# Revisiting randomised time integration for differential equations

Han C. Lie

Joint work with M. Stahn, T.J. Sullivan, A. Teckentrup, O. Teymur

2024 ProbNum Spring School, Southampton

## Motivation: Bayesian inference for initial value problems

Observation model:

$$y = G(x^*) + \varepsilon$$

$y$ - data
$G : \mathcal{X} \to \mathcal{Y}$ - parameter-to-observable map or 'forward model'
$x^*$ - true data-generating parameter
$\varepsilon$ - obs. noise

Observation model:

$$y = G(x^*) + \varepsilon$$

$y$ - data
$G : \mathcal{X} \to \mathcal{Y}$ - parameter-to-observable map or 'forward model'
$x^*$ - true data-generating parameter
$\varepsilon$ - obs. noise

Inverse problem: Given data $y$, infer 'truth' $x^*$

Bayesian approach:

1) model $x^*$ using RV $X$ with prior law $\mu_{\text{pri}}$
2) choose likelihood $\ell(x; y)$, e.g. if $\varepsilon \sim \mathcal{N}(0, C)$ for $C > 0$, then

$$\ell(x; y) = \frac{1}{Z(y)} \exp(-\tfrac{1}{2} \|y - G(x)\|^2_{C^{-1}}), \quad \|w\|^2_{C^{-1}} := w^\top C^{-1} w$$

3) solve inverse problem with posterior $\mu^y_{\text{pos}}$, obtained by Bayes:

$$\mu^y_{\text{pos}}(\,\mathrm{d}x) \propto \ell(x; y)\mu_{\text{pri}}(\,\mathrm{d}x)$$

Initial value problem (IVP) on $[0, T]$ with unique solution $z(\cdot\,, x)$

$$\frac{\mathrm{d}}{\mathrm{d}t} z(t, x) = f(z(t, x), x), \qquad z(0, x) = z_0(x)$$

- vector field $f$, init. cond. $z_0$ may depend on $x$

Initial value problem (IVP) on $[0, T]$ with unique solution $z(\cdot\,, x)$

$$\frac{\mathrm{d}}{\mathrm{d}t} z(t, x) = f(z(t, x), x), \qquad z(0, x) = z_0(x)$$

- vector field $f$, init. cond. $z_0$ may depend on $x$

Associated solution operator $S$:

$$S : \mathcal{X} \to C([0, T]; H) \qquad x \mapsto z(\cdot\,, x)$$

Fix time points $\{t_1, \dots, t_J\} \subset [0, T]$, define observation operator

$$O : C([0, T]; H) \to H^J, \quad z \mapsto O(z) = [z(t_1)^\top, \dots, z(t_J)^\top]^\top \in H^J$$

Parameter-to-observable map $G := O \circ S : \mathcal{X} \to H^J$

Initial value problem (IVP) on $[0, T]$ with unique solution $z(\cdot\,, x)$

$$\frac{\mathrm{d}}{\mathrm{d}t} z(t, x) = f(z(t, x), x), \qquad z(0, x) = z_0(x)$$

- vector field $f$, init. cond. $z_0$ may depend on $x$

Associated solution operator $S$:

$$S : \mathcal{X} \to C([0, T]; H) \qquad x \mapsto z(\cdot\,, x)$$

Fix time points $\{t_1, \dots, t_J\} \subset [0, T]$, define observation operator

$$O : C([0, T]; H) \to H^J, \quad z \mapsto O(z) = [z(t_1)^\top, \dots, z(t_J)^\top]^\top \in H^J$$

Parameter-to-observable map $G := O \circ S : \mathcal{X} \to H^J$

Problem: true solution map $S : \mathcal{X} \to C([0, T]; H)$ not available

$\Rightarrow$ True forward model $G := O \circ S$ not available

Example: parametrised IVP on $[0, T]$

$$\frac{\mathrm{d}}{\mathrm{d}t} z(t,x) = f(z(t,x),x), \qquad z(0,x) = z_0(x)$$

Cannot evaluate map $x \mapsto S(x) = z(\cdot\,, x)$ exactly

Solution: Approximate solution map

$$S_h : \mathcal{X} \to C([0,T];H), \qquad x \mapsto (z_h(t,x))_{t\in[0,T]}$$

$h$ - resolution, e.g. time step $h > 0$

Example: parametrised IVP on $[0, T]$

$$\frac{\mathrm{d}}{\mathrm{d}t} z(t,x) = f(z(t,x),x), \qquad z(0,x) = z_0(x)$$

Cannot evaluate map $x \mapsto S(x) = z(\cdot\,,x)$ exactly

Solution: Approximate solution map

$$S_h : \mathcal{X} \to C([0,T];H), \qquad x \mapsto (z_h(t,x))_{t\in[0,T]}$$

$h$ - resolution, e.g. time step $h > 0$

Resulting approximate objects:

- forward model $G_h := O \circ S_h$
- likelihood $\ell_h(x;y) = \frac{1}{Z_h(y)} \exp(-\frac{1}{2}\|y - G_h(x)\|^2_{C^{-1}})$
- posterior $\mu^y_{\mathrm{pos},h}(\mathrm{d}x) \propto \ell_h(x;y)\mu_{\mathrm{pri}}(\mathrm{d}x)$

Example: parametrised IVP on $[0, T]$

$$\frac{\mathrm{d}}{\mathrm{d}t} z(t, x) = f(z(t, x), x), \qquad z(0, x) = z_0(x)$$

Cannot evaluate map $x \mapsto S(x) = z(\cdot\,, x)$ exactly

Solution: Approximate solution map

$$S_h : \mathcal{X} \to C([0, T]; H), \qquad x \mapsto (z_h(t, x))_{t \in [0,T]}$$

$h$ - resolution, e.g. time step $h > 0$

Resulting approximate objects:

forward model $G_h := O \circ S_h$

likelihood $\ell_h(x; y) = \frac{1}{Z_h(y)} \exp(-\frac{1}{2} \| y - G_h(x) \|^2_{C^{-1}})$

posterior $\mu^y_{\mathrm{pos},h}(\mathrm{d}x) \propto \ell_h(x; y) \mu_{\mathrm{pri}}(\mathrm{d}x)$

**Questions**: 1) How to account for error in $G_h$?
2) How does error in $G_h$ propagate to error in $\mu^y_{\mathrm{pos},h}$?

