence for the Δx_ℓ to apply Lemma 3.1 in order to estimate deviations in u, we need the following criterion on a linear combination of the X_i.

Lemma 3.2. *For any random variables $X_0, X_1, X_2, \ldots, X_n$ and integer $k < n$, we can write*

$$X_n - \alpha_1 X_{n-1} - \cdots - \alpha_k X_{n-k} = \Delta x_n + \sum_{i=1}^{k-1} \gamma_i \Delta x_{n-i},$$

for constants γ_i if and only if $\sum_{i=1}^{k} \alpha_i = 1$, where $\Delta x_\ell = X_\ell - X_{\ell-1}$.

Proof. The proof is straightforward. ■

Lemma 3.2 gives us a recurrence $\Delta x_n = \sum_{j=0}^{n} \gamma_j \Delta x_{n-j+1} + e_n$, where

$$e_n = h \sum_{j=0}^{k} \beta_j f(X_{\ell-j}, t_{\ell-j}) + \sum_{j=1}^{k} \Gamma_j (X_{\ell-j}, t_{\ell-j}) I^{t_{\ell-j}, t_{\ell-j+1}}.$$

Note that, because multiple stochastic integrals have mean 0,

$$e_n \overset{\mathbb{E}}{=} h \sum_{j=0}^{k} \beta_j f(X_{\ell-j}, t_{\ell-j}).$$

Therefore, by Lemma 3.1 we may obtain a decomposition

$$\Delta x_n = d_n + \sum_{j=0}^{n} G_{n-j} e_j.$$

Hence, applying expectations,

$$\Delta x_n^b \overset{\mathbb{E}}{=} \left(d_n + \sum_{j=0}^{n} G_{n-j} e_j \right)^b .$$

This expression can be computed with an explicit expansion. All that is needed is to compute moments of e_j, which boils down to computing moments of multiple Itô integrals, which can be done by the following theorem.

Theorem 3.3 (Corollary 5.2.4 of [2])**.** *The multiple Itô integral*

$$I_{1,1,\ldots,1}^{t,t+h} = \mathcal{H}_\alpha (I_1^{t,t+h}),$$

where there are α ones in the subscript of the I on the left hand side of the equation, and $H_\alpha(x) = (-1)^\alpha e^{x^2/2} \frac{d^\alpha}{dx^\alpha} e^{-x^2/2}$ is a Hermite polynomial.

Computing these moments leads us to our main theorem.

Theorem 3.4. *A weak error expansion of any SLMM with error term $O(h^{\gamma+r+1})$ for arbitrarily high r may be computed explicitly algorithmically.*

Proof. As described by (2.1), we may express local deviations in u in terms of linear combinations of differentials of u. Truncating this infinite sum at finite bounds gives us an asymptotic expansion. One can simplify the sum of differentials using $\partial_t u = -\mathcal{L}u$ to obtain the exact expansion, noting that in general there are no other relations between the partial derivatives of u besides what is given by $\partial_t u = -\mathcal{L}u$ and the partial derivatives of this equation. The local deviations may then be summed to result in a weak order expansion of $\mathrm{Err}(T, h)$, as required. ■

Note that this method has the added advantage that it does not require smoothness conditions on f and g like the results in [3] do.

The computation of moments $\mathbb{E}[\Delta x_n^b]$ increases exponentially in complexity as b increases, so one may choose to instead bound these moments from above. This can be done through the use of Minkowski's inequality. Each e_i can be decomposed as follows

$$e_i = \underbrace{h\sum_{j=0}^{k}\beta_j f(X_{i-j}, t_{i-j})}_{e_{i,0}} + \underbrace{I_1^{0,h}\sum_{j=1}^{k} q_j(X_{i-j}, t_{i-j})}_{e_{i,1}} + \underbrace{I_{1,1}^{0,h}\sum_{j=1}^{k} q_j(X_{i-j}, t_{i-j}) I_{1,1}^{t_{i-j},t_{i-j}+h}}_{e_{i,2}} + \cdots$$

Hence

$$\Delta x_n = c_{n,0}h + c_{n,1}I_1^{0,h} + c_{n,2}I_{1,1}^{0,h} + \cdots$$

for some constants $c_{n,\ell}$ depending on the sequence G and $e_{i,r}$. Therefore, by Minkowski's inequality,

$$(\mathbb{E}[\Delta x_n^b])^{1/b} \le \left(\sum_{\ell\ge 0} c_{n,\ell}\left(\mathbb{E}|I_{1,1,\ldots,1}^{0,h}|^b\right)^{1/b}\right).$$

By Theorem 3.3, each of the moments $\mathbb{E}|I_{1,1,\ldots,1}^{0,h}|^b$ can be computed using the moments of $I_1^{0,h}$, a normal random variable with mean 0 and variance h. This gives one an upper bound on the moments of Δx_n, which results in an upper bound on the coefficients of the weak order expansion.

4 Weak Order Expansion of Certain Multistep Schemes

We now apply theory developed in Section 3 to previously studied SLMMs and derive weak error behavior for a specific class of SLMMs tailored to small noise SDEs.

4.1 2-Step Schemes

We will use Theorem 3.4 to explicitly compute the weak error expansions of the schemes in [3]: MAB2 and two-step Maruyama schemes. These methods were presented in Examples 2.1 and 2.2.

In the case of MAB2, we will denote $f^*(X_n, t_n) = \frac{3}{2}f(X_n, t_n) - \frac{1}{2}f(X_{n-1}, t_{n-1})$ and $g^* = g\partial_x g$. We can define $f^*(x, t)$ on other values of x, t such that f^* is piecewise linear. Before we compute the weak order expansion, we will need to compute the moments of Δx_n.

