Random Batch Methods for the Efficient Solution of Optimal Control Problems on Networks

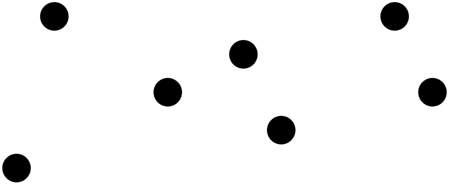
Daniël Veldman
(joint work with Enrique Zuazua)

Chair in Dynamics, Control, and Numerics
Friedrich-Alexander Universität Erlangen-Nürnberg

August 24, 2022

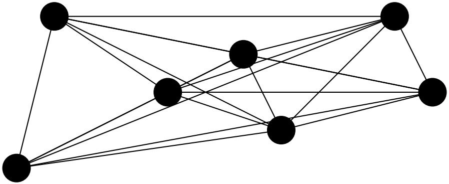
Origins of the Random Batch Method

Initial Motivation: Simulation and control of large *interacting particle systems* can be computationally demanding.



Origins of the Random Batch Method

Initial Motivation: Simulation and control of large *interacting particle systems* can be computationally demanding.

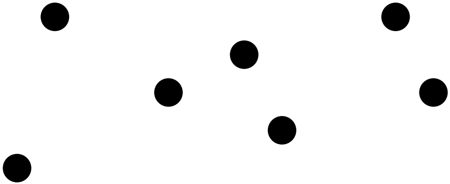


There are $N(N-1)/2$ interaction forces between N particles.
$\Rightarrow$ Computational cost grows rapidly when N is large.

Origins of the Random Batch Method

Proposed simulation method: The Random Batch Method

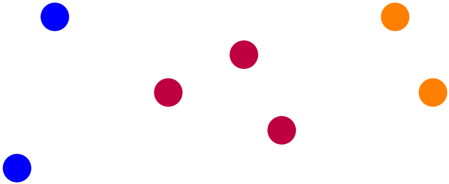
[Shi Jin, Lei Li, Jian-Guo Liu, J. of Computational Physics 400, 108877, 2020]



Origins of the Random Batch Method

Proposed simulation method: The Random Batch Method

[Shi Jin, Lei Li, Jian-Guo Liu, J. of Computational Physics 400, 108877, 2020]

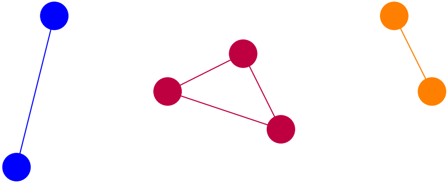


- Divide the N particles randomly into batches of size $P \geq 2$.

Origins of the Random Batch Method

Proposed simulation method: The Random Batch Method

[Shi Jin, Lei Li, Jian-Guo Liu, J. of Computational Physics 400, 108877, 2020]

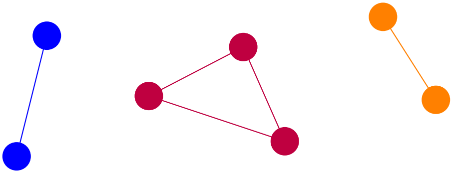


- Divide the N particles randomly into batches of size $P \geq 2$.
- Consider only interactions between particles in the same batch.

Origins of the Random Batch Method

Proposed simulation method: The Random Batch Method

[Shi Jin, Lei Li, Jian-Guo Liu, J. of Computational Physics 400, 108877, 2020]

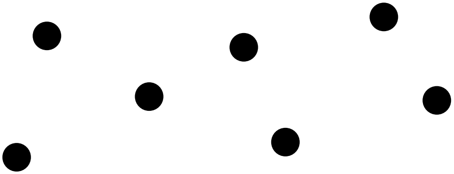


- Divide the N particles randomly into batches of size $P \geq 2$.
- Consider only interactions between particles in the same batch.
- Do a simulation over a short time interval of length h.

Origins of the Random Batch Method

Proposed simulation method: The Random Batch Method

[Shi Jin, Lei Li, Jian-Guo Liu, J. of Computational Physics 400, 108877, 2020]

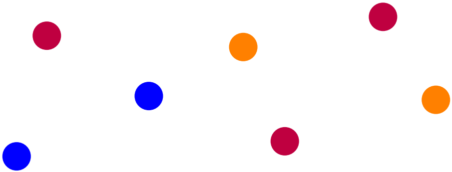


- Divide the N particles randomly into batches of size $P \geq 2$.
- Consider only interactions between particles in the same batch.
- Do a simulation over a short time interval of length h.
- Repeat.

Origins of the Random Batch Method

Proposed simulation method: The Random Batch Method

[Shi Jin, Lei Li, Jian-Guo Liu, J. of Computational Physics 400, 108877, 2020]

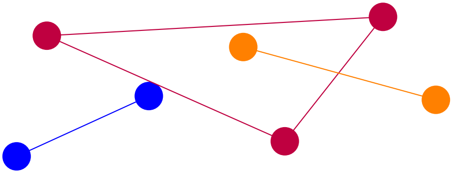


- Divide the N particles randomly into batches of size $P \geq 2$.
- Consider only interactions between particles in the same batch.
- Do a simulation over a short time interval of length h.
- Repeat.

Origins of the Random Batch Method

Proposed simulation method: The Random Batch Method

[Shi Jin, Lei Li, Jian-Guo Liu, J. of Computational Physics 400, 108877, 2020]



- Divide the N particles randomly into batches of size $P \geq 2$.
- Consider only interactions between particles in the same batch.
- Do a simulation over a short time interval of length h.
- Repeat.

In formulas

First-order particle dynamics: For each particle $i \in \{1, 2, \dots, N\}$

$$\dot{x}_i(t) = \frac{1}{N-1} \sum_{\substack{j=1 \\ j \neq i}}^{N} f_{ij}(x_j(t) - x_i(t)).$$

Random Batch Method.

1. Set $k = 0$ and set $\tilde{x}_i(0)$ for all $i \in \{1, 2, \dots, N\}$.
2. Partition $\{1, 2, \dots, N\}$ into batches of size $P \geq 2$, i.e.

$$\{1, 2, \dots, N\} = \dot{\bigcup}_{b=1}^{N/P} \mathcal{B}_b^k, \qquad |\mathcal{B}_b^k| = P.$$

In formulas

First-order particle dynamics: For each particle $i \in \{1, 2, \dots, N\}$

$$\dot{x}_i(t) = \frac{1}{N-1} \sum_{\substack{j=1 \\ j \neq i}}^{N} f_{ij}(x_j(t) - x_i(t)).$$

Random Batch Method.

