Error analysis of approximation methods for stochastically modeled chemical reaction systems

David F. Anderson*

with Arnab Ganguly and Thomas Kurtz

*anderson@math.wisc.edu

Department of Mathematics

University of Wisconsin - Madison

SAMSI Program on Stochastic Dynamics

September 1st, 2009

Outline

1. Describe the model of interest: stochastically modeled population processes.

2. Discuss possible numerical algorithms: exact and approximate.

3. Discuss error analysis and the importance of proper scalings.

Chemical reactions

- Standard notation for chemical reactions:

$$A + B \to C$$

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

- Each instance of the reaction $A + B \to C$ changes the state of the system by the vector:

$$\nu_k = \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}.$$

The model and notation

- Have d chemical species $\{S_1, S_2, \ldots, S_d\}$ undergoing a series of reactions, indexed by k.

- If reaction k occurs at time t, then the state of the system, $X(t)$, is updated via addition of the reaction vector $\nu_k \in \mathbb{Z}^d$:

$$X(t) = X(t-) + \nu_k.$$

- The waiting times for the reactions are exponentially distributed with intensity (propensity) functions $\lambda_k : \mathbb{R}^d \to \mathbb{R}_{\geq 0}$:

$$P(\, X(t + \Delta t) = X(t) + \nu_k \,|\, X(t)\,) = \lambda_k(X(t))\Delta t + o(\Delta t).$$

The model

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

$$(\mathcal{A}f)(x) = \sum_k \lambda_k(x)(f(x+\nu_k) - f(x)).$$

Kolmogorov's forward equation ("Chemical Master Equation")

$$\frac{d}{dt}P(x,t) = \sum_k \lambda_k(x-\nu_k)P(x-\nu_k,t) - \sum_k \lambda_k(x)P(x,t).$$

describes how the distribution of the process changes in time.

The model

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

$$(\mathcal{A}f)(x) = \sum_k \lambda_k(x)(f(x+\nu_k) - f(x)).$$

Kolmogorov's forward equation ("Chemical Master Equation")

$$\frac{d}{dt}P(x,t) = \sum_k \lambda_k(x-\nu_k)P(x-\nu_k,t) - \sum_k \lambda_k(x)P(x,t).$$

describes how the distribution of the process changes in time.

One intuitive representation for path-wise solutions is given by a random time-change

$$X(t) = X(0) + \sum_k Y_k\left(\int_0^t \lambda_k(X(s))ds\right)\nu_k,$$

where the Y_k are independent, unit-rate Poisson processes.

Mass-action kinetics

The standard intensity function chosen is mass-action kinetics:

$$\lambda_k(x) = \kappa_k \prod_{i=1}^{d} \frac{x_i!}{(x_i - \nu_{ki}^s)!}.$$

Example: If $S_1 \to$ anything, then $\lambda_k(x) = \kappa_k x_1$.

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

Example: If $S_1 + 2S_2 \to$ anything, then $\lambda_k(x) = \kappa_k x_1 x_2 (x_2 - 1)$.

Mass-action kinetics

The standard intensity function chosen is mass-action kinetics:

$$\lambda_k(x) = \kappa_k \prod_{i=1}^{d} \frac{x_i!}{(x_i - \nu_{ki}^s)!}.$$

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

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

Example: If $S_1 + 2S_2 \rightarrow$ anything, then $\lambda_k(x) = \kappa_k x_1 x_2 (x_2 - 1)$.

- Recall pure birth process ($X \rightarrow 2X$) described in Mark Alber's tutorial lecture.
- This type of model is used ubiquitously for cellular biochemical processes and is filling the pages of *Science* and *PNAS*.
- Typically referred to as a "chemical master equation" (Kolmogorov's forward equation) model in this literature.
- Models can be simple (single Gene-mRNA-Protein) or extremely complicated (E. coli heat shock response)

Methods of investigation: numerical simulation

$$X(t) = X(0) + \sum_k Y_k \left(\int_0^t \lambda_k(X(s))ds \right) \nu_k,$$

(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 one must find :

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

(ii) which reaction takes place at that time.

