International Conference on Multiscale Modeling and Simulation based on Physics and Data
in celebration of the 70th birthday of Russel Caflisch

Gradient-based Monte Carlo methods for hyperbolic and kinetic equations

Lorenzo Pareschi

School of Mathematical and Computer Sciences
Maxwell Institute for Mathematical Sciences, Heriot-Watt University, Edinburgh, UK

Department of Mathematics and Computer Science
University of Ferrara, Italy

joint research with:

G. Bertaglia (Ferrara), R. Caflisch (NYU), G. Dimarco (Ferrara)

IPAM, UCLA - April 25-26, 2024

MAXWELL INSTITUTE FOR MATHEMATICAL SCIENCES

The Wolfson Foundation THE ROYAL SOCIETY

HERIOT WATT UNIVERSITY

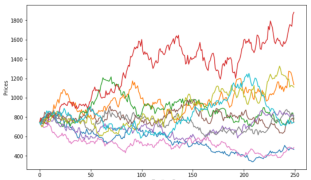
Our joint scientific journey on Monte Carlo methods

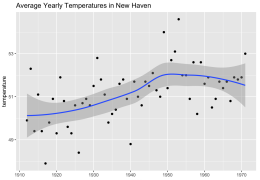
- L. P., R. Caflisch, J. Comput. Phys. (1999).
- R. Caflisch, L. P., IMA Vol. Math. Appl. (2004).
- R. Caflisch, C. Han, L. Erding, L. P., 44th AIAA Aerospace Sciences Meeting, (2006).
- G. Dimarco, R. Caflisch, L .P., Commun. Appl. Ind. Math. (2010).
- L. P. G. Russo '01; P. Degond, G. Dimarco, L. P. '11-'13; W. Ren, H. Liu, S. Jin '14; A. Crestetto, N. Crouseilles, G. Dimarco, M. Lemou '20; F. Fei '23
- G. Bertaglia, L. P., R. Caflisch, J. Sci. Comp., arXiv:2308.02904, (2023).

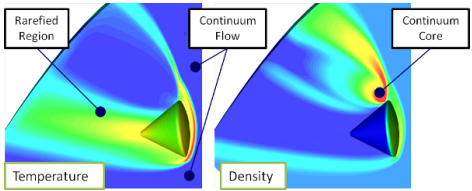
In spite of their importance in applications, Monte Carlo methods receive relatively little attention from numerical analysts and applied mathematicians [...] innovations [...] are developed mainly by practitioners, including physicists, systems engineers and statisticians (R.E. Caflisch, Acta Numerica, '98).

Monte Carlo methods and PDEs

- Monte Carlo methods are widely used in many scientific and industrial applications, but their systematic design as a numerical analysis tool to solve PDEs is still limited to specific contexts compared to deterministic approaches.





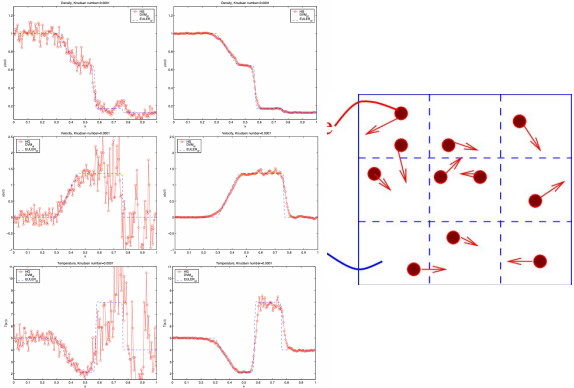


- Examples are applications involving stochastic terms, like in finance or uncertainty quantification, diffusion problems or when other numerical methods may not be feasible due to complexity aspects, like in collisional plasma physics and rarefied gas dynamics.
- Despite its conceptual simplicity, the computational effort to obtain a *good approximation* can be *very high*, since generally many samples are required to reduce statistical fluctuations.

Variance reduction techniques

Several variance reduction Monte Carlo techniques aiming to minimize the stochastic fluctuations in the simulation results have been developed by employing methods such as

- antithetic variables
- importance sampling
- stratified sampling
- moment matching
- control variate
- multi-level methods
- hybrid methods
- ...



$\Rightarrow$ One field in which Monte Carlo methods are particularly relevant is that of collisional kinetic theory, starting with Bird's famous DSMC algorithm[1].