1. Set $k = 0$ and set $\tilde{x}_i(0)$ for all $i \in \{1, 2, \dots, N\}$.
2. Partition $\{1, 2, \dots, N\}$ into batches of size $P \geq 2$, i.e.

$$\{1, 2, \dots, N\} = \dot{\bigcup}_{b=1}^{N/P} \mathcal{B}_b^k, \qquad |\mathcal{B}_b^k| = P.$$

3. For each i, solve on $[kh, (k+1)h]$

$$\dot{x}_{h,i}(t) = \frac{1}{P-1} \sum_{j \in \mathcal{B}_{b(i)}^k,\ j \neq i} f_{ij}(x_{h,j}(t) - x_{h,i}(t)), \qquad b(i) \text{ s.t. } i \in \mathcal{B}_{b(i)}^k.$$

4. Set $k \leftarrow k + 1$ and go to step 2.

The Random Batch Method in Optimal Control

Instead of computing the minimizer $u^*(t)$ of

$$J = \int_0^T f_0(x(t), u(t)) \,\mathrm{d}t,$$

subject to

$$\dot{x}_i(t) = \frac{1}{N-1} \sum_{\substack{j=1 \\ j \neq i}}^{N} f_{ij}(x_j(t) - x_i(t)) + \sum_{k=1}^{M} g_{ik}(x_i(t)) u_k(t),$$

it is faster to compute the minimizer $u_h^*(t)$ of J subject to

$$\dot{x}_{h,i}(t) = \frac{1}{P-1} \sum_{\substack{j \in \mathcal{B}_{b(i)}^k \\ j \neq i}} f_{ij}(x_{h,j}(t) - x_{h,i}(t)) + \sum_{k=1}^{M} g_{ik}(x_i(t)) u_{h,k}(t).$$

[D. Ko, E. Zuazua, Math. Models Methods Appl. Sci., Vol. 31, No. 8, 2021].

The RBM in linear-quadratic optimal control (1/2)

RBM to approximate the minimizer $u^*(t)$ of

$$\min_{u \in L^2(0,T;\mathbb{R}^q)} J(u) = \int_0^T \left(|x(t) - x_d(t)|_Q^2 + |u(t)|_R^2 \right) dt,$$

$$\dot{x}(t) = Ax(t) + Bu(t), \qquad x(0) = x_0.$$

The RBM in linear-quadratic optimal control (1/2)

RBM to approximate the minimizer $u^*(t)$ of

$$\min_{u \in L^2(0,T;\mathbb{R}^q)} J(u) = \int_0^T \left(|x(t) - x_d(t)|_Q^2 + |u(t)|_R^2 \right) dt,$$

$$\dot{x}(t) = Ax(t) + Bu(t), \qquad x(0) = x_0.$$

Step 1 Split the matrix A as

$$A = \sum_{m=1}^{M} A_m.$$

The RBM in linear-quadratic optimal control (1/2)

RBM to approximate the minimizer $u^*(t)$ of

$$\min_{u\in L^2(0,T;\mathbb{R}^q)} J(u) = \int_0^T \left(|x(t)-x_d(t)|_Q^2 + |u(t)|_R^2\right) dt,$$

$$\dot{x}(t) = Ax(t) + Bu(t), \qquad x(0) = x_0.$$

Step 1 Split the matrix A as

$$A = \sum_{m=1}^{M} A_m.$$

Step 2 Enumerate the 2^M subsets of $\{1,2,\ldots,M\}$ as $S_1, S_2, \ldots S_{2^M}$. Assign to each subset S_ω a probability p_ω.

The RBM in linear-quadratic optimal control (1/2)

RBM to approximate the minimizer $u^*(t)$ of

$$\min_{u \in L^2(0,T;\mathbb{R}^q)} J(u) = \int_0^T \left(|x(t) - x_d(t)|_Q^2 + |u(t)|_R^2 \right) dt,$$

$$\dot{x}(t) = Ax(t) + Bu(t), \qquad x(0) = x_0.$$

Step 1 Split the matrix A as

$$A = \sum_{m=1}^{M} A_m.$$

Step 2 Enumerate the 2^M subsets of $\{1, 2, \ldots, M\}$ as $S_1, S_2, \ldots S_{2^M}$. Assign to each subset S_ω a probability p_ω.

Step 3 Divide $[0, T]$ into K subintervals $[t_{k-1}, t_k)$ of length $\leq h$. For each $[t_{k-1}, t_k)$, randomly choose an index $\omega_k \in \{1, 2, \ldots, 2^M\}$ according to the probabilities p_ℓ. Set $\boldsymbol{\omega} = (\omega_1, \omega_2, \ldots, \omega_K)$.

The RBM in linear-quadratic optimal control (2/2)

Step 4 Define the matrix $A_h(\boldsymbol{\omega}, t)$

$$A_h(\boldsymbol{\omega}, t) = \sum_{m \in S_{\omega_k}} \frac{A_m}{\pi_m}, \qquad t \in [t_{k-1}, t_k),$$

where π_m is the probability that m is an element of the selected subset, i.e.

$$\pi_m = \sum_{\omega \in \{\omega | m \in S_\omega\}} p_\omega.$$

The RBM in linear-quadratic optimal control (2/2)

Step 4 Define the matrix $A_h(\omega, t)$

$$A_h(\omega, t) = \sum_{m \in S_{\omega_k}} \frac{A_m}{\pi_m}, \qquad t \in [t_{k-1}, t_k),$$

where π_m is the probability that m is an element of the selected subset, i.e.

$$\pi_m = \sum_{\omega \in \{\omega | m \in S_\omega\}} p_\omega.$$

Step 5 Compute the minimizer $u_h^*(\omega, t)$ of the 'simpler' LQR problem

$$\min_{u \in L^2(0,T;\mathbb{R}^q)} J_h(\omega, u) = \int_0^T \left(|x_h(\omega, t) - x_d(t)|_Q^2 + |u(t)|_R^2 \right) dt,$$

$$\dot{x}_h(\omega, t) = A_h(\omega, t) x_h(\omega, t) + Bu(t), \qquad x_h(\omega, 0) = x_0.$$

