Simulation methods for stochastically modeled population processes

David F. Anderson

anderson@math.wisc.edu

Department of Mathematics
University of Wisconsin - Madison

Joint work with: Arnab Ganguly, Tom Kurtz, and Jonathan Mattingly

BIRS

January 20th, 2010
Overview

1. Notation and terminology.

2. Discuss numerical approximation methods.

3. Question: how well do different methods approximate exact process.

4. Discuss diffusion approximations
   - Accuracy versus simulation methods.
   - Methods of approximating the diffusion approximation.
Stochastic models of (bio)chemical reactions

- Standard notation for chemical reactions:

\[ A + B \rightarrow C \]

is interpreted as “a molecule of A combines with a molecule of B to give a molecule of C.”

- We consider a network of reactions involving \( d \) chemical species, \( S_1, \ldots, S_d \):

\[
\sum_{i=1}^{d} \nu_{ik} S_i \rightarrow \sum_{i=1}^{d} \nu'_{ik} S_i
\]

- Each instance of the reaction \( A + B \rightarrow C \) changes the state of the system by the vector:

\[
\nu'_k - \nu_k = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}.
\]
Markov chain models

- The state of the system, \( X(t) \in \mathbb{Z}_{\geq 0}^d \), gives the number of molecules of each species in the system at time \( t \).

- \( \nu_k \): number of molecules of each chemical species consumed in the \( k \)th reaction.

- \( \nu'_k \): number of molecules of each chemical species created in the \( k \)th reaction.

- If \( k \)th reaction occurs at time \( t \), the new state becomes
  \[
  X(t) = X(t-) + \nu'_k - \nu_k.
  \]

- The waiting times for the reactions are exponentially distributed with rate (intensity/propensity) functions \( \lambda_k : \mathbb{R}^d \rightarrow \mathbb{R}_{\geq 0} \).
Markov chain models

This model is a continuous time Markov chain in $\mathbb{Z}_d^{\geq 0}$ with generator

$$(Af)(x) = \sum_k \lambda_k(x)(f(x + \nu'_k - \nu_k) - f(x)).$$

Kolmogorov’s forward equation ("Chemical Master Equation") describes the evolution of the distribution of the state of the system

$$\frac{d}{dt} P(x, t \mid \mu) = \sum_k \lambda_k(x - \nu_k')P(x - \nu_k' + \nu_k, t \mid \mu) - \sum_k \lambda_k(x)P(x, t \mid \mu),$$

where $P(x, t \mid \mu)$ is probability $X(t) = x$ conditioned on initial distribution $\mu$. 
A path-wise representation is given by random time-changes of Poisson processes

\[ X(t) = X(0) + \sum_k Y_k \left( \int_0^t \lambda_k(X(s))ds \right) (\nu'_k - \nu_k), \]

where the \( Y_k \) are independent, unit-rate Poisson processes. Note:

\[ R_k(t) = Y_k \left( \int_0^t \lambda_k(X(s))ds \right) \]

is a counting process giving number of times \( k \)th reaction has fired.
Pathwise Representations – Random time changes

A path-wise representation is given by random time-changes of Poisson processes

\[ X(t) = X(0) + \sum_k Y_k \left( \int_0^t \lambda_k(X(s)) \, ds \right) (\nu'_k - \nu_k), \]

where the \( Y_k \) are independent, unit-rate Poisson processes. Note:

\[ R_k(t) = Y_k \left( \int_0^t \lambda_k(X(s)) \, ds \right) \]

is a counting process giving number of times \( k \)th reaction has fired.

This representation is equivalent to the “chemical master equation” (Kolmogorov’s forward equation) representation found in biology/chemistry literature.
Mass-action kinetics

The standard intensity function chosen is mass-action kinetics:

$$\lambda_k(x) = \kappa_k \left( \prod_i \nu_{ik}! \right) \binom{x}{\nu_k} = \kappa_k \prod_i \frac{x_i!}{(x_i - \nu_{ik})!}.$$

Rate is proportional to the number of distinct subsets of the molecules present that can form inputs for the reaction. (this assumes vessel is “well-stirred”.)

Example: If $S_1 \rightarrow \text{anything}$, then $\lambda_k(x) = \kappa_k x_1$.

