UCLA

Recent Innovative Numerical Methods

Russel Caflisch
IPAM
Mathematics Department, UCLA

math changes everything.

Overview

- Multi Level Monte Carlo (MLMC)
- Hybrid Direct Simulation Monte Carlo (DSMC)
- Accelerated Molecular Dynamics (AMD)
- Multiscale Finite Element Method (MFEM)
- In-Painting for Scientific Computation
- Split Bregman
- Empirical Mode Decomposition (EMD)
- Stochastic Gradient Descent (SGD)

Overview

- Multi Level Monte Carlo (MLMC)
- Hybrid Direct Simulation Monte Carlo (DSMC)
- Accelerated Molecular Dynamics (AMD)
- Multiscale Finite Element Method (MFEM)
- In-Painting for Scientific Computation
- Split Bregman
- Empirical Mode Decomposition (EMD)
- Stochastic Gradient Descent (SGD)

Multi-Level Monte Carlo for SDEs

Stochastic differential equation (SDE)

$$dv_i = F_i\,dt + D_{ij}\,dW_j, \tag{11}$$

where f is probability density of $\mathbf{v}$ and i, j are component indices

- $W = W(t)$ is Brownian motion in velocity
- dW is white noise in velocity
-
-

Objective is an average of f:

$$\frac{1}{\rho}\int P(\mathbf{v})f(\mathbf{v},t)\,d\mathbf{v} \equiv \mathbb{E}[P(\mathbf{v}(t))] \tag{12}$$

Discretization of SDEs

Euler-Maruyama discretization in time:

$$v_{i,n+1} = v_{i,n} + F_{i,n}\Delta t + D_{ij,n}\,\Delta W_{j,n}, \tag{13}$$

$$\Delta \mathbf{W}_n = \mathbf{W}_{n+1} - \mathbf{W}_n \tag{14}$$

in which $v_{i,n} = v_i(t_n)$ and $\mathbf{F}_n = \mathbf{F}(\mathbf{v}_n)$

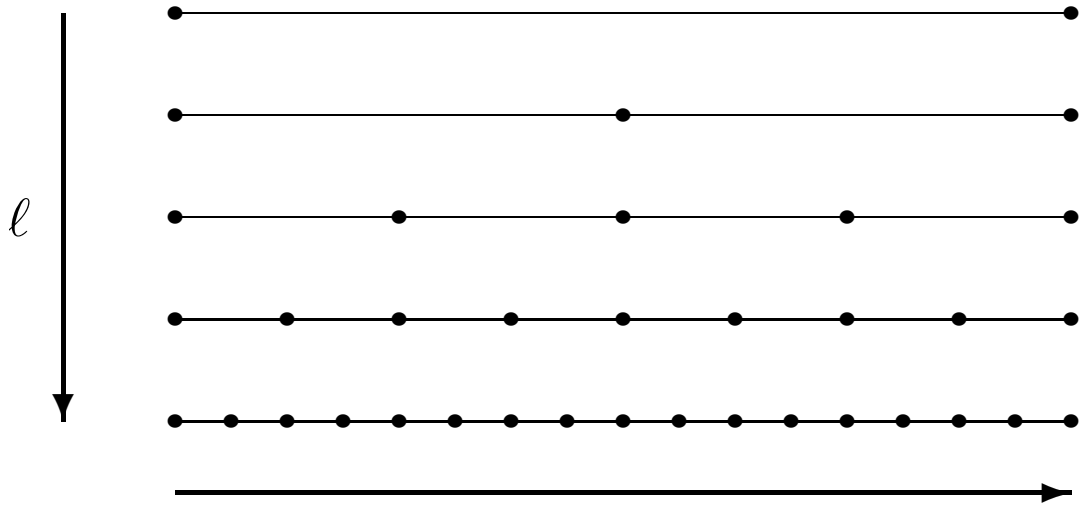
- Choose N Brownian paths to get N values of $P(\mathbf{v}(T))$
- Average to approximate $\mathbb{E}[P(\mathbf{v}(T))]$

Computational cost vs. Error ε:

- Statistical error is $O(N^{-1/2})$
- Δt error is $O(\Delta t)$, since $\Delta W = O(\sqrt{\Delta t})$ and random
- Optimal choice is $\varepsilon = N^{-1/2} = \Delta t$
- Cost $= N\Delta t^{-1} = \varepsilon^{-3}$

MLMC Basics

- Introduce time step levels, $\Delta t_\ell = T2^{-\ell}$, for $\ell = 0, ..., L$



- Let $P_\ell = P(\mathbf{v}_{\Delta t_\ell})$. Then

$$\mathbb{E}[P_L] = \mathbb{E}[P_0] + \sum_{\ell=1}^{L} \mathbb{E}\left[P_\ell - P_{\ell-1}\right] \tag{24}$$

- When computed using same Brownian path, the variance of $(P_\ell - P_{\ell-1})$ is $O(\text{strong error})^2$

MLMC Scaling

Optimal number of samples used to compute each $\mathbb{E}[P_\ell - P_{\ell-1}]$, constrained by RMSE $< \varepsilon$. The complexity now scales like[11]

$$\text{Cost} = \begin{cases} O\left(\varepsilon^{-2}(\log \varepsilon)^2\right) & \text{for Euler-Maruyama} \\ O\left(\varepsilon^{-2}\right) & \text{for Milstein} \end{cases} \tag{25}$$

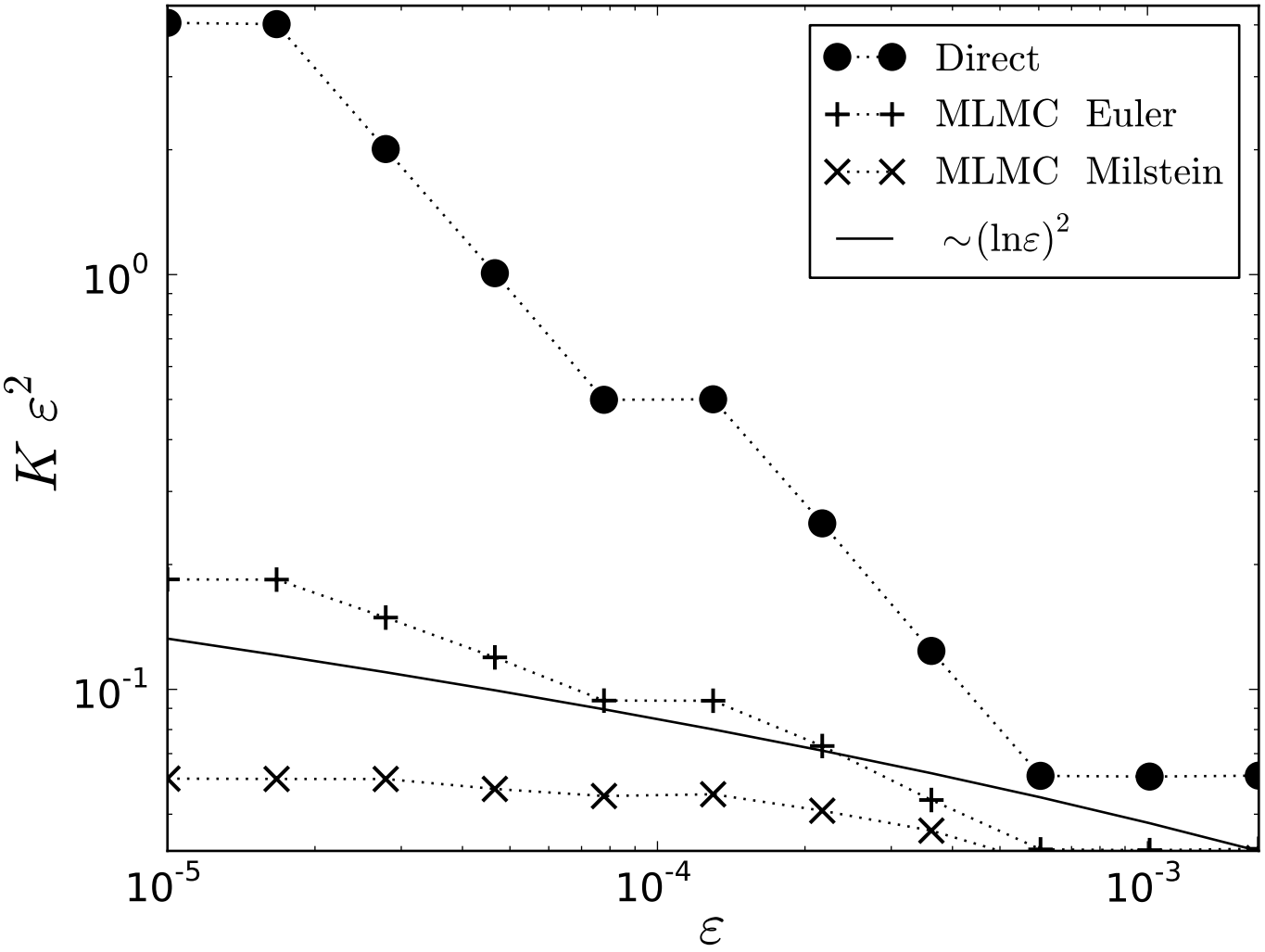
Notes:

- MLMC-Euler-Maruyama scales better than standard MC
- MLMC-Milstein is even better
  - Restricted to $d = 1, 2$ due to difficulty with Levy areas
- $O\left(\varepsilon^{-2}\right)$ scaling is possible without Milstein, using antithetic sampling method [12]

[11] Giles, *Operations Research*, 56(3):607, 2008

[12] Giles & Szpruch, *arXiv:1202.6283*, 2012

A Sample Plasma Problem



Rosin, Ricketson, et. al., submitted to JCP, 2013

Overview

- Multi Level Monte Carlo (MLMC)
- Hybrid Direct Simulation Monte Carlo (DSMC)
- Accelerated Molecular Dynamics (AMD)
- Multiscale Finite Element Method (MFEM)
- In-Painting for Scientific Computation
- Split Bregman
- Empirical Mode Decomposition (EMD)
- Stochastic Gradient Descent (SGD)

Hybrid DSMC for Plasmas

Vlasov-Poisson-Landau (VPL) system for non-equilibrium plasma

$$\begin{cases} \partial_t f + \mathbf{v} \cdot \nabla_{\mathbf{x}} f - \mathbf{E} \cdot \nabla_{\mathbf{v}} f = Q_L(f,f), \\ -\nabla_{\mathbf{x}} \cdot \mathbf{E} = \rho(t,\mathbf{x}) = \int f(t,\mathbf{x},\mathbf{v})\,\mathrm{d}\mathbf{v}, \end{cases}$$

Landau-Fokker-Planck operator for Coulomb interactions is

$$Q_L(g,f)(\mathbf{v}) = \frac{A}{4}\frac{\partial}{\partial v_i}\int_{\mathbb{R}^3} u\sigma_{tr}(u)(u^2\delta_{ij} - u_i u_j)\left(\frac{\partial}{\partial v_j} - \frac{\partial}{\partial w_j}\right) g(\mathbf{w}) f(\mathbf{v})\,\mathrm{d}\mathbf{w}$$

- $\mathbf{u} = \mathbf{v} - \mathbf{u}$
- $\sigma_{tr}(u) \approx u^{-3}$
- A is proportional to Coulomb logarithm
- Magnetic field omitted here
- $Q_L(g,f)$ is asymmetric, and describes the change in f due to collisions with g (used in negative particle section below)

DSMC for Plasma Kinetics

The PIC-DSMC method is widely used in plasma simulation

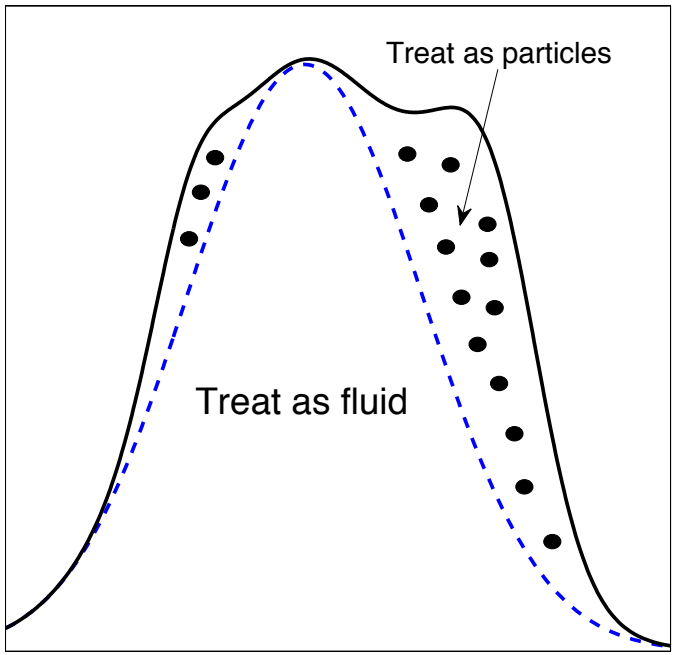
- Particle-In-Cell method (PIC) for collisionless plasma. Dawson 83, Birdsall-Langdon 85
- Direct Simulation Monte Carlo (DSMC) for binary collisions. Takizuka-Abe 77, Nanbu 97, Bobylev-Nanbu 2000

In each time step of DSMC, perform collisions between N_c randomly chosen pairs of particles

- For rarefied gas (charge neutral, short range), $N_c = O(\Delta t N)$.
  - Each particle collides at the physically correct rate.
  - Collisions are physical collisions
- For Coulomb gas (charged, long range), $N_c = N/2$.
  - Every particle collides once in every time step
  - Collisions are aggregates depending on Δt

Hybrid Scheme

Combine fluid and particle simulation methods[1]:



- Separate f into Maxwellian and non-Maxwellian (particle) components: $f = m + f_p$
- Treat m as fluid
- Simulate f_p by Monte Carlo
- Interaction of m and f_p: sample particles from m; collide with particles from f_p
- Similar to δf methods, but fully nonlinear
- Limited to $f_p \geq 0$

