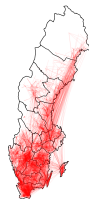
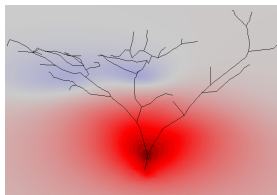


Stability and strong convergence in multiscale methods for spatial stochastic kinetics



Stefan Engblom

Div of Scientific Computing, Dept of Information Technology, Uppsala University

Spatially Distributed Stochastic Dynamical Systems in Biology

Isaac Newton Institute for mathematical Sciences, University of Cambridge, June 20, 2016

Outline

1. The computational framework
 - Stochastic reaction-transport modeling
 - A reminder: *why?*
2. Analysis
 - Assumptions and *a priori* results
 - (Multiscale) variable splitting methods
3. Applications
 - Multiscale neuronal model
 - National-scale epidemics

Summary

Local physics + transport mechanism

= Event-based mesoscopic & stochastic **computational framework**

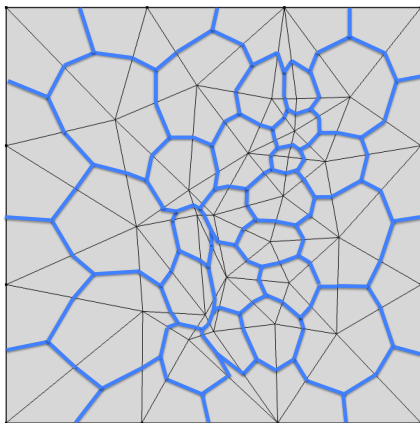
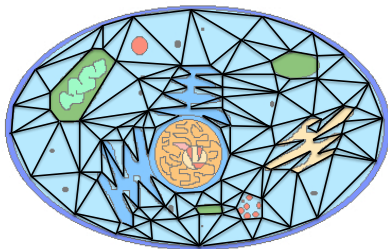


Figure: Primal mesh (thin), dual mesh (blue). The *state* is the # of agents (eg. molecules) in each dual voxel.

Local physics within each voxel, *connected* through transport mechanisms (eg. diffusion).

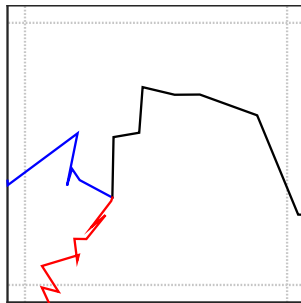
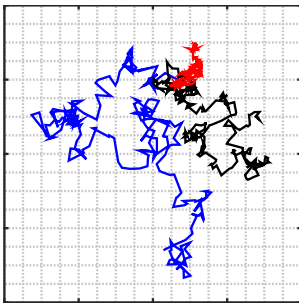
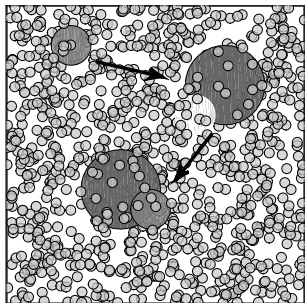


Like PDEs, but better!

“Local physics” first...

Well-stirred kinetics

Example: Bimolecular reaction $X + Y \rightarrow Z$. Or infection spread
 $S + I \rightarrow 2I$. Or...

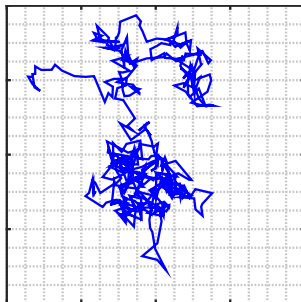
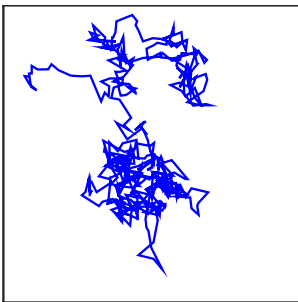
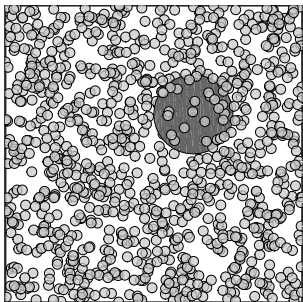


-When counting individual species/agents, a **continuous-time Markov chain** is the most immediate model of the physics in the zoomed in situation.