[1] G. Bird '94; K. Nanbu '80; L. P., G. Russo '01; R. Caflisch '98; S. Rjasanow, W. Wagner '05; R. Caflisch, D. Silantyev, Y. Yang '21

Outline

Monte Carlo for hyperbolic problems

The starting point is the following 2×2 relaxation system[2]

$$\partial_t u + \partial_x v = 0,$$
$$\partial_t v + a^2 \partial_x u = -\frac{1}{\varepsilon}\left(v - F(u)\right).$$

As $\varepsilon \to 0$ ($a > |F'(u)|$) we get the local equilibrium $v = F(u)$ and thus the scalar conservation law

$$\partial_t u + \partial_x F(u) = 0.$$

The system can be written in diagonal (kinetic) form

$$\partial_t f^+ + a\partial_x f^+ = -\frac{1}{\varepsilon}\left(f^+ - E^+(u)\right),$$
$$\partial_t f^- - a\partial_x f^- = -\frac{1}{\varepsilon}\left(f^- - E^-(u)\right),$$

where

$$f^+ = \frac{au + v}{2a}, \quad f^- = \frac{au - v}{2a}, \quad E^+(u) = \frac{au + F(u)}{2a}, \quad E^-(u) = \frac{au - F(u)}{2a}.$$

[2] S.Jin, Z. Xin '95; R. Natalini '96

A direct Monte Carlo approach

The solution in a time interval $[0, \Delta t]$ is approximated by a splitting between transport and relaxation

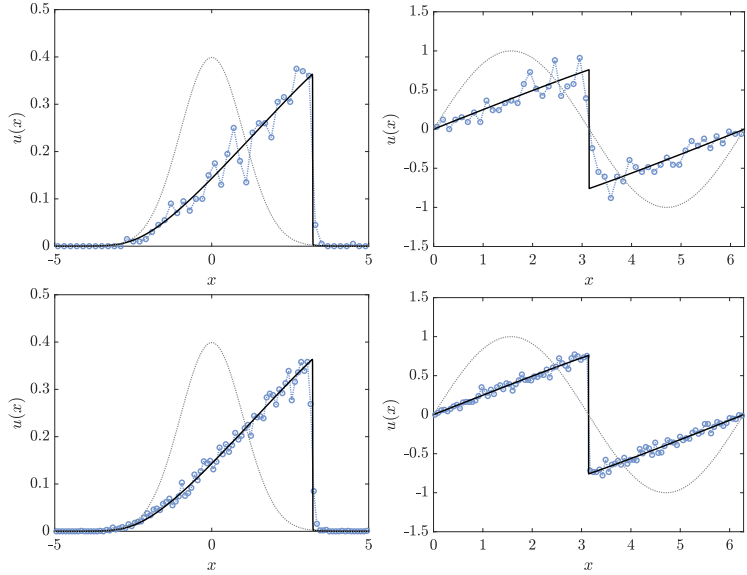
$$\partial_t f^+ + a\partial_x f^+ = 0, \qquad \partial_t f^+ = -\frac{1}{\varepsilon}\left(f^+ - E^+(u)\right),$$
$$\partial_t f^- - a\partial_x f^- = 0, \qquad \partial_t f^- = -\frac{1}{\varepsilon}\left(f^- - E^-(u)\right).$$

Given a set of samples $X_1, \dots, X_N$ with velocities $V_1, \dots, V_N$, $V_i \in \{-a, a\}$ the first step corresponds to a free particle motion whereas the second step modifies velocities accordingly to

$$f^+(x, \Delta t) = e^{-\Delta t/\varepsilon} f^+(x, \Delta t) + \left(1 - e^{-\Delta t/\varepsilon}\right) E^+(u(x, \Delta t)),$$
$$f^-(x, \Delta t) = e^{-\Delta t/\varepsilon} f^-(x, \Delta t) + \left(1 - e^{-\Delta t/\varepsilon}\right) E^-(u(x, \Delta t)).$$

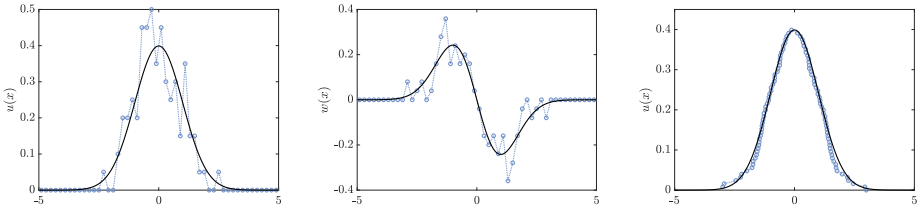
- In order to estimate the probability of a velocity change, we must reconstruct the mass density u in a neighborhood of the particle position. This requires the introduction of a grid in space.
- Negative solutions can also be considered by introducing samples with negative weights.

A numerical example: $F(u) = u^2/2$, $\varepsilon = 10^{-6}$, $N = 1000$, $N = 10000$, $M = 50$



Gradient based Monte Carlo methods

- A Monte Carlo approach, inspired by the vortex method for the Navier-Stokes equations, had considerable success in the '80s and '90s for reaction-diffusion problems[3].
- The idea of the method is to use the spatial derivative of the solution (i.e., the gradient in multiple dimensions) as the unknown variable.



- The statistical solution is not reconstructed from the histogram, as in standard Monte Carlo approaches, but directly as cumulative distribution function of the samples of the derivative.
- The Gradient Random Walk (GRW) method, besides reducing fluctuations, offers adaptivity by sampling based on space derivatives and features a grid-free structure.