An example

Split A into $M = 2$ pieces as

$$A = A_1 + A_2.$$

An example

Split A into $M = 2$ pieces as

$$A = A_1 + A_2.$$

Then there are $2^M = 4$ subsets S_ω, $\omega \in \{1,2,3,4\}$.

$$S_1 = \{1\}, \qquad S_2 = \{2\}, \qquad S_3 = \{1,2\}, \qquad S_4 = \emptyset.$$

An example

Split A into $M = 2$ pieces as

$$A = A_1 + A_2.$$

Then there are $2^M = 4$ subsets S_ω, $\omega \in \{1, 2, 3, 4\}$.

$$S_1 = \{1\}, \qquad S_2 = \{2\}, \qquad S_3 = \{1, 2\}, \qquad S_4 = \emptyset.$$

We assign the probabilities p_ω, $\omega \in \{1, 2, 3, 4\}$.

$$p_1 = \tfrac{1}{2}, \qquad p_2 = \tfrac{1}{2}, \qquad p_3 = 0, \qquad p_4 = 0.$$

An example

Split A into $M = 2$ pieces as

$$A = A_1 + A_2.$$

Then there are $2^M = 4$ subsets S_ω, $\omega \in \{1,2,3,4\}$.

$$S_1 = \{1\}, \qquad S_2 = \{2\}, \qquad S_3 = \{1,2\}, \qquad S_4 = \emptyset.$$

We assign the probabilities p_ω, $\omega \in \{1,2,3,4\}$.

$$p_1 = \tfrac{1}{2}, \qquad p_2 = \tfrac{1}{2}, \qquad p_3 = 0, \qquad p_4 = 0.$$

The probabilities π_m, $m \in \{1,2\}$ are then

$$\pi_1 = p_1 + p_3 = \tfrac{1}{2}, \qquad \pi_2 = p_2 + p_3 = \tfrac{1}{2}.$$

An example

Split A into $M = 2$ pieces as

$$A = A_1 + A_2.$$

Then there are $2^M = 4$ subsets S_ω, $\omega \in \{1, 2, 3, 4\}$.

$$S_1 = \{1\}, \qquad S_2 = \{2\}, \qquad S_3 = \{1, 2\}, \qquad S_4 = \emptyset.$$

We assign the probabilities p_ω, $\omega \in \{1, 2, 3, 4\}$.

$$p_1 = \tfrac{1}{2}, \qquad p_2 = \tfrac{1}{2}, \qquad p_3 = 0, \qquad p_4 = 0.$$

The probabilities π_m, $m \in \{1, 2\}$ are then

$$\pi_1 = p_1 + p_3 = \tfrac{1}{2}, \qquad \pi_2 = p_2 + p_3 = \tfrac{1}{2}.$$

The randomized matrix A-matrix is thus given by

$$A_h(\boldsymbol{\omega}, t) = 2A_{\omega_k}, \qquad t \in [t_{k-1}, t_k), \omega_k \in \{1, 2\}.$$

An intuitive convergence argument

In addition to the previous example, assume that

- A_1 and A_2 commute,
- that the time grid is uniform, i.e., $t_k = kh$, where $h = T/K$.

Then,

$$\begin{aligned} e^{\int_0^T A_h(\boldsymbol{\omega},t)\,\mathrm{d}t} &= e^{2A_{\omega_1}h} e^{2A_{\omega_2}h} e^{2A_{\omega_3}h} \cdots e^{2A_{\omega_K}h} \\ &= e^{2A_1 T_1 + 2A_2 T_2} \end{aligned}$$

where

$$T_1 = h\sum_{i=1}^{K} \chi_1(\omega_i), \qquad T_2 = h\sum_{i=1}^{K} \chi_2(\omega_i).$$

By the law of large numbers, T_1, $T_2 \to T/2$ for $K \to \infty$.

Properties of the matrix $A_h(\boldsymbol{\omega}, t)$

- The construction of $A_h(\boldsymbol{\omega}, t)$ assures that

$$\mathbb{E}[A_h(t)] = A.$$

Proof: For $t \in [t_{k-1}, t_k)$, $A_h(\boldsymbol{\omega}, t) = \sum_{m \in S_{\omega_k}} \frac{A_m}{\pi_m}$ only depends on ω_k. Therefore,

$$\mathbb{E}[A_h(t)] = \sum_{\omega=1}^{2^M} \sum_{m \in S_\omega} \frac{A_m}{\pi_m} p_\omega = \sum_{m=1}^{M} \sum_{\omega \in \{\omega | m \in S_\omega\}} \frac{A_m}{\pi_m} p_\omega$$

$$= \sum_{m=1}^{M} \frac{A_m}{\pi_m} \sum_{\omega \in \{\omega | m \in S_\omega\}} p_\omega = \sum_{m=1}^{M} \frac{A_m}{\pi_m} \pi_m = \sum_{m=1}^{M} A_m = A.$$

Properties of the matrix $A_h(\omega, t)$

- The construction of $A_h(\omega, t)$ assures that

$$\mathbb{E}[A_h(t)] = A.$$

Proof: For $t \in [t_{k-1}, t_k)$, $A_h(\omega, t) = \sum_{m \in S_{\omega_k}} \frac{A_m}{\pi_m}$ only depends on ω_k. Therefore,

$$\mathbb{E}[A_h(t)] = \sum_{\omega=1}^{2^M} \sum_{m \in S_\omega} \frac{A_m}{\pi_m} p_\omega = \sum_{m=1}^{M} \sum_{\omega \in \{\omega | m \in S_\omega\}} \frac{A_m}{\pi_m} p_\omega$$

$$= \sum_{m=1}^{M} \frac{A_m}{\pi_m} \sum_{\omega \in \{\omega | m \in S_\omega\}} p_\omega = \sum_{m=1}^{M} \frac{A_m}{\pi_m} \pi_m = \sum_{m=1}^{M} A_m = A.$$

- We also define

$$\mathrm{Var}[A_h] = \mathbb{E}[\|A_h(t) - A\|^2] = \sum_{\omega=1}^{2^M} \left\| \sum_{m \in S_\omega} \frac{A_m}{\pi_m} - A \right\|^2 p_\omega.$$

Theorem (D.V., E. Zuazua, Numerische Mathematik, 2022)

- *For a deterministic control $u(t)$*

$$\mathbb{E}[|x_h(t) - x(t)|^2] \leq h\mathrm{Var}[A_h](\|A\|t^2 + 2t)(|x_0| + |Bu|_{L^1})^2.$$