Numerical methods and statistical inference:

Ignoring numerical errors leads to overconfidence

## Q1: How to account for error in $G_h$?

Answer of Conrad et al. (2017): Use random variables (RVs)

## Q1: How to account for error in $G_h$?

Answer of Conrad et al. (2017): Use random variables (RVs)

Example: parametrised IVP

$$\frac{\mathrm{d}}{\mathrm{d}t} z(t,x) = f(z(t,x),x), \qquad z(0,x) = z_0(x)$$

Deterministic time integrator $\psi : H \to H$ with time step $h = \frac{T}{N} > 0$:

$$z_h((k+1)h,x) := \psi(z_h(kh,x)), \quad k = 0,\dots,N-1$$

Define $S_h(x)$ by interpolating between $(z_h(kh,x))_{k=0}^{N}$

**Q1: How to account for error in $G_h$?**

Answer of Conrad et al. (2017): Use random variables (RVs)

Example: parametrised IVP

$$\frac{\mathrm{d}}{\mathrm{d}t} z(t,x) = f(z(t,x),x), \qquad z(0,x) = z_0(x)$$

Deterministic time integrator $\psi : H \to H$ with time step $h = \frac{T}{N} > 0$:

$$z_h((k+1)h, x) := \psi(z_h(kh, x)), \quad k = 0, \dots, N-1$$

Define $S_h(x)$ by interpolating between $(z_h(kh,x))_{k=0}^{N}$

**Randomised time integrator** with RVs $(\xi_k(h))_k$, not necessarily i.i.d.

$$\widehat{Z}_h((k+1)h, x, \omega) := \psi(\widehat{Z}_h(kh, x, \omega)) + \xi_k(h, \omega), \quad k = 0, \dots, N-1$$

- $\xi_k(h,\omega)$ models error in time evolution map $\psi$ over $[kh, (k+1)h]$

Define random soln. operator $\widehat{S}_h(x,\omega)$ by interpolation

Example: Fitzhugh–Nagumo IVP

Solution $z=(z_1,z_2)\in C([0,T];\mathbb{R}^2)$, parameter $\theta=(\theta_1,\theta_2,\theta_3)\in\mathbb{R}^3$

$$\begin{aligned}\frac{\mathrm{d}z_1}{\mathrm{d}t} &=\theta_3\left(z_1-\frac{z_1^3}{3}+z_2\right)\\ \frac{\mathrm{d}z_2}{\mathrm{d}t} &=-\frac{1}{\theta_3}\left(z_1-\theta_1+\theta_2 z_2\right)\end{aligned}$$

Solution map $S$: $\theta\mapsto z(\cdot\,,\theta)\in C([0,T];\mathbb{R}^2)$

Deterministic integrator: implicit Euler $\psi:\mathbb{R}^2\to\mathbb{R}^2$, time step $h=\frac{T}{N}$

$\rightsquigarrow$ approx. solution operator $S_h$

Randomised integrator uses

$$\widehat{Z}_h((k+1)h,\theta,\omega):=\psi(\widehat{Z}_h(kh,\theta,\omega))+\xi_k(h,\omega),\quad k=0,\dots,N-1$$

with RVs $(\xi_k(h))_k$ with joint law $\nu_h$

$\rightsquigarrow$ random approx. solution operator $\widehat{S}_h$

# Ensemble of solutions to IVP

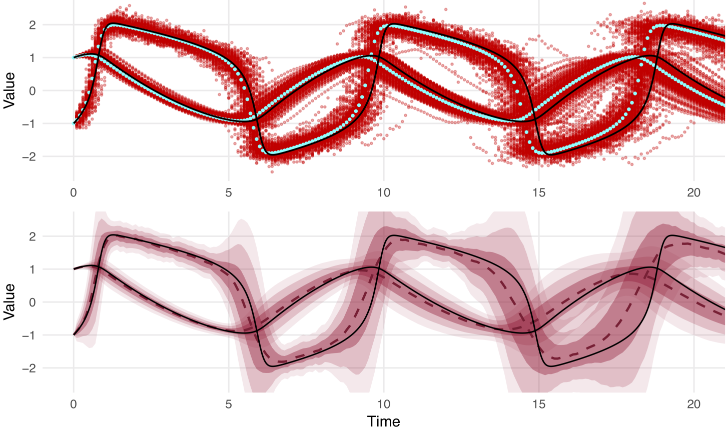


Figure: Ensemble of 500 realisations of $\widehat{Z}_h$ from randomised implicit Euler with $h = 0.1$, $T = 20$.

**Top**: Ensemble of $\widehat{Z}_h$ without interpolation (red)
Ensemble mean (light blue);
Deterministic solution $z$ (black solid line)

**Bottom**: Ensemble of $\widehat{Z}_h$ with interpolation

Image credit: Teymur et al., 2018

# Tackling overconfidence using randomisation

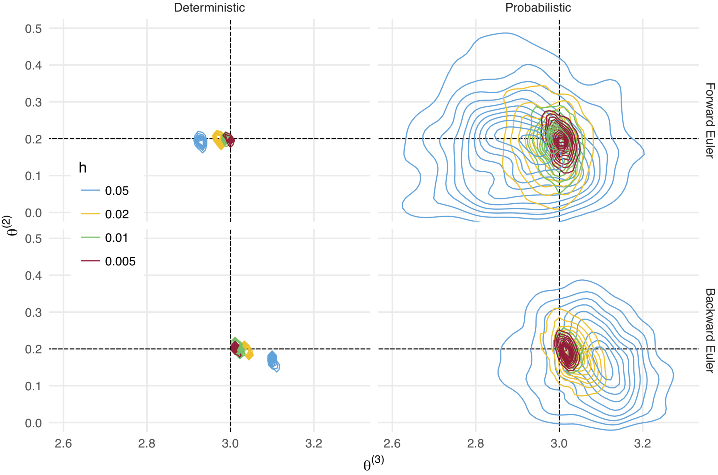


Figure: **Left column**: Density contour plots of approx. posterior $\mu^{y}_{\text{pos},h}$ on $(\theta_2, \theta_3)$

**Right column**: Density contour plots of approx. posterior $\widehat{\mu}^{y}_{\text{pos},h}$ on $(\theta_2, \theta_3)$

Step sizes for ODE solver: $h = 0.005$ (red), 0.01 (green), 0.02 (yellow), 0.05 (blue).