[3] A.J. Chorin '73; A. Ghoniem, F.S. Sherman '85; A.S. Sherman, C.S. Peskin '86; S. Roberts '89; A.S. Sherman, M. Mascagni '94

Gradient based Monte Carlo for Jin-Xin relaxation system

By introducing the auxiliary variables $w = \partial_x u$ and $z = \partial_x v$, we can rewrite the system as

$$\partial_t w + \partial_x z = 0,$$
$$\partial_t z + a^2 \partial_x w = -\frac{1}{\varepsilon}\left(z - F'(u)w\right).$$

Again introducing the new variables

$$g^+ = \frac{(aw+z)}{2a},\ g^- = \frac{(aw-z)}{2a},\ D^+(u,w) = \frac{w(a+F'(u))}{2a},\ D^-(u,w) = \frac{w(a-F'(u))}{2a},$$

we obtain the corresponding diagonal form

$$\partial_t g^+ + a\partial_x g^+ = -\frac{1}{\varepsilon}(g^+ - D^+(u,w)),$$
$$\partial_t g^- - a\partial_x g^- = -\frac{1}{\varepsilon}(g^- - D^-(u,w)).$$

The Gradient based Monte Carlo method

After splitting, the transport step and the relaxation step can be solved using a Monte Carlo approach, but there is no need to reconstruct $u(x,t)$ on a space grid during the relaxation step.

Given a set of samples in $X_1(t), \dots, X_N(t)$ with weights $m_i \in \{-m, m\}$ we compute $u(X_i(t))$ as

$$w_N(x,t) = \frac{1}{N}\sum_{k=1}^{N} m_k \delta(x - X_k(t)), \qquad u_N(x,t) = \frac{1}{N}\sum_{k=1}^{N} m_k H(x - X_k(t)),$$

where $H(\cdot)$ is the Heaviside function and the weight values $m_i \in \{-m, m\}$, $m > 0$.

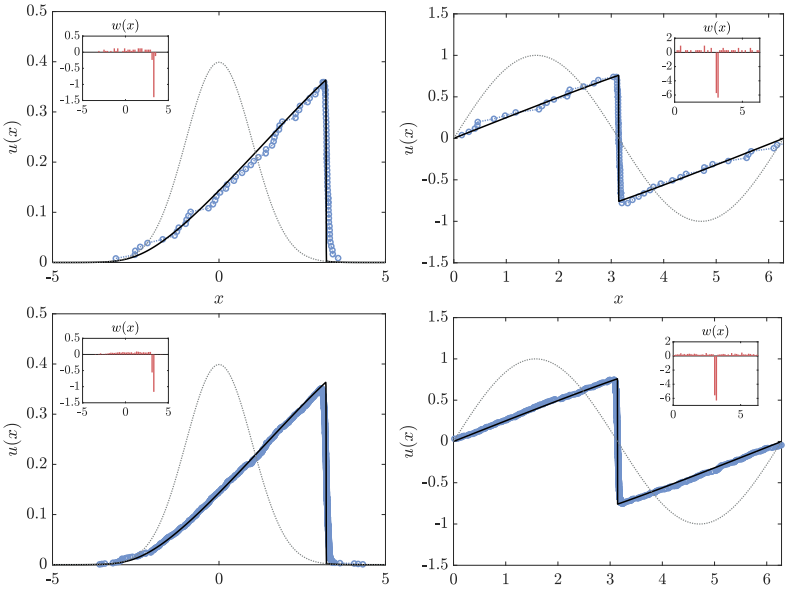
Since $D^+(u,w) + D^-(u,w) = w$, under the subcharacteristic condition $a > |F'(u)|$, the probabilities of a random velocity change read

$$p_D^+(x,\Delta t) = \frac{D^+(u,w)}{w} = \frac{a + F'(u)}{2a}, \qquad p_D^-(x,\Delta t) = \frac{D^-(u,w)}{w} = \frac{a - F'(u)}{2a}.$$

Probabilities depend only on u which is known at each particle position.

$\Rightarrow$ Boundary conditions can be exactly satisfied on one side, while on the other side, they are satisfied on average. Left and right reconstructions can be combined to increase accuracy.

A numerical example: $F(u) = u^2/2$, $\varepsilon = 10^{-6}$, $N = 100$, $N = 1000$



Reconstruction error: Monte Carlo (MC)

Given $X_1(t), \dots, X_N(t)$ i.i.d. as $u(x,t)$ we compute the empirical density function

$$u_N(x,t) = \frac{1}{N}\sum_{k=1}^{N} \delta(x - X_k(t)), \qquad x \in \Omega \subseteq \mathbb{R}, \quad t \geq 0.$$