Example: If $S_1 + S_2 \rightarrow \text{anything}$, then $\lambda_k(x) = \kappa_k x_1 x_2$.

Example: If $S_1 + 2S_2 \rightarrow \text{anything}$, then $\lambda_k(x) = \kappa_k x_1 x_2 (x_2 - 1)$. 
Methods of investigation: numerical simulation

\[ X(t) = X(0) + \sum_k Y_k \left( \int_0^t \lambda_k(X(s))ds \right) (\nu'_k - \nu_k), \]

**GOOD NEWS** There are a number of numerical methods that produce statistically exact sample paths:

1. Gillespie’s algorithm.
2. The first reaction method.
3. The next reaction method.

For each step of these methods must find:

*(i)* the amount of time that passes until the next reaction takes place: \( \Delta_n \).

(the minimum of exponential RVs).

*(ii)* which reaction takes place at that time.

(BAD NEWS) If \( P_k \lambda_k(X(t)) \gg 1 \), then \( \Delta_n \approx 1 \) and the time needed to produce a single exact sample path over an interval \([0, T]\) can be prohibitive.
Methods of investigation: numerical simulation

\[ X(t) = X(0) + \sum_k Y_k \left( \int_0^t \lambda_k(X(s)) ds \right) (\nu'_k - \nu_k), \]

(GOOD NEWS) There are a number of numerical methods that produce statistically exact sample paths:

1. Gillespie’s algorithm.
2. The first reaction method.
3. The next reaction method.

For each step of these methods must find:

(i) the amount of time that passes until the next reaction takes place: \( \Delta_n \).
   (the minimum of exponential RVs).

(ii) which reaction takes place at that time.

(BAD NEWS) If \( \sum_k \lambda_k(X(t)) \gg 1 \), then

\[ \Delta_n \approx \frac{1}{\sum_k \lambda_k(X(t))} \ll 1 \]

and the time needed to produce a single exact sample path over an interval \([0, T]\) can be prohibitive.
Tau-leaping

Standard “τ-leaping”\(^1\) was developed by Dan Gillespie in an effort to overcome the problem that \(\Delta_n\) may be prohibitively small.

**Tau-leaping** is essentially an **Euler approximation** of \(\int_0^t \lambda_k(X(s))ds\):

\[
Z(\tau) = Z(0) + \sum_k Y_k \left( \int_0^\tau \lambda_k(Z(s)) \, ds \right) (\nu'_k - \nu_k) \\
\approx Z(0) + \sum_k Y_k \left( \lambda_k(Z(0)) \tau \right) (\nu'_k - \nu_k) \\
\overset{d}{=} Z(0) + \sum_k \text{Poisson} \left( \lambda_k(Z(0)) \tau \right) (\nu'_k - \nu_k).
\]

Algorithm: For each time step do the following:

1. For each reaction, calculate \( P_{k,n} = \text{Poisson}(\lambda_k(Z(t_n)) \tau) \).

2. Update, \( Z(t_{n+1}) = Z(t_n) + \sum_k P_{k,n} \times (\nu'_k - \nu_k) \).

A pathwise representation for \( Z(t) \) generated by \( \tau \)-leaping is

\[
Z(t) = X(0) + \sum_k Y_k \left( \int_0^t \lambda_k(Z \circ \eta(s)) ds \right) (\nu'_k - \nu_k),
\]

where

\[
\eta(s) = t_n, \quad \text{if} \quad t_n \leq s < t_{n+1} = t_n + \tau
\]

is a step function giving left endpoints of time discretization.
Another algorithm: A midpoint method

For a time discretization $0 = t_0 < t_1 < \cdots < t_N = T$, with $\tau = t_n - t_{n-1}$, let

$$\rho(z) = z + \frac{1}{2} \tau \sum_k \lambda_k(z)(\nu'_k - \nu_k),$$

be a “deterministic” midpoint approximation

Algorithm: For each time step do the following:

1. Let $y^*_n = \rho(\mathcal{Z}(t_n))$.
2. For each reaction, calculate $P_{k,n} = \text{Poisson}(\lambda_k(y^*_n) \tau)$.
3. Update, $\mathcal{Z}(t_{n+1}) = \mathcal{Z}(t_n) + \sum_k P_{k,n} \times (\nu'_k - \nu_k)$. 
Another algorithm: A midpoint method