... “transport physics” next

Space-discrete, time-continuous model of moving particle

Example: Brownian motion.



(micro) \rightarrow (stoch) The stochastic model is **simpler** but random (*error*: microscale effects in a statistical sense only).

(stoch) \rightarrow (meso) Discrete space approximation (*error*: finite $h > 0$).

Why stochastic? Why discrete? Why space?

SE '06 *“The situation is clearly different when biological systems inside living cells are considered. [—] It is intuitively clear that under such circumstances the inherent **stochasticity** of the system **plays a vital role**”*

SE & others '09 *“**Intrinsic noise** in biochemical networks can have a **large impact** [—] The extremely **complex** ... microscopic behavior paired with the fact that the copy number is a small nonnegative integer make a **discrete, stochastic** description of the system **necessary**”*

SE & others '15 *“...spatial stochastic models based on a Markov process formalism are popular due to their high level of **biological realism** compared to [PDEs], with only a moderate increase in computational complexity...”*

-And a great many similar remarks have been made by several many others...! (everybody “knows”)

The main message

Just to rub it in...

Terms & conditions. Want to use these models when either one of

- ▶ stochasticity
- ▶ nonlinearities
- ▶ species discreteness
- ▶ spatial inhomogeneities

make a big, or at least an interesting difference. Hence the physical model itself is *sensitive to perturbations* in anyone of these.

The main message

Just to rub it in...

Terms & conditions. Want to use these models when either one of

- ▶ stochasticity
- ▶ nonlinearities
- ▶ species discreteness
- ▶ spatial inhomogeneities

make a big, or at least an interesting difference. Hence the physical model itself is *sensitive to perturbations* in anyone of these.

Designing/understanding computational models: either we do

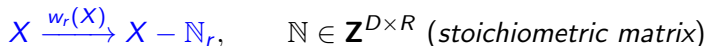
- ▶ An analysis by analogy/fingerspitzengefühl...
- ▶ Or, using the **Lax principle**: *if the numerical physics \approx the wanted "true" physics (consistency), then the numerical solution \rightarrow the true solution (convergence) IFF the numerical physics is stable*

Notation

Local physics

-State $X \in \mathbf{Z}_+^D$, counting the number of each of D species/agents/compartments.

-Events/reactions are transitions between these states,



with propensity $w_r : \mathbf{Z}_+^D \rightarrow \mathbf{R}_+$, $r = 1 \dots R$.

-Poisson representation

$$X(t) = X(0) - \sum_r \mathbb{N}_r \Pi_r \left(\int_0^t w_r(X(s)) ds \right),$$

each Π_r a unit-rate Poisson process.

Notation (cont)

Mesoscopic spatial kinetics

Total volume Ω subdivided into small enough computational cells Ω_j such that the local physics is an accurate model.

- ▶ The state of the system is now an array \mathbb{X} with $D \times K$ elements; D species \mathbb{X}_{ij} , $i = 1, \dots, D$, counted separately in K cells, $j = 1, \dots, K$.
- ▶ This state is changed by local physics events (**vertically** in \mathbb{X}) and by transport into adjacent cells (**horizontally** in \mathbb{X}).

Local physics

(eg. reactions)

Same model in K cells, $j = 1, \dots, K$,

$$\mathbb{X}_{ij}(t) = \mathbb{X}_{ij}(0) - \sum_r \mathbb{N}_{ri} \Pi_{rj} \left(\int_0^t w_{rj}(\mathbb{X}_{\cdot,j}(s)) ds \right),$$

for $i = 1, \dots, D$ species.