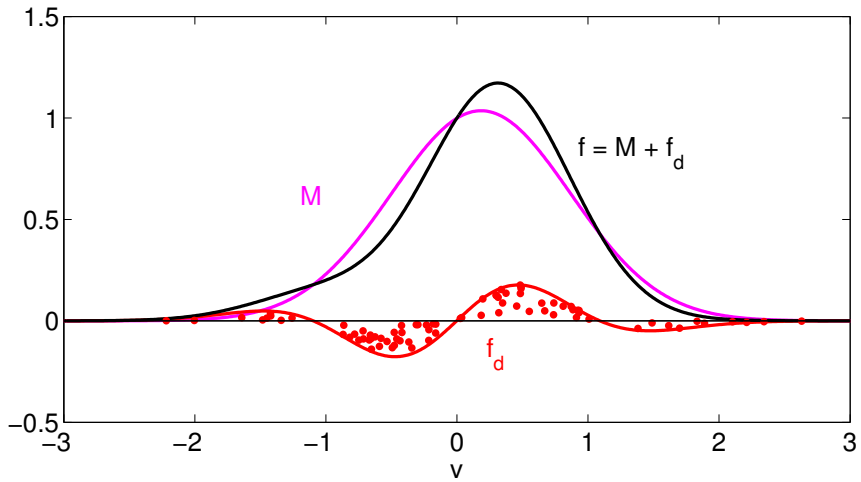
[1] Caflisch et. al, Multiscale Model. Simul. 2008

Motivation for Negative Particles

Apply decomposition

$$f(t, \mathbf{x}, \mathbf{v}) = M(t, \mathbf{x}, \mathbf{v}) + f_d(t, \mathbf{x}, \mathbf{v}),$$

with $f_d(t, \mathbf{x}, \mathbf{v})$ alowed to be positive or negative, so that f_d is minimized.



$(f_d)_+$ and $(f_d)_-$ are represented by positive and negative deviational particles. Developed here for LFP. Similar work for RGD.[5]

[5] Baker & Hadjiconstantinou, Phys Fl (2005)

Meaning of Negative Particles

- A negative particle w_- in f_d cancels a (positive) particle w_+ in m or f_d
- So a $P-N$ collision (v_+, w_-) cancels a corresponding $P-P$ collision (v_+, w_+).
  - The $P-P$ collision removes v_+, w_+ and adds v'_+, w'_+:

$$\text{P-P:} \quad v_+, w_+ \rightarrow v'_+, w'_+$$

  - So the $P-N$ collision adds v_+, removes w_- (i.e., adds w_+), and adds v'_-, w'_- (i.e., removes v'_+, w'_+)

$$\text{P-N:} \quad v_+, w_- \rightarrow 2v_+, v'_-, w'_-$$

Derived from the Boltzmann equation.[7]

Particle number can grow!
New method controls this growth.

[7] Baker & Hadjiconstantinou, Phys Fl (2005)

Nonlinear Landau Damping in VPL system

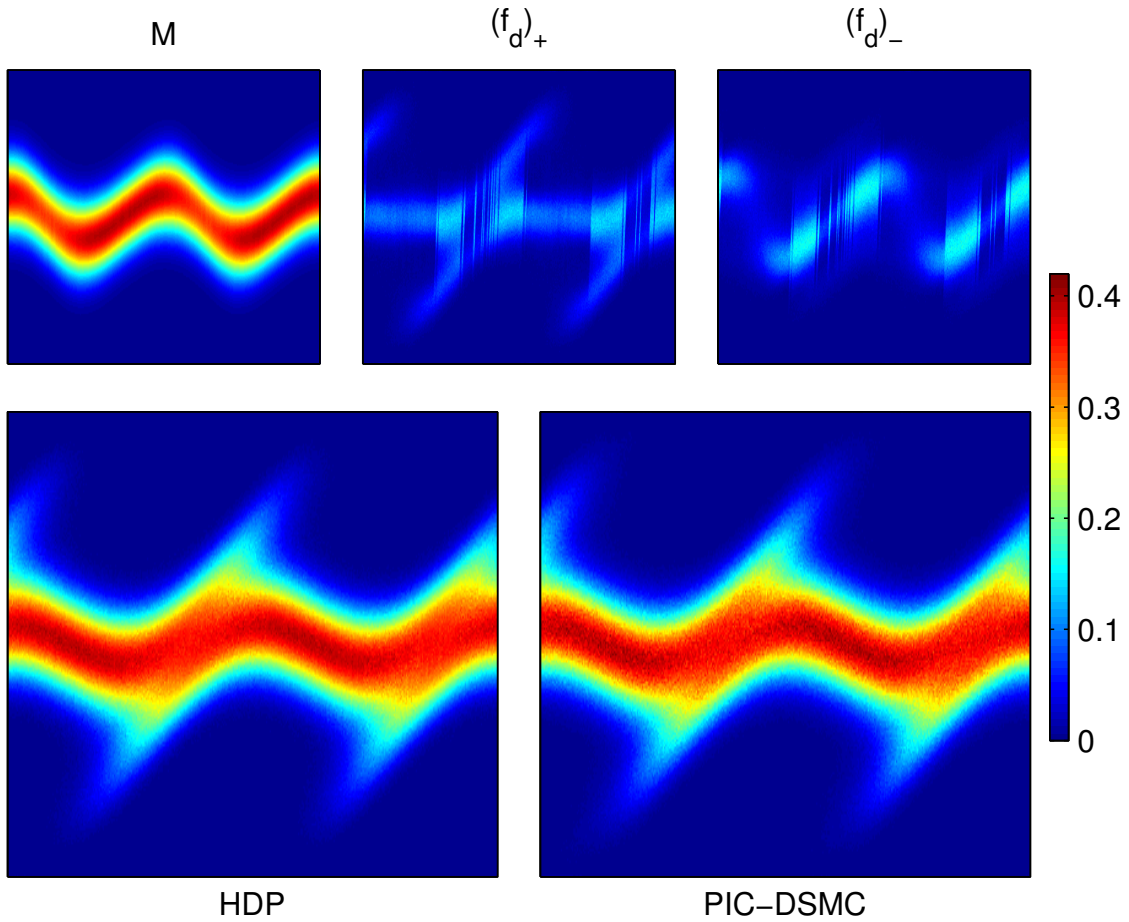


Figure: The distribution in the $x - v_1$ phase space at time $t = 1.25$ in the nonlinear Landau damping problem of the VPL system.

