# Efficient Sampling Algorithms for Non-Gaussian Data Assimilation

Ahmed Attia[1], Vishwas Rao[1], and Adrian Sandu[1]

[1] Computational Science Laboratory (CSL)
Department of Computer Science
Virginia Tech
{attia;sandu}@cs.vt.edu

Workshop on Meteorological Sensitivity Analysis and Data Assimilation
1-5 June 2015 Roanoke, West Virginia

# Outline

- Motivation
- Sampling approach
- Sampling filter+smoother
- Experiments and results
- Conclusion

# Motivation:

- **Data Assimilation**:

$$\underbrace{\textbf{Model + Prior + Observations}}_{\text{with associated uncertainties}} \rightarrow \underbrace{\textbf{Best Description of the posterior}}_{\textit{Ensemble}+\textit{Variational}}$$

# Motivation:

- **Data Assimilation**:

**Model + Prior + Observations** (with associated uncertainties) $\rightarrow$ **Best Description of the posterior** (*Ensemble+Variational*)

- **Goal**: Ensemble representation of the posterior PDF for the general non-Gaussian and non-linear cases.

**Sample the posterior PDF**

- **MCMC** (Gold Standard) : popular and guaranteed to converge BUT:

Transition Kernel, R-W behaviour , Convergence Rate, Acceptance Rate , Poor Mixing , ...

- Accelerated MCMC: **Hybrid Monte Carlo (HMC)**

Duane et. al. (1987); Neal (1993); Bennett, and Chua (1994)

**Recursively use HMC for Filtering and Smoothing**

# Hybrid MC:

+ The Hamiltonian:

$$H(\mathbf{p},\mathbf{x}) = \frac{1}{2}\mathbf{p}^T\mathbf{M}^{-1}\mathbf{p} - \log(\pi(\mathbf{x})) = \underbrace{\frac{1}{2}\mathbf{p}^T\mathbf{M}^{-1}\mathbf{p}}_{\text{kinetic energy}} + \underbrace{\mathcal{J}(\mathbf{x})}_{\text{potential energy}} \tag{1}$$

+ The Hamiltonian dynamics:

$$\frac{d\mathbf{x}}{dt} = \nabla_{\mathbf{p}} H = \mathbf{M}^{-1}\mathbf{p}, \qquad \frac{d\mathbf{p}}{dt} = -\nabla_{\mathbf{x}} H = -\nabla_{\mathbf{x}}\mathcal{J}(\mathbf{x}) \tag{2}$$

+ Symplectic integrator(e.g. Verlet, two-stage, three-stage,...) is used:

$$\begin{aligned} \mathbf{x}_{k+1/2} &= \mathbf{x}_k + \frac{h}{2}\mathbf{M}^{-1}\mathbf{p}_k, \qquad \mathbf{p}_{k+1} = \mathbf{p}_k - h\nabla_{\mathbf{x}}\mathcal{J}(\mathbf{x}_{k+1/2}), \\ \mathbf{x}_{k+1} &= \mathbf{x}_{k+1/2} + \frac{h}{2}\mathbf{M}^{-1}\mathbf{p}_{k+1}. \end{aligned} \tag{3}$$

+ The canonical PDF of $(\mathbf{p},\mathbf{x})$:

$$\propto \exp\left(-H(\mathbf{p},\mathbf{x})\right) = \exp\left(-\frac{1}{2}\mathbf{p}^T\mathbf{M}^{-1}\mathbf{p} - \mathcal{J}(\mathbf{x})\right) = \exp\left(-\frac{1}{2}\mathbf{p}^T\mathbf{M}^{-1}\mathbf{p}\right)\pi(\mathbf{x}) \tag{4}$$

# HMC Sampling Algorithm to sample from $\propto \pi(\mathbf{x})$:

- View state vector ($\mathbf{x}$ ) as a position variable in an extended phase space,
- Add "*momentum*" $\mathbf{p} \sim \mathcal{N}(0, \mathbf{M})$ and sample from the joint (canonical) PDF.
  $\rightarrow$ Generate a **MC** with invariant distribution $\propto \exp\left(-H(\mathbf{p}, \mathbf{x})\right)$.