Methods of investigation: numerical simulation

$$X(t) = X(0) + \sum_k Y_k \left(\int_0^t \lambda_k(X(s))ds \right) \nu_k,$$

(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 one must find :

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

(ii) which reaction takes place at that time.

(BAD NEWS) If $\sum_k \lambda_k(X(t)) \gg 0$, 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 Δ_n may be prohibitively small.

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

$$
\begin{aligned}
Z(h) &= Z(0) + \sum_k Y_k \left(\int_0^h \lambda_k(Z(s))\, ds \right) \nu_k \\
&\approx Z(0) + \sum_k Y_k \Big(\lambda_k(Z(0))\, h \Big) \nu_k \\
&\stackrel{d}{=} Z(0) + \sum_k \text{Poisson} \Big(\lambda_k(Z(0))\, h \Big) \nu_k .
\end{aligned}
$$

[1] D. T. Gillespie, J. Chem. Phys., **115**, 1716 – 1733.

Tau-leaping

One "intuitive" representation for $Z(t)$ generated by τ-leaping is

$$Z(t) = X(0) + \sum_k Y_k \left(\int_0^t \lambda_k(Z \circ \eta(s)) ds \right) \nu_k,$$

where

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

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 $h = t_n - t_{n-1}$, let

$$\rho(z) = z + \frac{1}{2} h \sum_k \lambda_k(z) \nu_k,$$

be a "deterministic" midpoint approximation

and let $\mathcal{Z}(t)$ solve:

$$\mathcal{Z}(t) = X(0) + \sum_k Y_k \left(\int_0^t \lambda_k \circ \rho(\mathcal{Z} \circ \eta(s)) ds \right) \nu_k,$$

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

Previous error analysis

Under the scaling $h \to 0$:

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

2. Li[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:

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

$$\left|\mathbb{E}f(Z(T)) - \mathbb{E}f(X(T))\right| \leq Ch,$$

   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

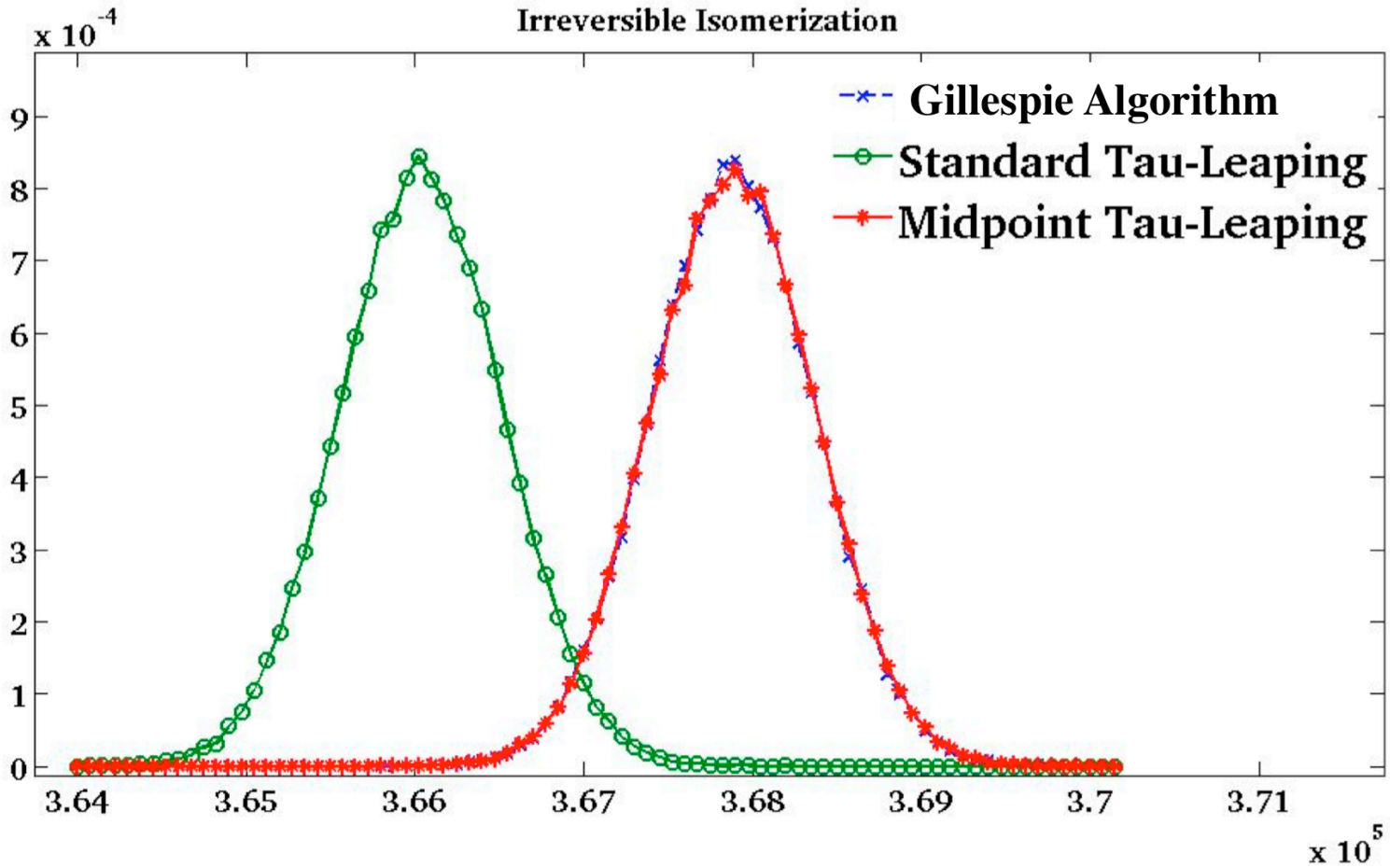
[2] M. Rathinam et al., SIAM Multi. Model. Simul., **4**, 2005, 867 – 895.
[3] T. Li, SIAM Multi. Model. Simul., **6**, 2007, 417 – 436.

An example that needs explaining: part I

Consider the simple example

$$X \xrightarrow{1} 0, \qquad X(0) = 10^6.$$

Letting $h = 1/10^2$ and simulating 100,000 sample paths with each method yields the following approximate distributions



Problem with the $h \to 0$ scaling

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

$$h \gg \frac{1}{\sum_k \lambda_k(X(t))} \approx \Delta_n \qquad \text{or} \qquad h \sum_k \lambda_k(X(t)) \gg 1.$$

So $h \to 0$ does not seem to be an appropriate scaling to describe the accuracy of these methods.

The “classical scaling”

Summary: τ-leaping is only useful if

$$h\sum_k \lambda_k(X(t)) \gg 1 \quad \text{while} \quad \frac{1}{\sum_k \lambda_k(X(t))} \ll 1.$$

We will make the following natural assumptions on our model:

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

(*ii*) The rate constants of reactions of the form

$$S_1 \to *, \qquad S_1 + S_2 \to *, \qquad S_1 + S_2 + S_3 \to *,$$

scale with V like:

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

A key property of this scaling

For each k and $x \in \mathbb{R}^d$

$$\lambda_k(Vx) = VA_k^V(x) = O(V).$$

Example:

- For the reaction $S_1 + S_2 \to *$,

$$\lambda_k(Vx) = \frac{d_k}{V}(Vx_1)(Vx_2) = V(d_k x_1 x_2).$$

Time-step

Recall: τ-leaping is only useful if

$$h \sum_k \lambda_k(X(t)) \gg 1 \quad \text{while} \quad \frac{1}{\sum_k \lambda_k(X(t))} \ll 1.$$