Transport mechanism

(eg. diffusion)

Linear model (convection/diffusion, but also crowding): transport from one cell Ω_j to another cell Ω_k according to

$$\mathbb{X}_{ij} \xrightarrow{q_{ijk} \mathbb{X}_{ij}} \mathbb{X}_{ik},$$

where q_{ijk} is non-zero only for connected cells.

$$\mathbb{X}_{ij}(t) = \mathbb{X}_{ij}(0) - \sum_k \Pi'_{ijk} \left(\int_0^t q_{ijk} \mathbb{X}_{ij}(s) ds \right) + \sum_k \Pi'_{ikj} \left(\int_0^t q_{ikj} \mathbb{X}_{ik}(s) ds \right).$$

Stochastic reaction-transport framework

“RDME”

Combining reactions with transport events we arrive at

$$\begin{aligned} \mathbb{X}_{ij}(t) = & \mathbb{X}_{ij}(0) - \sum_r \mathbb{N}_{ri} \Pi_{rj} \left(\int_0^t w_{rj}(\mathbb{X}_{\cdot j}(s)) ds \right) \\ & - \sum_k \Pi'_{ijk} \left(\int_0^t q_{ijk} \mathbb{X}_{ij}(s) ds \right) + \sum_k \Pi'_{ikj} \left(\int_0^t q_{ikj} \mathbb{X}_{ik}(s) ds \right). \end{aligned}$$

-Formulated in already discrete space! The limit when the cell size $\rightarrow 0$ is not straightforward.

Assumptions & *a priori*: well-stirred case

Local physics first...

Recall: CTMC $X(t) \in \mathbf{Z}_+^D$ governed by transitions

$$X \xrightarrow{w_r(X)} X - \mathbb{N}_r, \quad r = 1 \dots R, \quad \mathbb{N} \in \mathbf{Z}^{D \times R},$$

or, to get some ODE-feeling, “ $X'(t) = -\mathbb{N}w(X)$ ”.

Norm $\|x\|_l := \mathbf{1}^T x$, $x \in \mathbf{Z}_+^D$, normalized so $\min_i l_i = 1$.

Assumptions & *a priori*: well-stirred case

Local physics first...

Recall: CTMC $X(t) \in \mathbf{Z}_+^D$ governed by transitions

$$X \xrightarrow{w_r(X)} X - \mathbb{N}_r, \quad r = 1 \dots R, \quad \mathbb{N} \in \mathbf{Z}^{D \times R},$$

or, to get some ODE-feeling, “ $X'(t) = -\mathbb{N}w(X)$ ”.

Norm $\|x\|_I := I^T x$, $x \in \mathbf{Z}_+^D$, normalized so $\min_i I_i = 1$.

Assumptions: $x, y \in \mathbf{Z}_+^D$,

(i) $-I^T \mathbb{N}w(x) \leq A + \alpha \|x\|_I$ (“*I*-outward bound”)

(ii) $(-I^T \mathbb{N})^2 w(x) / 2 \leq B + \beta_1 \|x\|_I + \beta_2 \|x\|_I^2$ (“*I*-outward absolute bound”)

(iii) $|w_r(x) - w_r(y)| \leq L_r(P) \|x - y\|$, $r = 1, \dots, R$, and $\|x\|_I \vee \|y\|_I \leq P$

The “blue assumptions”.

Assumptions & *a priori*: local physics

Summary of results

With suitable initial data...

- ▶ This $\mathbb{E}[\sup_{s \in [0, t]} \|X(s)\|_l^p]$ bounded, any $p \geq 1$
- ▶ if $X(0) = Y(0)$ almost surely, then $\mathbb{E}[\|X(t) - Y(t)\|^2] = 0$
- ▶ if $\alpha + \beta_2(p - 1) < 0$, then $\mathbb{E}[\|X(t)\|_l^p]$ bounded as $t \rightarrow \infty$

Assumptions & *a priori*: local physics

Summary of results

With suitable initial data...

- ▶ This $\mathbb{E}[\sup_{s \in [0, t]} \|X(s)\|_l^p]$ bounded, any $p \geq 1$
- ▶ if $X(0) = Y(0)$ almost surely, then $\mathbb{E}[\|X(t) - Y(t)\|^2] = 0$
- ▶ if $\alpha + \beta_2(p - 1) < 0$, then $\mathbb{E}[\|X(t)\|_l^p]$ bounded as $t \rightarrow \infty$

-In fact, if $X(0) = Y(0)$ almost surely, and if $Y(t)$ is obtained by δ -perturbing the transition intensities ($w_r \rightarrow (1 \pm \delta)w_r$), then $\lim_{\delta \rightarrow 0} \mathbb{E}[\|X(t) - Y(t)\|^2] = 0$.