---

- Initialize the MC $\leftarrow$ $\mathbf{x}_0$
- For $k = 0, 1, \ldots$

  1- Draw $\mathbf{p}_k \sim \mathcal{N}(0, \mathbf{M})$

  2- Use (Verlet, two-stage, three-stage,...) to propose a new state:

$$\underbrace{(\mathbf{p}^*, \mathbf{x}^*) = \Phi_T(\mathbf{p}_k, \mathbf{x}_k)}_{\text{Acts as a TRANSITION KERNEL of the MC}} \quad ; \quad T = m\,h\,.$$

  3- Acceptance Probability: $a^{(k)} = 1 \wedge e^{-\Delta H}, \quad \Delta H = H(\mathbf{p}^*, \mathbf{x}^*) - H(\mathbf{p}_k, \mathbf{x}_k)$

  4- Discard both $\mathbf{p}^*$, $\mathbf{p}_k$

  5-

$$\mathbf{x}_{k+1} = \begin{cases} \mathbf{x}^* & \text{with probability} \quad a^{(k)} \\ \mathbf{x}_k & \text{with probability } 1 - a^{(k)} \end{cases}$$

- Collect samples after CONVERGENCE

# Sequential DA:

- Baye's Theorem:

$$\mathcal{P}^{\mathrm{a}}(\mathbf{x}) = \mathcal{P}(\mathbf{x}|\mathbf{y}) = \frac{\mathcal{P}(\mathbf{y}|\mathbf{x})\mathcal{P}^{\mathrm{b}}(\mathbf{x})}{\mathcal{P}(\mathbf{y})}, \quad (5a)$$

$$\propto \mathcal{P}(\mathbf{y}|\mathbf{x})\mathcal{P}^{\mathrm{b}}(\mathbf{x}) = \pi(\mathbf{x}) \quad (5b)$$

- Gaussian framework:

$$\mathcal{P}^{\mathrm{b}}(\mathbf{x}) \propto \exp\left(-\frac{1}{2}(\mathbf{x}-\mathbf{x}^{\mathrm{b}})^T \mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}^{\mathrm{b}})\right), \quad (6)$$

$$\mathcal{P}(\mathbf{y}|\mathbf{x}) \propto \exp\left(-\frac{1}{2}(\mathcal{H}(\mathbf{x})-\mathbf{y})^T \mathbf{R}^{-1}(\mathcal{H}(\mathbf{x})-\mathbf{y})\right). \quad (7)$$

- Posterior PDF:

$$\mathcal{P}^{\mathrm{a}}(\mathbf{x}) \propto \overbrace{\exp\left(-\mathcal{J}(\mathbf{x})\right)}^{\pi(\mathbf{x})}, \quad (8)$$

$$\mathcal{J}(\mathbf{x}) = \frac{1}{2}\|\mathbf{x}-\mathbf{x}^{\mathrm{b}}\|_{\mathbf{B}^{-1}} + \frac{1}{2}\|\mathcal{H}(\mathbf{x})-\mathbf{y}\|_{\mathbf{R}^{-1}}, \quad (9)$$

where

$$\nabla_{\mathbf{x}}\mathcal{J}(\mathbf{x}) = \mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}^{\mathrm{b}}) + \mathbf{H}^T\mathbf{R}^{-1}(\mathcal{H}(\mathbf{x})-\mathbf{y}). \quad (10)$$

# HMC Sampling Filter:

Given observation $\mathbf{y}_k$ at time point $t_k$,

**Assimilation cycle**

- **Forecast:** given an analysis ensemble $\{\mathbf{x}^{\mathrm{a}}_{k-1}(e)\}_{e=1,2,\ldots,\mathrm{N_{ens}}}$ at time $t_{k-1}$:

$$\mathbf{x}^{\mathrm{b}}_k(e) = \mathcal{M}_{t_{k-1}\rightarrow t_k}\left(\mathbf{x}^{\mathrm{a}}_{k-1}(e)\right), \quad e = 1, 2, \ldots, \mathrm{N_{ens}} \tag{11}$$

- **Analysis:** use HMC to sample from $\propto \pi(\mathbf{x}) = \exp\left(-\mathcal{J}(\mathbf{x})\right)$:

  1- Set the elements of Hamiltonian and the integrator:

  - **M** (e.g. constant diagonal, diagonal of $\mathbf{B}_k^{-1}$)
  - (*Pseudo*) time step settings ($T = hm$) of the symplectic integrator

  2- Initialize the chain ($\mathbf{x}_0$): (e.g. Forecast, EnKF, 3D-Var ,... )

  4- Generate the MC and **sample** $\{\mathbf{x}^{\mathrm{a}}_k(e)\}_{e=1,2,\ldots,\mathrm{N_{ens}}}$ after convergence.