Efficiency Test on VPL System

$$\rho(t = 0, x) = 1 + \alpha \sin(x)$$

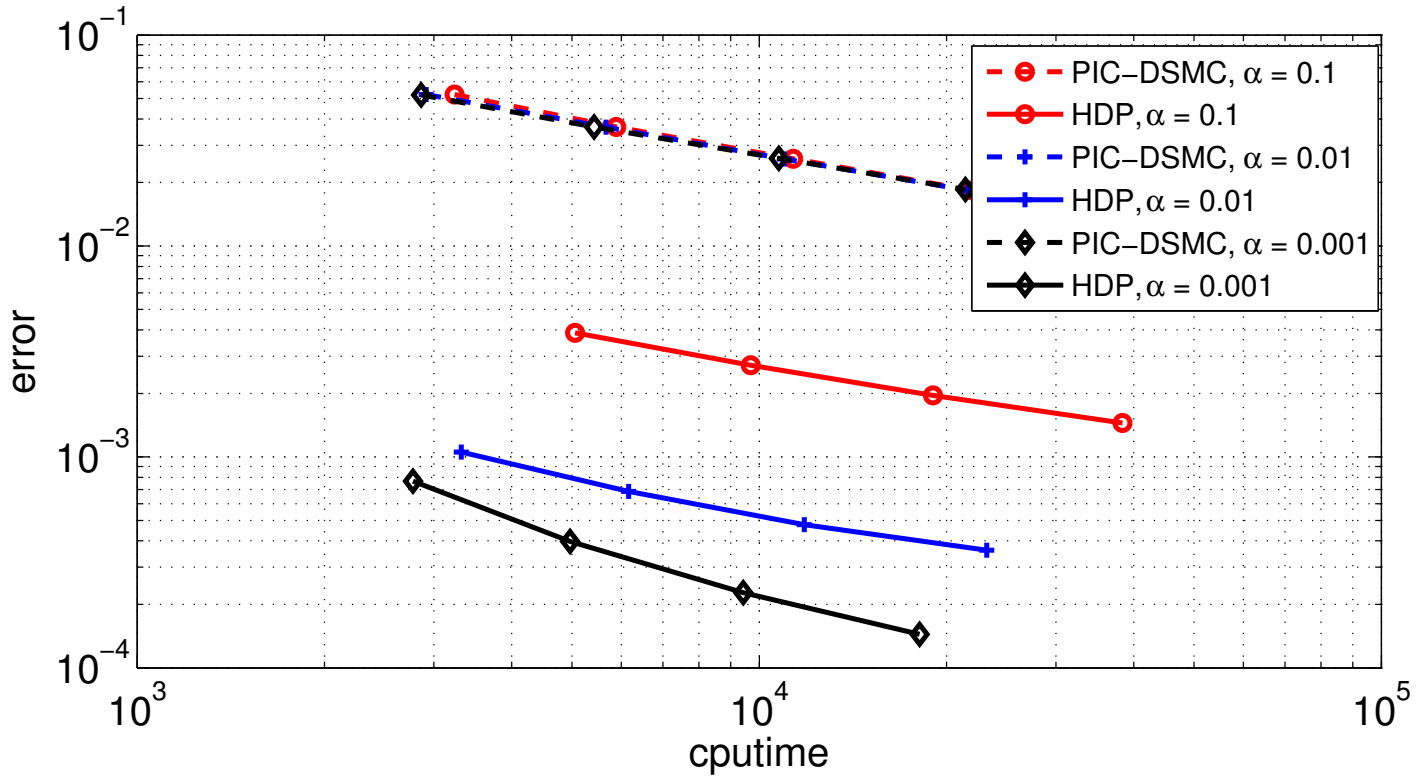


Figure: The efficiency test of the HDP method on the VPL system for different α in the initial density.

Overview

- Multi Level Monte Carlo (MLMC)
- Hybrid Direct Simulation Monte Carlo (DSMC)
- Accelerated Molecular Dynamics (AMD)
- Multiscale Finite Element Method (MFEM)
- In-Painting for Scientific Computation
- Split Bregman
- Empirical Mode Decomposition (EMD)
- Stochastic Gradient Descent (SGD)

Accelerated Molecular Dynamics Methods

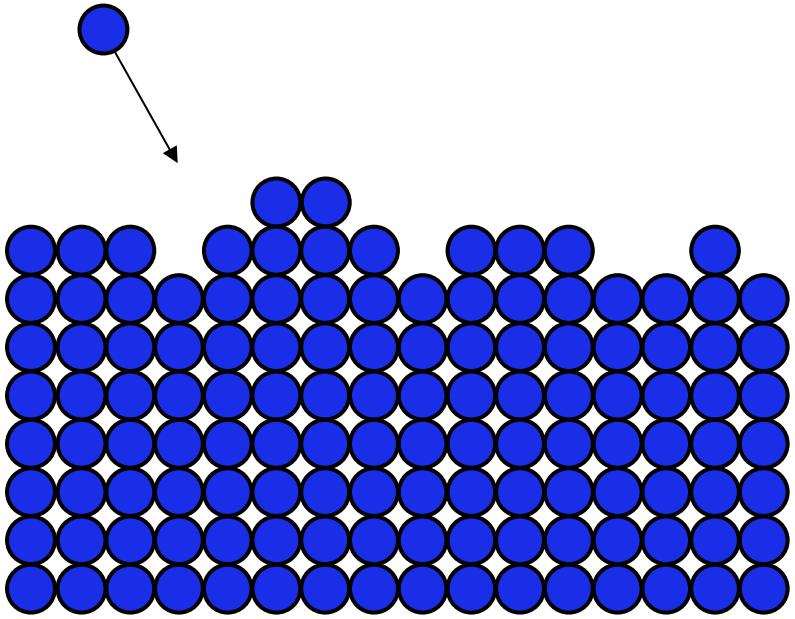
A very brief introduction

Arthur F. Voter
Theoretical Division
Los Alamos National Laboratory
Los Alamos, New Mexico USA

Work supported by
DOE/BES
Los Alamos LDRD program
DOE/ASCR, DOE/SciDAC

Los Alamos

Example: Film or Crystal Growth



Deposition event takes ~2 ps

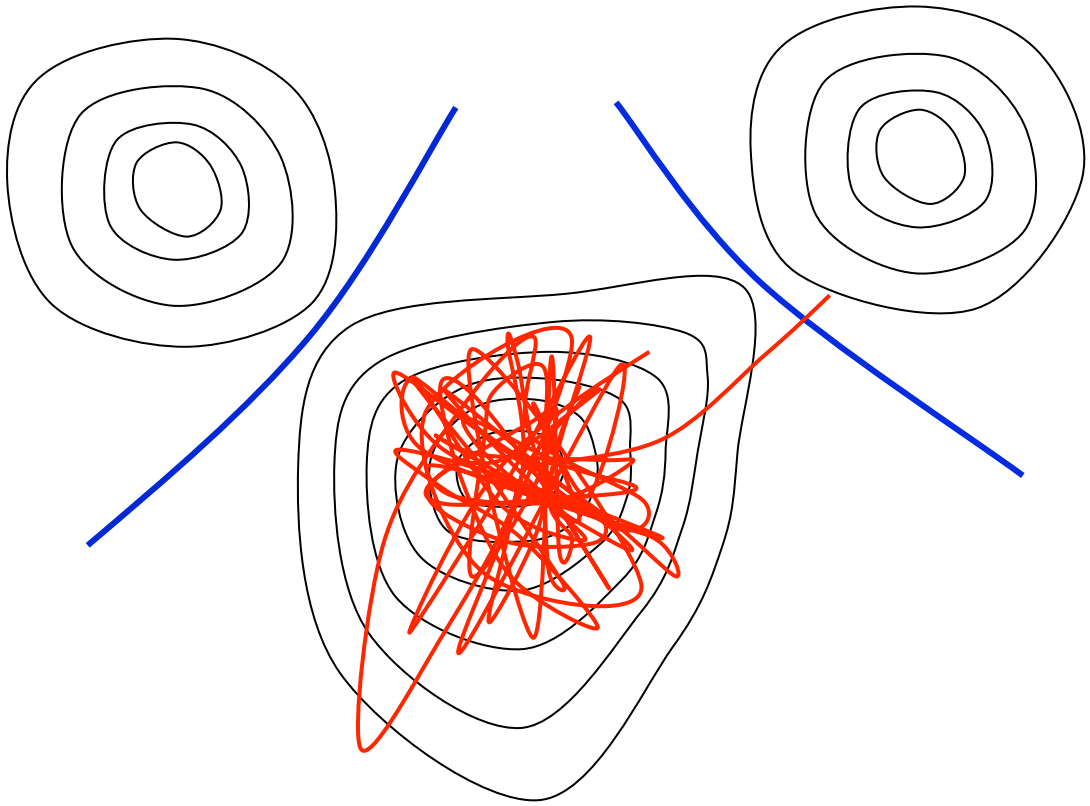
– use molecular dynamics (can reach ns)

Time to next deposition is ~1 s

- diffusion events affect the film morphology
- mechanisms can be surprisingly complex

--> need another approach to treat these

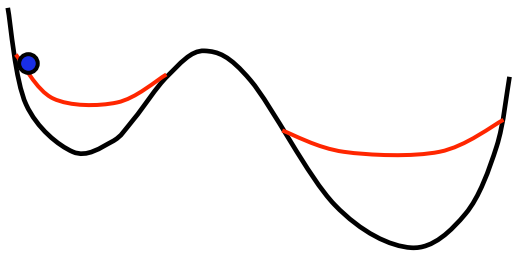
Infrequent-Event System



The system vibrates in 3-N dimensional basin many times before finding an escape path. If we could afford to run molecular dynamics long enough (perhaps millions of vibrations), the trajectory would find an appropriate way out of the state. It is interesting that the trajectory can do this without ever knowing about any of the other possible escape paths.