- *For a stochastic control $u_h(\omega, t)$ satisfying $|Bu_h(\omega)|_{L^2} \leq U$*

$$\mathbb{E}[|x_h(t) - x(t)|^2] \leq C_{[T, \|A\|]} h\mathrm{Var}[A_h](x_0 + U\sqrt{t}).$$

- *Optimality gap*

$$\mathbb{E}[|J_h(u_h^*) - J(u^*)|] \leq C\left(\sqrt{h\mathrm{Var}[A_h]} + h\mathrm{Var}[A_h]\right).$$

- *Convergence in the controls*

$$\mathbb{E}[|u_h^* - u^*|_{L^2}^2] \leq Ch\mathrm{Var}[A_h].$$

Structure of the proof (1/2)

1 We give a bound for $\mathbb{E}[|x_h(t) - x(t)|^2]$ where

$$\dot{x}_h(\omega, t) = A_h(\omega, t) x_h(\omega, t) + Bu(t), \qquad x_h(\omega, 0) = x_0,$$
$$\dot{x}(t) = Ax(t) + Bu(t), \qquad x(0) = x_0.$$

Note: the control $u(t)$ is deterministic (independent of ω).

Structure of the proof (1/2)

1 We give a bound for $\mathbb{E}[|x_h(t) - x(t)|^2]$ where

$$\begin{aligned} \dot{x}_h(\omega, t) &= A_h(\omega, t)x_h(\omega, t) + Bu(t), & x_h(\omega, 0) &= x_0, \\ \dot{x}(t) &= Ax(t) + Bu(t), & x(0) &= x_0. \end{aligned}$$

Note: the control $u(t)$ is deterministic (independent of ω).

2 We give a bound for $\mathbb{E}[|x_h(t) - x(t)|^2]$ where

$$\begin{aligned} \dot{x}_h(\omega, t) &= A_h(\omega, t)x_h(\omega, t) + Bu_h(\omega, t), & x_h(\omega, 0) &= x_0, \\ \dot{x}(\omega, t) &= Ax(\omega, t) + Bu_h(\omega, t), & x(\omega, 0) &= x_0, \end{aligned}$$

Note: the control $u(\omega, t)$ is stochastic.

Structure of the proof (1/2)

1 We give a bound for $\mathbb{E}[|x_h(t) - x(t)|^2]$ where

$$\begin{aligned} \dot{x}_h(\omega, t) &= A_h(\omega, t) x_h(\omega, t) + Bu(t), & x_h(\omega, 0) &= x_0, \\ \dot{x}(t) &= Ax(t) + Bu(t), & x(0) &= x_0. \end{aligned}$$

Note: the control $u(t)$ is deterministic (independent of ω).

2 We give a bound for $\mathbb{E}[|x_h(t) - x(t)|^2]$ where

$$\begin{aligned} \dot{x}_h(\omega, t) &= A_h(\omega, t) x_h(\omega, t) + Bu_h(\omega, t), & x_h(\omega, 0) &= x_0, \\ \dot{x}(\omega, t) &= Ax(\omega, t) + Bu_h(\omega, t), & x(\omega, 0) &= x_0, \end{aligned}$$

Note: the control $u(\omega, t)$ is stochastic.

3a Using 2 we obtain a bound for $\mathbb{E}[|J_h(u_h) - J(u_h)|]$ for any stochastic control u_h. **A Γ-convergence argument** then gives

$$\mathbb{E}[|J_h(u_h^*) - J(u^*)|].$$

Structure of the proof (2/2)

3b Using 2 we obtain a bound for

$$\mathbb{E}[|\delta J_h(u; v_h) - \delta J(u, v_h)|] \leq C\sqrt{h\mathrm{Var}[A_h]\mathbb{E}[|v_h|^2_{L^2(0,T;\mathbb{R}^q)}]},$$

for any deterministic control u and any stochastic control v_h.

Setting $u(t) = u^*(t)$ and $v_h(\omega, t) = u_h^*(\omega, t) - u^*(t)$, the **convexity** of the functional $J_h(\omega, \cdot)$ yields a bound for

$$\mathbb{E}[|u_h^* - u^*|^2_{L^2(0,T;\mathbb{R}^q)}].$$

Structure of the proof (2/2)

3b Using 2 we obtain a bound for

$$\mathbb{E}[|\delta J_h(u; v_h) - \delta J(u, v_h)|] \leq C\sqrt{h\mathrm{Var}[A_h]\mathbb{E}[|v_h|^2_{L^2(0,T;\mathbb{R}^q)}]},$$

for any deterministic control u and any stochastic control v_h.

Setting $u(t) = u^*(t)$ and $v_h(\omega, t) = u_h^*(\omega, t) - u^*(t)$, the **convexity** of the functional $J_h(\omega, \cdot)$ yields a bound for

$$\mathbb{E}[|u_h^* - u^*|^2_{L^2(0,T;\mathbb{R}^q)}].$$

As a consequence, we also get a bound for

$$\mathbb{E}[|J(u_h^*) - J(u^*)|].$$

Sketch of proof for Step 1 (1/4)

$$\begin{aligned} \dot{x}_h(\omega,t) &= A_h(\omega,t)x_h(\omega,t) + Bu(t), & x_h(\omega,0) &= x_0, \\ \dot{x}(t) &= Ax(t) + Bu(t), & x(0) &= x_0. \end{aligned}$$

Sketch of proof for Step 1 (1/4)

$$\begin{aligned} \dot{x}_h(\omega,t) &= A_h(\omega,t)x_h(\omega,t) + Bu(t), & x_h(\omega,0) &= x_0,\\ \dot{x}(t) &= Ax(t) + Bu(t), & x(0) &= x_0. \end{aligned}$$

Define $e_h(\omega,t) := x_h(\omega,t) - x(t)$. Then

$$\begin{aligned} \dot{e}_h(\omega,t) &= A_h(\omega,t)x_h(\omega,t) - Ax(t)\\ &= A_h(\omega,t)e_h(\omega,t) + (A_h(\omega,t) - A)x(t), \quad e_h(\omega,0) = 0. \end{aligned}$$

Sketch of proof for Step 1 (1/4)

$$\dot{x}_h(\omega,t) = A_h(\omega,t)x_h(\omega,t) + Bu(t), \qquad x_h(\omega,0) = x_0,$$
$$\dot{x}(t) = Ax(t) + Bu(t), \qquad x(0) = x_0.$$