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

$$h = 1/V^{\beta}.$$

This scaling satisfies the above requirements:

$$h \sum_k \lambda(X(t)) = V^{-\beta} O(V) = O(V^{1-\beta}) \gg 1.$$

Error analysis on scaled processes

We consider the scaled processes

$$X^V(t) \stackrel{\text{def}}{=} \frac{X(t)}{V}, \quad Z^V(t) \stackrel{\text{def}}{=} \frac{Z(t)}{V}, \quad \mathcal{Z}^V(t) \stackrel{\text{def}}{=} \frac{\mathcal{Z}(t)}{V}.$$

which satisfy

$$X^V(t) = X^V(0) + \frac{1}{V}\sum_k Y_k\left(V\int_0^t A_k^V(X^V(s))ds\right)\nu_k$$

$$Z^V(t) = X^V(0) + \frac{1}{V}\sum_k Y_k\left(V\int_0^t A_k^V(Z^V\circ\eta(s))ds\right)\nu_k$$

$$\mathcal{Z}^V(t) = X^V(0) + \frac{1}{V}\sum_k Y_k\left(V\int_0^t A_k^V\circ\rho^V(\mathcal{Z}^V\circ\eta(s))ds\right)\nu_k,$$

where

$$\rho^V(z) \stackrel{\text{def}}{=} z + \frac{1}{2}h\sum_k A_k^V(z)\nu_k,$$

and prove theorems in this setting.

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 $h = V^{-\beta}$. There exists a constant $C = C(T) > 0$ such that

$$\sup_{t \leq T} \mathbb{E}|X^V(t) - Z^V(t)| \leq \frac{C}{V^\beta} = Ch.$$

A new coupling of $X^V(t)$ and $Z^V(t)$

To sharpen the result, let

$$X^V(t) = X^V(0) + \frac{1}{V}\sum_k \left[Y_{k,1}\left(V\int_0^t A_k^V(X^V(s)) \wedge A_k^V(Z^V \circ \eta(s))\, ds \right) + Y_{k,2}\left(V\int_0^t A_k^V(X^V(s)) - A_k^V(X^V(s)) \wedge A_k^V(Z^V \circ \eta(s))\, ds \right) \right] \nu_k$$

$$Z^V(t) = X^V(0) + \frac{1}{V}\sum_k \left[Y_{k,1}\left(V\int_0^t A_k^V(X^V(s)) \wedge A_k^V(Z^V \circ \eta(s))\, ds \right) + Y_{k,3}\left(V\int_0^t A_k^V(Z^V \circ \eta(s)) - A_k^V(X^V(s)) \wedge A_k^V(Z^V \circ \eta(s))\, ds \right) \right] \nu_k,$$

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

Note: distribution of the "marginal processes" $X^V(t)$ and $Z^V(t)$ are unchanged.

Exact asymptotics for standard 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 $V \to \infty$

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

where $\mathcal{E}(t)$ is deterministic and satisfies

$$\mathcal{E}(t) = \int_0^t DF(x(s))\mathcal{E}(s)ds + \frac{1}{2}\int_0^t DF(x(s))F(x(s))ds, \quad \mathcal{E}(0) = 0,$$

and where $f^V \to f$ means either

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

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

Idea of proof

Can show:

$$X^V(t) - Z^V(t) \approx M^V(t) + \int_0^t DF^V(Z^V \circ \eta(s))(X^V(s) - Z^V(s))ds$$
$$+ V^{-\beta}\frac{1}{2}\int_0^t DF(Z^V \circ \eta(s))F^V(Z \circ \eta(s))ds$$

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.

Weak error of standard tau-leaping

The following is now immediate.