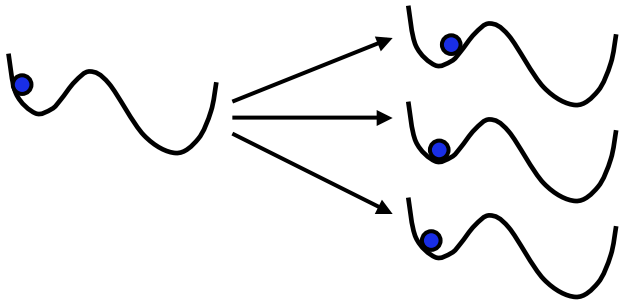
Accelerated Molecular Dynamics Methods

Hyperdynamics



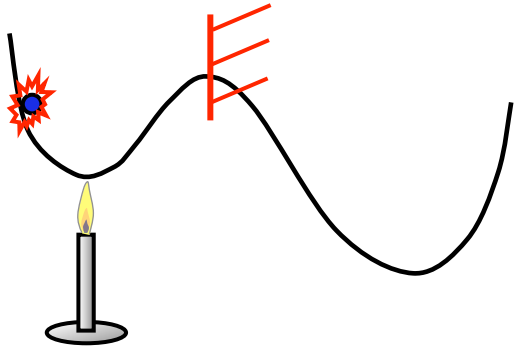
Builds on transition state theory and importance sampling to hasten the escape from each state in a true dynamical way. The boosted time is calculated as the simulation proceeds.
(AFV, J. Chem. Phys., 1997)

Parallel Replica Dynamics



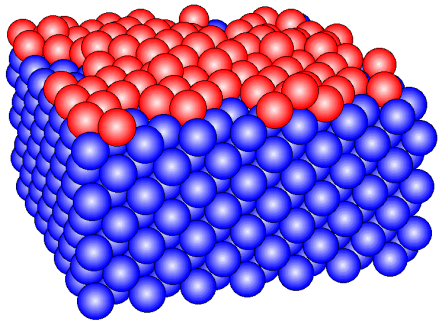
Harnesses parallel power to boost the time scale. Very simple and very general; exact for any infrequent event system obeying exponential escape statistics.
(AFV, Phys. Rev. B, 1998)

Temperature Accelerated Dynamics

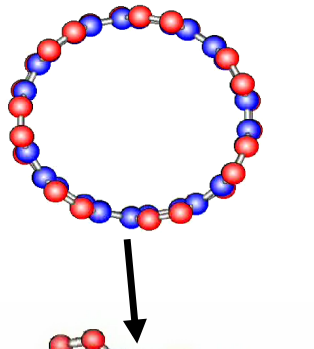


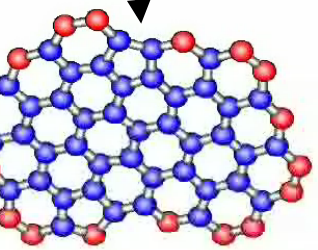
Raise T to make events happen more quickly. Filter out events that should not have happened at correct T. More approximate, but more powerful.
(M.R. Sorensen and AFV, J. Chem. Phys., 2000)

Wide range of systems can be studied

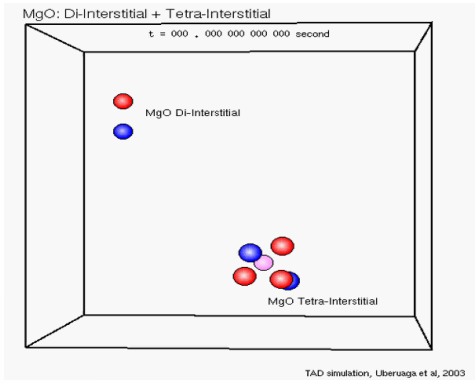


Cu/Ag(100), 1 ML/25 s
T=77K, Sprague et al, 2002

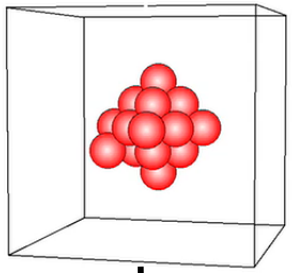


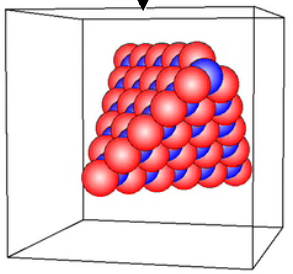


Annealing nanotube slices, μs, Uberuaga et al, 2011.

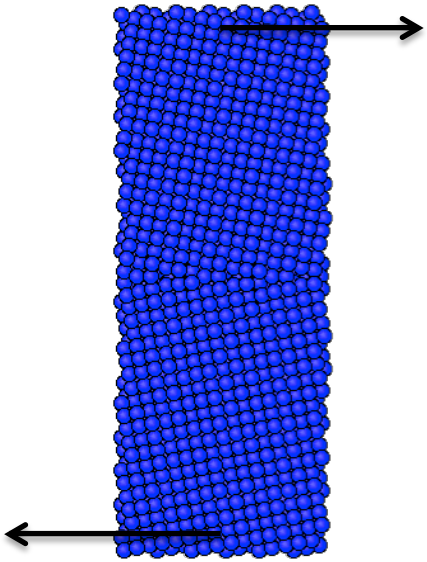


Interstitial defects in MgO, ps – s, Uberuaga et al, 2004.

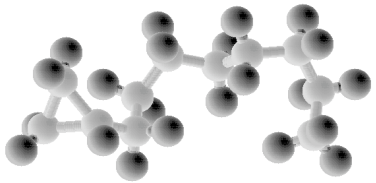




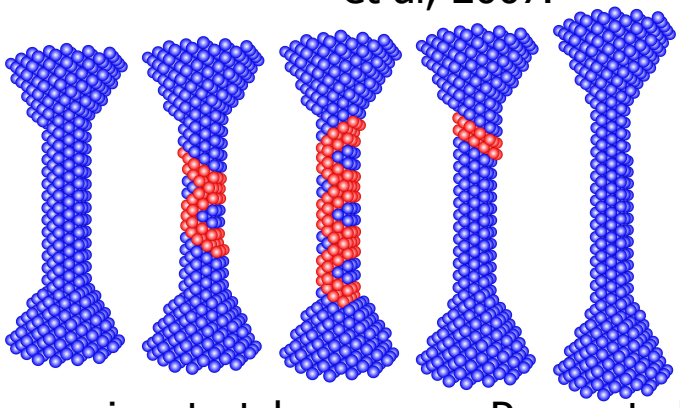
Cu void collapse to SFT, μs, Uberuaga et al, 2007.