The above function needs to be reconstructed on a mesh with width $\Delta x > 0$

$$u_{N,\Delta x}(x,t) = \frac{1}{N}\sum_{k=1}^{N} S_{\Delta x}(x - X_k(t)).$$

For example using the rectangular kernel $S_{\Delta x}(x) = \chi(|x| \leq \Delta x/2)/\Delta x$.
The reconstruction error can be estimated from

$$\|u(\cdot,t) - u_{N,\Delta x}(\cdot,t)\|_{L^p(\Omega,L^2(\Omega))} \leq \underbrace{\|u(\cdot,t) - u_{\Delta x}(\cdot,t)\|_{L^p(\Omega)}}_{\text{deterministic error}} + \underbrace{\|u_{\Delta x}(\cdot,t) - u_{N,\Delta x}(\cdot,t)\|_{L^p(\Omega,L^2(\Omega))}}_{\text{stochastic error}},$$

where $\|g\|_{L^p(\Omega,L^2(\Omega))} = \|\mathbb{E}\left[g^2\right]^{1/2}\|_{L^p(\Omega)}$ and $u_{\Delta x}(x,t) = \int_\Omega S_{\Delta x}(x-y)u(y,t)\,dy$.

For the stochastic error we have the classical root mean squared error estimate

$$\mathbb{E}\left[(u_{\Delta x}(x,t) - u_{N,\Delta x}(x,t))^2\right]^{1/2} = \frac{\sigma_S(x,t)}{N^{1/2}}$$

where

$$\sigma_S^2(t,x) = \int_\Omega \left(S_{\Delta x}(x-y) - u_{\Delta x}(x,t)\right)^2 u(y,t)\, dy.$$

Next, by assuming that the first term satisfies

$$\|u(\cdot,t) - u_{\Delta x}(\cdot,t)\|_{L^p(\Omega)} \leq C_q(\Delta x)^q,$$

according to the order of accuracy $q \geq 1$ used in the reconstruction, we have

Theorem

For a sufficiently smooth function $u(x,t)$ the reconstruction error satisfies

$$\|u(\cdot,t) - u_{N,\Delta x}(\cdot,t)\|_{L^p(\Omega,L^2(\Omega))} \leq \frac{\|\sigma_S\|_{L^p(\Omega)}}{N^{1/2}} + C_q(\Delta x)^q,$$

where C_q depends on the q derivative of $u(x,t)$ and the domain Ω.

More in general, the error is affected by the numerical solution of the PDE. In this latter case for a first order (in space and time) approximation we can prove

Theorem (MC error)

Let us denote by $\tilde{u}_{N,\Delta x}(\cdot,t)$ the reconstructed MC solution then

$$\|u(\cdot,t)-\tilde{u}_{N,\Delta x}(\cdot,t)\|_{L^p(\Omega,L^2(\Omega))} \leq \frac{\|u\|^{1/2}_{L^{p/2}(\Omega)}}{(\Delta x N)^{1/2}} + \frac{C_2}{N^{1/2}} + C_1\Delta x$$

where the constants C_1, C_2 depend on the first order derivative of $u(x,t)$ and Ω.

If we want to minimize the error with respect to Δx we should take

$$\Delta x = \left(\frac{\|u\|_{L^{p/2}(\Omega)}}{4C_1^2 N}\right)^{1/3}.$$

This will ensure an optimal error decay $O(N^{-1/3})$. Therefore, the optimal mesh will scale as $\Delta x \approx N^{-1/3}$ for a convergence rate of $O(N^{-1/3})$, where the precise value of C_1 depends on the particular PDE under consideration.

Reconstruction error: Gradient-based Monte Carlo (GBMC)

Assume, for simplicity, that $u(x,t)$ is nondecreasing so that

$$w(x,t) = \frac{\partial u(x,t)}{\partial x} \geq 0.$$

Therefore, we can consider $w(x,t)$ as a probability density and $u(x,t)$ its cumulative distribution function (CDF).

Given N samples $X_1(t), \ldots, X_N(t)$ i.i.d. as $w(x,t)$, we compute the empirical CDF as

$$u_N(x,t) = \frac{1}{N} \sum_{k=1}^{N} H(x - X_k(t)).$$

The empirical CDF does not need any further regularization over a mesh grid.

Now, let's consider the problem of estimating the convergence rate in L^p spaces. Since

$$u(x,t) = \int_\Omega H(x-y)w(y)\,dy, \qquad u_N(x,t) = \int_\Omega H(x-y)w_N(y)\,dy,$$

the numerical error of the empirical CDF can be estimated from

$$\mathbb{E}[(u(x,t) - u_N(x,t))^2]^{1/2} = \frac{\sigma_H(x,t)}{N^{1/2}}$$

where

$$\sigma_H^2(x,t) = \int_\Omega \left(H(x-y) - u(x,t)\right)^2 w(y,t)\,dy.$$

Theorem

The empirical CDF satisfies

$$\|u(\cdot,t) - u_N(\cdot,t)\|_{L^p(\Omega,L^2(\Omega))} \leq \frac{\|u\|^{1/2}_{L^{p/2}(\Omega)}}{N^{1/2}}.$$

The proof follows immediately since $\sigma_H^2(x,t) = u(x,t)(1-u(x,t)) \leq u(x,t)$.