Define $e_h(\omega,t) := x_h(\omega,t) - x(t)$. Then

$$\begin{aligned}\dot{e}_h(\omega,t) &= A_h(\omega,t)x_h(\omega,t) - Ax(t)\\ &= A_h(\omega,t)e_h(\omega,t) + (A_h(\omega,t) - A)x(t), \quad e_h(\omega,0) = 0.\end{aligned}$$

and also (for $t \in [t_{k-1}, t_k)$)

$$\begin{aligned}\frac{d}{dt}|e_h(\omega,t)|^2 = e_h(\omega,t)^\top A_h(\omega,t)e_h(\omega,t)\\ + e_h(\omega,t_k)^\top (A_h(\omega,t) - A)x(t)\\ + (e_h(\omega,t) - e_h(\omega,t_k))^\top (A_h(\omega,t) - A)x(t).\end{aligned}$$

Sketch of proof for Step 1 (2/4)

$$\begin{aligned}\frac{d}{dt}|e_h(\omega,t)|^2 = {} & e_h(\omega,t)^\top A_h(\omega,t) e_h(\omega,t) \\ & + e_h(\omega,t_k)^\top (A_h(\omega,t) - A)x(t) \\ & + (e_h(\omega,t) - e_h(\omega,t_k))^\top (A_h(\omega,t) - A)x(t).\end{aligned}$$

Sketch of proof for Step 1 (2/4)

$$\begin{aligned}\frac{d}{dt}|e_h(\omega,t)|^2 = e_h(\omega,t)^\top A_h(\omega,t)e_h(\omega,t) \\ + e_h(\omega,t_k)^\top (A_h(\omega,t)-A)x(t) \\ + (e_h(\omega,t)-e_h(\omega,t_k))^\top (A_h(\omega,t)-A)x(t).\end{aligned}$$

The **first term** is bounded by a (quasi-)dissipative assumption:

- Assumption: the matrices A_m are dissipative, i.e.

$$\langle x, A_m x\rangle \leq 0, \qquad \forall m \in \{1,2,\dots,M\}, x \in \mathbb{R}^N.$$

Then $A_h(\omega,t) = \sum_{m\in S_{\omega_k}} \frac{1}{\pi_m} A_m$ is also dissipative.

Sketch of proof for Step 1 (2/4)

$$\begin{aligned}\frac{d}{dt}|e_h(\boldsymbol{\omega},t)|^2 = e_h(\boldsymbol{\omega},t)^\top A_h(\boldsymbol{\omega},t)e_h(\boldsymbol{\omega},t) \\ + e_h(\boldsymbol{\omega},t_k)^\top (A_h(\boldsymbol{\omega},t)-A)x(t) \\ + (e_h(\boldsymbol{\omega},t)-e_h(\boldsymbol{\omega},t_k))^\top (A_h(\boldsymbol{\omega},t)-A)x(t).\end{aligned}$$

The **first term** is bounded by a (quasi-)dissipative assumption:

- Assumption: the matrices A_m are dissipative, i.e.

  $$\langle x, A_m x\rangle \leq 0, \qquad \forall m \in \{1,2,\dots,M\}, x \in \mathbb{R}^N.$$

  Then $A_h(\boldsymbol{\omega},t) = \sum_{m\in S_{\omega_k}} \frac{1}{\pi_m} A_m$ is also dissipative.
- Alternative: there exists a $a \in \mathbb{R}$ such that

  $$\sum_{m\in S_\omega} \frac{1}{\pi_m}\langle x, A_m x\rangle \leq a|x|^2, \qquad \forall \omega \in \{1,2,\dots,2^M\}, x \in \mathbb{R}^N.$$

Sketch of proof for Step 1 (3/4)

$$
\begin{aligned}
\frac{d}{dt}|e_h(\omega,t)|^2 \leq 0 & \\
& + e_h(\omega,t_k)^\top (A_h(\omega,t) - A)x(t) \\
& + (e_h(\omega,t) - e_h(\omega,t_k))^\top (A_h(\omega,t) - A)x(t).
\end{aligned}
$$

Sketch of proof for Step 1 (3/4)

$$\frac{d}{dt}|e_h(\boldsymbol{\omega},t)|^2 \leq 0$$
$$+ e_h(\boldsymbol{\omega},t_k)^\top (A_h(\boldsymbol{\omega},t)-A)x(t)$$
$$+ (e_h(\boldsymbol{\omega},t)-e_h(\boldsymbol{\omega},t_k))^\top (A_h(\boldsymbol{\omega},t)-A)x(t).$$

The **second term** vanishes in expectation.

Observe: For $t \in [t_k, t_{k+1})$,

- $A_h(\boldsymbol{\omega},t)$ depends only on ω_k
- $e_h(\boldsymbol{\omega},t_k)$ depends only on $\omega_1, \omega_2, \ldots, \omega_{k-1}$.

Therefore,

$$\mathbb{E}[e_h(t_k)^\top (A_h(t)-A)] = \mathbb{E}[e_h(t_k)^\top]\mathbb{E}[A_h(t)-A] = \mathbb{E}[e_h(t_k)^\top]\cdot 0 = 0.$$

Sketch of proof for Step 1 (3/4)

$$\begin{aligned}\frac{d}{dt}\mathbb{E}[|e_h(t)|^2] \leq 0 \\ &+ \mathbb{E}[e_h(t_k)^\top (A_h(t) - A)]x(t) \\ &+ \mathbb{E}[(e_h(t) - e_h(t_k))^\top (A_h(t) - A)]x(t).\end{aligned}$$

The **second term** vanishes in expectation.

Observe: For $t \in [t_k, t_{k+1})$,

- $A_h(\boldsymbol{\omega}, t)$ depends only on ω_k
- $e_h(\boldsymbol{\omega}, t_k)$ depends only on $\omega_1, \omega_2, \dots, \omega_{k-1}$.