Driven Cu GB sliding, 500 μm/s Mishin et al, 2007.



Hexadecane pyrolysis, μs, Kum et al, 2004.



Ag nanowire stretch, μs - ms, Perez et al, TBP.

Recent brief review: Perez et al, Ann. Rep. Comp. Chem. **5**, 79 (2009).

Overview

- Multi Level Monte Carlo (MLMC)
- Hybrid Direct Simulation Monte Carlo (DSMC)
- Accelerated Molecular Dynamics (AMD)
- Multiscale Finite Element Method (MFEM)
- In-Painting for Scientific Computation
- Split Bregman
- Empirical Mode Decomposition (EMD)
- Stochastic Gradient Descent (SGD)

Multiscale Finite Element Method

- PDEs with multiscale solutions.
  - Goal: obtain the large scale solutions, without resolving small scales. Method: construct finite element base functions which capture the small scale information within each element.
  - Small scale information correctly influences large scales global stiffness matrix.
  - Base functions are constructed from the leading order homogeneous elliptic equation in each element.
- Difficulty: resonance can lead to larger errors
  - Resonance is between the small and large scales
  - Solution: Choose BCs for the base function to cancel resonance errors
  - Convergence independent of the small scales

Hou & Wu, JCP 134 (1997) 169-189

Multiscale Finite Element Method (MFEM)

- Steady flow in porous medium
  - Pressure u from rapidly varying conductivity tensor a(x)

$$-\nabla \cdot a(x)\nabla u = f$$

  - Velocity field q is related to the pressure u through Darcy's law:

$$q = -a\nabla u$$

  - MFEM-L, MFEM-O refer to BC choices, MFEM-os refers to oversampling

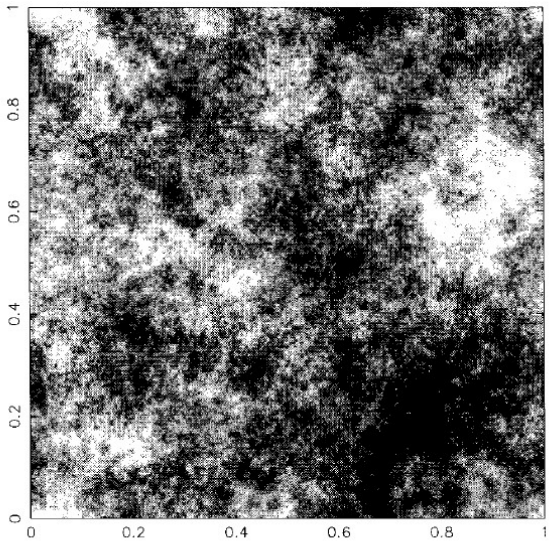


FIG. 5.3. Porosity field with fractal dimension of 2.8 generated using the spectral method.

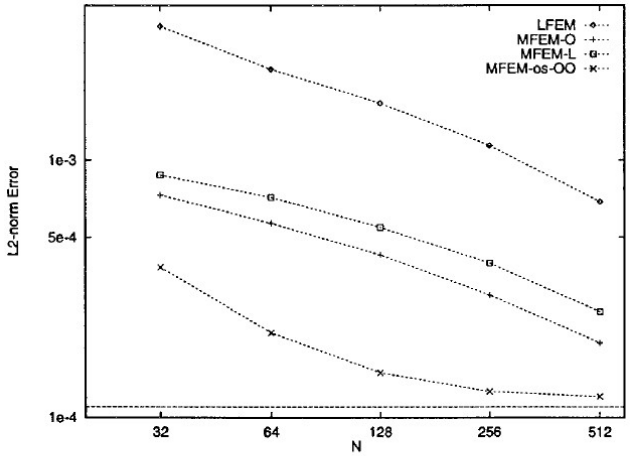


FIG. 5.5. The l^2-norm error of the solutions using various schemes for a fractally distributed permeability field. The horizontal dash line indicates the error of the LFEM solution with $N = 2048$.

Overview

- Multi Level Monte Carlo (MLMC)
- Hybrid Direct Simulation Monte Carlo (DSMC)
- Accelerated Molecular Dynamics (AMD)
- Multiscale Finite Element Method (MFEM)
- In-Painting for Scientific Computation
- Split Bregman
- Empirical Mode Decomposition (EMD)
- Stochastic Gradient Descent (SGD)

In-Painting for Scientific Computation

- Image inpainting
- Extend image to region $D \subset \Omega$ where info is missing or corrupted
  - TV inpainting model: given image u_0 and region D, inpainted image u minimizes

$$V(u \mid u_0, D) = \int_\Omega |\nabla u| dx + \lambda \int_{\Omega \setminus D} (u - u_0)^2 dx$$

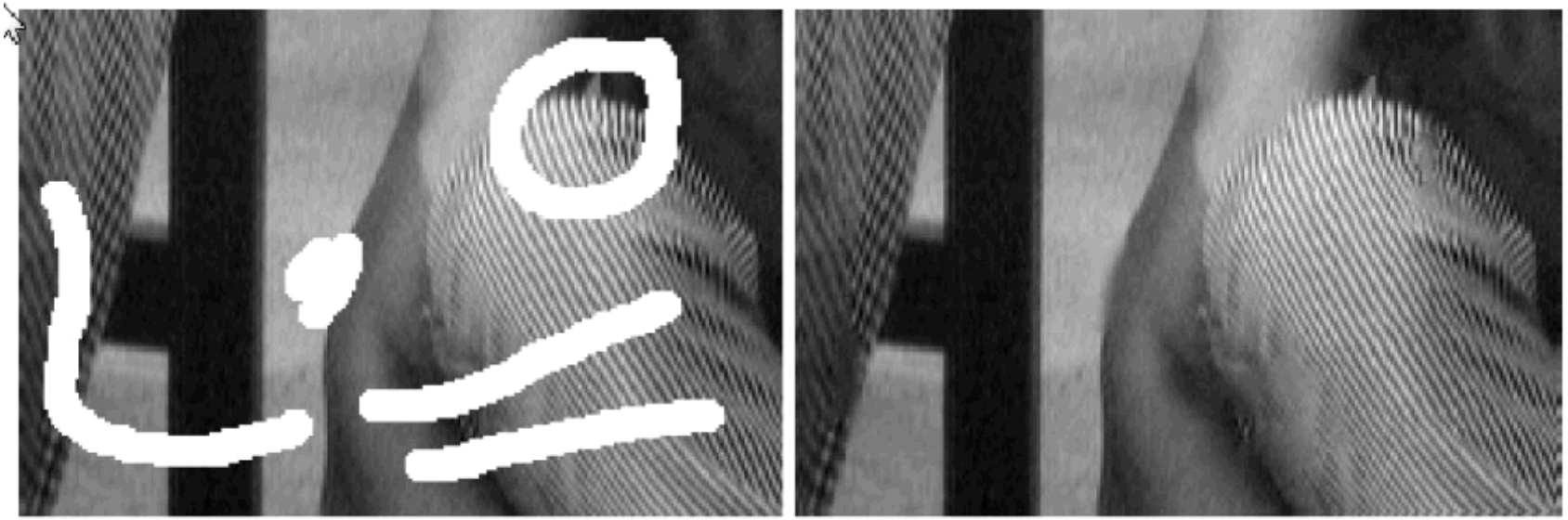
  - T. Chan and J. Shen (2002)