For a time discretization $0 = t_0 < t_1 < \cdots < t_N = T$, with $\tau = t_n - t_{n-1}$, let

$$\rho(z) = z + \frac{1}{2} \tau \sum_k \lambda_k(z)(\nu'_k - \nu_k),$$

be a "deterministic" midpoint approximation

Algorithm: For each time step do the following:

1. Let $y_n^* = \rho(Z(t_n))$.
2. For each reaction, calculate $P_{k,n} = \text{Poisson}(\lambda_k(y_n^*) \tau)$.
3. Update, $Z(t_{n+1}) = Z(t_n) + \sum_k P_{k,n} \times (\nu'_k - \nu_k)$.

Then $Z(t)$ solves:

$$Z(t) = X(0) + \sum_k Y_k \left( \int_0^t \lambda_k(\rho(Z \circ \eta(s))) ds \right) (\nu'_k - \nu_k),$$

where $\eta(s) = t_n$, if $t_n \leq s < t_{n+1} = t_n + \tau$. 
Previous error analysis

Under the scaling $\tau \to 0$:

1. Rathinam, Petzold, Cao, and Gillespie\textsuperscript{2} showed, among other things (implicit methods, etc.), that tau-leaping is first order accurate in a weak sense if the intensity functions $\lambda_k$ are linear.

2. Li\textsuperscript{3} extended this result by showing that standard tau-leaping has a strong error (in the $L^2$ norm) of order $1/2$ and a weak error of order one:

$$\sup_{n\leq N} \sqrt{\mathbb{E} |Z(t_n) - X(t_n)|^2} \leq C \tau^{1/2}$$

$$|\mathbb{E} f(Z(T)) - \mathbb{E} f(X(T))| \leq C \tau,$$

where $0 = t_0 < t_1 < \cdots < t_N = T$ is a partition of $[0, T]$.

3. The midpoint method is no more accurate than standard tau-leaping


Example

Consider $A$ as a prey and $B$ as a predator.

$$A \overset{\kappa_1}{\rightarrow} 2A, \quad A + B \overset{\kappa_2}{\rightarrow} 2B, \quad B \overset{\kappa_3}{\rightarrow} \emptyset,$$

with $A(0) = B(0) = 1000$ and $\kappa_1 = 2$, $\kappa_2 = 0.02$, $\kappa_3 = 2$.

Letting $\tau = 1/20$ and simulating 30,000 sample paths with each method yields the following approximate distributions for $B(10)$:
Problem with the $\tau \to 0$ scaling

**Recall**, tau-leaping methods are only useful if $\tau \gg \Delta_n$, for otherwise an exact method would be performed. Therefore, we should require that

$$\tau \gg \frac{1}{\sum_k \lambda_k(X(t))} \approx \Delta_n \quad \text{while} \quad \sum_k \lambda_k(X(t)) \gg 1.$$

So $\tau \to 0$ does not seem to be an appropriate scaling to describe the accuracy of these methods.
What is an appropriate scaling?

I can not answer in generality. It will depend upon *how* the system satisfies

\[ \sum_k \lambda_k(X(t)) \gg 1 \]

Multiple possibilities:

1. Could be that \( X_i = O(V^\alpha_i) \) for \( V \gg 1 \) with \( k_k = O(1) \).

2. Could be that \( \kappa_k = O(V^\beta_k) \) for \( V \gg 1 \) with \( X_i = O(1) \).

3. Could be multiple scales.

4. Infinite number of possibilities.
The “classical scaling”

We will make the following natural assumptions on our model:

(i) $X_i(0) = \mathcal{O}(V)$ for some $V \gg 1$ (think of as proportional to “volume”).

(ii) The rate constants of reactions of the form

$$S_1 \rightarrow *, \quad S_1 + S_2 \rightarrow *, \quad S_1 + S_2 + S_3 \rightarrow *,$$

scale with $V$ like:

$$\kappa_k = \mathcal{O}(1), \quad \kappa_k = \mathcal{O}\left(\frac{1}{V}\right), \quad \kappa_k = \mathcal{O}\left(\frac{1}{V^2}\right).$$
The “classical scaling”

We will make the following natural assumptions on our model:

(i) $X_i(0) = O(V)$ for some $V \gg 1$ (think of as proportional to “volume”).