Lemma 4.1. *For the MAB2 scheme, the moments are*

$$
\begin{aligned}
\mathbb{E}[\Delta x_n] &= hf^*(X_{n-1}, t_{n-1}),\\
\mathbb{E}[\Delta x_n^2] &= h^2 f^*(X_{n-1}, t_{n-1})^2 + hg(X_{n-1}, t_{n-1}) + \tfrac{1}{2}h^2 g^*(X_{n-1}, t_{n-1})^2,\\
\mathbb{E}[\Delta x_n^3] &= 3h^2 f^*(X_{n-1}, t_{n-1})g(X_{n-1}, t_{n-1})^2 + O(h^3),\\
\mathbb{E}[\Delta x_n^4] &= 3h^2 g^*(X_{n-1}, t_{n-1})^4 + O(h^3),\\
\mathbb{E}[\Delta x_n^r] &= O(h^3)
\end{aligned}
$$

for any natural number $r \geq 5$.

The proof follows by straightforward properties of the moment generating function of the normal distribution, so we omit it.

This lemma allows us to compute the weak error expansion using Theorem 3.4.

Theorem 4.2. *A weak error expansion for the MAB2 scheme is*

$$
\mathrm{Err}^{MAB2}(T, h) = h/2f(X_{T/h}, T) + h^2 \int_0^{T/h} \mathbb{E}[\Phi^{MAB2}(s, X_s)]\ ds + O(h^2),
$$

where

$$
\Phi^{MAB2}(t, x) = f^*\partial_{tx}u + \frac{1}{2}\partial_{tt}u + \frac{1}{2}(f^{*2} + g^{*2})\partial_{xx}u + \frac{1}{2}g\partial_{txx}u + \frac{1}{2}f^*g^2\partial_{xxx}u + \frac{1}{8}g^{*4}\partial_{xxxx}u.
$$

Proof. Let $u(t, x)$ be the solution to the KB-PDE for some function ϕ. The weak error is then

$$
\sum_{n=1}^{T/h} u((n-1)h + h, X_{n-1} + (X_n - X_{n-1})) - u((n-1)h, X_{n-1}).
$$

By a Taylor series expansion, we can simplify local deviations in u:

$$
\begin{aligned}
u(t+h, x+\Delta x_n) - u(t,x) &\overset{\mathbb{E}}{=} \sum_w \frac{\Delta w}{|w|!}\partial_w u \\
&\overset{\mathbb{E}}{=} \sum_{a\geq 1}\sum_{b\geq 1} \frac{1}{a!b!} h^a (\Delta x)^b \partial_t^a \partial_x^b u \\
&\overset{\mathbb{E}}{=} h\partial_t u + \Delta x_n \partial_x u + \frac{1}{2}(\Delta x_n)^2 \partial_{xx} u + \Delta x_n h \partial_{tx} u \\
&\quad + \frac{1}{2}h^2\partial_{tt}u + + \frac{1}{2}h\Delta x^2 \partial_{txx}u + \frac{1}{6}\Delta x^3 \partial_{xxx} u + \frac{1}{8}\Delta x^4 \partial_{xxxx} u + O(h^3),
\end{aligned}
$$

where $O(h^3)$ is the error term by Lemma 4.1. Substituting in the moments of Δx_n, we get

$$
u(t+h, x+\Delta x_n) - u(t,x) \overset{\mathbb{E}}{=} h(\partial_t u + \frac{1}{2}g^2 \partial_{xx} u + f^* \partial_x u) + h^2 \Phi^{MAB2}(t,x) + O(h^3).
$$

Notice that because $\partial_t u + \frac{1}{2}g^2\partial_{xx}u + f\partial_x u = \partial_t u + \mathcal{L}u = 0$, we have

$$
\begin{aligned}
\operatorname{Err}(T,h) &= \sum_{n=1}^{T/h} u(nh, X_{n-1}+\Delta x_n) - u((n-1)h, X_{n-1}) \\
&= O(h^2) + \sum_{n=1}^{T/h} h/2(f(X_{n-1},t_{n-1}) - f(X_{n-2},t_{n-2})) + \sum_{n=1}^{T/h} h^2 \Phi^{MAB2}(t_n, X_n) \\
&= h/2 f(X_{T/h}, T) + h^2 \int_0^{T/h} \Phi^{MAB2}(s, X_s)\, ds + O(h^2),
\end{aligned}
$$

as desired. ■

Due to the boundedness of partial derivatives of u, we have the following corollary.

Corollary 4.3. *The MAB2 scheme has weak order* 1.

We have a similar result for two-step Maruyama schemes assuming appropriate consistency conditions. First, as for MAB2, we compute the moments of Δx_n. Notice that $X_n = \alpha_1 X_{n-1} + \alpha_2 X_{n-2} + e_n$ for a sequence e_n. We may thus express X_n explicitly in terms of e_n, a sequence G_n, and a sequence d_n in the form

$$
X_n = d_n + \sum_{j=0}^{n} G_{n-j} e_j,
$$

by Lemma 3.2. Let $G_n^* = G_n - G_{n-1}$, and define $G_0^* = 1$. Denote $f_n = f(X_n, t_n)$ and $g_n = g(X_n, t_n)$. Additionally, let

$$
\Lambda_n^f = \sum_{j=0}^{n} G_{n-j}^* (\beta_1 f_{j-1} + \beta_2 f_{j-2}),
$$

$$\Lambda_n^g = \sum_{j=0}^{n} G_{n-j}^*(\gamma_1 g_{j-1} + \gamma_2 g_{j-2}).$$

We define f and g at negative indices to be 0. With this notation, we may express Δx_n explicitly with the following lemma.

Lemma 4.4. *For two-step Maruyama schemes, we have*

$$\Delta x_n = h\Lambda_n^f + \Delta W_n \Lambda_n^g.$$

Proof. This follows by Lemma 3.2 and the fact that $e_n = h(\beta_1 f_{n-1} + \beta_2 f_{n-2}) + \Delta W_n(\gamma_1 g_{n-1} + \gamma_2 g_{n-2})$, as well as the assumption that $d_n - d_{n-1} = 0$ (which one can make by specifying initial values of X to be the same). ■

This allows us to concisely calculate moments of Δx_n with the following lemma.