TV Impainting

Chan, Shen (2005)

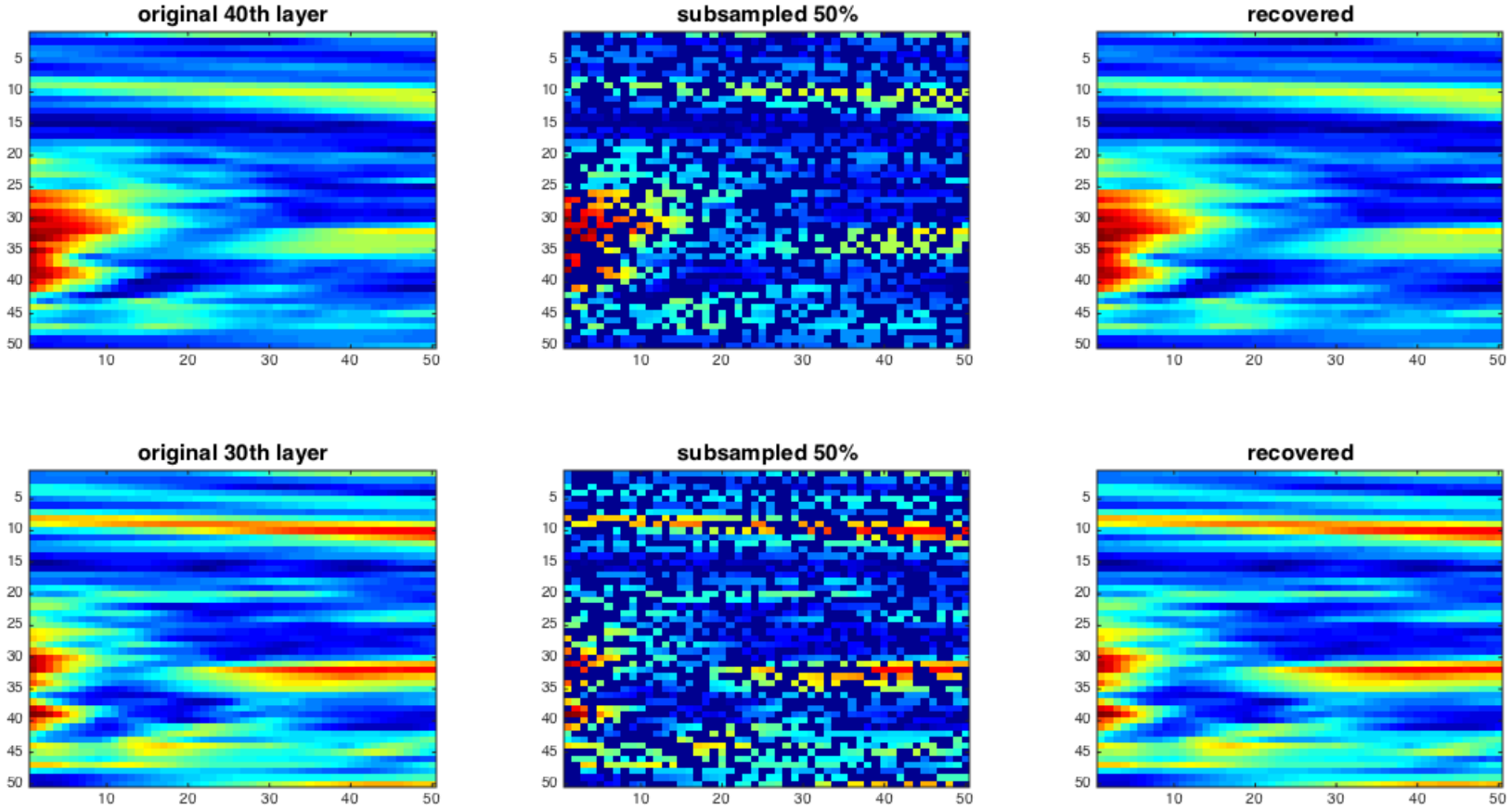
Figure 6.10. *Inpainting for text removal.*

Impainting of Texture



Bertalmio, Vese, Sapiro, Osher (2003)

In-Painting for Plasma Computations



- Doesn't work well for continuation of PDE solution
- Jenko, Osher, Zhu

Overview

- Multi Level Monte Carlo (MLMC)
- Hybrid Direct Simulation Monte Carlo (DSMC)
- Accelerated Molecular Dynamics (AMD)
- Multiscale Finite Element Method (MFEM)
- In-Painting for Scientific Computation
- Split Bregman
- Empirical Mode Decomposition (EMD)
- Stochastic Gradient Descent (SGD)

Split Bregman

- An optimization method for compressed sensing and related fields

$$\min_u \|\Phi(u)\|_1 + H(u)$$

$$\Phi(u) = u$$

$$H(u) = \mu \|Au - f\|_2^2$$

  - Norms are L^1 and L^2, respectively.
  - Difficulty: L^1 term isn't smooth

- Relax the first term by

$$\min_{u,d} \|d\|_1 + H(u) + \lambda \|d - \Phi(u)\|_2^2$$

Split Bregman

- Starting from

$$\min_{u,d} \|d\|_1 + H(u) + \lambda \|d - \Phi(u)\|_2^2$$

- Improve iteration by "feeding back the noise" through term b:

$$(u^{k+1}, d^{k+1}) = \arg\min_{u,d} \|d\|_1 + H(u) + \lambda \|d - \Phi(u) - b^k\|_2^2$$

$$b^{k+1} = b^k + \Phi(u^{k+1}) - d^{k+1}$$

- Split the first line into two pieces to get split Bregman

$$u^{k+1} = \arg\min_u H(u) + \lambda \|d^k - \Phi(u) - b^k\|_2^2$$

$$d^{k+1} = \arg\min_d \|d\|_1 + \lambda \|d - \Phi(u^{k+1}) - b^k\|_2^2$$

$$b^{k+1} = b^k + \Phi(u^{k+1}) - d^{k+1}$$

  - Easily solved: u problem is smooth, d problem is soft-thresholding
  - Widely used for problems involving sparsity

Split Bregman Results

- Convergence results

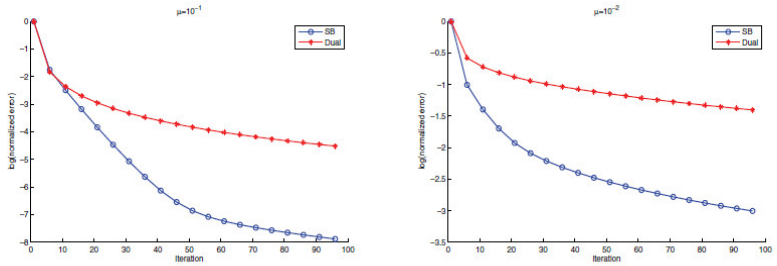


Figure 1. *Error vs. iteration number for the split Bregman ROF minimization algorithm and for an iterative method based on the dual problem [8]. The error at iteration k is defined as $\|u^k - u^*\|$, where u^k is the approximation at iteration k, and u^* is the exact solution. Convergence results are for the test image "Lena" with Gaussian noise ($\sigma = 15$). Left: Results for $\mu = 0.1$. Right: Results for $\mu = 0.01$.*