(ii) The rate constants of reactions of the form

$$S_1 \rightarrow *, \quad S_1 + S_2 \rightarrow *, \quad S_1 + S_2 + S_3 \rightarrow *,$$

scale with $V$ like:

$$\kappa_k = O(1), \quad \kappa_k = O\left(\frac{1}{V}\right), \quad \kappa_k = O\left(\frac{1}{V^2}\right).$$

- For each $k$ and $x \in \mathbb{R}^d$: $\lambda_k(Vx) = VA_k^V(x) = O(V)$.

- In the limit as $V \to \infty$, $X(V, t)/V \to c(t)$, which satisfies deterministic mass action kinetics.
Important to recognize

1. It is clear that the choice of scaling laid out above will not explicitly cover all cases of interest:
   
   i. Multiple scales
   
   ii. Low copy numbers of constituent molecules.

2. However, the purpose is to give a more accurate picture of how tau-leaping methods approximate the exact solution, both strongly and weakly, in at least one plausible setting.
**Recall:** \( \tau \)-leaping is only useful if

\[
\tau \sum_k \lambda_k(X(t)) \gg 1 \quad \text{while} \quad \frac{1}{\sum_k \lambda_k(X(t))} \ll 1.
\]

For \( \beta \in (0, 1) \), we let

\[
\tau = \frac{1}{V^\beta}.
\]

This scaling satisfies the above requirements:

\[
\tau \sum_k \lambda(X(t)) = V^{-\beta} \mathcal{O}(V) = \mathcal{O}(V^{1-\beta}) \gg 1.
\]
Error analysis on scaled processes

We consider the scaled processes

\[
X^V(t) \triangleq \frac{X(t)}{V}, \quad Z^V(t) \triangleq \frac{Z(t)}{V}, \quad Z^V(t) \triangleq \frac{Z(t)}{V}.
\]

which satisfy

\[
X^V(t) = X^V(0) + \frac{1}{V} \sum_k Y_k \left( V \int_0^t A_k^V(X^V(s)) ds \right) (\nu'_k - \nu_k)
\]

\[
Z^V(t) = X^V(0) + \frac{1}{V} \sum_k Y_k \left( V \int_0^t A_k^V(Z^V \circ \eta(s)) ds \right) (\nu'_k - \nu_k)
\]

\[
Z^V(t) = X^V(0) + \frac{1}{V} \sum_k Y_k \left( V \int_0^t A_k^V \circ \rho^V(Z^V \circ \eta(s)) ds \right) (\nu'_k - \nu_k),
\]

where \( \rho^V(z) \triangleq z + \frac{1}{2} \tau \sum_k A_k^V(z)(\nu'_k - \nu_k), \)

and consider the relationship of the normalized approximate processes, \( Z^V(t) \) and \( Z^V(t) \), to the original process \( X^V(t) \), normalized similarly.
Strong error of standard tau-leaping

Theorem (Anderson, Ganguly, Kurtz – 2009)

Let $X^V(t)$ and $Z^V(t)$ denote the scaled processes for $t \in [0, T]$ with $\tau = V^{-\beta}$. There exists a constant $C = C(T) > 0$ such that

$$\sup_{t \leq T} \mathbb{E}|X^V(t) - Z^V(t)| \leq \frac{C}{V^\beta} = C\tau.$$
A new coupling of $X^V(t)$ and $Z^V(t)$

To sharpen the result, let

$$X^V(t) = X^V(0) + \frac{1}{V} \sum_k \left[ Y_{k,1} \left( V \int_0^t A_k^V(X^V(s)) \wedge A_k^V(Z^V \circ \eta(s)) \, ds \right) \\
+ Y_{k,2} \left( V \int_0^t A_k^V(X^V(s)) - A_k^V(X^V(s)) \wedge A_k^V(Z^V \circ \eta(s)) \, ds \right) \right] (\nu'_k - \nu_k)$$

$$Z^V(t) = X^V(0) + \frac{1}{V} \sum_k \left[ Y_{k,1} \left( V \int_0^t A_k^V(X^V(s)) \wedge A_k^V(Z^V \circ \eta(s)) \, ds \right) \\
+ Y_{k,3} \left( V \int_0^t A_k^V(Z^V \circ \eta(s)) - A_k^V(X^V(s)) \wedge A_k^V(Z^V \circ \eta(s)) \, ds \right) \right] (\nu'_k - \nu_k),$$

where the $Y_{k,i}$ are mutually independent Poisson processes.