If we now consider the error introduced by the GBMC solution of the PDE, we can assume

$$\|u(\cdot,t) - \tilde{u}(\cdot,t)\|_{L^p(\Omega)} \leq \tilde{C}_1 \Delta t,$$

where $\tilde{u}(\cdot,t) = \mathbb{E}[\tilde{u}_N(\cdot,t)]$, $\tilde{u}_N(\cdot,t)$ is the GBMC solution of the PDE and $\tilde{C}_1$ depends on the first order time derivative of the solution and the domain Ω.

Theorem (GBMC error)

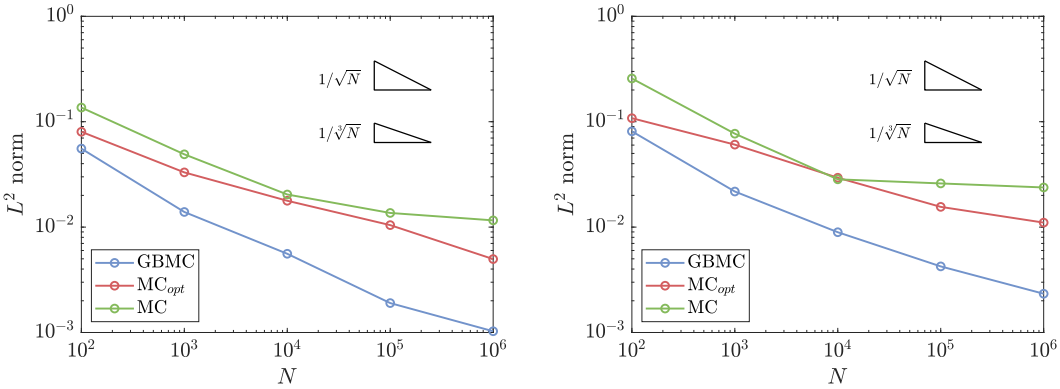
Let us denote by $\tilde{u}_N(\cdot,t)$ the GBMC solution, then

$$\|u(\cdot,t) - \tilde{u}_N(\cdot,t)\|_{L^p(\Omega,L^2(\Omega))} \leq \frac{\|u\|^{1/2}_{L^{p/2}(\Omega)}}{N^{1/2}} + \tilde{C}_1 \Delta t.$$

Clearly the time step should be taken as $\Delta t \approx N^{-1/2}$ to optimize the error in the GBMC solution and achieve an optimal convergence rate of $O(N^{-1/2})$.

The GBMC error is always smaller than the MC error. More precisely, the two errors have the same decay for a fixed (non optimal) Δx, whereas for an optimal Δx the GBMC method has a faster convergence rate as $O(N^{-1/2})$ against $O(N^{-1/3})$ of the MC method.

Numerical convergence rate



Comparison of the relative L^2 error norms of the Monte Carlo (MC) and the Gradient-based Monte Carlo (GBMC) with respect to the number of particles N for $F(u) = u^2/2$, $\varepsilon = 10^{-6}$ at $t = 2.5$ with normal distribution as initial datum (left) and at $t = 0.5$ with sinusoidal distribution as initial datum (right).

Generalization to hyperbolic relaxation systems

The relaxation approximation for systems of conservation laws now reads[4]

$$\partial_t \mathbf{u} + \partial_x \mathbf{v} = 0,$$
$$\partial_t \mathbf{v} + A^2 \partial_x \mathbf{u} = -\frac{1}{\varepsilon}(\mathbf{v} - \mathbf{F}(\mathbf{u})),$$

where $\mathbf{v} \in \mathbb{R}^n$ and $A^2 = \mathrm{diag}\{a_1^2, \dots, a_n^2\}$ must satisfy the condition $A^2 > \mathbf{F}'(\mathbf{u})^2$.
The diagonal variables and their equilibrium states are

$$\mathbf{f}^{\pm} = A^{-1}\frac{A\mathbf{u} \pm \mathbf{v}}{2}, \qquad \mathbf{E}^{\pm}(\mathbf{u}) = A^{-1}\frac{A\mathbf{u} \pm \mathbf{F}(\mathbf{u})}{2}$$

which yield the system

$$\partial_t \mathbf{f}^+ + A\partial_x \mathbf{f}^+ = -\frac{1}{\varepsilon}(\mathbf{f}^+ - \mathbf{E}^+(\mathbf{u}))$$
$$\partial_t \mathbf{f}^- - A\partial_x \mathbf{f}^- = -\frac{1}{\varepsilon}(\mathbf{f}^- - \mathbf{E}^-(\mathbf{u})),$$

[4] G.Q. Chen, D. Levermore, T.P. Liu '94

The Gradient-based formulation

To extend the GBMC method we introduce the auxiliary vectors $\mathbf{w} = \partial_x \mathbf{u}$, $\mathbf{z} = \partial_x \mathbf{v}$ and the diagonal variables