- Reconstruction results

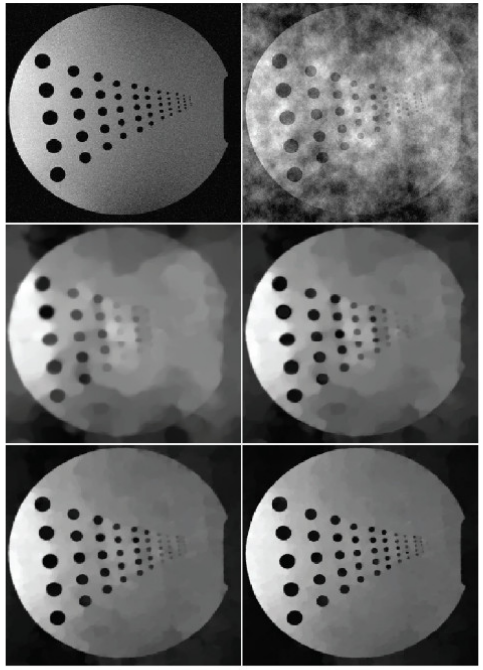


Figure 5. *Split Bregman CS reconstruction of an MR image using 30% of the k-space data. Top left: Original image reconstructed using the full k-space. Top right: Image reconstructed using 30% of k-space data. Reconstruction was done using the conventional method, which fills in missing samples with zeros. Middle left: Results of split Bregman algorithm after 10 inner iterations (2 outer iterations). Middle right: Results of split Bregman algorithm after 20 inner iterations (4 outer iterations). Bottom left: Results after 30 inner iterations. Bottom right: Optimal results obtained after 40 inner iterations.*

Overview

- Multi Level Monte Carlo (MLMC)
- Hybrid Direct Simulation Monte Carlo (DSMC)
- Accelerated Molecular Dynamics (AMD)
- Multiscale Finite Element Method (MFEM)
- In-Painting for Scientific Computation
- Split Bregman
- Empirical Mode Decomposition (EMD)
- Stochastic Gradient Descent (SGD)

Empirical Mode Decomposition (EMD)

- Adaptive data analysis method
  - Determine trend and instantaneous frequency of time series f(t)
  - Find sparsest representation of f(t) within a dictionary of intrinsic mode functions (IMFs)
- Dictionary construction

$$D = \{a(t)\cos\theta(t) : \theta'(t) \geq 0,\ a(t) \textit{ smoother than } \cos\theta(t)\}$$

$$\min M \ \ s.t.\ f(t) = \sum_{k=1}^{M} a_k(t)\cos\theta_k(t), \qquad a_k\cos\theta_k \in D$$

- Smoothness measured through TV^3 norm

$$TV^3(a) = \int \left| a^{(4)}(t) \right| dt$$

Huang, Proc Roy Soc, 1989
Hou & Shi, Adv Adaptive Data Anal, 2011

Empirical Mode Decomposition

- Example

$$f(t) = 6t + \cos(8\pi t) + 0.5\cos(40\pi t)$$

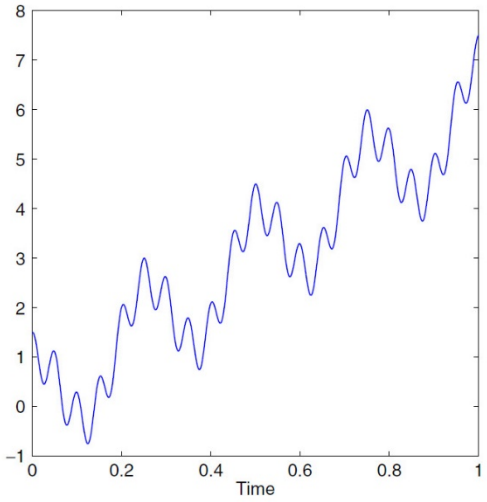


Fig. 1. Original data in Example 1.

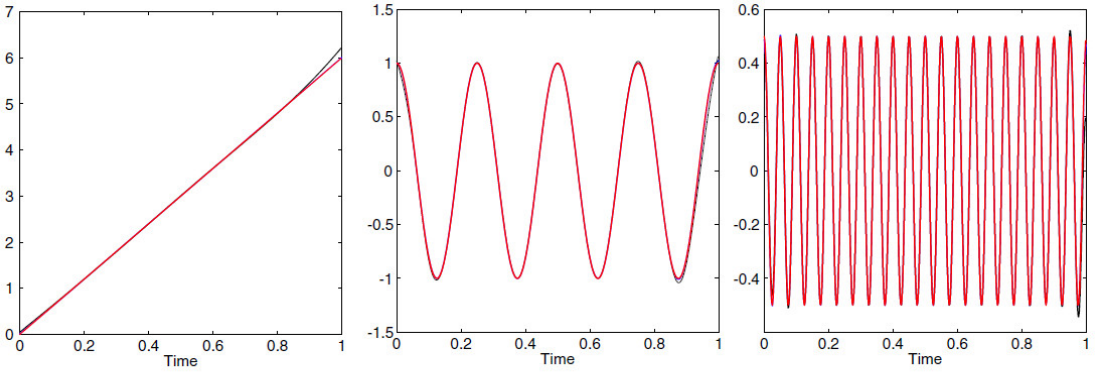


Fig. 2. IMFs and trend in Example 1. Red: analytical results; Blue: our method; and Black: EMD method.

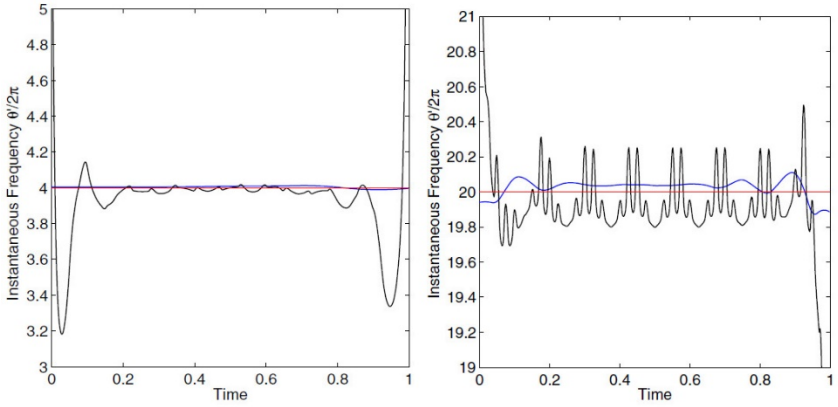


Fig. 3. Instantaneous frequency in Example 1. Red: analytical results; Blue: our method; and Black: EMD method.

- IMFs captured almost exactly
- Frequencies captured very well

Overview

- Multi Level Monte Carlo (MLMC)
- Hybrid Direct Simulation Monte Carlo (DSMC)
- Accelerated Molecular Dynamics (AMD)
- Multiscale Finite Element Method (MFEM)
- In-Painting for Scientific Computation
- Split Bregman
- Empirical Mode Decomposition (EMD)
- Stochastic Gradient Descent (SGD)

Stochastic Gradient Descent

- Optimization for functions of the form

$$Q(w) = \sum_{i=1}^{n} Q_i(w)$$

  - Widely used for machine learning and internet computations
  - Each i represents piece of data used for training of a learning method
- Randomly choose data for ith step
  - Perform gradient descent with step size η

$$w := w - \eta \nabla Q_i(w)$$

- Many variants
  - Series of batches with η constant within batch, decreasing between batches

Wikipedia

Stochastic Gradient Descent

- Convergence of mini-batch method