Lemma 4.5. *For two-step Maruyama schemes, we have the moments*

$$\begin{aligned}
\mathbb{E}\Delta x_n &= h\Lambda_n^f, \\
\mathbb{E}\Delta x_n^2 &= h^2\Lambda_n^f + h\Lambda_n^g, \\
\mathbb{E}\Delta x_n^3 &= 3h^2{\Lambda_n^f}^2 + O(h^3), \\
\mathbb{E}\Delta x_n^4 &= 3h^2{\Lambda_n^g}^4 + O(h^3),
\end{aligned}$$

and $\mathbb{E}\Delta x_n^r = O(h^3)$ *for all* $r \geq 5$.

The proof is identical to the proof of Lemma 4.1 and follows by Lemma 4.4.

Finally, we obtain the weak error expansion.

Theorem 4.6. *The error for two-step Maruyama schemes is*

$$\mathrm{Err}(T, h) = O(h^2) + \sum_{n=0}^{T/h} R_{h,n},$$

where

$$\begin{aligned}
R_{h,n} &= h\partial_t u + h\Lambda_n^f \partial_x u + \frac{1}{2}(h^2\Lambda_n^f + h\Lambda_n^g)\partial_{xx} u + h^2\Lambda_n^f \partial_{tx} u \\
&+ \frac{1}{2}h^2\partial_{tt} u + \frac{1}{2}h^2\Lambda_n^g \partial_{txx} u + \frac{1}{2}h^2{\Lambda_n^f}^2 \partial_{xxx} u + \frac{1}{8}h^2{\Lambda_n^g}^4 \partial_{xxxx} u.
\end{aligned}$$

Proof. We proceed in a similar manner as in Theorem 4.2. Let $u(t, x)$ be the solution to the KB-PDE for some function ϕ. The weak error is then

$$\mathrm{Err}(T,h)=\sum_{n=1}^{T/h}u((n-1)h+h,X_{n-1}+(X_n-X_{n-1}))-u((n-1)h,X_{n-1}).$$

By a Taylor series expansion and Lemma 4.5,

$$\begin{aligned}u(t+h,x+\Delta x_n)-u(t,x)&\overset{\mathbb{E}}{=}h\partial_t u+\Delta x_n\partial_x u+\frac{1}{2}(\Delta x_n)^2\partial_{xx}u+\Delta x_n h\partial_{tx}u\\&\quad+\frac{1}{2}h^2\partial_{tt}u+\frac{1}{2}h\Delta x_n^2\partial_{txx}u+\frac{1}{6}\Delta x_n^3\partial_{xxx}u+\frac{1}{24}\Delta x_n^4\partial_{xxxx}u+O(h^3).\end{aligned}$$

The result follows by summing and applying the moments from Lemma 4.5. ■

Corollary 4.7. *There exist two linear relations in the $\alpha_i, \beta_i, \gamma_i$ such that any two-step Maruyama scheme satisfying Dahlquist's root condition has weak order* 1.

Proof Sketch. The idea is that G_n^* are small if and only if the Maruyama scheme satisfies Dahlquist's root condition by standard results from the theory of linear recurrences. Then the conditions required are that two sums telescope to $O(1)$ values, which occurs if and only if linear relations in the coefficients hold.

Note that these consistency conditions are similar to the consistency conditions that are required for mean square convergence of two-step Maruyama given in [3].

4.2 Multistep Adams-Bashforth

We now analyze a particular class of SLMMs tailored to the small-noise regime. Buckwar and Winkler show in [3] that if the noise term in the original SDE is small, SLMMs provide asymptotic improvement over classical one-step methods in mean square convergence error. More precisely, we consider SDEs of the form

$$dX_t=f(X_t,t)dt+g(X_t,t)dW_t,$$

where $g=\varepsilon\hat{g}$ for some small constant ε and some function $\hat{g}\in C_b^r$ for some given r. Buckwar and Winkler only consider 2-step Maruyama schemes, which we resolve with Theorem 4.6. We instead study SLMMs that adapt the Adams-Bashforth method to better capture the drift term. We obtain small noise asymptotics for such schemes used a continuous version of the algorithm in Theorem 3.4.

We may define the k-step Adams-Bashforth SLMM, k-ABM for short, as follows.

Definition 4.8. Let $P_k(t)$ and $Q_k(t)$ be the Lagrange Interpolating Polynomials for the points $(hj,f(X_{n-j},t_{n-j}))$ and $(hj,g(X_{n-j},t_{n-j}))$ for $1\le j\le k$, respectively. The k-ABM scheme is the

SLMM defined by

$$X_{n+1} = X_n + \int_{[t_n,t_{n+1}]} P_k(t)dt + \int_{[t_n,t_{n+1}]} Q_k(t)dW_t.$$

Our main result is the following estimate of the weak error $\mathrm{Err}(T,h)$ for the k-ABM.

Theorem 4.9. *For the k-ABM,* $\mathrm{Err}(T,h) = O(h\varepsilon^2 + h^k)$ *assuming that* $f, \hat{g} \in C_b^{k+1}$*. More precisely,*

$$\begin{aligned}\mathrm{Err}(T,h) \leq{} & ||\partial_x u||_\infty \left(\frac{h\varepsilon^2}{4} ||f''||_\infty ||\hat{g}^2||_\infty + \frac{1}{k+1} ||f^{(k+1)}||_\infty h^k \right) \\ & + \frac{1}{2} ||\partial_{xx} u||_\infty \left(\frac{h\varepsilon^2}{4} ||(\hat{g}^2)''||_\infty ||\hat{g}^2||_\infty + \frac{1}{k+1} ||(\hat{g}^2)^{(k+1)}||_\infty h^k \right) \\ & + O(h^2\varepsilon^4 + h^{k+1}).\end{aligned}$$