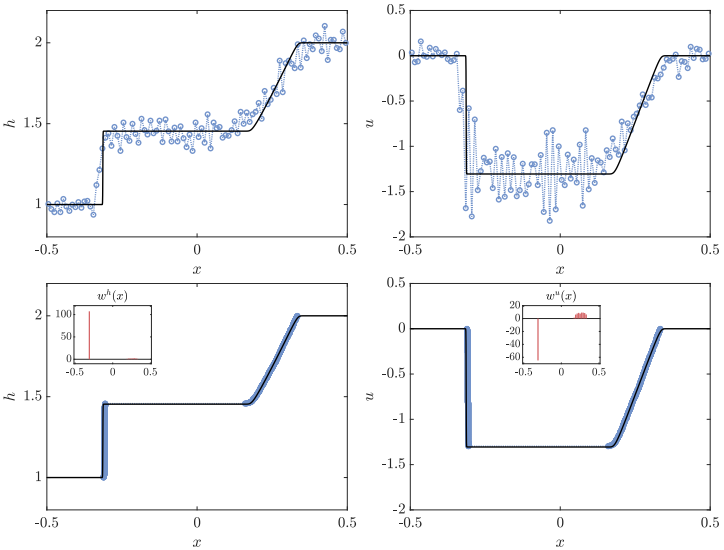
$$\mathbf{g}^{\pm} = A^{-1}\frac{A\mathbf{w} \pm \mathbf{z}}{2}, \qquad \mathbf{D}^{\pm}(\mathbf{u}, \mathbf{w}) = A^{-1}\frac{(A \pm \mathbf{F}'(\mathbf{u}))\mathbf{w}}{2},$$

to obtain

$$\partial_t \mathbf{g}^+ + A\partial_x \mathbf{g}^+ = -\frac{1}{\varepsilon}(\mathbf{g}^+ - \mathbf{D}^+(\mathbf{u}, \mathbf{w}))$$
$$\partial_t \mathbf{g}^- - A\partial_x \mathbf{g}^- = -\frac{1}{\varepsilon}(\mathbf{g}^- - \mathbf{D}^-(\mathbf{u}, \mathbf{w})).$$

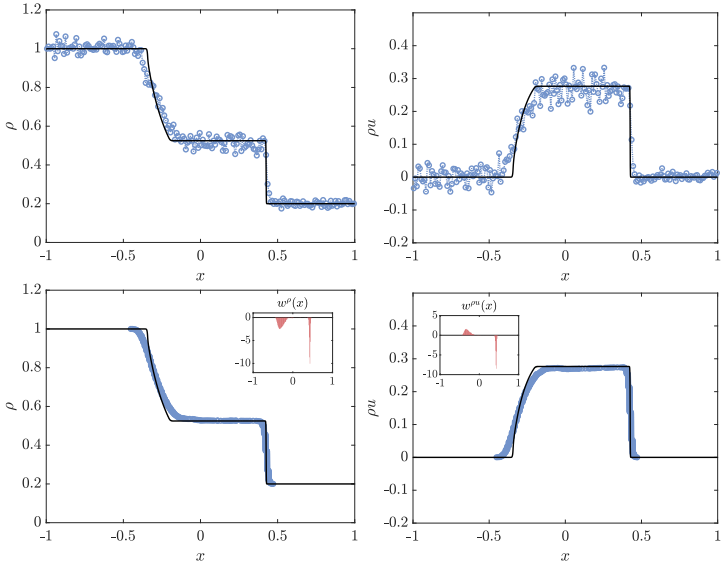
- In general, we cannot make the probabilities of velocity switches in the relaxation process independent on the vector $\mathbf{w}$, unless we can diagonalize the Jacobian matrix $\mathbf{F}'(\mathbf{u})$.
- However, equilibrium states are still defined for each individual particle, $\mathbf{w}$ being reconstructed in the grid cells but $\mathbf{u}$ being particle-dependent.

Shallow water equations: $\mathbf{u} = (h,\, hu)^T$, $\mathbf{F}(\mathbf{u}) = (hu,\, gh^2/2 + hu^2)^T$, $N = 10^5$, $N = 2000$



Initial data: $h(x,0) = 1$, $u(x,0) = 0$, for $x < 0$ and $h(x,0) = 2$, $u(x,0) = 0$, for $x > 0$.

Isentropic Euler equations: $\mathbf{u} = (\rho,\, \rho u)^T$, $\mathbf{F}(\mathbf{u}) = (\rho u,\, (\rho + \rho u^2)/2)^T$, $N = 2 \times 10^5$, $N = 2000$



Initial data: $\rho(x,0) = 1$, $u(x,0) = 0$, for $x < 0$ and $\rho(x,0) = 0.2$, $u(x,0) = 0$, for $x > 0$.