Image credit: Teymur et al., NeurIPS 2018

Analytical example of posterior overconfidence

See also Abdulle and Garegnani (2020)

# Posterior consistency for linear Gaussian BIP

Initial value problem (IVP) for $t \in [0, h]$, self-adjoint, linear, positive $A$:

$$\frac{\mathrm{d}}{\mathrm{d}t} z(t, x) = -Az(t, x), \qquad z(0, x) = x$$

Linear forward model $G : \mathcal{X} \to \mathcal{Y}, \quad x \mapsto e^{-Ah}x = z(h, x)$

Data: $y = G(x^*) + \sigma\varepsilon, \quad \varepsilon \sim \mathcal{N}(0, \Gamma_{\mathrm{obs}})$

Inverse problem: infer true parameter $x^*$ from $y$

# Posterior consistency for linear Gaussian BIP

Initial value problem (IVP) for $t \in [0, h]$, self-adjoint, linear, positive $A$:

$$\frac{\mathrm{d}}{\mathrm{d}t} z(t,x) = -Az(t,x), \qquad z(0,x) = x$$

Linear forward model $G : \mathcal{X} \to \mathcal{Y}, \quad x \mapsto e^{-Ah}x = z(h,x)$

Data: $y = G(x^*) + \sigma\varepsilon, \quad \varepsilon \sim \mathcal{N}(0, \Gamma_{\text{obs}})$

Inverse problem: infer true parameter $x^*$ from $y$

If prior is $\mu_{\text{pri}} = \mathcal{N}(m_{\text{pri}}, \Gamma_{\text{pri}})$, then posterior is $\mu^y_{\text{pos}} = \mathcal{N}(m, \mathcal{C})$

$$m = m_{\text{pri}} + \Gamma_{\text{pri}} G(\sigma\Gamma_{\text{obs}} + G\Gamma_{\text{pri}} G)^{-1}(y - Gm_{\text{pri}})$$
$$\mathcal{C} = \Gamma_{\text{pri}} - \Gamma_{\text{pri}} G(\sigma\Gamma_{\text{obs}} + G\Gamma_{\text{pri}} G)^{-1} G\Gamma_{\text{pri}}$$

# Posterior consistency for linear Gaussian BIP

Initial value problem (IVP) for $t \in [0, h]$, self-adjoint, linear, positive $A$:

$$\frac{\mathrm{d}}{\mathrm{d}t} z(t, x) = -Az(t, x), \qquad z(0, x) = x$$

Linear forward model $G : \mathcal{X} \to \mathcal{Y}, \quad x \mapsto e^{-Ah}x = z(h, x)$

Data: $y = G(x^*) + \sigma\varepsilon, \quad \varepsilon \sim \mathcal{N}(0, \Gamma_{\text{obs}})$

Inverse problem: infer true parameter $x^*$ from $y$

If prior is $\mu_{\text{pri}} = \mathcal{N}(m_{\text{pri}}, \Gamma_{\text{pri}})$, then posterior is $\mu^y_{\text{pos}} = \mathcal{N}(m, \mathcal{C})$

$$m = m_{\text{pri}} + \Gamma_{\text{pri}} G(\sigma\Gamma_{\text{obs}} + G\Gamma_{\text{pri}} G)^{-1}(y - Gm_{\text{pri}})$$
$$\mathcal{C} = \Gamma_{\text{pri}} - \Gamma_{\text{pri}} G(\sigma\Gamma_{\text{obs}} + G\Gamma_{\text{pri}} G)^{-1} G\Gamma_{\text{pri}}$$

Behaviour in **small noise limit** $\sigma \to 0$:

$\mathcal{C} \xrightarrow[\sigma \to 0]{} 0$ (vanishing variance), $\quad m \xrightarrow[\sigma \to 0]{} x^*$ (vanishing bias)

Hence, $\mu^y_{\text{pos}} \Rightarrow \delta_{x^*}$ (consistency of $\mu^y_{\text{pos}}$)

# Lack of consistency for approximate posterior

Initial value problem (IVP) for $t \in [0, h]$

$$\frac{\mathrm{d}}{\mathrm{d}t} z(t,x) = -Az(t,x), \qquad z(0,x) = x$$

Implicit Euler with time step $h > 0$: $x \mapsto (I + hA)^{-1} x$

Approximate forward model $\widetilde{G}_h : \mathcal{X} \to \mathcal{Y}, \quad x \mapsto (I + hA)^{-1} x$

Use same data, same prior as before

# Lack of consistency for approximate posterior

Initial value problem (IVP) for $t \in [0, h]$

$$\frac{\mathrm{d}}{\mathrm{d}t} z(t,x) = -Az(t,x), \qquad z(0,x) = x$$

Implicit Euler with time step $h > 0$: $x \mapsto (I + hA)^{-1} x$

Approximate forward model $\widetilde{G}_h : \mathcal{X} \to \mathcal{Y}, \quad x \mapsto (I + hA)^{-1} x$

Use same data, same prior as before

Approximate posterior $\widetilde{\mu}^y_{\text{pos},h} = \mathcal{N}(\widetilde{m}_h, \widetilde{\mathcal{C}}_h)$

$$\widetilde{m}_h = m_{\text{pri}} + \Gamma_{\text{pri}} \widetilde{G}_h (\sigma \Gamma_{\text{obs}} + \widetilde{G}_h \Gamma_{\text{pri}} \widetilde{G}_h)^{-1} (y - \widetilde{G}_h m_{\text{pri}})$$

$$\widetilde{\mathcal{C}}_h = \Gamma_{\text{pri}} - \Gamma_{\text{pri}} \widetilde{G}_h (\sigma \Gamma_{\text{obs}} + \widetilde{G}_h \Gamma_{\text{pri}} \widetilde{G}_h)^{-1} \widetilde{G}_h \Gamma_{\text{pri}}$$