Therefore,

$$\mathbb{E}[e_h(t_k)^\top (A_h(t) - A)] = \mathbb{E}[e_h(t_k)^\top]\mathbb{E}[A_h(t) - A] = \mathbb{E}[e_h(t_k)^\top] \cdot 0 = 0.$$

Sketch of proof for Step 1 (4/4)

$$
\begin{aligned}
\frac{d}{dt}\mathbb{E}[|e_h(t)|^2] \leq 0 & \\
&+ 0 \\
&+ \mathbb{E}[(e_h(t) - e_h(t_k))^\top (A_h(t) - A)]x(t).
\end{aligned}
$$

The **third term** can be bounded as follows.

$$
\begin{aligned}
\mathbb{E}[(e_h(t) - e_h(t_k))^\top (A_h(t) - A)]x(t) & \\
&\leq \mathbb{E}[|e_h(t) - e_h(t_k)| \|A_h(t) - A\|]|x(t)| \\
&\leq \sqrt{\mathbb{E}[|e_h(t) - e_h(t_k)|^2]\mathbb{E}[\|A_h(t) - A\|^2]}|x(t)| \\
&\leq C_{[\|A\|,T,u]} h \mathrm{Var}[A_h].
\end{aligned}
$$

Sketch of proof for Step 1 (4/4)

$$\frac{d}{dt}\mathbb{E}[|e_h(t)|^2] \leq 0$$
$$+\,0$$
$$+\,\mathbb{E}[(e_h(t) - e_h(t_k))^\top (A_h(t) - A)]x(t).$$

The **third term** can be bounded as follows.

$$\begin{aligned}
\mathbb{E}[(e_h(t) - e_h(t_k))^\top (A_h(t) - A)]x(t) \\
&\leq \mathbb{E}[|e_h(t) - e_h(t_k)| \|A_h(t) - A\|]|x(t)| \\
&\leq \sqrt{\mathbb{E}[|e_h(t) - e_h(t_k)|^2]\mathbb{E}[\|A_h(t) - A\|^2]}|x(t)| \\
&\leq C_{[\|A\|,T,u]} h \mathrm{Var}[A_h].
\end{aligned}$$

Conclusion:

$$\mathbb{E}[|x_h(t) - x(t)|^2] \leq C_{[\|A\|,T,u]} h \mathrm{Var}[A_h].$$

Numerical example

Consider a heat equation on $V = [-L, L]^3$,

$$
\begin{aligned}
y_t(t,\boldsymbol{\xi}) &= \Delta y(t,\boldsymbol{\xi}), && \boldsymbol{\xi} \in [-L, L]^3, \\
\nabla y(t,\boldsymbol{\xi}) \cdot \mathrm{n} &= u(t), && \boldsymbol{\xi} \in S_{\mathrm{top}}, \\
\nabla y(t,\boldsymbol{\xi}) \cdot \mathrm{n} &= 0, && \boldsymbol{\xi} \in \partial V \backslash S_{\mathrm{top}}, \\
y(0,\boldsymbol{\xi}) &= e^{-|\boldsymbol{\xi}|^2/(8L^2)},
\end{aligned}
$$

where $S_{\mathrm{top}} = \{(\xi_1, \xi_2, \xi_3) \in [-L, L]^3 \mid \xi_3 = L\}$. Minimize

$$
J = 1000 \int_0^T \iint_{S_{\mathrm{side}}} (y(t,\boldsymbol{\xi}))^2 \, \mathrm{d}\boldsymbol{\xi} \, \mathrm{d}t + \int_0^T (u(t))^2 \, \mathrm{d}t,
$$

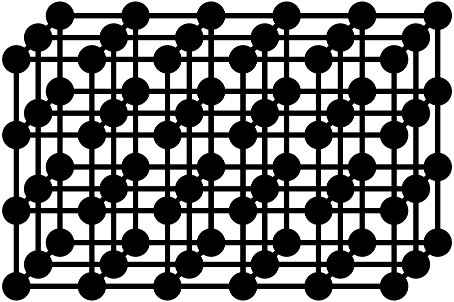
where $S_{\mathrm{side}} = \{(\xi_1, \xi_2, \xi_3) \in [-L, L]^3 \mid \xi_1 = -L\}$.

Spatial discretization and splitting

The PDE is discretized by finite differences.

Spatial discretization and splitting

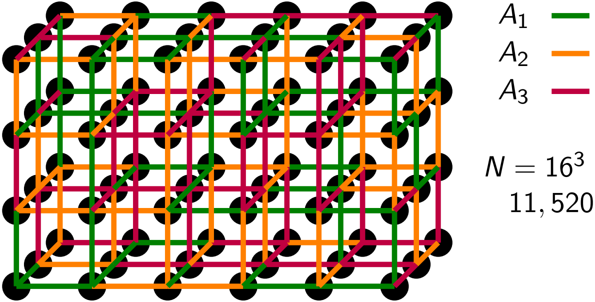
The PDE is discretized by finite differences.
The A-matrix is the graph-Laplacian for



$N = 16^3 = 4,096$
$11,520$ edges

Spatial discretization and splitting

The PDE is discretized by finite differences.
The A-matrix is the graph-Laplacian for

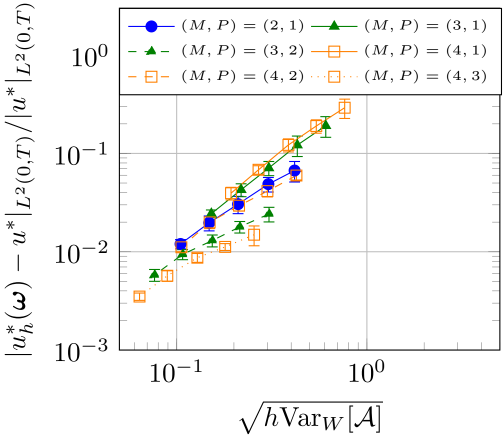


$N = 16^3 = 4,096$
$11,520$ edges

The splitting $A = \sum_{m=1}^{M} A_m$ is obtained by (randomly) dividing the edges in the graph into $M (\ll N)$ subgroups of equal size. (These subgroups are fixed during the application of the RBM)

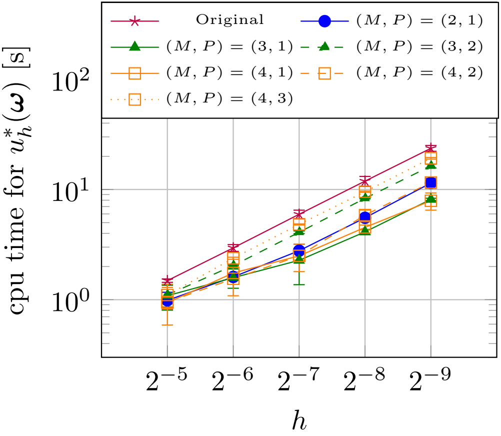
Convergence rate

$$\mathbb{E}[|u_h^* - u^*|^2_{L^2(0,T;\mathbb{R}^q)}] \leq Ch\mathrm{Var}[A_h]$$



In each time interval $[t_{k-1}, t_k)$, we P of the M submatrices A_m are used simulteneously.

