KU LEUVEN

Reversible random number generators for adjoint Monte Carlo

E. Løvbak

F. Blondeel, A. Lee, L. Vanroye, A. Van Barel, S. Vandewalle, G. Samaey

KU Leuven, Department of Computer Science, NUMA Section

Research Foundation Flanders
Opening new horizons

NUMA

Adjoint-based optimization

- PDE-constrained optimization

$$\min_{\hat{\Omega}} \hat{\mathcal{J}}(\hat{q}, \hat{\Omega}), \quad \text{subject to} \quad \hat{\mathcal{B}}(\hat{q}; \hat{\Omega}) = 0$$

- Naive application of chain rule using $\hat{q}' = \hat{q}(\hat{\Omega})$

$$\frac{\mathsf{d}\hat{\mathcal{J}}}{\mathsf{d}\hat{\Omega}}(\hat{q}', \hat{\Omega}) = \frac{\partial \hat{\mathcal{J}}}{\partial \hat{\Omega}}(\hat{q}', \hat{\Omega}) + \frac{\partial \hat{\mathcal{J}}}{\partial \hat{q}}(\hat{q}', \hat{\Omega}) \frac{\mathsf{d}\hat{q}}{\mathsf{d}\hat{\Omega}}(\hat{\Omega})$$

Adjoint-based optimization

- PDE-constrained optimization

$$\min_{\hat{\Omega}} \hat{\mathcal{J}}(\hat{q}, \hat{\Omega}), \quad \text{subject to} \quad \hat{\mathcal{B}}(\hat{q}; \hat{\Omega}) = 0$$

- Naive application of chain rule using $\hat{q}' = \hat{q}(\hat{\Omega})$

$$\frac{\mathrm{d}\hat{\mathcal{J}}}{\mathrm{d}\hat{\Omega}}(\hat{q}', \hat{\Omega}) = \frac{\partial \hat{\mathcal{J}}}{\partial \hat{\Omega}}(\hat{q}', \hat{\Omega}) + \frac{\partial \hat{\mathcal{J}}}{\partial \hat{q}}(\hat{q}', \hat{\Omega}) \frac{\mathrm{d}\hat{q}}{\mathrm{d}\hat{\Omega}}(\hat{\Omega})$$

Adjoint-based optimization

- PDE-constrained optimization

$$\min_{\hat{\Omega}} \hat{\mathcal{J}}(\hat{q}, \hat{\Omega}), \quad \text{subject to} \quad \hat{\mathcal{B}}(\hat{q}; \hat{\Omega}) = 0$$

- Naive application of chain rule using $\hat{q}' = \hat{q}(\hat{\Omega})$

$$\frac{\mathrm{d}\hat{\mathcal{J}}}{\mathrm{d}\hat{\Omega}}(\hat{q}', \hat{\Omega}) = \frac{\partial \hat{\mathcal{J}}}{\partial \hat{\Omega}}(\hat{q}', \hat{\Omega}) + \frac{\partial \hat{\mathcal{J}}}{\partial \hat{q}}(\hat{q}', \hat{\Omega}) \frac{\mathrm{d}\hat{q}}{\mathrm{d}\hat{\Omega}}(\hat{\Omega})$$

- Solution: Lagrangian $\mathcal{L}(\hat{q}, \hat{q}^*, \hat{\Omega}) = \hat{\mathcal{J}}(\hat{q}, \hat{\Omega}) + \hat{q}^{*\top} \hat{\mathcal{B}}(\hat{q}, \hat{\Omega})$

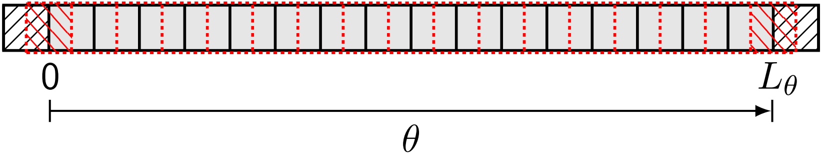
$$\hat{\mathcal{B}}(\hat{q}, \hat{\Omega}) \approx \hat{\mathcal{B}}(\hat{q}', \hat{\Omega}) + \frac{\partial \hat{\mathcal{B}}}{\partial \hat{q}}(\hat{q}', \hat{\Omega})(\hat{q} - \hat{q}') = 0$$ State equation

$$\frac{\partial \hat{\mathcal{J}}}{\partial \hat{q}}^\top (\hat{q}, \hat{\Omega}) + \frac{\partial \hat{\mathcal{B}}}{\partial \hat{q}}^\top (\hat{q}, \hat{\Omega}) \hat{q}^* = 0$$ Adjoint equation

$$\frac{\partial \hat{\mathcal{J}}}{\partial \hat{\Omega}}^\top (\hat{q}, \hat{\Omega}) + \frac{\partial \hat{\mathcal{B}}}{\partial \hat{\Omega}}^\top (\hat{q}, \hat{\Omega}) \hat{q}^* = 0$$ Design equation