**Note:** distribution of $X^V(t)$ and $Z^V(t)$ are unchanged.
Theorem (Anderson, Ganguly, Kurtz – 2009)

Let $X^V(t)$ and $Z^V(t)$ satisfy the new coupling for $t \in [0, T]$. Then as $V \to \infty$

$$V^\beta (X^V - Z^V) \to \mathcal{E},$$

where

- $\mathcal{E}(t)$ is deterministic and independent of $V$ and $\beta$.
- And where $f^V \to f$ means strong path-wise convergence:

$$\lim_{V \to \infty} P\{ \sup_{t \leq T} |f^V(t) - f(t)| > \epsilon \} = 0, \text{ for all } \epsilon > 0$$

$$\lim_{V \to \infty} \mathbb{E} \sup_{t \leq T} \left| f^V(t) - f(t) \right| = 0.$$
Idea of proof

Can show:

\[
X^V(t) - Z^V(t) \approx M^V(t) + \int_0^t DF^V(Z^V \circ \eta(s))(X^V(s) - Z^V(s))ds
\]

\[
+ V^{-\beta} \frac{1}{2} \int_0^t DF^V(Z^V \circ \eta(s))F^V(Z^V \circ \eta(s))ds
\]

where \( M^V(t) \) is a martingale and \( V^{(1+\beta)/2}M^V \Rightarrow M \), where \( M \) is a mean-zero Gaussian process.

- \( (1 + \beta)/2 > \beta \), so done.
Theorem (Anderson, Ganguly, Kurtz – 2009)

Let $X^V(t)$ and $Z^V(t)$ be as before. Then, for any continuously differentiable function $f$

$$
\lim_{V \to \infty} \frac{1}{V^\beta} \left( \mathbb{E} f(X^V(t)) - \mathbb{E} f(Z^V(t)) \right) = \mathcal{E}(t) \cdot \nabla f(x(t)).
$$
Theorem (Anderson, Ganguly, Kurtz – 2009)

Let $X^V(t)$ and $Z^V(t)$ denote the scaled processes for $t \in [0, T]$. Then there exists a constant $C = C(T) > 0$ such that

$$\sup_{t \leq T} \mathbb{E} |X^V(t) - Z^V(t)| \leq \frac{C}{V^{\kappa(\beta)}}, \quad \text{where} \quad \kappa(\beta) = \min \left\{ 2\beta, \frac{1 + \beta}{2} \right\}. $$
A new coupling of $X^V(t)$ and $Z^V(t)$

To sharpen the result, let

$$X^V(t) = X^V(0) + \frac{1}{V} \sum_k \left[ Y_{k,1} \left( V \int_0^t A_k^V(X^V(s)) \wedge A_k^V(\rho^V \circ Z^V \circ \eta(s)) \, ds \right) \right.$$ 

$$+ Y_{k,2} \left( V \int_0^t A_k^V(X^V(s)) - A_k^V(X^V(s)) \wedge A_k^V(\rho^V \circ Z^V \circ \eta(s)) \, ds \right) \right] (\nu'_k - \nu_k)$$

$$Z^V(t) = X^V(0) + \frac{1}{V} \sum_k \left[ Y_{k,1} \left( V \int_0^t A_k^V(X^V(s)) \wedge A_k^V(\rho^V \circ Z^V \circ \eta(s)) \, ds \right) \right.$$ 

$$+ Y_{k,3} \left( V \int_0^t A_k^V(\rho^V \circ Z^V \circ \eta(s)) - A_k^V(X^V(s)) \wedge A_k^V(\rho^V \circ Z^V \circ \eta(s)) \, ds \right) \right] (\nu'_k - \nu_k),$$

where the $Y_{k,i}$ are mutually independent Poisson processes.
Exact asymptotics for midpoint tau-leaping

Theorem (Anderson, Ganguly, Kurtz – 2009)