# Results using Lorenz-96

- **Lorenz-96 with 40-variables:**

$$\frac{dx_i}{dt} = x_{i-1}\left(x_{i+1} - x_{i-2}\right) - x_i + F\,, \tag{12}$$

where $\mathbf{x} = (x_1, x_2, \ldots, x_{40})^T \in \mathbb{R}^{40}$; $x_0 \equiv x_{40}$.

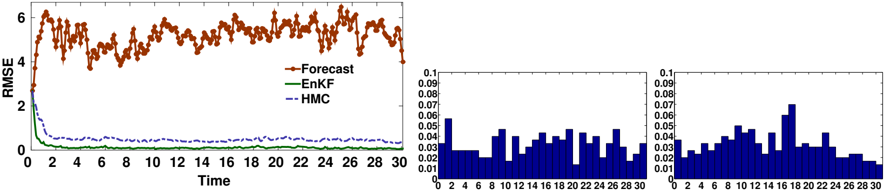
The forcing parameter $F = 8$.

Observation uncertainty: 5%

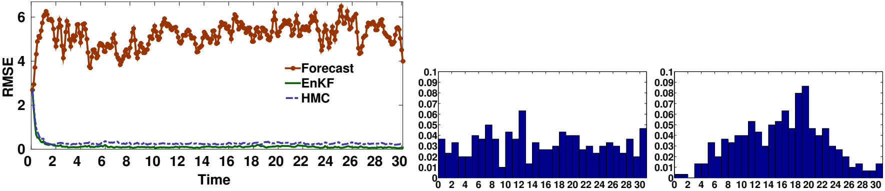
Background uncertainty: 8%

Every third component is observed

# Results using Lorenz-96: linear $\mathcal{H}$



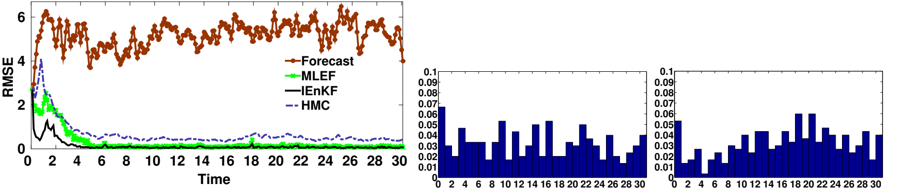

(a) Verlet integrator (b) $\mathbf{x}_1$ (c) $\mathbf{x}_2$




(d) Two-stage integrator (e) $\mathbf{x}_1$ (f) $\mathbf{x}_2$

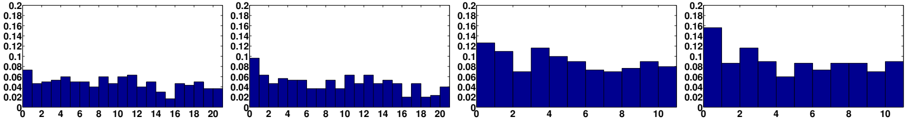
**Figure :** $h = 0.01$, $m = 10$. 50 burn-in steps, and 10 mixing steps. $\mathrm{N_{ens}} = 30$.

# Results using Lorenz-96: discontinuous quadratic $\mathcal{H}$




(a) Three-stage integrator (b) $\mathbf{x}_1$ (c) $\mathbf{x}_2$

**Figure :** $h = 0.01$, $m = 10$. 50 burn-in steps, and 10 mixing steps. $\mathrm{N_{ens}} = 30$.



(a) Three-stage integrator; $\mathrm{N_{ens}} = 20$ (b) Three-stage integrator; $\mathrm{N_{ens}} = 20$ (c) Three-stage integrator; $\mathrm{N_{ens}} = 10$ (d) Three-stage integrator; $\mathrm{N_{ens}} = 10$