# Lack of consistency for approximate posterior

Initial value problem (IVP) for $t \in [0, h]$

$$\frac{\mathrm{d}}{\mathrm{d}t} z(t, x) = -Az(t, x), \qquad z(0, x) = x$$

Implicit Euler with time step $h > 0$: $x \mapsto (I + hA)^{-1} x$

Approximate forward model $\widetilde{G}_h : \mathcal{X} \to \mathcal{Y}, \quad x \mapsto (I + hA)^{-1} x$

Use same data, same prior as before

Approximate posterior $\widetilde{\mu}^y_{\mathrm{pos},h} = \mathcal{N}(\widetilde{m}_h, \widetilde{\mathcal{C}}_h)$

$$\widetilde{m}_h = m_{\mathrm{pri}} + \Gamma_{\mathrm{pri}} \widetilde{G}_h (\sigma \Gamma_{\mathrm{obs}} + \widetilde{G}_h \Gamma_{\mathrm{pri}} \widetilde{G}_h)^{-1} (y - \widetilde{G}_h m_{\mathrm{pri}})$$
$$\widetilde{\mathcal{C}}_h = \Gamma_{\mathrm{pri}} - \Gamma_{\mathrm{pri}} \widetilde{G}_h (\sigma \Gamma_{\mathrm{obs}} + \widetilde{G}_h \Gamma_{\mathrm{pri}} \widetilde{G}_h)^{-1} \widetilde{G}_h \Gamma_{\mathrm{pri}}$$

Behaviour in small noise limit: $\widetilde{\mathcal{C}}_h \xrightarrow[\sigma \to 0]{} 0$ (vanishing variance)

$\widetilde{m}_h \xrightarrow[\sigma \to 0]{} \widetilde{G}_h^{-1} G x^* = (I + hA) e^{-hA} x^*$

# Lack of consistency for approximate posterior

Initial value problem (IVP) for $t \in [0, h]$

$$\frac{\mathrm{d}}{\mathrm{d}t} z(t, x) = -Az(t, x), \qquad z(0, x) = x$$

Implicit Euler with time step $h > 0$: $x \mapsto (I + hA)^{-1} x$

Approximate forward model $\widetilde{G}_h : \mathcal{X} \to \mathcal{Y}, \quad x \mapsto (I + hA)^{-1} x$

Use same data, same prior as before

Approximate posterior $\widetilde{\mu}^y_{\text{pos},h} = \mathcal{N}(\widetilde{m}_h, \widetilde{\mathcal{C}}_h)$

$$\widetilde{m}_h = m_{\text{pri}} + \Gamma_{\text{pri}} \widetilde{G}_h (\sigma \Gamma_{\text{obs}} + \widetilde{G}_h \Gamma_{\text{pri}} \widetilde{G}_h)^{-1} (y - \widetilde{G}_h m_{\text{pri}})$$
$$\widetilde{\mathcal{C}}_h = \Gamma_{\text{pri}} - \Gamma_{\text{pri}} \widetilde{G}_h (\sigma \Gamma_{\text{obs}} + \widetilde{G}_h \Gamma_{\text{pri}} \widetilde{G}_h)^{-1} \widetilde{G}_h \Gamma_{\text{pri}}$$

Behaviour in small noise limit: $\widetilde{\mathcal{C}}_h \xrightarrow[\sigma \to 0]{} 0$ (vanishing variance)

$\widetilde{m}_h \xrightarrow[\sigma \to 0]{} \widetilde{G}_h^{-1} G x^* = (I + hA) e^{-hA} x^* \neq x^*$ (nonvanishing bias)

# Lack of consistency for approximate posterior

Initial value problem (IVP) for $t \in [0, h]$

$$\frac{\mathrm{d}}{\mathrm{d}t} z(t, x) = -Az(t, x), \qquad z(0, x) = x$$

Implicit Euler with time step $h > 0$: $x \mapsto (I + hA)^{-1} x$

Approximate forward model $\widetilde{G}_h : \mathcal{X} \to \mathcal{Y}, \quad x \mapsto (I + hA)^{-1} x$

Use same data, same prior as before

Approximate posterior $\widetilde{\mu}^y_{\text{pos},h} = \mathcal{N}(\widetilde{m}_h, \widetilde{\mathcal{C}}_h)$

$$\widetilde{m}_h = m_{\text{pri}} + \Gamma_{\text{pri}} \widetilde{G}_h (\sigma \Gamma_{\text{obs}} + \widetilde{G}_h \Gamma_{\text{pri}} \widetilde{G}_h)^{-1} (y - \widetilde{G}_h m_{\text{pri}})$$
$$\widetilde{\mathcal{C}}_h = \Gamma_{\text{pri}} - \Gamma_{\text{pri}} \widetilde{G}_h (\sigma \Gamma_{\text{obs}} + \widetilde{G}_h \Gamma_{\text{pri}} \widetilde{G}_h)^{-1} \widetilde{G}_h \Gamma_{\text{pri}}$$

Behaviour in small noise limit: $\widetilde{\mathcal{C}}_h \xrightarrow[\sigma \to 0]{} 0$ (vanishing variance)

$\widetilde{m}_h \xrightarrow[\sigma \to 0]{} \widetilde{G}_h^{-1} G x^* = (I + hA) e^{-hA} x^* \neq x^*$ (nonvanishing bias)

Hence, $\widetilde{\mu}^y_{\text{pos},h} \not\Rightarrow \delta_{x^*}$ as $\sigma \to 0$ (inconsistency of approx. posterior)

# Some observations

# Some observations

Approximate posterior $\widetilde{\mu}^{y}_{\mathrm{pos},h}$ is **overconfident** in $\sigma \to 0$ limit

- $\widetilde{\mu}^{y}_{\mathrm{pos},h}$ has vanishing variance: $\widetilde{\mathcal{C}}_h \xrightarrow[\sigma\to 0]{} 0$
- $\widetilde{\mu}^{y}_{\mathrm{pos},h}$ has nonvanishing bias: $\widetilde{m}_h \xrightarrow[\sigma\to 0]{} \widetilde{G}_h^{-1} G x^* \neq x^*$