Let $X^V(t)$ and $Z^V(t)$ satisfy the new coupling for $t \in [0, T]$. Then as $V \to \infty$

$$V^{2\beta}(X^V - Z^V) \to \mathcal{E}_1, \quad \text{if} \quad \beta < 1/3.$$  

$$V^{2\beta}(X^V - Z^V) \Rightarrow \mathcal{E}_2, \quad \text{if} \quad \beta = 1/3.$$  

$$V^{(1+\beta)/2}(X^V - Z^V) \Rightarrow \mathcal{E}_3, \quad \text{if} \quad \beta > 1/3.$$  

Where $\mathcal{E}_1$ is deterministic and $\mathcal{E}_2, \mathcal{E}_3$ are stochastic.
Idea of proof

Can show:

\[ X^V(t) - Z^V(t) \approx M^V(t) + \int_0^t DF^V(Z^V \circ \eta(s))(X^V(s) - Z^V(s))ds + V^{-2\beta}H^V(t) \]

where \( V^{(1+\beta)/2}M^V \Rightarrow M(t) \) and \( H^V \to H \).
Weak error of midpoint tau-leaping

Theorem (Anderson, Ganguly, Kurtz – 2009)

Let $X^V(t)$ and $Z^V(t)$ be as before. Then, for any $C^3$ function $f$, there exists a constant $C = C(f, T) > 0$ such that

$$V^{2/3} \sup_{t \leq T} \left| \mathbb{E}f(X^V(t)) - \mathbb{E}f(Z^V(t)) \right| \leq C.$$
Example
Consider $A$ as a prey and $B$ as a predator.

$$A \xrightarrow{\kappa_1} 2A, \quad A + B \xrightarrow{\kappa_2} 2B, \quad B \xrightarrow{\kappa_3} \emptyset,$$

with $A(0) = B(0) = 1000$ and $\kappa_1 = 2$, $\kappa_2 = .002$, $\kappa_3 = 2$.

$V = 1,000$ and $\tau = 1/20 = 1/V^{.434}$ (so $\beta = .434$). Simulating 30,000 sample paths yields the following approximate distributions for $B(10)$:
Diffusion Approximation

Under the classical scaling, can use diffusion approximation

\[
X^V(t) = X^V(0) + \frac{1}{V} \sum_k Y_k \left( V \int_0^t A^V_k(X^V(s))ds \right) (\nu'_k - \nu_k)
\]

\[
\approx X^V(0) + \sum_k \int_0^t A^V_k(X^V(s))ds (\nu'_k - \nu_k)
\]

\[
+ \frac{1}{\sqrt{V}} \sum_k W_k \left( \int_0^t A^V_k(X^V(s))ds \right) (\nu'_k - \nu_k).
\]

Hence can approximate \( X^V \) with solution to

\[
dZ^V(t) = \sum_k A^V_k(Z^V(t)) (\nu'_k - \nu_k) \, dt + \frac{1}{\sqrt{V}} \sum_k (\nu'_k - \nu_k) \sqrt{A^V_k(Z^V(t))} \, dW_k(t).
\]
Error of Diffusion Approximation

Accuracy$^1$:

- Pathwise, the error is $O\left(\frac{\log(V)}{V}\right)$.
- For fixed $t$, weak error is $O\left(\frac{1}{V^2}\right)$.

1. Pathwise, tau-leaping methods can be more or less accurate, depending upon choice of $\beta$ and $V$.

<table>
<thead>
<tr>
<th>$V$</th>
<th>Euler $\tau$-leap more accurate if $\beta &gt; \beta_0$</th>
<th>Midpoint $\tau$-leap more accurate if $\beta &gt; \beta_0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>$\beta &gt; .6684$</td>
<td>$\beta &gt; .3368$</td>
</tr>
<tr>
<td>1,000</td>
<td>$\beta &gt; .7202$</td>
<td>$\beta &gt; .4404$</td>
</tr>
<tr>
<td>10,000</td>
<td>$\beta &gt; .7589$</td>
<td>$\beta &gt; .5178$</td>
</tr>
<tr>
<td>100,000</td>
<td>$\beta &gt; .7878$</td>
<td>$\beta &gt; .5756$</td>
</tr>
</tbody>
</table>

Approximating the diffusion approximation

- For fixed $t$, even midpoint $\tau$-leap will never be as accurate. However, now you must approximate the diffusion approximation.