**Figure :** $h = 0.01$, $m = 10$. 50 burn-in steps, and 10 mixing steps.

# Results using SWE

- Shallow water equations on a sphere:

$$\frac{\partial u}{\partial t} + \frac{1}{a\cos\theta}\left(u\frac{\partial u}{\partial \lambda} + v\cos\theta\frac{\partial u}{\partial \theta}\right) - \left(f + \frac{u\tan\theta}{a}\right)v + \frac{g}{a\cos\theta}\frac{\partial h}{\partial \lambda} = 0, \tag{13a}$$

$$\frac{\partial v}{\partial t} + \frac{1}{a\cos\theta}\left(u\frac{\partial v}{\partial \lambda} + v\cos\theta\frac{\partial v}{\partial \theta}\right) + \left(f + \frac{u\tan\theta}{a}\right)u + \frac{g}{a}\frac{\partial h}{\partial \theta} = 0, \tag{13b}$$

$$\frac{\partial h}{\partial t} + \frac{1}{a\cos\theta}\left(\frac{\partial\,(hu)}{\partial \lambda} + \frac{\partial(hv\cos\theta)}{\partial \theta}\right) = 0\,. \tag{13c}$$

State vector: $\mathbf{x} = [u, v, h]^T \in \mathbb{R}^{7776}$

$h$ is the height; $u,\ v$ are zonal and meridional wind.

All components are observed

# Results using SWE: linear $\mathcal{H}$ I

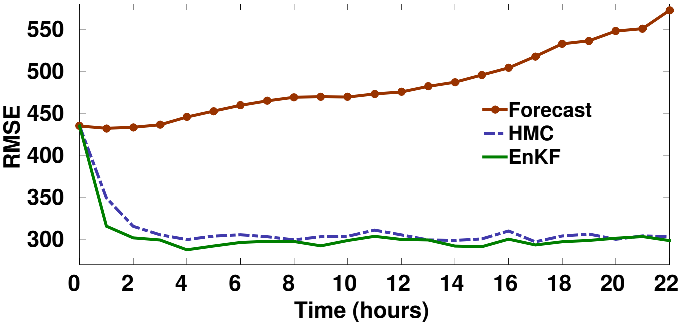
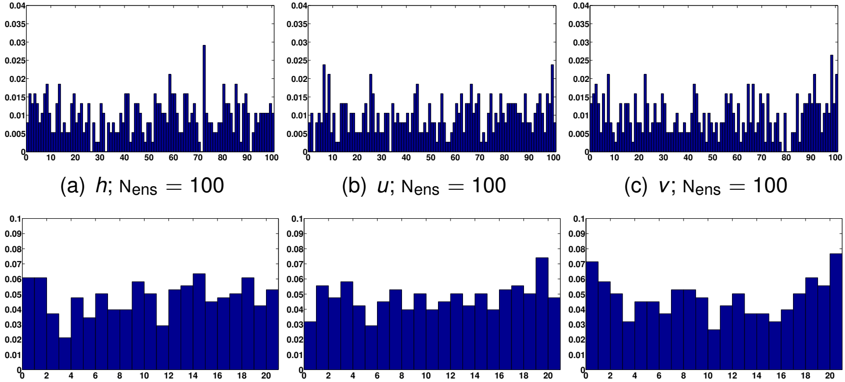



**Figure :** Two-stage, $h = 0.01$, $m = 10$. 50 burn-in steps, 10 mixing steps. $\mathrm{N_{ens}} = 100$

# Results using SWE: linear $\mathcal{H}$ II



(a) $h$; $\mathrm{N_{ens}} = 100$ (b) $u$; $\mathrm{N_{ens}} = 100$ (c) $v$; $\mathrm{N_{ens}} = 100$

(d) $h$; $\mathrm{N_{ens}} = 20$ (e) $u$; $\mathrm{N_{ens}} = 20$ (f) $v$; $\mathrm{N_{ens}} = 20$

**Figure :** Two-stage, $h = 0.01,\ m = 10$. 50 burn-in steps, 10 mixing steps.

# HMC Filter for replenishing purposes:

- Parallel implementations of operational filters (e.g. EnKF, MLEF, IEnKF,...) face complications due to the death of one or more of the nodes.

- HMC filter can be used to replenish the ensemble with new independent members.