Computational cost

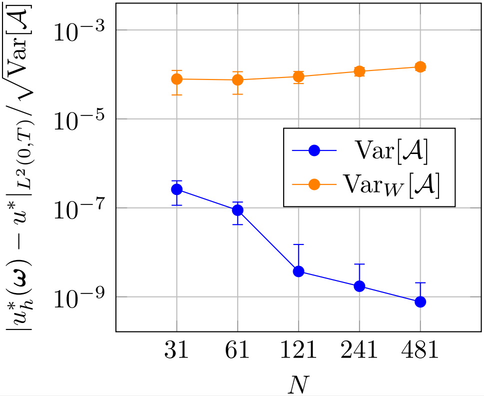


In each time interval $[t_{k-1}, t_k)$, we P of the M submatrices A_m are used simulteneously.

The computational cost can be reduced by a factor 3.

Open questions

- **Extension to infinite dimensional systems**
  Main obstruction in Step 1:
  $\|A\|$ appears when estimating $\mathbb{E}[|e_h(t) - e_h(t_k)|^2]$



Open questions

- **Extension to infinite dimensional systems**
  Main obstruction in Step 1:
  $\|A\|$ appears when estimating $\mathbb{E}[|e_h(t) - e_h(t_k)|^2]$
- **Extension to nonlinear systems**
  Relevant for interacting-particle systems and neural ODEs.
  Main obstruction in Step 2: stochastic controls
  Also convexity is lost $\Rightarrow$ Step 3b does not apply anymore.

Open questions

- **Extension to infinite dimensional systems**
  Main obstruction in Step 1:
  $\|A\|$ appears when estimating $\mathbb{E}[|e_h(t) - e_h(t_k)|^2]$
- **Extension to nonlinear systems**
  Relevant for interacting-particle systems and neural ODEs.
  Main obstruction in Step 2: stochastic controls
  Also convexity is lost $\Rightarrow$ Step 3b does not apply anymore.
- **Combination with Model Predictive Control**
  Novel analysis of MPC in the LQ-setting was developed.
  [Veldman, Zuazua, https://arxiv.org/abs/2206.01097, 2022]
  Convergence of RBM-MPC in the LQ setting in the MSc internship Alexandra Borkowski. Nonlinear setting still open.

The RBM for hyperbolic systems

Consider the transport equation

$$\begin{aligned} &y_t(t,x) + v(x) y_x(t,x) = 0, && t \in (0,T), x \in \mathbb{R}, \\ &y(0,x) = y_0(x), && x \in \mathbb{R}, \end{aligned}$$

where $v(x)$ is bounded an Lipschitz and y_0 is C^1.

Note: the operator $v(x)\frac{\partial}{\partial x}$ is unbounded.

$\Rightarrow$ Convergence proof from before breaks down.

The RBM for hyperbolic systems

Consider the transport equation

$$\begin{aligned} y_t(t,x) + v(x) y_x(t,x) &= 0, && t \in (0,T), x \in \mathbb{R}, \\ y(0,x) &= y_0(x), && x \in \mathbb{R}, \end{aligned}$$

where $v(x)$ is bounded an Lipschitz and y_0 is C^1.

Note: the operator $v(x)\frac{\partial}{\partial x}$ is unbounded.

$\Rightarrow$ Convergence proof from before breaks down.

The solution $y(t,x)$ is given by

$$y(t,x) = y_0(\xi(0;t,x)),$$

where $\xi(s;t,x)$ is the solution of the ODE

$$\tfrac{d}{ds}\xi(s;t,x) = v(\xi(s;t,x)), \qquad \xi(t;t,x) = x.$$

The RBM for hyperbolic systems

We split the generator of the semi-group as

$$\underbrace{-v(x)\frac{\partial}{\partial x}}_{A} = \sum_{m=1}^{M} \underbrace{-v_m(x)\frac{\partial}{\partial x}}_{A_m},$$

where the $v_m(x)$ are Lipschitz and bounded and consider as before

$$A_h(\boldsymbol{\omega}, t) = \sum_{m \in S_{\omega_k}} \frac{A_m}{\pi_m}, \qquad t \in [t_{k-1}, t_k).$$

The RBM for hyperbolic systems

We thus consider the randomized transport equation

$$\begin{aligned} &y_t(\omega,t,x)+v_h(\omega,x)y_x(\omega,t,x)=0, && t\in(0,T), x\in\mathbb{R},\\ &y(\omega,0,x)=y_0(x), && x\in\mathbb{R}, \end{aligned}$$

where

$$v_h(\omega,x)=\sum_{m\in S_{\omega_k}}\frac{v_m(x)}{\pi_m}, \qquad t\in[t_{k-1},t_k).$$

The RBM for hyperbolic systems

We thus consider the randomized transport equation

$$\begin{aligned} &y_t(\omega,t,x)+v_h(\omega,x)y_x(\omega,t,x)=0, && t\in(0,T), x\in\mathbb{R},\\ &y(\omega,0,x)=y_0(x), && x\in\mathbb{R}, \end{aligned}$$

where

$$v_h(\omega,x)=\sum_{m\in S_{\omega_k}}\frac{v_m(x)}{\pi_m}, \qquad t\in[t_{k-1},t_k).$$

The solution $y_h(\omega,t,x)$ is given by

$$y_h(\omega,t,x)=y_0(\xi(\omega,0;t,x)),$$

where $\xi_h(\omega,s;t,x)$ is the solution of the ODE

$$\frac{d}{ds}\xi_h(\omega,s;t,x)=v_h(\omega,\xi_h(\omega,s;t,x)), \qquad \xi(\omega,t;t,x)=x.$$

The RBM for hyperbolic systems

Consider the two systems of characteristics

$$\frac{d}{ds}\xi(s;t,x) = v(\xi(s;t,x)), \qquad \xi(t;t,x) = x,$$

$$\frac{d}{ds}\xi_h(\omega,s;t,x) = v_h(\omega,\xi_h(\omega,s;t,x)), \qquad \xi(\omega,t;t,x) = x.$$