Multiscale kinetic equations

Consider a kinetic equation in the scaled form[5]

$$\partial_t f + \frac{1}{\varepsilon^{\alpha-1}} v \nabla_x f = \frac{1}{\varepsilon^{\alpha}} Q(f,f)$$

where $f = f(x,v,t)$, $(x,v) \in \mathbb{R}^d \times \Omega$, $\varepsilon > 0$. The term $Q(f,f)$ characterizes the changes in the velocity field. Here $\alpha \in [1,2]$, $\alpha = 1$ fluid-dynamic scaling, $\alpha = 2$ diffusive scaling.

- Neutron transport ($\alpha = 2$)

$$Q(f,f) = \left(\frac{\sigma_S(x)}{|\Omega|} \int_\Omega f \, dv - \sigma(x) f \right).$$

- BGK-type models ($\alpha = 1$)

$$Q(f,f) = M[f] - f, \qquad M[f] = \frac{\rho}{(2\pi T)^{d/2}} \exp\left(-\frac{(v-u)^2}{2T} \right),$$

where

$$\rho = \int_{\mathbb{R}^d} f \, dv, \quad u = \frac{1}{\rho} \int_{\mathbb{R}^d} v f \, dv, \quad T = \frac{1}{d\rho} \int_{\mathbb{R}^d} (v-u)^2 f \, dv.$$

[5] C. Cercignani, R. Illner, M. Pulvirenti '94

A direct Monte Carlo for the BGK model

A direct MC method is based on splitting in a time step Δt of transport

$$\partial_t f + v \nabla_x f = 0,$$

and collision

$$\partial_t f = \frac{1}{\varepsilon}(M[f] - f).$$

After the free transport we can consider the exact solution of the collision step

$$f(x, v, t + \Delta t) = e^{-\frac{\Delta t}{\varepsilon}} f(x, v, t) + (1 - e^{-\frac{\Delta t}{\varepsilon}}) M[f](x, v, t).$$

$\Rightarrow$ At each point $x \in \mathbb{R}$ with probability $1 - e^{-\frac{\Delta t}{\varepsilon}}$ we replace the particle velocities with samples from the Maxwellian distribution $M[f]$.

The Gradient based Monte Carlo

We restrict to the one-dimensional case $d = 1$ and introduce the auxiliary function

$$g(x, v, t) = \partial_x f(x, v, t)$$

which satisfies the equation

$$\partial_t g + v\partial_x g = \frac{1}{\varepsilon}(\tilde{M}[f, g] - g),$$

where $\tilde{M}[f, g] = \partial_x M[f]$.

We can easily compute

$$\tilde{M}[f, g] = h[f, g] \exp\left(-\frac{(v-u)^2}{2T}\right),$$

with

$$h[f, g] = \frac{T\partial_x\rho - \rho\partial_x T}{T\sqrt{2\pi T}} + \frac{(v-u)\partial_x u}{T} + \frac{(v-u)^2\partial_x T}{2T^2}$$

where derivatives of moments are computed from

$$\partial_x\rho = \int_{\mathbb{R}^d} g\, dv, \quad \partial_x(\rho u) = \int_{\mathbb{R}^d} vg\, dv, \quad \partial_x(\rho E) = \frac{1}{2}\int_{\mathbb{R}^d} v^2 g\, dv.$$

Moments

We apply a Monte Carlo method to compute the evolution of $g(x, v, t)$. Using N samples $(X_1, V_1), \dots, (X_N, V_N)$, $X_i \in \mathbb{R}$, $V_i \in \mathbb{R}$ we have

$$g_N(x, v, t) = \frac{1}{N} \sum_{k=1}^{N} m_k \delta(x - X_k) \delta(v - V_k).$$

The weights $m_k \in \{-1, 1\}$ takes into account the presence of negative values, and since

$$f(x, v, t) = \int_{-\infty}^{x} g(y, v, t)\, dy$$

we get

$$f_N(x, v, t) = \frac{1}{N} \sum_{k=1}^{N} m_k H(x - X_k) \delta(v - V_k).$$

The moments can be recovered at each position x from

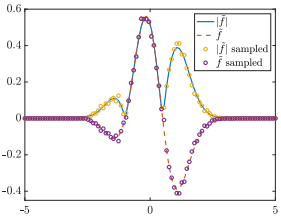
$$\int_{\mathbb{R}} f_N(x, v, t) \varphi(v)\, dv = \frac{1}{N} \sum_{k=1}^{N} m_k H(x - X_k) \varphi(V_k), \qquad \phi(v) = 1, v, |v|^2/2.$$

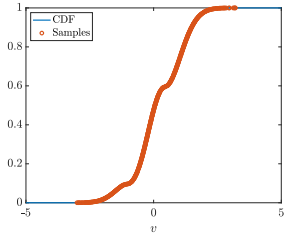
Sampling the collision process

The collision step is solved by Monte Carlo, observing that the moments are defined at each particle position, while the derivatives of the moments require a mesh to be evaluated.