Theorem (Anderson, Ganguly, Kurtz – 2009)

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

$$\lim_{V\to\infty} V^{\beta}\left(\mathbb{E}f(X^V(t)) - \mathbb{E}f(Z^V(t))\right) = \mathcal{E}(t)\cdot\nabla f(x(t)).$$

Strong error of midpoint tau-leaping

Theorem (Anderson, Ganguly, Kurtz – 2009)

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

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

A new coupling of $X^V(t)$ and $\mathcal{Z}^V(t)$

To sharpen the result, let

$$X^V(t) = X^V(0) + \frac{1}{V}\sum_k \left[Y_{k,1}\left(V\int_0^t A_k^V(X^V(s)) \wedge A_k^V(\rho^V \circ \mathcal{Z}^V \circ \eta(s))\, ds \right) \right.$$
$$\left. + Y_{k,2}\left(V\int_0^t A_k^V(X^V(s)) - A_k^V(X^V(s)) \wedge A_k^V(\rho^V \circ \mathcal{Z}^V \circ \eta(s))\, ds \right) \right] \nu_k$$

$$\mathcal{Z}^V(t) = X^V(0) + \frac{1}{V}\sum_k \left[Y_{k,1}\left(V\int_0^t A_k^V(X^V(s)) \wedge A_k^V(\rho^V \circ \mathcal{Z}^V \circ \eta(s))\, ds \right) \right.$$
$$\left. + Y_{k,3}\left(V\int_0^t A_k^V(\rho^V \circ \mathcal{Z}^V \circ \eta(s)) - A_k^V(X^V(s)) \wedge A_k^V(\mathcal{Z}^V \circ \eta(s))\, ds \right) \right] \nu_k,$$

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

$$V^{2\beta}(X^V - \mathcal{Z}^V) \to \mathcal{E}_1, \qquad \text{if} \qquad \beta < 1/3.$$
$$V^{2\beta}(X^V - \mathcal{Z}^V) \Rightarrow \mathcal{E}_2, \qquad \text{if} \qquad \beta = 1/3.$$
$$V^{(1+\beta)/2}(X^V - \mathcal{Z}^V) \Rightarrow \mathcal{E}_3, \qquad \text{if} \qquad \beta > 1/3.$$

Where $\mathcal{E}_1$ solves

$$\mathcal{E}_1(t) = \int_0^t DF(x(s))\mathcal{E}_1(s)ds + \mathcal{H}(t),$$

$\mathcal{E}_2$ solves

$$\mathcal{E}_2(t) = M(t) + \int_0^t DF(x(s))\mathcal{E}_2(s)ds + \mathcal{H}(t),$$

$\mathcal{E}_3$ solves

$$\mathcal{E}_3(t) = M(t) + \int_0^t DF(x(s))\mathcal{E}_3(s)ds.$$

Idea of proof

Can show:

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

where $V^{(1+\beta)/2}M^V \Rightarrow M(t)$.

For those who must know:

$$\mathcal{H}(t) = \frac{1}{6}\int_0^t DF(x(s))^2 F(x(s))ds + \frac{1}{24}\int_0^t F(x(s))^T HF(x(s))F(x(s))ds.$$

Weak error of midpoint tau-leaping

Theorem (Anderson, Ganguly, Kurtz – 2009)