Proof. As in Theorem 3.4, we let $u(t,x)$ be the solution to the KB-PDE and study the local deviation $e_i = u(t_{i+1}, X_{t_{i+1}}) - u(t_i, X_{t_i})$. We interpolate our discrete approximations in the following manner:

$$\tilde{X}_t = X_n + \int_{t_n}^t P_k(s)ds + \int_{t_n}^t Q_k(s)dW_s, \ t \in [t_n, t_{n+1}].$$

Therefore

$$e_i = \int_{[t_i,t_{i+1}]} du(t,\tilde{X}_t) \overset{\mathbb{E}}{=} \int_{[t_i,t_{i+1}]} \left[\partial_t u(t,\tilde{X}_t) + \partial_x u(t,\tilde{X}_t) P_k(t) + \frac{1}{2} \partial_{xx} u(t,\tilde{X}_t) Q_k(t)^2 \right] dt,$$

where the second equality follows from Itô's Lemma. Notice that, by the definition of u, one has

$$(\partial_t + \mathcal{L}) u(t,\tilde{X}_t) = 0,$$

so that

$$\partial_t u(t,\tilde{X}_t) + \partial_x u(t,\tilde{X}_t) f(\tilde{X}_t,t) + \frac{1}{2} \partial_{xx} u(t,\tilde{X}_t) g(\tilde{X}_t,t)^2 = 0.$$

Therefore, dropping expectations,

$$\begin{aligned} e_i &= \int_{[t_i,t_{i+1}]} \partial_x u(t,\tilde{X}_t)(f(\tilde{X}_t,t) - P_k(t))dt + \frac{1}{2} \int_{[t_i,t_{i+1}]} \partial_{xx} u(t,\tilde{X}_t)(g(\tilde{X}_t,t)^2 - Q_k(t)^2)dt \\ &\leq \max_{t\in[t_i,t_{i+1}]} |\partial_x u(t,\tilde{X}_t)| \int_{[t_i,t_{i+1}]} f(\tilde{X}_t,t) - P_k(t)dt \\ &+ \frac{1}{2} \max_{t\in[t_i,t_{i+1}]} |\partial_{xx} u(t,\tilde{X}_t)| \int_{[t_i,t_{i+1}]} g(\tilde{X}_t,t)^2 - Q_k(t)^2 dt. \end{aligned}$$

Let

$$e_i^f = \int_{[t_i,t_{i+1}]} f(\tilde{X}_t,t) - P_k(t)dt,$$

$$e_i^g = \int_{[t_i,t_{i+1}]} g^2(\tilde{X}_t, t) - Q_k^2(t)dt.$$

Because partial derivatives of u are bounded, $e_i = O(e_i^f + e_i^g)$. The strategy to bound e_i^f will be to decompose

$$e_i^f = \underbrace{\int_{[t_i,t_{i+1}]} f(\tilde{X}_t, t) - f(\mathbb{E}\tilde{X}_t, t)dt}_{e_{i,1}^f} + \underbrace{\int_{[t_i,t_{i+1}]} f(\mathbb{E}\tilde{X}_t, t) - P_k(t)dt}_{e_{i,2}^f}. \tag{4.1}$$

We can bound the deviation $e_{i,1}^f$ using a Taylor expansion of $f(\tilde{X}_t, t)$ about $\mathbb{E}\tilde{X}_t$, noting that

$$\tilde{X}_t - \mathbb{E}\tilde{X}_t = \int_{t_n}^t Q_k(s)dW_s.$$

Dropping the redundant t parameter, we get

$$\begin{aligned} f(\tilde{X}_t) &= f(\mathbb{E}\tilde{X}_t) + f'(\mathbb{E}\tilde{X}_t)\int_{t_n}^t Q_k(s)dW_s + \frac{1}{2}f''(\mathbb{E}\tilde{X}_t)\left[\int_{t_n}^t Q_k(s)dW_s\right]^2 + \cdots \\ &\overset{\mathbb{E}}{=} f(\mathbb{E}\tilde{X}_t) + \frac{1}{2}f''(\mathbb{E}\tilde{X}_t)\int_{t_n}^t Q_k(s)^2 ds + O((t-t_n)^2 \sup_{[t_n,t_{n+1}]} Q_k(t)^4), \end{aligned}$$

where we applied the Itô isometry and facts about moments of stochastic integrals. Therefore,

$$\int_{[t_i,t_{i+1}]} f(\tilde{X}_t, t) - f(\mathbb{E}\tilde{X}_t, t)dt \le \frac{h^2}{4}\max_{t\in[t_i,t_{i+1}]} |f''(\mathbb{E}\tilde{X}_t)| \max_{t\in[t_i,t_{i+1}]} |Q_k(t)^2| + O(h^3\varepsilon^4),$$

because $Q_k(t)$ interpolates g. We conclude that $e_{i,1}^f = O(h^2\varepsilon^2)$.

The quantity $e_{i,2}^f$ is deterministic, and to bound it one can simply use the Lagrange Interpolation error bound. Given the filtration $\mathcal{F}_{t_i}$, the polynomial $P_k(t)$ fits values of $f(\mathbb{E}\tilde{X}_t, t)$. Therefore,

$$f(\mathbb{E}\tilde{X}_t, t) - P_k(t) \le \frac{\max_{\xi\in I} |f^{(k+1)}(\mathbb{E}\tilde{X}_\xi, \xi)|}{(k+1)!}\prod_{j=1}^k (t-(n-j)h) \le \frac{\max_{\xi\in I} |f^{(k+1)}(\mathbb{E}\tilde{X}_\xi, \xi)|}{k+1}h^k.$$

Thus $e_{i,2}^f \overset{\mathbb{E}}{=} O(h^{k+1})$ and $e_i^f \overset{\mathbb{E}}{=} O(h^2\varepsilon^2 + h^{k+1})$ by (4.1).