Research Foundation Flanders
Opening new horizons
NUMA

Domain length optimization[1]



$$\mathcal{J}(q, L_\theta) = \frac{1}{2}\left(n_{\mathbf{i}} b_\theta u_{\parallel} - \Gamma_d\right)^2\Big|_{L_\theta} + \frac{\kappa}{2}\left(L_\theta - L_0\right)^2$$

$$q = \left(n_{\mathbf{i}}, u_{\parallel}, f_{\mathbf{n}}\right)^\top$$

$n_{\mathbf{i}}$	plasma density
$u_{\parallel}$	plasma velocity
$f_{\mathbf{n}}$	neutral position-velocity

1: W. Dekeyser, *Optimal Plasma Edge Configurations for Next-Step Fusion Reactors.* PhD thesis (2014)

Plasma edge model

- Plasma $\Rightarrow$ Finite volume

$$\frac{\partial}{\partial\theta}\left(n_{\mathbf{i}} b_\theta u_{\|}\right) = S_{n_{\mathbf{i}}} - K_{\mathsf{d}} n_{\mathbf{i}}$$

$$\frac{\partial}{\partial\theta}\left(m n_{\mathbf{i}} b_\theta {u_{\|}}^2 - \nu_{\mathsf{d}} \frac{\partial u_{\|}}{\partial\theta}\right) = S_{u_{\|}} - b_\theta \frac{\partial p}{\partial\theta}$$

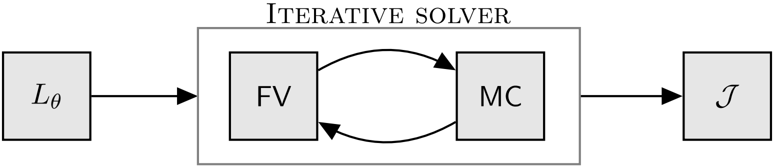
$$S_\psi = \int f_{\mathbf{n}}(\theta, v)\Psi(\theta, v)\mathrm{d}v, \quad \psi \in \{n_{\mathbf{i}}/u_{\|}\} \quad \Rightarrow \quad \text{Low-dimensional}$$

- Neutrals $\Rightarrow$ Monte Carlo

$$v\frac{\partial}{\partial\theta} f_{\mathbf{n}}(\theta, v) + K_{\mathsf{i}} f_{\mathbf{n}}(\theta, v) = S_{f_{\mathbf{n}}}\left(n_{\mathbf{i}}, u_{\|}\right) + K_{\mathsf{cx}}\left(n_{\mathbf{n}}\mathcal{V}(v; u_{\|}) - f_{\mathbf{n}}(\theta, v)\right)$$

$$n_{\mathbf{n}} = \int f_{\mathbf{n}}(\theta, v)\mathrm{d}v$$

Coupled discretization

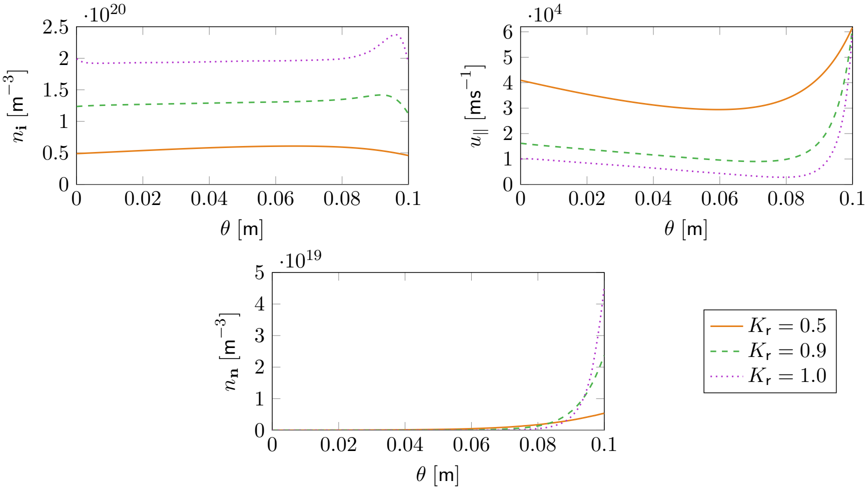


$$\hat{q} = \begin{bmatrix} \hat{n}_{\mathbf{i}} \\ \hat{u}_{\|} \\ \hat{S}_{n_{\mathbf{i}}} \\ \hat{S}_{u_{\|}} \\ Q \end{bmatrix}$$