-Actually, if both X and Y are bounded, then $\mathbb{E}[\|X(t) - Y(t)\|^2] = O(\delta)$.

Assumptions & *a priori*: spatial case

Recall: CTMC $\mathbb{X}(t) \in \mathbf{Z}_+^{D \times K}$ with transitions

$$\mathbb{X}_{\cdot,k} \xrightarrow{w_{rk}(\mathbb{X}_{\cdot,k})} \mathbb{X}_{\cdot,k} - \mathbb{N}_r, \quad \mathbb{X}_{ij} \xrightarrow{q_{ijk} \mathbb{X}_{ij}} \mathbb{X}_{ik},$$

$k = 1 \dots K, i = 1 \dots D, r = 1 \dots R$. To get “PDE-feeling”,

$$\mathbf{v}_t = -\mathbb{N}u(\mathbf{v}) + \underbrace{Q}_{\text{eg. } \approx \nabla \cdot \Sigma \nabla} \mathbf{v}.$$

Assumptions & *a priori*: spatial case

Recall: CTMC $\mathbb{X}(t) \in \mathbf{Z}_+^{D \times K}$ with transitions

$$\mathbb{X}_{\cdot,k} \xrightarrow{w_{rk}(\mathbb{X}_{\cdot,k})} \mathbb{X}_{\cdot,k} - \mathbb{N}_r, \quad \mathbb{X}_{ij} \xrightarrow{q_{ijk} \mathbb{X}_{ij}} \mathbb{X}_{ik},$$

$k = 1 \dots K, i = 1 \dots D, r = 1 \dots R$. To get “PDE-feeling”,

$$\mathbf{v}_t = -\mathbb{N}u(\mathbf{v}) + \underbrace{Q}_{\text{eg. } \approx \nabla \cdot \Sigma \nabla} \mathbf{v}.$$

Assumptions:

- ▶ on the mesh, some natural and quite weak assumptions (...)
- ▶ reactions, as before, *plus*
 - (iv) $w_{rk}(x) = \Omega_k u_r(\Omega_k^{-1}x)$, “density dependent”

Assumptions & *a priori*: spatial case

Summary of results

Norm $\|\mathbb{X}\|_{l,1} \equiv \sum_{k=1}^K \|\mathbb{X}_{\cdot,k}\|_l = l^T \mathbb{X} \mathbf{1}.$

With suitable initial data...

- ▶ only reactions: as before
- ▶ pure transport: $\|\mathbb{X}(t)\|_{l,1} = \|\mathbb{X}(0)\|_{l,1}$, so bounded by initial data
- ▶ coupled spatial model: $\mathbb{E}[\sup_{s \in [0,t]} \|\mathbb{X}(s)\|_{l,1}^p]$ bounded, any $p \geq 1$
- ▶ (strong) continuous dependence on parameters as before

Application: Multiscale variable splitting

Set-up: ϵ, h

Consider the separation of scales:

- ▶ species are either abundant $\sim \epsilon^{-1}$, or appear in low copy numbers ~ 1 (*on a per voxel basis!*)
- ▶ rate constants are either fast ~ 1 , or slow ϵ (...)

\implies rescaled variable $\bar{X}(t) = \bar{X}_{ij}(t) \sim 1$.

Application: Multiscale variable splitting

Set-up: ϵ, h

Consider the separation of scales:

- ▶ species are either abundant $\sim \epsilon^{-1}$, or appear in low copy numbers ~ 1 (*on a per voxel basis!*)
- ▶ rate constants are either fast ~ 1 , or slow ϵ (...)

\implies rescaled variable $\bar{X}(t) = \bar{X}_{ij}(t) \sim 1$.