Consider SDEs of the form

$$dX(t) = b(X(t))dt + \sum_k \sigma_k(X(t)) \nu_k \, dW_k(t),$$

where $X: \mathbb{R} \rightarrow \mathbb{R}^d$, $b: \mathbb{R}^d \rightarrow \mathbb{R}^d$, $\sigma_k: \mathbb{R}^d \rightarrow \mathbb{R}_{\geq 0}$, $\nu_k \in \mathbb{R}^d$, and $W_k(t)$ are one-dimensional Wiener processes.
Approximating the diffusion approximation

Let approximate path be $Y_i$. The following is 2nd order accurate in weak sense.

**Algorithm** \(^1\) (D. Anderson and J. Mattingly, 2010). Fixing a $\theta \in (0, 1)$, we define

$$
\alpha_1 \overset{\text{def}}{=} \frac{1}{2} \frac{1}{\theta(1 - \theta)} \quad \text{and} \quad \alpha_2 \overset{\text{def}}{=} \frac{1}{2} \frac{(1 - \theta)^2 + \theta^2}{\theta(1 - \theta)}.
$$

(1)

Next fixing a discretization step $h$, for each $i \in \{1, 2, 3, \ldots\}$ we repeat the following steps in which we first compute a $\theta$-midpoint $y^*$ and then the new value $Y_i$:

**Step 1.** Set

$$
y^* = Y_{i-1} + b(Y_{i-1})\theta h + \sum_k \sigma_k(Y_{i-1}) \nu_k \eta^{(i)}_{1k} \sqrt{\theta h}
$$

**Step 2.** Set

$$
Y_i = y^* + (\alpha_1 b(y^*) - \alpha_2 b(Y_{i-1}))(1 - \theta)h + \sum_k \sqrt{[\alpha_1 \sigma_k^2(y^*) - \alpha_2 \sigma_k^2(Y_{i-1})]^+} \nu_k \eta^{(i)}_{2k} \sqrt{(1 - \theta)h}.
$$

Why this works

System

\[ dX(t) = b(X(t))dt + \sum_{k} \sigma_k(X(t)) \nu_k \, dW_k(t), \]

is equivalent to

\[ X(t) = X(0) + \int_{0}^{t} b(X(s))ds + \sum_{k=1}^{M} \nu_k \int_{0}^{\infty} \int_{0}^{t} 1_{[0, \sigma_k^2(X(s))]}(u) \tilde{W}_k(ds \times du), \]

where the \( \tilde{W}_k \) are independent space-time white noise processes.

Challenge is in approximating diffusion term.
Why this works: $\theta = 1/2$

\[
\sigma_k^2(X(0)) \leq V \leq \sigma_k^2(X(t))
\]
Why this works: $\theta = 1/2$

Region 1

$\sigma^2_k(y^*)$

$\sigma^2_k(X(0))$

Region 2

$\sigma^2_k(X(t))$

Region 3

$\sigma^2_k(X(0))$

Region 4

$\sigma^2_k(X(t))$

Region 5

$|V| = \sigma^2_k(y^*) - \sigma^2_k(X(0))$

$|V| = \sigma^2_k(X(t)) - \sigma^2_k(X(0))$

$|V| = \sigma^2_k(X(t)) - \sigma^2_k(X(t))$

Region 6

$|V| = \sigma^2_k(X(0)) - \sigma^2_k(X(0))$

Region 7

$|V| = \sigma^2_k(X(t)) - \sigma^2_k(X(t))$

Region 8

$|V| = \sigma^2_k(X(0)) - \sigma^2_k(X(0))$
Why this works: $\theta = 1/2$

\[ \sigma_k^2(X(0)) \]

\[ \sigma_k^2(y^*) \]

Region 1

\[ V = \sigma_k^2(y^*) - \sigma_k^2(X(0)) \]

Region 2

\[ V = \sigma_k^2(y^*) - \sigma_k^2(X(0)) \]

Region 3

\[ V = \sigma_k^2(y^*) - \sigma_k^2(X(0)) \]

Region 4

\[ 2|V| \]

Region 5

\[ |V| \]

$\sigma_k^2(X(t))$