The RBM for hyperbolic systems

Consider the two systems of characteristics

$$\frac{d}{ds}\xi(s;t,x) = v(\xi(s;t,x)), \qquad \xi(t;t,x) = x,$$

$$\frac{d}{ds}\xi_h(\omega,s;t,x) = v_h(\omega,\xi_h(\omega,s;t,x)), \qquad \xi(\omega,t;t,x) = x.$$

Lemma

If $v(x)$ and the $v_m(x)$ are bounded and Lipschitz, $\exists$ a C s.t.

$$\mathbb{E}[|\xi_h(s;t,x) - \xi(s;t,x)|^2] \leq C_{[t-s]}h.$$

The RBM for hyperbolic systems

Consider the two systems of characteristics

$$\frac{d}{ds}\xi(s;t,x) = v(\xi(s;t,x)), \qquad \xi(t;t,x) = x,$$
$$\frac{d}{ds}\xi_h(\omega,s;t,x) = v_h(\omega,\xi_h(\omega,s;t,x)), \qquad \xi(\omega,t;t,x) = x.$$

Lemma

If $v(x)$ and the $v_m(x)$ are bounded and Lipschitz, $\exists$ a C s.t.

$$\mathbb{E}[|\xi_h(s;t,x) - \xi(s;t,x)|^2] \leq C_{[t-s]}h.$$

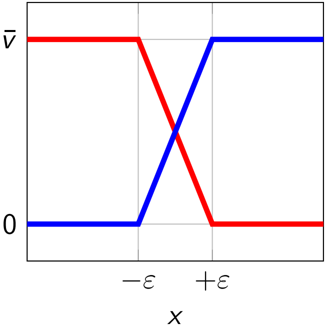
As a consequence

$$\begin{aligned} |y_h(\omega,t,x) - y(t,x)| &= |y_0(\xi_h(\omega,0;t,x)) - y_0(\xi(0;t,x))| \\ &\leq |y_0|_{C^1}|\xi_h(\omega,0;t,x) - \xi(0;t,x)|. \end{aligned}$$

$$\mathbb{E}[|y_h(t,x) - y(t,x)|^2] \leq |y_0|^2_{C^1} C_{[t]}h.$$

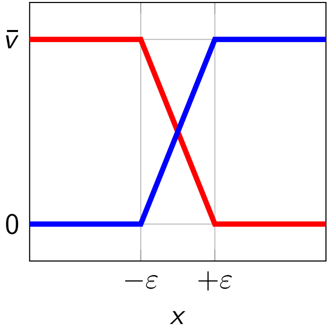
Visualization

$$v(x) = \bar{v} = v_1(x) + v_2(x)$$

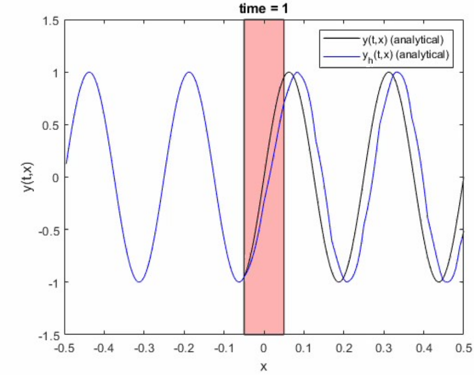


Visualization

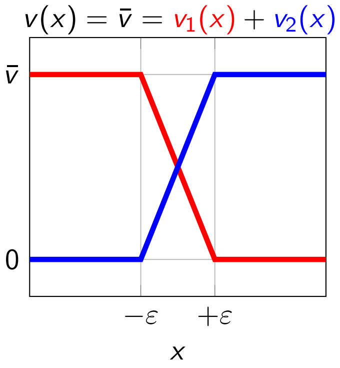
$$v(x) = \bar{v} = v_1(x) + v_2(x)$$



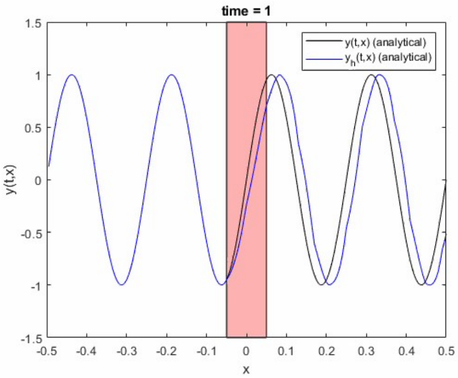
$h = 0.05$



Visualization



$h = 0.02$



Remarks and outlook

'Easy' extensions:

- bounded domains $x \in (0, L)$.
- n-dimensional spatial domains

$$y_t(t, \vec{x}) = \vec{v}(\vec{x}) \cdot \nabla_{\vec{x}} y(t, \vec{x}).$$

- systems of linear transport equations

$$\vec{y}_t(t, x) = A(x)\vec{y}_x(t, x).$$

Remarks and outlook

'Easy' extensions:

- bounded domains $x \in (0, L)$.
- n-dimensional spatial domains

$$y_t(t, \vec{x}) = \vec{v}(\vec{x}) \cdot \nabla_{\vec{x}} y(t, \vec{x}).$$

- systems of linear transport equations

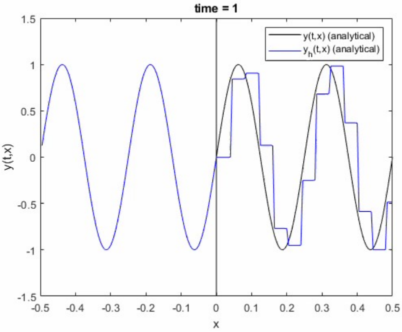
$$\vec{y}_t(t, x) = A(x) \vec{y}_x(t, x).$$

(Much) more challenging:

- Removing the Lipschitz conditions (networks of transport equations)
- Weaker notions of solutions.
- Nonlinear transport equations / conservation laws
  In particular, networks of incompressible Euler equations

Remarks and outlook

$\varepsilon \approx 0, h = 0.02$



'Easy' extensions:

- bounded domains $x \in (0, L)$.
- n-dimensional spatial domains

$$y_t(t, \vec{x}) = \vec{v}(\vec{x}) \cdot \nabla_{\vec{x}} y(t, \vec{x}).$$

- systems of linear transport equations

$$\vec{y}_t(t, x) = A(x)\vec{y}_x(t, x).$$

(Much) more challenging:

- Removing the Lipschitz conditions (networks of transport equations)
- Weaker notions of solutions.
- Nonlinear transport equations / conservation laws In particular, networks of incompressible Euler equations