Multiscale splitting methods:

“Exact”, $\bar{Y}(t)$ all Poisson processes driving an abundant species are replaced with mean drift terms, $\Pi(t) \approx t$

“Numerical”, $\bar{Y}^{(h)}(t)$ discrete steps h ; low copy number variables are first simulated in $[t, t + h)$ letting abundant species be frozen at time t , next abundant species are integrated in $[t, t + h)$

Scale separation

Details

Scale vector $S \in \mathbf{R}^D$,

$$\mathbb{X}_{i,\cdot}(t) = S_i \bar{\mathbb{X}}_{i,\cdot}(t), \quad S_i = 1 \text{ or } \epsilon^{-1}.$$

The rates are assumed to obey the scaling laws

$$\begin{aligned} q_{ijk}x &= \epsilon^{-\mu_i} \bar{q}_{ijk} S^{-1}x, \\ u_r(x) &= \epsilon^{-\nu_r} \bar{u}_r(S^{-1}x). \end{aligned}$$

The scaled rates $\{\bar{q}_{ijk}, \bar{u}_r(\cdot)\}$ are understood to be $O(1)$ with respect to ϵ .

Scale separation

Existence

If the following **scaled assumptions** hold,

$$-I^T S^{-1} \mathbb{N} u(x) \leq A + \alpha \|S^{-1} x\|_I \quad (1)$$

$$(-I^T S^{-1} \mathbb{N})^2 u(x) / 2 \leq B + \beta_1 \|S^{-1} x\|_I + \beta_2 \|S^{-1} x\|_I^2 \quad (2)$$

$$|\bar{u}_r(x) - \bar{u}_r(y)| \leq L_r(P) \|x - y\|, \quad r = 1 \dots R, \text{ and } \|x\|_I \vee \|y\|_I \leq P \quad (3)$$

for $\{I, A, \alpha, B, \beta_1, \beta_2, L\}$ all *independent of ϵ* .

Then in an $O(1)$ interval of time, with $O(1)$ initial data, $\mathbb{E}[\sup_{s \in [0, t]} \|\bar{X}(s)\|_{I,1}^P]$ is also $O(1)$.

Scale separation

Existence (cont)

1. Replace (2) with

$$\left(-I_1^T \mathbb{N}^{(1)}\right)^2 u(x)/2 \leq B + \beta_1 \|S^{-1}x\|_I + \beta_2 \|S^{-1}x\|_I^2$$

(I -outward absolute bound for stochastic part only)

\implies Then $\bar{Y}(t)$ is also $O(1)$.

2. Additionally replace (1) with

$$\max\left(-I_1^T \mathbb{N}^{(1)} u(x), -I_2^T \epsilon \mathbb{N}^{(2)} u(x)\right) \leq A + \alpha \|S^{-1}x\|_I$$

(I -outward bound for deterministic/stochastic parts individually)

\implies Then $\bar{Y}^{(h)}(t)$ is also $O(1)$.

Multiscale split

Terms & conditions

Species in low numbers $i \in G_1$, in large numbers $i \in G_2$. Put

$$R(G_1) := \{r; \text{transition } r \text{ affects a species } i \in G_1\}$$

(and same for $R(G_2)$).

Define also

$$u := \min_{r \in R(G_1)} -\nu_r \wedge \min_{i \in G_1} -\mu_i \text{ ('worst' } \epsilon\text{-scaling of transition affecting } G_1)$$

$$v := 1 + \min_{r \in R(G_2)} -\nu_r \wedge \min_{i \in G_2} -\mu_i \text{ ('worst' } \epsilon\text{-scaling of transition affecting } G_2 \text{ plus 1)}$$

Errors

Convergence results

Under the (Assumptions) above, then

- ▶ $\mathbb{E}[\|\bar{Y}(t) - \bar{X}(t)\|^2] = O(\epsilon^{1+\nu} + \epsilon^{1/2+\nu/2+u})$
- ▶ Bounded/unbounded case: *almost* the same result...