# Some observations

Approximate posterior $\widetilde{\mu}^{y}_{\text{pos},h}$ is **overconfident** in $\sigma \to 0$ limit

- $\widetilde{\mu}^{y}_{\text{pos},h}$ has vanishing variance: $\widetilde{\mathcal{C}}_h \xrightarrow[\sigma\to 0]{} 0$
- $\widetilde{\mu}^{y}_{\text{pos},h}$ has nonvanishing bias: $\widetilde{m}_h \xrightarrow[\sigma\to 0]{} \widetilde{G}_h^{-1} G x^* \neq x^*$

• Bias is related to error $\widetilde{G}_h - G$:

$$\|x^* - \widetilde{G}_h^{-1} G x^*\| = \|\widetilde{G}_h^{-1}(\widetilde{G}_h - G)x^*\| \leq \|\widetilde{G}_h^{-1}\| \|\widetilde{G}_h - G\| \|x^*\|$$

$\widetilde{G}_h - G \leftrightarrow$ worst-case error of one $h$-step of implicit Euler

# Some observations

Approximate posterior $\widetilde{\mu}^{y}_{\mathrm{pos},h}$ is **overconfident** in $\sigma \to 0$ limit

- $\widetilde{\mu}^{y}_{\mathrm{pos},h}$ has vanishing variance: $\widetilde{\mathcal{C}}_h \xrightarrow[\sigma\to 0]{} 0$
- $\widetilde{\mu}^{y}_{\mathrm{pos},h}$ has nonvanishing bias: $\widetilde{m}_h \xrightarrow[\sigma\to 0]{} \widetilde{G}_h^{-1} G x^* \neq x^*$

• Bias is related to error $\widetilde{G}_h - G$:

$$\|x^* - \widetilde{G}_h^{-1} G x^*\| = \|\widetilde{G}_h^{-1}(\widetilde{G}_h - G)x^*\| \leq \|\widetilde{G}_h^{-1}\|\|\widetilde{G}_h - G\|\|x^*\|$$

$\widetilde{G}_h - G \leftrightarrow$ worst-case error of one $h$-step of implicit Euler

• Conrad et al.: Use randomisation to account for unknown error

# Randomised forward model

Initial value problem (IVP) for $t \in [0, h]$

$$\frac{\mathrm{d}}{\mathrm{d}t} z(t, x) = -Az(t, x), \qquad z(0, x) = x$$

# Randomised forward model

Initial value problem (IVP) for $t \in [0, h]$

$$\frac{\mathrm{d}}{\mathrm{d}t} z(t, x) = -Az(t, x), \qquad z(0, x) = x$$

Randomised forward model (implicit Euler) with scaling $h^{p+1}$:

$$\widehat{G}_h : \mathcal{X} \times \Omega \to \mathcal{Y}, \quad (x, \omega) \mapsto (I + hA)^{-1} x + h^{p+1} \zeta(\omega) \text{ for } \zeta \sim \mathcal{N}(0, \Gamma_1)$$

# Randomised forward model

Initial value problem (IVP) for $t \in [0, h]$

$$\frac{\mathrm{d}}{\mathrm{d}t} z(t,x) = -Az(t,x), \qquad z(0,x) = x$$

Randomised forward model (implicit Euler) with scaling $h^{p+1}$:

$$\widehat{G}_h : \mathcal{X} \times \Omega \to \mathcal{Y}, \quad (x,\omega) \mapsto (I + hA)^{-1} x + h^{p+1} \zeta(\omega) \text{ for } \zeta \sim \mathcal{N}(0, \Gamma_1)$$

- Use same data and prior. Use $h^{p+1}\zeta$ to model error $(\widetilde{G}_h - G)x$

# Randomised forward model

Initial value problem (IVP) for $t \in [0, h]$

$$\frac{\mathrm{d}}{\mathrm{d}t} z(t,x) = -Az(t,x), \qquad z(0,x) = x$$

Randomised forward model (implicit Euler) with scaling $h^{p+1}$:

$$\widehat{G}_h : \mathcal{X} \times \Omega \to \mathcal{Y}, \quad (x, \omega) \mapsto (I + hA)^{-1} x + h^{p+1} \zeta(\omega) \text{ for } \zeta \sim \mathcal{N}(0, \Gamma_1)$$

- Use same data and prior. Use $h^{p+1}\zeta$ to model error $(\widetilde{G}_h - G)x$

Approximate posterior $\widehat{\mu}^y_{\text{pos},h} = \mathcal{N}(\widehat{m}_h, \widehat{\mathcal{C}}_h)$: $\widehat{m}_h = \widetilde{m}_h$ and

$$\widehat{\mathcal{C}}_h = \Gamma_{\text{pri}} - \Gamma_{\text{pri}} \widetilde{G}_h (\sigma \Gamma_{\text{obs}} + h^{2p+2} \Gamma_1 + \widetilde{G}_h \Gamma_{\text{pri}} \widetilde{G})_h^{-1} \widetilde{G}_h \Gamma_{\text{pri}}$$

# Randomised forward model

Initial value problem (IVP) for $t \in [0, h]$

$$\tfrac{\mathrm{d}}{\mathrm{d}t} z(t,x) = -Az(t,x), \qquad z(0,x) = x$$

Randomised forward model (implicit Euler) with scaling $h^{p+1}$:

$$\widehat{G}_h : \mathcal{X} \times \Omega \to \mathcal{Y}, \quad (x,\omega) \mapsto (I + hA)^{-1} x + h^{p+1} \zeta(\omega) \text{ for } \zeta \sim \mathcal{N}(0, \Gamma_1)$$

- Use same data and prior. Use $h^{p+1}\zeta$ to model error $(\widetilde{G}_h - G)x$