We may perform a similar calculation for e_i^g by decomposing e_i^g as

$$e_i^g = \underbrace{\int_{[t_i,t_{i+1}]} g^2(\tilde{X}_t, t) - g^2(\mathbb{E}\tilde{X}_t, t)dt}_{e_{i,1}^g} + \underbrace{\int_{[t_i,t_{i+1}]} g^2(\mathbb{E}\tilde{X}_t, t) - Q_k^2(t)dt}_{e_{i,2}^g}.$$

An identical Taylor expansion computation yields $e^g_{i,1} \overset{\mathbb{E}}{=} O(h^2\varepsilon)$. Using the expression

$$g^2(\mathbb{E}\tilde{X}_t, t) - Q_k^2(t) = (g(\mathbb{E}\tilde{X}_t, t) - Q_k(t))(g(\mathbb{E}\tilde{X}_t, t) + Q_k(t))$$

and applying the Lagrange Interpolation error bound in the same manner as was done previously, we obtain $e^g_{i,2} = O(h^{k+1})$ and hence $e^g_i \overset{\mathbb{E}}{=} O(h^2\varepsilon^2 + h^{k+1})$.

Therefore, $e_i \overset{\mathbb{E}}{=} O(h^2\varepsilon^2 + h^{k+1})$, and one may finally conclude that

$$\mathrm{Err}(T, h) \overset{\mathbb{E}}{=} \sum_{i=1}^{T/h} e_i \overset{\mathbb{E}}{=} O(h\varepsilon^2 + h^k).$$

The precise error bound follows routinely from this proof and the fact that, over $[t_i, t_{i+1}]$, $|Q_k - g|$ is sufficiently small. ■

Note that Theorem 4.9 shows that SLMMs can asymptotically decrease the weak error of their analogous one-step schemes in the small noise case. It is worth noting that, unlike Adams-Bashforth and BDF schemes in the ODE case, the k-ABM is stable for all k by Theorem 3.2 of [3].

One should also note that the infinity norms in Theorem 4.9 can be replaced by infinity norms over I. Additionally, terms like $||\partial_x u||_\infty$ and $||\partial_{xx} u||_\infty$ can be upper bounded by functions independent of u by the work in [16].

5 Generalization to Lévy Processes

One may generalize the results we have proven to Lévy processes. The overall philosophy is the same: we describe a method to compute the weak error expansion, and we can show that in the small noise case, SLMMs provide improvement over their one-step analogues by deriving an error bound which is asymptotically smaller than $O(h)$ when noise and jump terms are small. We begin by adapting Theorem 3.4 to the Lévy case.

Theorem 5.1. *A weak error expansion of any SLMM for Lévy processes with error term* $O(h^{\gamma+r+1})$ *for arbitrarily high* r *may be computed explicitly algorithmically.*

Proof. The general idea is quite similar to Theorem 3.4. The difficulty is that we cannot Taylor expand $u(t+\Delta t, x+\Delta x_n)$ because the moments of Δx_n are no longer finite. Instead, one can follow the reasoning in [16]. For the Lévy-driven SDE $dX_t = h(X_{t-})dZ_t$ one can first approximate Z_t by Z_t^m, a process with bounded jumps. Write Δx_n as a linear combination of the past errors, which involve deterministic terms and increments of the Lévy process Z_t^m, by Lemma 3.1. Then one may construct a natural continuous path between $u(t, x + \Delta x_n)$ and $u(t, x)$, which may be written as a function of increments of Z_t^m. One can then expand this to an arbitrarily high order of accuracy by repeatedly applying Dynkin's formula [17]. We omit the details due to space constraints. ■

As we did in the continuous case, we can devise an SLMM for Lévy-driven SDEs that performs well in the small noise case by approximating the drift and diffusion terms with polynomial interpolation (as is done by k-ABM), and performing a naive Euler-Maruyama approximation of the jump terms. Letting $I_n = (t_n, t_{n+1}]$ we define the following.

Definition 5.2. Let $P_k(t)$ and $Q_k(t)$ be defined as in Definition 4.8. The Lévy k-ABM scheme is the SLMM defined by

$$\begin{aligned} X_{n+1} = X_n &+ \int_{[t_n,t_{n+1}]} P_k(t)dt + \int_{[t_n,t_{n+1}]} Q_k(t)dW_t \\ &+ \int_{|z|<1} \gamma(X_n, z)\tilde{J}(I_n, dz) + \int_{|z|\geq 1} \delta(X_n, z)J(I_n, dz). \end{aligned}$$

As before, we make these discrete approximations continuous with the following interpolation:

$$\begin{aligned} \tilde{X}_t = X_n &+ \int_{[t_n,t]} P_k(t)dt + \int_{[t_n,t]} Q_k(t)dW_t \\ &+ \int_{|z|<1} \gamma(X_n, z)\tilde{J}((t_n, t], dz) + \int_{|z|\geq 1} \delta(X_n, z)J((t_n, t], dz). \end{aligned}$$

When the jumps are small or γ, δ are small, one would expect this scheme to be roughly the k-ABM for SDEs without jump terms. Hence, one would expect a similar error as in Theorem 4.9. To derive precise asymptotics, we begin by defining the following quantities for $t \in [t_n, t_{n+1}]$:

$$\begin{aligned} I_{\gamma,n}(x) &= \int_{|z|<1} \int_{t_n}^{t} du(t, x + \gamma(\tilde{X}_t, z))\nu(dz)dt,\ I_{\gamma,n} = \int_{[t_n,t_{n+1}]} I_{\gamma,n}(\tilde{X}_t)dt,\ I_\gamma = \sum_{n=0}^{T/h} I_{\gamma,n}, \\ I_{\delta,n}(x) &= \int_{|z|\geq 1} \int_{t_n}^{t} du(t, x + \delta(\tilde{X}_t, z))\nu(dz)dt,\ I_{\delta,n} = \int_{[t_n,t_{n+1}]} I_{\delta,n}(\tilde{X}_t)dt,\ I_\delta = \sum_{n=0}^{T/h} I_{\delta,n}, \\ I^*_{\gamma,n} &= \int_{[t_n,t_{n+1}]} \int_{|z|<1} \int_{t_n}^{t} d\gamma(\tilde{X}_t, z)\nu(dz)dt,\ I^*_\gamma = \sum_{n=0}^{T/h} I^*_{\gamma,n}, \\ J^*_t &= \int_{|z|<1} \gamma(X_n, z)\tilde{J}((t_n, t], dz) + \int_{|z|\geq 1} \delta(X_n, z)J((t_n, t], dz) - h\int_{|z|\geq 1} \delta(X_n, z)\nu(dz), \\ J^{**}_t &= \int_{|z|<1} \gamma(X_n, z)^2\nu(dz) + \int_{|z|\geq 1} \delta(X_n, z)^2\nu(dz). \end{aligned}$$

With these definitions, we have the following theorem, generalizing Theorem 4.9 to Lévy processes, showing that the Lévy k-ABM asymptotically outperforms its one-step analog in the small noise case.

Theorem 5.3. *For the Lévy k-ABM, assuming $f, \hat{g} \in C_b^{k+1}$ and $\gamma \in C_b^1$,*

$$\mathrm{Err}(T,h) \leq O(h\varepsilon^2 + h^k + h||J_t^{**}||_\infty + I_\gamma + I_\delta + ||\partial_x u||_\infty I_\gamma^*).$$

Proof. Similarly to the proof of Theorem 4.9, we let $u(t,x)$ be the solution to the KB-PDE with generator $\mathcal{L}$ now involving jump terms and define e_i as before. The issue is that, in the Lévy case, Itô's Lemma no longer holds. Instead, we must use the Lévy-Itô Lemma to expand differentials:

$$\begin{aligned}
e_i &= \int_{[t_n,t_{n+1}]} du(t,\tilde{X}_t) \\
&\overset{\mathbb{E}}{=} \int_{[t_n,t_{n+1}]} (\partial_t u + P_k(t)\partial_x u + \frac{1}{2}Q_k(t)^2\partial_{xx}u)dt \\
&+ \int_{[t_n,t_{n+1}]}\int_{|z|<1} u(t,\tilde{X}_{t-}+\gamma(X_n,z)) - u(t,\tilde{X}_{t-}) - \gamma(X_n,z)\partial_x u(t,\tilde{X}_{t-})\nu(dz)dt \\
&+ \int_{[t_n,t_{n+1}]}\int_{|z|\geq 1} u(t,\tilde{X}_{t-}+\delta(X_n,z)) - u(t,\tilde{X}_{t-})\nu(dz)dt.
\end{aligned}$$

Comparing this to the equation given by $(\partial_t + \mathcal{L})u = 0$, one obtains an estimate

$$\mathbb{E}[e_i] \leq \mathbb{E}[e_i^f||\partial_x u||_\infty + e_i^g||\partial_{xx}u||_\infty + e_i^\gamma + e_i^\delta],$$

with

$$\begin{aligned}
e_n^f &= \int_{[t_n,t_{n+1}]} f(\tilde{X}_t,t)dt - \int_{[t_n,t_{n+1}]} P_k(t)dt, \\
e_n^g &= \int_{[t_n,t_{n+1}]} g^2(\tilde{X}_t,t)dt - \int_{[t_n,t_{n+1}]} Q_k^2(t)dt, \\
e_n^\gamma &= \int_{[t_n,t_{n+1}]}\int_{|z|<1} u(t,\tilde{X}_{t-}+\gamma(X_n,z)) - u(t,\tilde{X}_{t-}) - \gamma(X_n,z)\partial_x u(t,\tilde{X}_{t-})\nu(dz)dt \\
&- \int_{[t_n,t_{n+1}]}\int_{|z|<1} u(t,\tilde{X}_{t-}+\gamma(\tilde{X}_{t-},z)) - u(t,\tilde{X}_{t-}) - \gamma(\tilde{X}_{t-},z)\partial_x u(t,\tilde{X}_{t-})\nu(dz)dt, \\
e_n^\delta &= \int_{[t_n,t_{n+1}]}\int_{|z|\geq 1} u(t,\tilde{X}_{t-}+\delta(X_n,z)) - u(t,\tilde{X}_{t-})\nu(dz)dt \\
&- \int_{[t_n,t_{n+1}]}\int_{|z|\geq 1} u(t,\tilde{X}_{t-}+\delta(\tilde{X}_{t-},z)) - u(t,\tilde{X}_{t-})\nu(dz)dt.
\end{aligned}$$

We must bound each term individually. We will bound e_i^f and e_i^g in a similar manner as was done in Theorem 4.9, although there are some key differences. We will then bound e_i^γ and e_i^δ using $I_{\gamma,i}, I_{\delta,i}, I_{\gamma,i}^*$ and sum over i.

To bound e_i^f we decouple as was done in (4.1). We bound the deviation $e_{i,1}^f$ using a Taylor expansion again. However, in this case,

$$\tilde{X}_t - \mathbb{E}\tilde{X}_t = \underbrace{\int_{t_n}^{t} Q_k(s)dW_s}_{\text{diffusion error}} + \underbrace{J_t^*}_{\text{jump error}}.$$

The fact that the jump error is precisely J_t^* follows from properties of Lévy integrals. Hence, in the Taylor expansion,

$$f(\tilde{X}_t) - f(\mathbb{E}\tilde{X}_t) = \underbrace{f'(\mathbb{E}\tilde{X}_t)\int_{t_n}^{t} Q_k(s)dW_s + \frac{1}{2}f''(\mathbb{E}\tilde{X}_t)\left[\int_{t_n}^{t} Q_k(s)dW_s\right]^2 + \cdots}_{\text{diffusion error}}$$
$$+ \underbrace{J_t^* f'(\mathbb{E}\tilde{X}_t) + J_t^{*2} f''(\mathbb{E}\tilde{X}_t) + O(J_t^{*3})}_{\text{jump error}}.$$