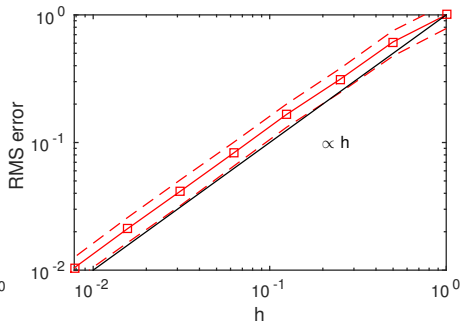
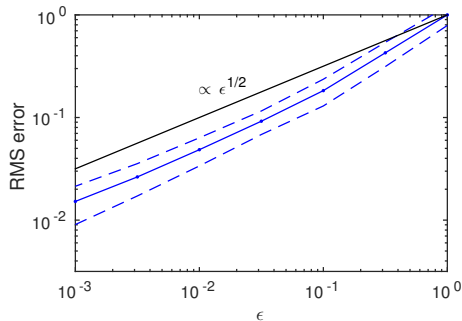
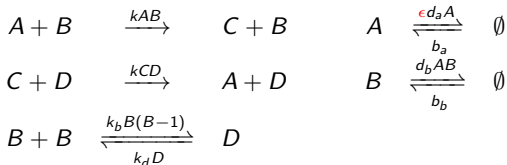
Under the (Assumptions) above, then *if the processes are bounded*,

- ▶ $\mathbb{E}[\|\bar{Y}^{(h)}(t) - \bar{Y}(t)\|^2] = O(h(\epsilon^{2u} + \epsilon^{u+\nu})) + O(h^2\epsilon^{2\nu})$
- ▶ Unbounded case: only convergence as $h \rightarrow 0$ remains...

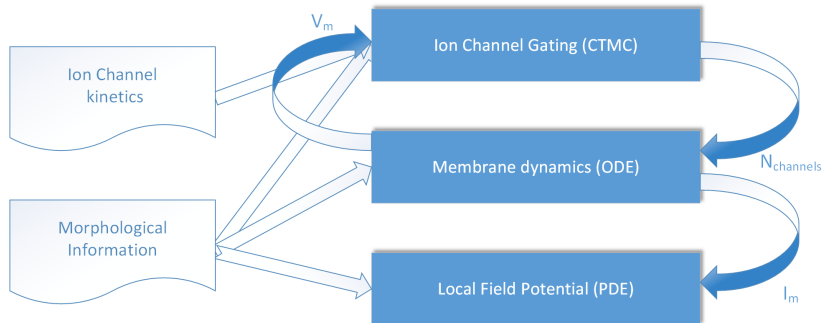
Example: catalytic process

“Stress test” of theory

$(A, C) \sim \epsilon^{-1}$, $(B, D) \sim 1$, $\text{diffusion}_{A,C} \sim \epsilon$, $\text{diffusion}_{B,D} \sim 1$.

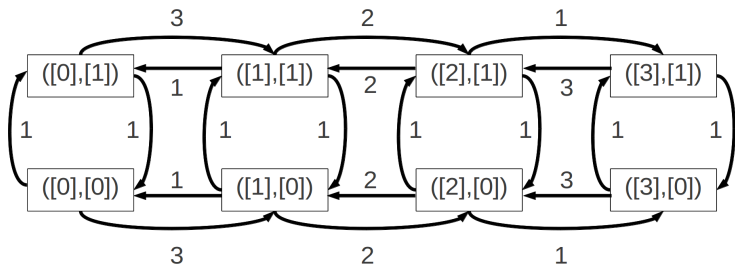


Application: multiscale neuronal model



Bottom level

Ion channel gating

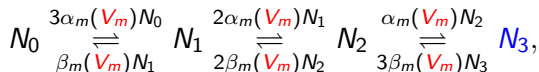


Gating process: sodium channels.

Bottom level

Ion channel gating

The gating process of ion channels can be mesoscopically described as



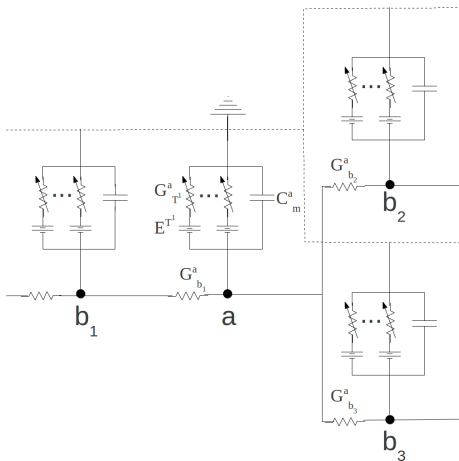
again a *continuous-time Markov chain*. Output: N_3 , the number of open gates.