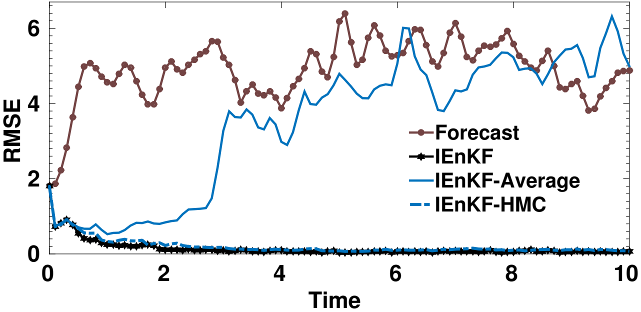



**Figure :** Lorenz-96. Quadratic discontinuous $\mathcal{H}$. Two-stage with $h = 0.01$, and $m = 10$. 4 mixing steps.

# Four-dimensional DA:

- From Baye's Theorem: $\mathcal{P}^{\mathrm{a}}(\mathbf{x}_0) = \mathcal{P}(\mathbf{x}_0|\mathbf{y}_{0:m}) \propto \mathcal{P}(\mathbf{y}_{0:m}|\mathbf{x}_0)\mathcal{P}^{\mathrm{b}}(\mathbf{x}_0)$
- Gaussian framework:

$$\mathcal{P}^{\mathrm{b}}(\mathbf{x}_0) \propto \exp\left(-\frac{1}{2}(\mathbf{x}_0 - \mathbf{x}_0^{\mathrm{b}})^T \mathbf{B}_0^{-1}(\mathbf{x}_0 - \mathbf{x}_0^{\mathrm{b}})\right), \quad (14)$$

$$\mathcal{P}(\mathbf{y}_{0:m}|\mathbf{x}_0) \propto \exp\left(-\frac{1}{2}\sum_{k=0}^{m}\left((\mathcal{H}_k(\mathbf{x}_k) - \mathbf{y}_k)^T \mathbf{R}_k^{-1}(\mathcal{H}_k(\mathbf{x}_k) - \mathbf{y}_k)\right)\right), \quad (15)$$

$$\mathbf{x}_k = \mathcal{M}_{t_0 \rightarrow t_k}(\mathbf{x}_0)$$

- Posterior PDF:

$$\mathcal{P}^{\mathrm{a}}(\mathbf{x}_0) \propto \overbrace{\exp\left(-\mathcal{J}(\mathbf{x}_0)\right)}^{\pi(\mathbf{x}_0)}, \quad (16a)$$

$$\mathcal{J}(\mathbf{x}_0) = \frac{1}{2}\|\mathbf{x}_0 - \mathbf{x}_0^{\mathrm{b}}\|_{\mathbf{B}_0^{-1}} + \frac{1}{2}\sum_{k=0}^{m}\|\mathcal{H}_k(\mathbf{x}_k) - \mathbf{y}_k\|_{\mathbf{R}_k^{-1}}, \quad (16b)$$

$$\nabla_{\mathbf{x}_0}\mathcal{J}(\mathbf{x}_0) = \mathbf{B}^{-1}(\mathbf{x}_0 - \mathbf{x}_0^{\mathrm{b}}) + \sum_{k=0}^{m}\mathbf{M}_{0,k}^T\mathbf{H}_k^T\mathbf{R}_k^{-1}(\mathcal{H}_k(\mathbf{x}_k) - \mathbf{y}_k), \quad (17)$$

# HMC Sampling Smoother:

Given Forecast ensemble $\mathbf{x}_0(e)_{e=1,2,\ldots,\mathrm{N_{ens}}}$, and observations $\mathbf{y}_0,\ \mathbf{y}_1,\ \ldots\mathbf{y}_m$:

## Assimilation cycle

- **Analysis:** use HMC to sample from $\propto \pi(\mathbf{x}_0) = \exp\left(-\mathcal{J}(\mathbf{x}_0)\right)$:

  1- Set the elements of Hamiltonian and the integrator:
  - **M** (e.g. constant diagonal, diagonal of $\mathbf{B}_0^{-1}$)
  - Time (*artificial*) step settings of the symplectic integrator

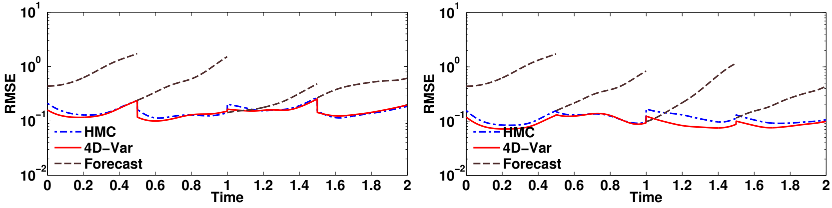
  2- Initialize the chain ($\mathbf{x}_0$): (e.g. Forecast, EnKS, 4D-Var ,... )

  4- Generate the MC and **sample** $\{\mathbf{x}_0^{\mathrm{a}}(e)\}_{e=1,2,\ldots,\mathrm{N_{ens}}}$ after convergence.