The diffusion error is the same error as in Theorem 4.9, which equates to $O(h^2\varepsilon^2)$ after taking expectations and integrating. A straightforward calculation shows that $J_t^* \overset{\mathbb{E}}{=} 0$ and $J_t^{*2} \overset{\mathbb{E}}{=} hJ_t^{**}$. Higher moments of J_t^* can be shown to have terms that are $o(hJ_t^{**})$ assuming regularity hypotheses on γ, δ or ν. In particular, higher moments consist of terms like $h\int_{|z|<1}\gamma(X_n, z)^k\nu(dz)$ for $k > 2$. Therefore, the jump error is $O(h^2||J_t^{**}||_\infty)$ (after integrating). Hence $e_{i,1}^f = O(h^2\varepsilon^2 + h^2||J_t^{**}||_\infty)$.

We get that $e_{i,2}^f = O(h^{k+1})$ by the Lagrange Interpolation error bound, because $P_k(t)$ fits values of the deterministic function $f(\mathbb{E}\tilde{X}_t, t)$, given the filtration $\mathcal{F}_{t_i}$.

These bounds give, in expectation,

$$e_i^f = O(h^2\varepsilon^2 + h^{k+1} + h^2||J_t^{**}||_\infty).$$

One can repeat this argument for e_i^g, mimicking the argument in Theorem 4.9, to get that

$$e_i^g = O(h^2\varepsilon^2 + h^{k+1} + h^2||J_t^{**}||_\infty)$$

as well.

Straightforward calculations show that, in expectation,

$$e_i^\gamma \leq I_{\gamma,i} + ||\partial_x u||_\infty I_{\gamma,i}^*,$$
$$e_i^\delta = I_{\delta,i}.$$

Summing over i completes the proof. ■

As before, infinity norms can be replaced with infinity norms over I. Precise constant factors may be obtained in the same way as Theorem 4.9. The quantities $I_\gamma, I_\delta, I_\gamma^*$ are all $O(h)$ without regularity conditions on γ and ν, so one recovers a bound that is essentially equivalent to the main bound in [16] as a corollary. With regularity conditions on γ and ν, one can expand I_γ, I_δ, and I_γ^* with the Lévy-Itô Lemma to get "small noise" results for Lévy-driven SDEs. In particular, SLMMs can asymptotically improve upon weak error in small noise cases over one-step analogues for Lévy-driven SDEs.

6 Numerical Simulations

We can additionally simulate numerically the weak convergence of various schemes. Figure 1 is a simulation with $\phi(x) = x$ and $f(X_t, t) = 0.8X_t$, $g(X_t, t) = 0.3X_t$ (the Black-Scholes equation with particular parameters).

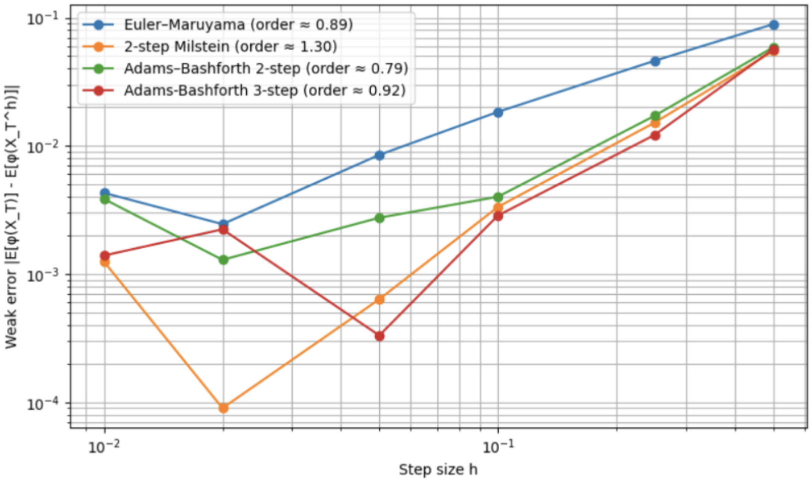

Figure 1: Comparison of weak errors of SLMMs.

In Figure 1 we plot the weak error $\text{Err}(T, h)$ as a function of h. The true expected value is calculated explicitly as X_t follows a Geometric Brownian Motion (see [1]), and the approximated expected value is estimated through Monte Carlo simulation. We include Adams-Bashforth 2-step and Adams-Bashforth 3-step schemes, which are variants of Euler-Maruyama that use Adams-Bashforth to approximate $\int f dt$. Notice that the behavior is less accurate for small h, as simulation requires more computational power. Nonetheless, we see that k-ABM is more accurate as k increases (noting that the 1-ABM is Euler-Maruyama). This empirically verifies Theorem 4.9. One may also note that 2-step Milstein performs well, suggesting that approximating higher order stochastic integrals helps decrease weak error, along with increasing k. Approximations of the weak order of each scheme is estimated given the simulated data points.

We observe the performance of multistep schemes as a function of the size of the noise term.

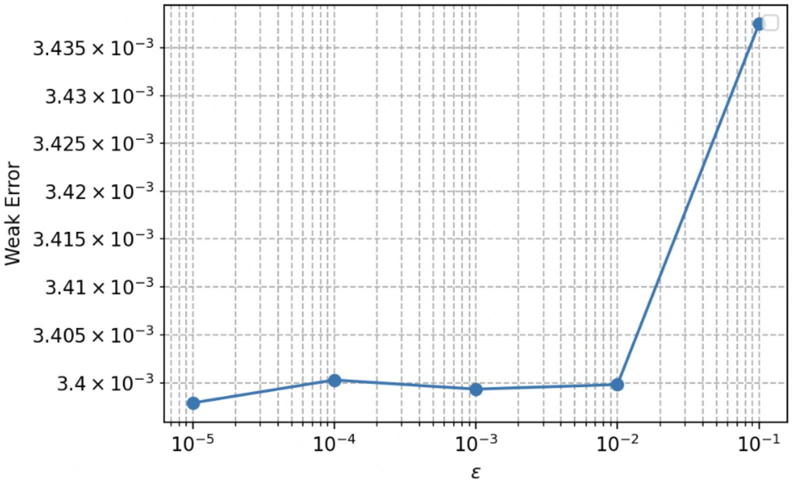

Figure 2: 2-Step Adams-Bashforth performance as a function of noise size.