Approximate posterior $\widehat{\mu}^y_{\text{pos},h} = \mathcal{N}(\widehat{m}_h, \widehat{\mathcal{C}}_h)$: $\widehat{m}_h = \widetilde{m}_h$ and

$$\widehat{\mathcal{C}}_h = \Gamma_{\text{pri}} - \Gamma_{\text{pri}} \widetilde{G}_h (\sigma \Gamma_{\text{obs}} + h^{2p+2} \Gamma_1 + \widetilde{G}_h \Gamma_{\text{pri}} \widetilde{G})_h^{-1} \widetilde{G}_h \Gamma_{\text{pri}}$$

- Variance inflation term $h^{2p+2}\Gamma_1$ counteracts overconfidence: for fixed $h > 0$,

$$\lim_{\sigma \to 0} \widehat{\mathcal{C}}_h = \Gamma_{\text{pri}} - \Gamma_{\text{pri}} \widetilde{G}_h (h^{2p+2} \Gamma_1 + \widetilde{G}_h \Gamma_{\text{pri}} \widetilde{G}_h)^{-1} \widetilde{G}_h \Gamma_{\text{pri}} \neq 0$$

True posterior associated to true $G(\cdot)$:

$$\mu^y_{\text{pos}}(\,\mathrm{d}x) \propto \exp(-\tfrac{1}{2}\,\|y - G(x)\|^2_{C^{-1}})\mu_{\text{pri}}(\,\mathrm{d}x)$$

Random approx. posterior associated to random $\widehat{G}_h(\cdot\,,\omega) \sim \nu_h$:

$$\widehat{\mu}^y_{\text{pos},h}(\,\mathrm{d}x,\omega) \propto \exp(-\tfrac{1}{2}\|y - \widehat{G}_h(x,\omega)\|^2_{C^{-1}})\mu_{\text{pri}}(\,\mathrm{d}x)$$

True posterior associated to true $G(\cdot)$:

$$\mu^y_{\text{pos}}(\,\mathrm{d}x) \propto \exp(-\tfrac{1}{2}\,\|y - G(x)\|^2_{C^{-1}})\mu_{\text{pri}}(\,\mathrm{d}x)$$

Random approx. posterior associated to random $\widehat{G}_h(\cdot\,,\omega) \sim \nu_h$:

$$\widehat{\mu}^y_{\text{pos},h}(\,\mathrm{d}x,\omega) \propto \exp(-\tfrac{1}{2}\|y - \widehat{G}_h(x,\omega)\|^2_{C^{-1}})\mu_{\text{pri}}(\,\mathrm{d}x)$$

**Q2: How does error in $\widehat{G}_h$ propagate to error in $\widehat{\mu}^y_{\mathbf{pos},h}$?**

A: Theorem (L., Sullivan and Teckentrup, 2018)
Under conditions on $G$ and $(\widehat{G}_h)_{h>0}$, there exist $s_1, s_2 \geq 1$ and $D > 0$ that do not depend on $h$, such that

$$\mathbb{E}_{\nu_N}\left[d_{\text{Hell}}(\mu^y_{\text{pos}}, \widehat{\mu}^y_{\text{pos},h})^2\right]^{\frac{1}{2}} \leq D\,\mathbb{E}_{\mu_{\text{pri}}}\Big[\,\mathbb{E}_{\nu_h}\Big[\Big\|G - \widehat{G}_h\Big\|^{2s_1}\Big]^{\frac{s_2}{s_1}}\Big]^{\frac{1}{2s_2}}$$

◦ Local Lipschitz stability of $\mu^y_{\text{pos}}$ (Stuart 2010; Sprungk 2020)
◦ Similar result for Kullback–Leibler divergence $d_{\text{KL}}$ can be proven

Local Lipschitz stability inequality

$$\mathbb{E}_{\nu_N}\left[d_{\text{Hell}}(\mu^y_{\text{pos}}, \widehat{\mu}^y_{\text{pos},h})^2\right]^{\frac{1}{2}} \leq D\, \mathbb{E}_{\mu_{\text{pri}}}[\, \mathbb{E}_{\nu_h}\left[\left\| G - \widehat{G}_h \right\|^{2s_1}\right]^{\frac{s_2}{s_1}}]^{\frac{1}{2s_2}}$$

Local Lipschitz stability inequality

$$\mathbb{E}_{\nu_N}\left[d_{\text{Hell}}(\mu^y_{\text{pos}}, \widehat{\mu}^y_{\text{pos},h})^2\right]^{\frac{1}{2}} \leq D\, \mathbb{E}_{\mu_{\text{pri}}}[\, \mathbb{E}_{\nu_h}\left[\left\|G - \widehat{G}_h\right\|^{2s_1}\right]^{\frac{s_2}{s_1}}]^{\frac{1}{2s_2}}$$

**Q3: How to control right-hand side?**

Recall time integration of IVP with $H$-valued state variable
If $O$ given by $J$ state observations at times $(t_k)_{k=1}^J$,

$$O : C([0,T]; H) \to H^J, \quad z \mapsto O(z) = [z(t_1)^\top, \dots, z(t_J)^\top]^\top \in H^J,$$

Forward models $G = O \circ S$ and $\widehat{G}_h = O \circ \widehat{S}_h$ take values in $H^J$,

$$\|G(x) - \widehat{G}_N(x,\omega)\|_{H^J} \leq J \sup_{0 \leq k \leq T/h} \|z(t_k, x) - \widehat{Z}_h(t_k, x, \omega)\|_H$$

Local Lipschitz stability inequality

$$\mathbb{E}_{\nu_N}\left[d_{\text{Hell}}(\mu^y_{\text{pos}}, \widehat{\mu}^y_{\text{pos},h})^2\right]^{\frac{1}{2}} \leq D\,\mathbb{E}_{\mu_{\text{pri}}}[\,\mathbb{E}_{\nu_h}\left[\left\|G - \widehat{G}_h\right\|^{2s_1}\right]^{\frac{s_2}{s_1}}]^{\frac{1}{2s_2}}$$

**Q3: How to control right-hand side?**

Recall time integration of IVP with $H$-valued state variable
If $O$ given by $J$ state observations at times $(t_k)_{k=1}^J$,

$$O : C([0, T]; H) \to H^J, \quad z \mapsto O(z) = [z(t_1)^\top, \dots, z(t_J)^\top]^\top \in H^J,$$