For efficient model coupling we freeze the voltage dependency for a short time-step τ (“split-step” or “1st order Strang split”):

$$X(t + \tau) = X(t) - \mathbb{N}\Pi \left(\int_t^{t+\tau} w(X(s), V_m(t)) ds \right).$$

Middle level

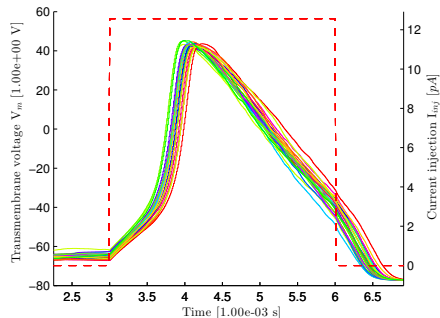
Membrane dynamics



Cable equation circuit.

Middle level

Membrane dynamics



- ▶ Morphological information extracted using the *Trees toolbox*
- ▶ System of current-balance and cable equations is solved for each time step τ

$$I_m = c_m \frac{dV_m}{dt} + \sum_{i \in C_v} \gamma_i N_3^i(t) [V_m(t) - E_i]$$

Top level

Maxwell's equations, potential form

Electric field intensity \mathbf{E} in terms of the *electric scalar potential* V ,

$$\mathbf{E} = -\nabla V.$$

Trans-membrane current I_m is scaled with the compartment surface area and coupled as a current source,

$$-\nabla \cdot \left(\sigma \nabla V + \varepsilon_0 \varepsilon_r \frac{\partial}{\partial t} \nabla V \right) = \frac{1}{\Omega_c} I_m,$$

with conductivity σ and permittivity ε . The time dependent potential V is solved via finite element methods.

Sample simulation

Application: national-scale epidemics

- ▶ Modeling the spread of verotoxinogenic *E. coli* O157:H7 (VTEC O157:H7) in the Swedish cattle population
- ▶ Important *zoonotic pathogen* (animal \rightarrow humans) of great public health interest, causing enterohemorrhagic colitis (EHEC) in humans (\sim 500 cases annually in Sweden).

Application: national-scale epidemics

- ▶ Modeling the spread of verotoxinogenic *E. coli* O157:H7 (VTEC O157:H7) in the Swedish cattle population
- ▶ Important *zoonotic pathogen* (animal → humans) of great public health interest, causing enterohemorrhagic colitis (EHEC) in humans (~500 cases annually in Sweden).
- ▶ In Germany during the summer 2011, a particularly aggressive variant emerged, with 3,816 reported cases and 54 deceased.
- ▶ *Infected animals show no signs of the disease!*
- ▶ Cattle is a main reservoir of the bacteria, ongoing research to better understand the epidemiology of VTEC O157:H7 in the cattle population
- ▶ Mixed event-based approach:
 - ▶ Data-driven simulation using all registered cattle events 2005-2013
 - ▶ Stochastic simulation of within-herd dynamics (i.e. **mesoscopic**)

Data-driven

REPORTER	WHERE	ABATTOIR	DATE	EVENT	ANIMALID	BIRTHDATE
83466	83958	0	2009-10-01	2	SE0834660433	1997-04-04
83958	83466	0	2009-10-01	1	SE0834660433	1997-04-04
83958	83829	0	2012-03-15	2	SE0834660433	1997-04-04
83829	83958	0	2012-03-15	1	SE0834660433	1997-04-04
83829	83958	0	2012-03-15	4	SE0834660433	1997-04-04
54234	83829	0	2012-04-11	1	SE0834660433	1997-04-04
83829	54234	0	2012-04-11	2	SE0834660433	1997-04-04
83829	83958	0	2012-04-11	5	SE0834660433	1997-04-04