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

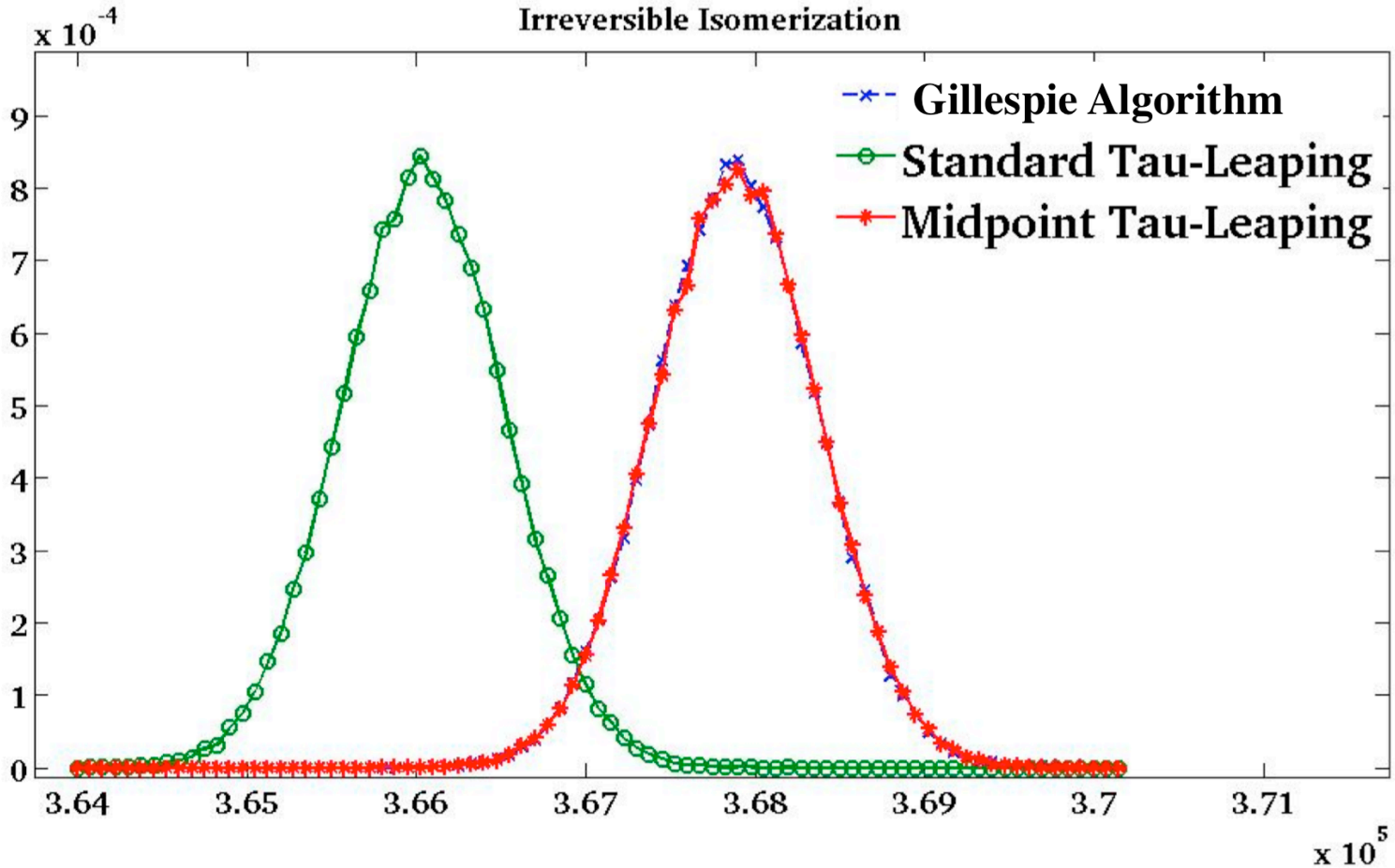
$$V^{2\beta} \sup_{t \leq T} \left| \mathbb{E} f(X^V(t)) - \mathbb{E} f(\mathcal{Z}^V(t)) \right| \leq C.$$

An example that needs explaining: part II

Consider the simple example

$$X \xrightarrow{1} 0, \qquad X(0) = 10^6.$$

$V = 10^6$. Letting $h = 1/10^2 = 1/V^{1/3}$ (so $\beta = 1/3$) and simulating 100,000 sample paths with each method yields the following approximate distributions



Take home messages

1. Scaling to get instructive error estimates is a tricky business and must be handled with care.

2. Can be informative in that it tells you what algorithms do (and which should be used).