Forward models $G = O \circ S$ and $\widehat{G}_h = O \circ \widehat{S}_h$ take values in $H^J$,

$$\|G(x) - \widehat{G}_N(x, \omega)\|_{H^J} \leq J \sup\nolimits_{0\leq k\leq T/h} \|z(t_k, x) - \widehat{Z}_h(t_k, x, \omega)\|_H$$

Rest of talk: bounding $\omega$-expectation of

$$\sup\nolimits_{k\in[N]_0} \|z(t_k, x) - \widehat{Z}_h(t_k, x, \omega)\|_H$$

## Randomised time integration

## for operator differential equations

# Randomised time integration

Exact flow map $\varphi$ maps $(t, u_s)$ to $(t+h, \varphi(h, t, u_s))$ using vector field $f$:

$$\varphi(h, t, u_s) = u_s + \int_t^{t+h} f(\tau, \varphi(\tau, t, u_s))\, \mathrm{d}\tau$$

Time grid on $[0, T]$ with spacing $(h_k)_k$:

$$0 =: t_0 < t_1 < \cdots < t_N := T, \quad h_k := t_{k+1} - t_k, \quad h := \max_{k \in [N-1]_0} h_k$$

Deterministic exact seq. $z(t_{k+1}) = \varphi(h_k, t_k, z(t_k)), \quad k \in [N-1]_0$

Inexact flow map $\psi$ (deterministic time integration method)

Randomised approximate solution sequence $(\widehat{Z}_h(t_k))_{k \in [N]_0}$

$$\widehat{Z}_h(t_{k+1}) := \psi(h_k, t_k, \widehat{Z}_h(t_k)) + \xi_k(h_k), \quad k \in [N-1]_0,$$

# Conrad et al. (2017)

Uniform time grid: $t_k := hk$, $k \in [N]_0$, $h = T/N$

Assumptions:

- globally Lipschitz vector field $f \Rightarrow$ globally Lipschitz true flow map $\varphi$
- decay of second moments: $\|\xi_k(h)\|_{L^2(\Omega;\mathbb{R}^d)} \lesssim h^{p+1/2}$
- there exist $q \geq 0$ and $h^* \in \mathbb{R}_{>0}$, such that if $0 < h \leq h^*$ then

$$\sup_{(t,v)\in[0,T-h]\times\mathbb{R}^d} |\varphi(h,t,v) - \psi(h,t,v)|_{\mathbb{R}^d} \lesssim h^{q+1}$$

  - uniform local truncation error of order $q+1$
  - valid if $f$ smooth enough, has bounded derivatives

Theorem (Conrad et al.): Given the assumptions, if $\widehat{Z}_0 = z(0)$, then

$$\sup_{k\in[N]_0} \left\| z(t_k) - \widehat{Z}_h(t_k) \right\|_{L^2(\Omega;\mathbb{R}^d)} \lesssim h^{\min\{p,q\}}.$$

# L., Stahn and Sullivan (2022)

Suppose unique solution $z$ of IVP belongs to $C([0,T];H)$

**Assumption.** Time integration method $\psi$ admits following:

1. Order parameter $q \geq 0$
   Truncation error function $C_{\varphi,\psi}\colon [0,T]\times H \to (0,\infty)$
   (bounded on bounded subsets)
   $\mathcal{D}\subset H$ dense, s.t. if $(t,x)$ lies on a solution $u\in C([0,T];H)$ with initial condition in $\mathcal{D}$, then

$$|\varphi(h,t,x)-\psi(h,t,x)|_H \leq C_{\varphi,\psi}(t,x)h^{q+1}$$

2. $L_\psi \in \mathbb{R}_{>0}$ such that $\forall (h,t)\in[0,h^*]\times[0,T]$, $\forall x,y\in H$,

$$|\psi(h,t,x)-\psi(h,t,y)|_H \leq (1+L_\psi h)\,|x-y|_H.$$

Interpretation:

1. - non-uniform local truncation error of $\psi$
2. - global Lipschitz continuity w.r.t. ‘spatial’ argument

# L., Stahn and Sullivan (2022)

Given Young function $\Upsilon : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$, define Orlicz norm $\| \cdot \|_{\Upsilon(\Omega;\mathbb{R})}$:

$$\|W\|_{\Upsilon(\Omega;\mathbb{R})} := \inf\{r \in \mathbb{R}_{>0} \; : \; \mathbb{E}[\Upsilon(|W|/r)] \leq 1\}.$$

$\| \cdot \|_{\Upsilon(\Omega;\mathbb{R})}$ includes as special case $\| \cdot \|_{L^R(\Omega;\mathbb{R})}$, $\forall R > 1$

If $\Upsilon(z) := \exp(z^2) - 1$, then $\|W\|_{\Upsilon(\Omega,\mathbb{R})} < \infty$ iff $W$ sub-Gaussian

# L., Stahn and Sullivan (2022)

Given Young function $\Upsilon : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$, define Orlicz norm $\|\cdot\|_{\Upsilon(\Omega;\mathbb{R})}$:

$$\|W\|_{\Upsilon(\Omega;\mathbb{R})} := \inf\{r \in \mathbb{R}_{>0} \ : \ \mathbb{E}[\Upsilon(|W|/r)] \leq 1\}.$$

$\|\cdot\|_{\Upsilon(\Omega;\mathbb{R})}$ includes as special case $\|\cdot\|_{L^R(\Omega;\mathbb{R})}$, $\forall R > 1$

If $\Upsilon(z) := \exp(z^2) - 1$, then $\|W\|_{\Upsilon(\Omega,\mathbb{R})} < \infty$ iff $W$ sub-Gaussian

**Assumption**. $H$-valued RVs $(\xi_k)_{k\in\mathbb{N}_0}$ admit $\|\cdot\|_{\Upsilon}$ and $p \geq 0$ such that

$$\|\xi_k(h)\|_{\Upsilon(\Omega;H)} \lesssim h^{p+1}, \quad \forall k \in \mathbb{N}_0, \ h > 0.$$