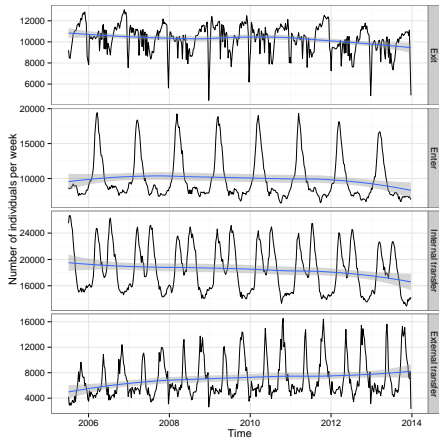
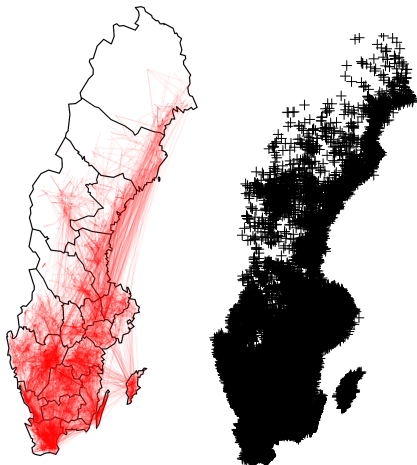
Total: 18 649 921 reports and 37 221 holdings

Events

- ▶ Exit (n=1 438 506)
- ▶ Enter (n=3 479 000)
- ▶ Internal transfer (n=6 593 921)
- ▶ External transfer (n=732 292)

Events

(Population UK:Sweden is $\sim 10:1$, area $\sim 5:9$)



Epidemic model

“Locally well-stirred” (SIS_E)

Model states: **S**usceptible, **I**nfected, in $i = 1, \dots, \sim 40,000$ holdings and in 3 age categories $j \in \{\text{calves}, \text{youngstock}, \text{adults}\}$.

State transitions at node i in the j th age category,

$$\text{Rate } I_{ij} \rightarrow S_{ij} = \gamma_j I_{ij}(t)$$

$$\text{Rate } S_{ij} \rightarrow I_{ij} = \nu_j S_{ij}(t) \varphi_i(t)$$

Environmental infectious pressure

$$\frac{d\varphi_i}{dt} = \frac{\alpha \sum_j I_{ij}(t)}{\sum_j S_{ij}(t) + I_{ij}(t)} - \beta(t) \varphi_i(t) + \epsilon$$

Epidemic model

“Locally well-stirred” (SIS_E)

Model states: **S**usceptible, **I**nfected, in $i = 1, \dots, \sim 40,000$ holdings and in 3 age categories $j \in \{\text{calves}, \text{youngstock}, \text{adults}\}$.

State transitions at node i in the j th age category,

$$\text{Rate } I_{ij} \rightarrow S_{ij} = \gamma_j I_{ij}(t)$$

$$\text{Rate } S_{ij} \rightarrow I_{ij} = \nu_j S_{ij}(t) \varphi_i(t)$$

Environmental infectious pressure

$$\frac{d\varphi_i}{dt} = \frac{\alpha \sum_j I_{ij}(t)}{\sum_j S_{ij}(t) + I_{ij}(t)} - \beta(t) \varphi_i(t) + \epsilon$$

Finding #1: $\beta = \beta(t)$ required in the Swedish climate.

Finding #2: finite-time extinction for $\epsilon = 0$, contrary to the corresponding ODE-model.

Sample simulation

Connected through ~ 9 years of actual transport data

Summary

- ▶ Mesoscopic stochastic reaction-transport, **event-based computational framework**: fairly intuitive modeling & coupling
- ▶ **Terms & conditions**. If used when required: accurately capturing a stochastic nonlinear phenomenon is a very hard constraint for method's development!
- ▶ The **Lax principle** \implies Well-posedness, stability, consistency, convergence
- ▶ Analysis of simple numerical methods
- ▶ Multiscale neuronal application solved in **URDME** (GitHub): proof of concept for coupling different types of models
- ▶ Epidemiological national-scale model solved in **SimInf** (GitHub): data-driven simulation, some findings when attempting to fit parameters to data

Thanks

Programs, Papers, and Preprints are available from my web-page.
Thank you for the attention!