Let us write the exact solution of the collision step in a time interval Δt

$$g(x, v, t + \Delta t) = e^{-\Delta t/\varepsilon} g(x, v, t) + \left(1 - e^{-\Delta t/\varepsilon}\right) \tilde{M}[f, g](x, v, t).$$



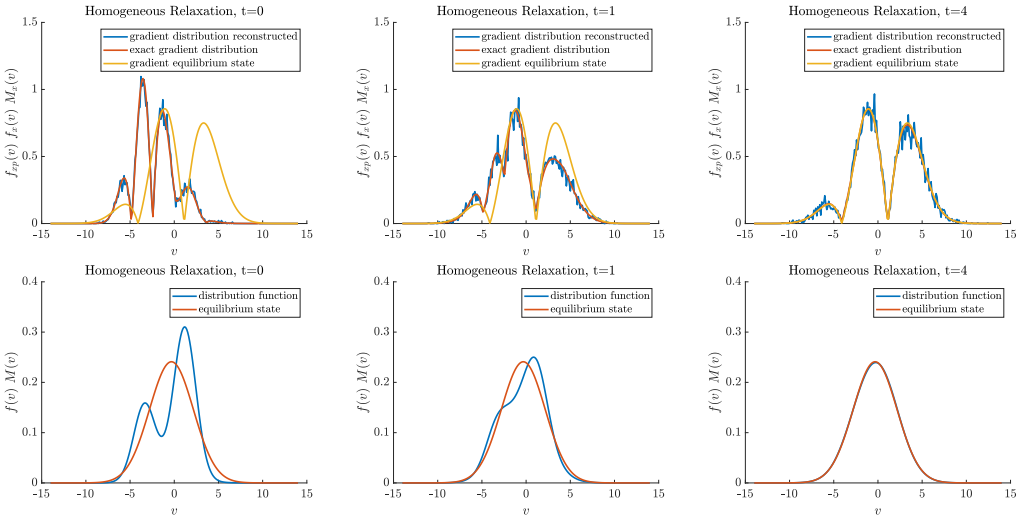


Sampling from $\tilde{M}[f, g](x, v, t)$ can be formalized as sampling from the absolute value of

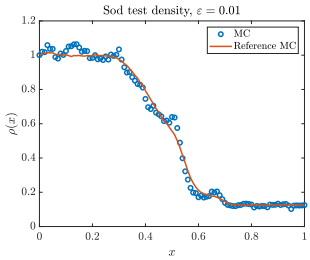
$$\tilde{f}(v) = (a + bv + cv^2)e^{-\beta v^2}, \qquad a, b, c \in \mathbb{R}, \quad \beta \in \mathbb{R}^+,$$

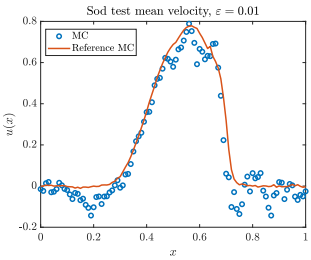
where we assign positive or negative weights to the samples accordingly to the sign of the polynomial $a + bv + cv^2$. An efficient sampling is obtained using the inverse CDF method.

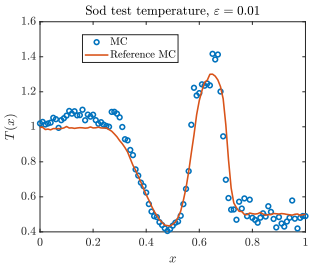
Homogeneous relaxation: $N = 10^4$

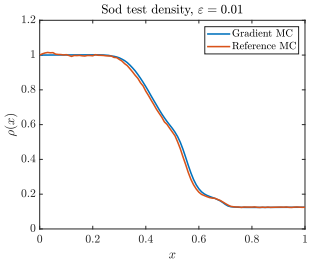


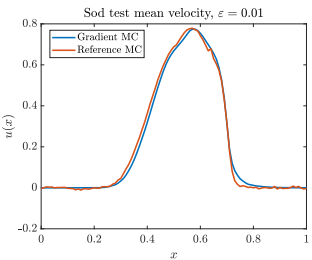
Sod shock tube: $\varepsilon = 0.01$, $N = 10^5$

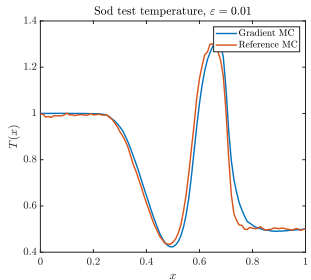












Concluding remarks

- By analogy with direct simulation methods in kinetic theory, we have shown how to extend Monte Carlo techniques to a broad class of problems, including systems of conservation laws.
- Based on previous ideas on reaction-diffusion equations, we designed new Gradient-based Monte Carlo methods both for hyperbolic relaxation systems and kinetic equations.
- This latter class of methods has shown great potential due to its ability to concentrate samples where the solution has large derivatives and its grid-free structure.
- Several research directions:
  - diffusion limits
  - extension to multi-dimensions
  - other kinetic equations
  - uncertainty quantification, control of particle systems, ...

Happy 70th birthday Russ!