**Theorem**. Suppose assumptions on $\psi$ and $(\xi_k)_{k\in\mathbb{N}_0}$ hold. If initial condition $\vartheta$ of IVP belongs to dense subset $\mathcal{D} \subset H$ from Asmp. 1 and if $z(0) = \widehat{Z}_0$, then

$$\left\| \sup_{k\in[N]_0} |z(t_k) - \widehat{Z}_h(t_k)|_H \right\|_{\Upsilon(\Omega;\mathbb{R})} \lesssim h^{\min\{p,q\}}.$$

# Comparison of convergence results

Conrad et al. (2017): $\psi$ has uniform local truncation error

$$\sup_{(t,v)\in[0,T-h]\times\mathbb{R}^d} |\varphi(h,t,v)-\psi(h,t,v)|_{\mathbb{R}^d} \lesssim h^{q+1}$$

- valid under strong assumptions on vector field $f$ of ODE

L., Stahn, Sullivan (2022): $\psi$ has non-uniform local truncation error

$$|\varphi(h,t,x)-\psi(h,t,x)|_H \leq C_{\varphi,\psi}(t,x)h^{q+1}$$

- valid under same assumptions as for time integration for PDEs

# Importance of error decomposition

Common step to proofs of both theorems:

$$|z(t_{k+1}) - \widehat{Z}_h(t_{k+1})| \leq |\varphi(h_k, t_k, u(t_k)) - \psi(h_k, t_k, U_k)| + |\xi_k(h_k)|$$

# Importance of error decomposition

Common step to proofs of both theorems:

$$|z(t_{k+1}) - \widehat{Z}_h(t_{k+1})| \leq |\varphi(h_k, t_k, u(t_k)) - \psi(h_k, t_k, U_k)| + |\xi_k(h_k)|$$

Error decomposition of Conrad et al. (2017):

$$\begin{aligned} &|\varphi(h_k, t_k, u(t_k)) - \psi(h_k, t_k, U_k)| \\ &\leq |\varphi(h_k, t_k, u(t_k)) - \varphi(h_k, t_k, U_k)| + |\varphi(h_k, t_k, U_k) - \psi(h_k, t_k, U_k)| \end{aligned}$$

- global Lipschitz property of $\varphi$ controls first term
- uniform local truncation error assumption controls second term

# Importance of error decomposition

Common step to proofs of both theorems:

$$|z(t_{k+1}) - \widehat{Z}_h(t_{k+1})| \leq |\varphi(h_k, t_k, u(t_k)) - \psi(h_k, t_k, U_k)| + |\xi_k(h_k)|$$

Error decomposition of Conrad et al. (2017):

$$\begin{aligned} &|\varphi(h_k, t_k, u(t_k)) - \psi(h_k, t_k, U_k)| \\ &\leq |\varphi(h_k, t_k, u(t_k)) - \varphi(h_k, t_k, U_k)| + |\varphi(h_k, t_k, U_k) - \psi(h_k, t_k, U_k)| \end{aligned}$$

- global Lipschitz property of $\varphi$ controls first term
- uniform local truncation error assumption controls second term

Error decomposition of L., Stahn, Sullivan (2022):

$$\begin{aligned} &|\varphi(h_k, t_k, u(t_k)) - \psi(h_k, t_k, U_k)| \\ &\leq |\varphi(h_k, t_k, u(t_k)) - \psi(h_k, t_k, u(t_k))| + |\psi(h_k, t_k, u(t_k)) - \psi(h_k, t_k, U_k)| \end{aligned}$$

- non-uniform local truncation error asmp. controls first term
- global Lipschitz property of $\psi$ controls second term

# So what?

Conrad et al. (2017):

- Finite-dimensional ODEs
- Uniform local truncation error
- Randomisation decays like $h^{p+1/2}$, error $\sup_{k\in[N]_0}\|z(t_k)-\widehat{Z}_h(t_k)\|_{L^2(\Omega;\mathbb{R}^d)}$ decays like $h^{\min\{p,q\}}$

L., Stahn, Sullivan (2022):

- Operator diff. eqs. on Gelfand triples $(V,H,V')\leftrightarrow$ time-dependent PDEs
- No uniform local truncation error assumption
- Randomisation decays like $h^{p+1}$, error $\|\sup_{k\in[N]_0}\|z(t_k)-\widehat{Z}_h(t_k)\|_H\|_{\Upsilon(\Omega;\mathbb{R})}$ decays like $h^{\min\{p,q\}}$

• Randomised time integration works in more general settings and under weaker assumptions than considered in Conrad et al. (2017)

• Choice of error decomposition affects assumptions

**Summary**:

Inverse problem: Given data $y = G(x^*) + \varepsilon$, solve for $x^*$

In practice, we only have approximation $G_h$ of true $G$

Error in $G_h$ propagates to error in resulting approx. posterior $\mu^y_{\text{pos},h}$

Modelling error $G_h - G$ by RVs results in random approx. $\widehat{\mu}^y_{\text{pos},h}$

Error bounds w.r.t. Hellinger / Kullback–Leibler for $\widehat{\mu}^y_{\text{pos},h}$

**Open question**: Good choices of randomisation $(\xi_k)_k$

## References

A. Abdulle, G. Garegnani, “Random time step probabilistic methods for uncertainty quantification in chaotic and geometric numerical integration”, Stat. Comput. (2020)

P. Conrad, M. Girolami, S. Särkka, A. Stuart, and K. Zygalakis, “Statistical analysis of differential equations: introducing probability measures on numerical solutions”, Stat. Comput. (2017)

H. C. Lie, M. Stahn, and T. J. Sullivan, “Randomised one-step time integration methods for deterministic operator differential equations”, *Calcolo* (2022)

H. C. Lie, T. J. Sullivan, and A. Teckentrup, “Random forward models and log-likelihoods in Bayesian inverse problems”, *SIAM/ASA J. Uncertain. Quantif.* (2019)

B. Sprungk, “On the local Lipschitz stability of Bayesian inverse problems”, *Inverse Probl.* (2020)

A. M. Stuart, “Inverse problems: a Bayesian perspective”, *Acta Numerica* (2010)

O. Teymur, H. C. Lie, T. J. Sullivan, and B. Calderhead, “Probabilistic Linear Multistep Methods”, *NeurIPS* (2018)