For all particles p, time steps n:

$$Q_{p,n} = \begin{bmatrix} X_{p,n} \\ V_{p,n} \\ W_{p,n} \end{bmatrix}$$

Reference simulation results



Monte Carlo as a matrix

- For each particle p and time step n

$$B_{p,n+1}(Q_{p,n+1}, Q_{p,n}; \hat{q}, \hat{\Omega}) = \begin{bmatrix} X_{p,n+1} - X_{p,n} - \Delta t V_{p,n} \\ V_{p,n+1} - \hat{\mathcal{Q}}(V_{p,n}, \hat{q}) \\ W_{p,n+1} - \exp(-\Delta t K_{\mathsf{i}} Z_{\mathsf{i}} \hat{n}_{\mathbf{i}}) W_{p,n} \end{bmatrix} = 0$$

- Particle Jacobian

$$\frac{\partial \hat{\mathcal{B}}_p}{\partial \hat{q}_p}(Q_p, \hat{q}, \hat{\Omega}) = \begin{bmatrix} \frac{\partial B_{p,1}}{\partial Q_{p,1}} & & & \\ \frac{\partial B_{p,2}}{\partial Q_{p,1}} & \frac{\partial B_{p,2}}{\partial Q_{p,2}} & & \\ & \ddots & \ddots & \\ & & \frac{\partial B_{p,T}}{\partial Q_{p,T-1}} & \frac{\partial B_{p,T}}{\partial Q_{p,T}} \end{bmatrix}$$

Research Foundation Flanders
Opening new horizons
NUMA

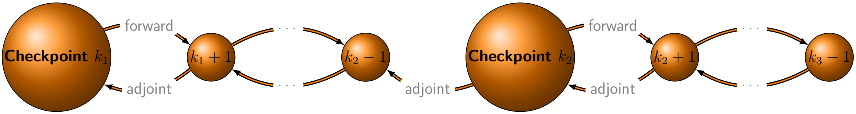
Monte Carlo adjoint

$$\frac{\partial \hat{\mathcal{B}}}{\partial \hat{\mathbf{q}}_{\mathrm{MC}}} = \begin{bmatrix} \frac{\partial \hat{\mathcal{B}}_{n_\mathbf{i}}}{\partial \hat{n}_\mathbf{i}} & \frac{\partial \hat{\mathcal{B}}_{n_\mathbf{i}}}{\partial \hat{u}_\parallel} & \frac{\partial \hat{\mathcal{B}}_{n_\mathbf{i}}}{\partial \hat{S}_{n_\mathbf{i}}} & \frac{\partial \hat{\mathcal{B}}_{n_\mathbf{i}}}{\partial \hat{S}_{u_\parallel}} & \frac{\partial \hat{\mathcal{B}}_{n_\mathbf{i}}}{\partial Q} \\ \frac{\partial \hat{\mathcal{B}}_{u_\parallel}}{\partial \hat{n}_\mathbf{i}} & \frac{\partial \hat{\mathcal{B}}_{u_\parallel}}{\partial \hat{u}_\parallel} & \frac{\partial \hat{\mathcal{B}}_{u_\parallel}}{\partial \hat{S}_{n_\mathbf{i}}} & \frac{\partial \hat{\mathcal{B}}_{u_\parallel}}{\partial \hat{S}_{u_\parallel}} & \frac{\partial \hat{\mathcal{B}}_{u_\parallel}}{\partial Q} \\ \frac{\partial \hat{\mathcal{B}}_{S_{n_\mathbf{i}}}}{\partial \hat{n}_\mathbf{i}} & \frac{\partial \hat{\mathcal{B}}_{S_{n_\mathbf{i}}}}{\partial \hat{u}_\parallel} & \frac{\partial \hat{\mathcal{B}}_{S_{n_\mathbf{i}}}}{\partial \hat{S}_{n_\mathbf{i}}} & \frac{\partial \hat{\mathcal{B}}_{S_{n_\mathbf{i}}}}{\partial \hat{S}_{u_\parallel}} & \frac{\partial \hat{\mathcal{B}}_{S_{n_\mathbf{i}}}}{\partial Q} \\ \frac{\partial \hat{\mathcal{B}}_{\hat{S}_{u_\parallel}}}{\partial \hat{n}_\mathbf{i}} & \frac{\partial \hat{\mathcal{B}}_{S_{u_\parallel}}}{\partial \hat{u}_\parallel} & \frac{\partial \hat{\mathcal{B}}_{S_{u_\parallel}}}{\partial \hat{S}_{n_\mathbf{i}}} & \frac{\partial \hat{\mathcal{B}}_{S_{u_\parallel}}}{\partial \hat{S}_{u_\parallel}} & \frac{\partial \hat{\mathcal{B}}_{S_{u_\parallel}}}{\partial Q} \\ \frac{\partial \hat{\mathcal{B}}_{Q}}{\partial \hat{n}_\mathbf{i}} & \frac{\partial \hat{\mathcal{B}}_{Q}}{\partial \hat{u}_\parallel} & \frac{\partial \hat{\mathcal{B}}_{Q}}{\partial \hat{S}_{n_\mathbf{i}}} & \frac{\partial \hat{\mathcal{B}}_{Q}}{\partial \hat{S}_{u_\parallel}} & \frac{\partial \hat{\mathcal{B}}_{Q}}{\partial Q} \end{bmatrix} \qquad \hat{q}^* = \begin{bmatrix} \hat{n}_\mathbf{i}^* \\ \hat{u}_\parallel^* \\ \hat{S}_{n_\mathbf{i}}^* \\ \hat{S}_{u_\parallel}^* \\ Q^* \end{bmatrix}$$