- **Forecast:** given an analysis ensemble $\{\mathbf{x}_0^{\mathrm{a}}(e)\}_{e=1,2,\ldots,\mathrm{N_{ens}}}$ at time $t_{k-1}$:

$$\mathbf{x}_k^{\mathrm{b}}(e) = \mathcal{M}_{t_{k-1}\rightarrow t_k}\left(\mathbf{x}_{k-1}^{\mathrm{a}}(e)\right), \quad e = 1, 2, \ldots, \mathrm{N_{ens}};\ k = 1, 2, \ldots, m \tag{18}$$

# Results using Lorenz-96:




(a) Linear observation operator

(b) Quadratic observation operator

**Figure :** Quadratic observation operator; every second component only observed. Two-stage integrator , $h = 0.01$, $m = 10$. 30 burn-in steps, 10 mixing steps.

# 4D DA results (SWE): linear $\mathcal{H}$

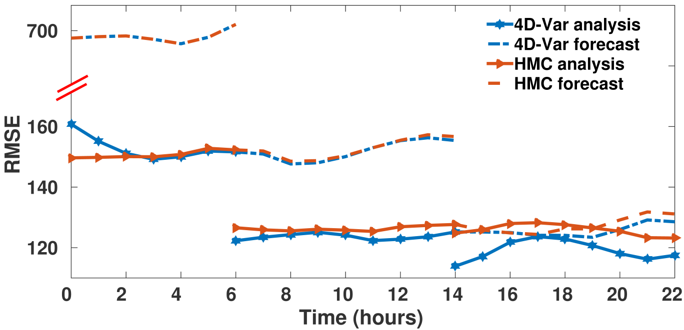



**Figure :** Linear observation operator. 8 observations on each assimilation window. Two-stage with $h = 0.01$, and $m = 10$. 30 burn-in steps, 4 mixing steps.

# Conclusion

- Non-Gaussian sampling filter and smoother, based on HMC, were presented.
- Better description of the prior PDF!
- Avoid using the adjoint of the full model,
- Filter and smoother enhancement and parallelization,
- Test with larger models in operational settings (SPEEDY, WRF,...)

**Thank You**

# Symplectic numerical integrators

One step advance of the solution of the Hamiltonian equations from time $t_k$ to time $t_{k+1} = t_k + h$ as follows:

1. **Position Verlet integrator**

$$\mathbf{x}_{k+1/2} = \mathbf{x}_k + \frac{h}{2}\mathbf{M}^{-1}\mathbf{p}_k\,, \tag{19a}$$
$$\mathbf{p}_{k+1} = \mathbf{p}_k - h\nabla_{\mathbf{x}}\mathcal{J}(\mathbf{x}_{k+1/2})\,, \tag{19b}$$
$$\mathbf{x}_{k+1} = \mathbf{x}_{k+1/2} + \frac{h}{2}\mathbf{M}^{-1}\mathbf{p}_{k+1}\,. \tag{19c}$$

2. **Two-stage integrator**

$$\mathbf{x}_1 = \mathbf{x}_k + (a_1 h)\mathbf{M}^{-1}\mathbf{p}_k\,, \tag{20a}$$
$$\mathbf{p}_1 = \mathbf{p}_k - (b_1 h)\nabla_{\mathbf{x}}\mathcal{J}(\mathbf{x}_1)\,, \tag{20b}$$
$$\mathbf{x}_2 = \mathbf{x}_1 + (a_2 h)\mathbf{M}^{-1}\mathbf{p}_1\,, \tag{20c}$$
$$\mathbf{p}_{k+1} = \mathbf{p}_1 - (b_1 h)\nabla_{\mathbf{x}}\mathcal{J}(\mathbf{x}_2)\,, \tag{20d}$$
$$\mathbf{x}_{k+1} = \mathbf{x}_2 + (a_2 h)\mathbf{M}^{-1}\mathbf{p}_{k+1}\,, \tag{20e}$$

$$a_1 = 0.21132\,,\ \ a_2 = 1 - 2a_1\,,\ \ \mathbf{b}_1 = 0.5\,.$$