Figure 2 is a simulation with $\phi(x) = x$ and $f(X_t, t) = 0.8X_t$, $g_\varepsilon(X_t, t) = \varepsilon X_t$ (the Black-Scholes equation with particular parameters). The weak error of 2-Step Adams-Bashforth is estimated with a Monte Carlo simulation. As expected, the weak error is roughly an increasing function of ε, with some variance due to error in Monte Carlo estimation. This also empirically verifies Theorem 4.9.

7 Conclusion and Discussion

In this work, we developed a general framework for computing weak error expansions of SLMMs applied to Itô and Lévy-driven SDEs. Using the Feynman-Kac formula, we turn questions about weak error of SDEs into questions about elliptic PDEs, which can be understood through local deviation analysis and linear recurrence theory. We applied this framework to analyze several SLMMs from the literature, obtaining their weak error expansions explicitly, which are new to the best of our knowledge.

We constructed novel SLMMs optimized for the small-noise regime and proved that they achieve asymptotic improvements in weak error over one-step methods. Numerical experiments support our theoretical findings and highlight the practical benefits of SLMMs when stochastic perturbations are minor.

Our results show that SLMMs can significantly outperform standard one-step methods in applications where accurately capturing expectations under small noise is essential—such as in particle dynamics, options pricing, and gene expression modeling. Beyond the methods themselves, our framework lays the groundwork for further analytical and computational exploration of SLMMs.

8 Acknowledgements

I am deeply grateful to my mentor, Joonsoo Lee (Massachusetts Institute of Technology), for his invaluable guidance, insightful feedback, and for suggesting this project. I would also like to thank my tutor, AnaMaria Perez, for her consistent support and thoughtful advice throughout the summer. Additionally, I would like to thank Dr. Jenny Sendova for her thoughtful comments and advice. I am especially thankful to Dr. Tanya Khovanova, Professor Roman Bezrukavnikov, and Dr. Jonathan Bloom for their support and guidance. This project was made possible through the support of the Center for Excellence in Education and the Massachusetts Institute of Technology. I also thank Ashwat Prassana, Marina Lin, Akilan Sankaran, and Alvaro Carballeira Mora for their helpful feedback on my work. Finally, I am sincerely thankful to my individual sponsors—Shaunak K. Deepak, Asaf Noy, Gang Lan, and Arvind Parthasarathi—for their generosity.

References

[1] E. Vanden-Eijnden. *Applied Stochastic Analysis*, volume 30 of *Courant Lecture Notes.* American Mathematical Society, Providence, RI, 2021.

[2] P. E. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations*, volume 23 of *Applications of Mathematics.* Springer, 1992.

[3] E. Buckwar and R. Winkler. Multistep methods for SDEs and their application to problems with small noise. *SIAM Journal on Numerical Analysis*, 44(2):779–803, 2006.

[4] S. Ninomiya and N. Victoir. Weak Approximation of Stochastic Differential Equations and Application to Derivative Pricing. *Applied Mathematical Finance*, 15(2):107–121, Apr. 2008.

[5] A. Kohatsu-Higa and H.-L. Ngo. Weak Approximations for SDE's Driven by Lévy Processes. In R. C. Dalang, M. Dozzi, and F. Russo, editors, *Seminar on Stochastic Analysis, Random Fields and Applications VII*, pages 131–169, Basel, 2013. Springer Basel.

[6] A. Kohatsu-Higa and P. Tankov. Jump-adapted discretization schemes for Lévy-driven SDEs. *Stochastic Processes and their Applications*, 120(11):2258–2285, November 2010.

[7] P. Tankov. High order weak approximation schemes for Lévy-driven SDEs, 2010.

[8] N. Fournier. Simulation and approximation of Levy-driven stochastic differential equations, 2009.

[9] D. Platonov and V. Knopova. Strong Convergence Rates for Euler Schemes of Levy-Driven SDE using Dynamic Cutting, 2025.

[10] D. Talay and L. Tubaro. Expansion of the global error for numerical schemes solving stochastic differential equations. *Stochastic Analysis and Applications*, 8(4):483–509, Jan. 1990.

[11] Y. Iguchi and T. Yamada. An extended Milstein scheme for effective weak approximation of diffusions, 2024.

[12] M. Kac. On Distributions of Certain Wiener Functionals. *Transactions of the American Mathematical Society*, 65(1):1–13, 1949.

[13] G. Dahlquist. Convergence and stability in the numerical integration of ordinary differential equations. *MATHEMATICA SCANDINAVICA*, 4:33–53, Dec. 1956.

[14] V. Zolotarev. *One-dimensional Stable Distributions.* Translations of mathematical monographs. American Mathematical Society, 1986.

[15] S. Asmussen and J. Rosiński. Approximations of Small Jumps of Lévy Processes with a View Towards Simulation. *Journal of Applied Probability*, 38(2):482–493, 2001.

[16] P. Protter and D. Talay. The Euler Scheme for Lévy driven Stochastic Differential Equations. (RR-2621):31, 1995.

[17] E. Dynkin, V. Greenberg, J. Fabius, A. Maitra, and G. Majone. *Markov Processes: Volume I.* Grundlehren der mathematischen Wissenschaften. Springer Berlin Heidelberg, 2012.