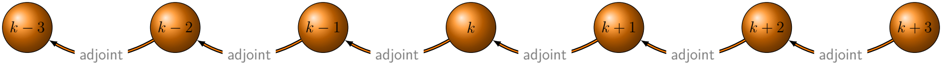
$$W_{p,n}^* = \exp\left(-\Delta t R_i Z_i \hat{n}_\mathbf{i}\left(X_{p,n}\right)\right) W_{p,n+1}^* + \frac{\partial \hat{\mathcal{B}}_{S_{n_\mathbf{i}}}}{\partial W_{p,n}}^\top \hat{n}_\mathbf{i}^* + \frac{\partial \hat{\mathcal{B}}_{S_{u_\parallel}}}{\partial W_{p,n}}^\top \hat{u}_\parallel^*.$$

Matching forward and adjoint simulations

- Same paths in forward/backward simulation
- Challenge: $P \times T$ large
- Solutions:
  - Checkpointing: 2 forward simulations + backward simulation



  - Generating the paths in reverse (this work)



Reversing a random number generator

- PCG: permuted congruential generator[2]
  - internal state ζ_z and constant vectors a, c, Z

$$\zeta_{z+1} = a\zeta_z + c \mod Z,$$

  - 1-way (permutation) function generates output from ζ_z
  - Passes `TestU01` with flying colors
- Reversing modular operations $\rightarrow$ reversed uniform sequence

$$\zeta_z = a^{-1}(\zeta_{z+1} - c) \mod Z,$$

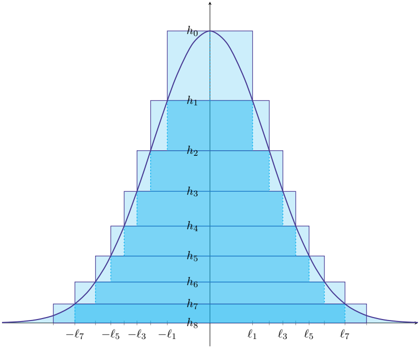
$$a^{-1} \equiv a^{m-2} \mod Z$$

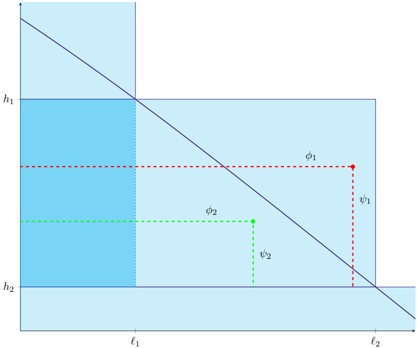
- Exponential distribution through inverse transform:

$$u \sim \mathcal{U}([0,1]) \Rightarrow -\lambda \ln(1-u) \sim \mathcal{E}(\lambda)$$

2: M.E. O'Neill, *PCG: A Family of Simple Fast Space-Efficient Statistically Good Algorithms for Random Number Generation.* Technical report HMC-CS-2014-0905, Harvey Mudd College (2014)

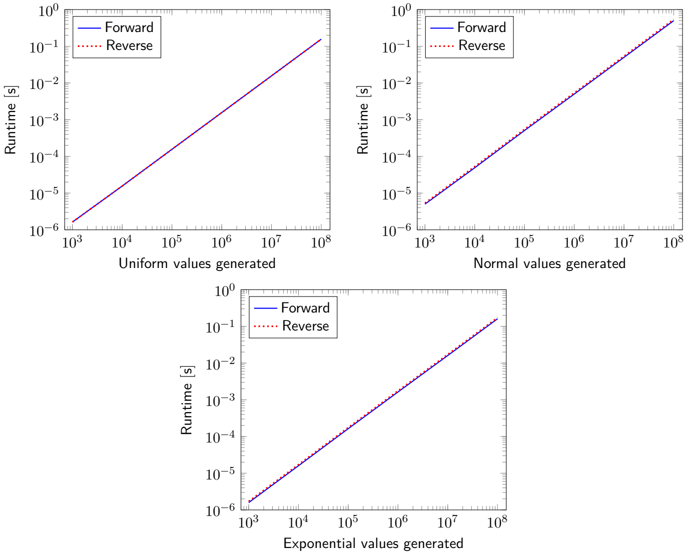
Normal distribution through Ziggurat



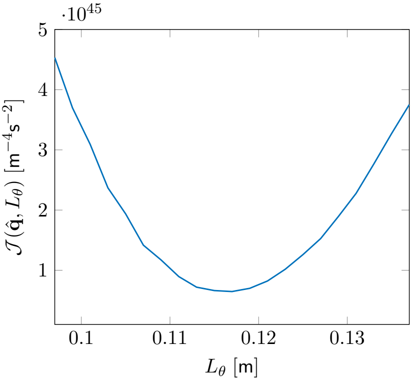


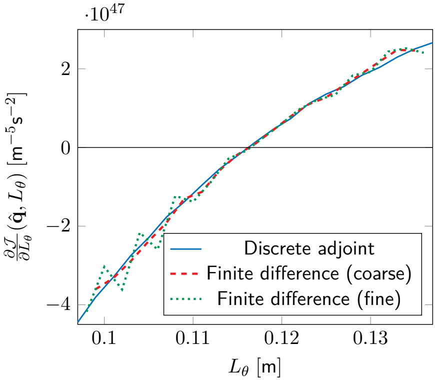
- Uses either 1, 2 or $1 + 2i,\ i = 1, 2, \ldots$ uniform values
- How many depends on the first value
- Solution: Seed second generator

Generator timings



Preliminary results



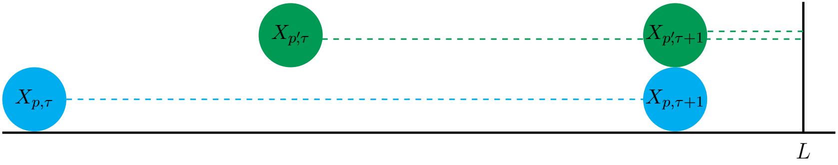


	$n_\mathbf{i}$-solver	$u_\parallel$-solver	Neutral solver
State	3.5×10^{-5} s	4.1×10^{-5} s	6.6 s
Adjoint	3.1×10^{-5} s	3.6×10^{-5} s	9.8 s

Research Foundation Flanders
Opening new horizons

Generality

- Works in principle for time stepping linear in $\hat{q}$
- ... but adjoint can be tricky to implement, e.g.
  - Reflections
  - Collisions in batched simulations
- Combination with AD still open question



EUROfusion
This work has been carried out within the framework of the EUROfusion Consortium, funded by the European Union via the Euratom Research and Training Programme (Grant Agreement No 101052200 — EUROfusion). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the European Commission can be held responsible for them.

Løvbak, E., Blondeel, F., Lee, A., Vanroye, L., Van Barel, A., Samaey, G., *Reversible random number generation for adjoint Monte Carlo simulation of the heat equation.* Monte Carlo and Quasi-Monte Carlo Methods - MCQMC 2022. **Submitted** (2023) *arXiv:2302.02778*

Løvbak, E., *Multilevel and adjoint Monte Carlo methods for plasma edge neutral particle models.* PhD thesis (2023)

Research Foundation Flanders
Opening new horizons
NUMA