Stochastic and Deterministic Simulations for Biological Problems

Petros KOUMOUTSAKOS

with : Basil BAYATI, Mattia GAZZOLA, Florian MILDE, Diego ROSSINELI, Michael BERGDORF, Philippe CHATELAIN

and : Urs Greber Laboratory (Uni ZH)

CSE LAB
COMPUTATIONAL SCIENCE & ENGINEERING
LABORATORY

HTTP://CSE-LAB.ETHZ.CH

ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Suddenly, the flock became suspicious:
How come the newcomer wasn't shorn?

Death - Birth and everything in between

The Lotka-Volterra Model

X : Sheeps, Y:predators, A: food

$$X + A \xrightarrow{k_1} 2X \qquad X + Y \xrightarrow{k_2} 2Y \qquad Y \xrightarrow{k_3} B$$

Deterministic ODE's

$$\frac{d[X]}{dt} = k_1[A][X] - k_2[X][Y] \qquad \frac{d[Y]}{dt} = k_2[X][Y] - k_3[Y]$$

Stochastic Interpretation

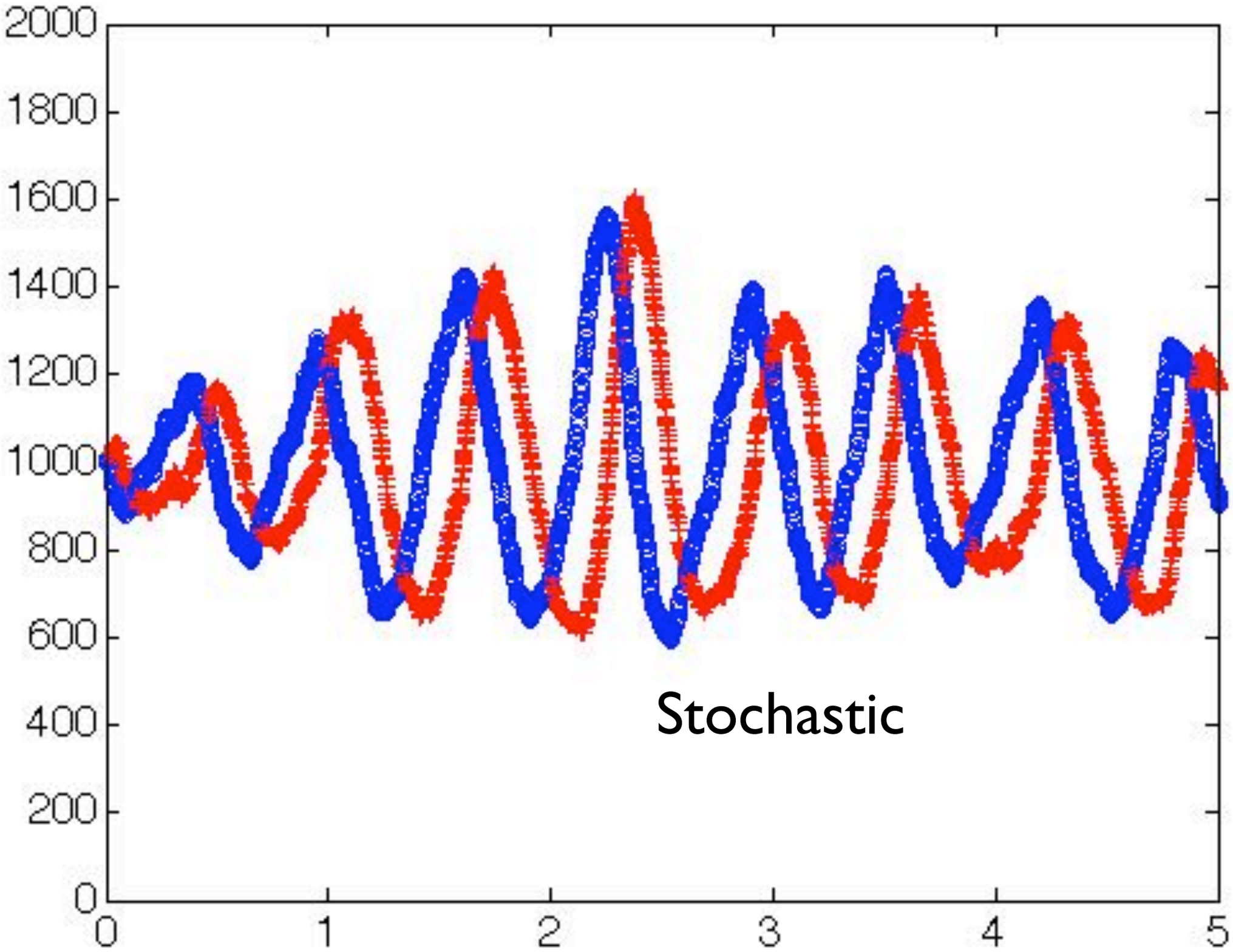
$$\mathcal{P}(x \to x+1; y \to y) = k_1 a\, x\, dt$$

$$\mathcal{P}(x \to x-1; y \to y+1) = k_2\, x\, y\, dt$$

$$\mathcal{P}(x \to x; y \to y-1) = k_3\, y\, dt$$

$$\mathcal{P}(x \to x; y \to y) = 1 - (k_1\, ax + k_2 xy + k_3\, y)\, dt$$

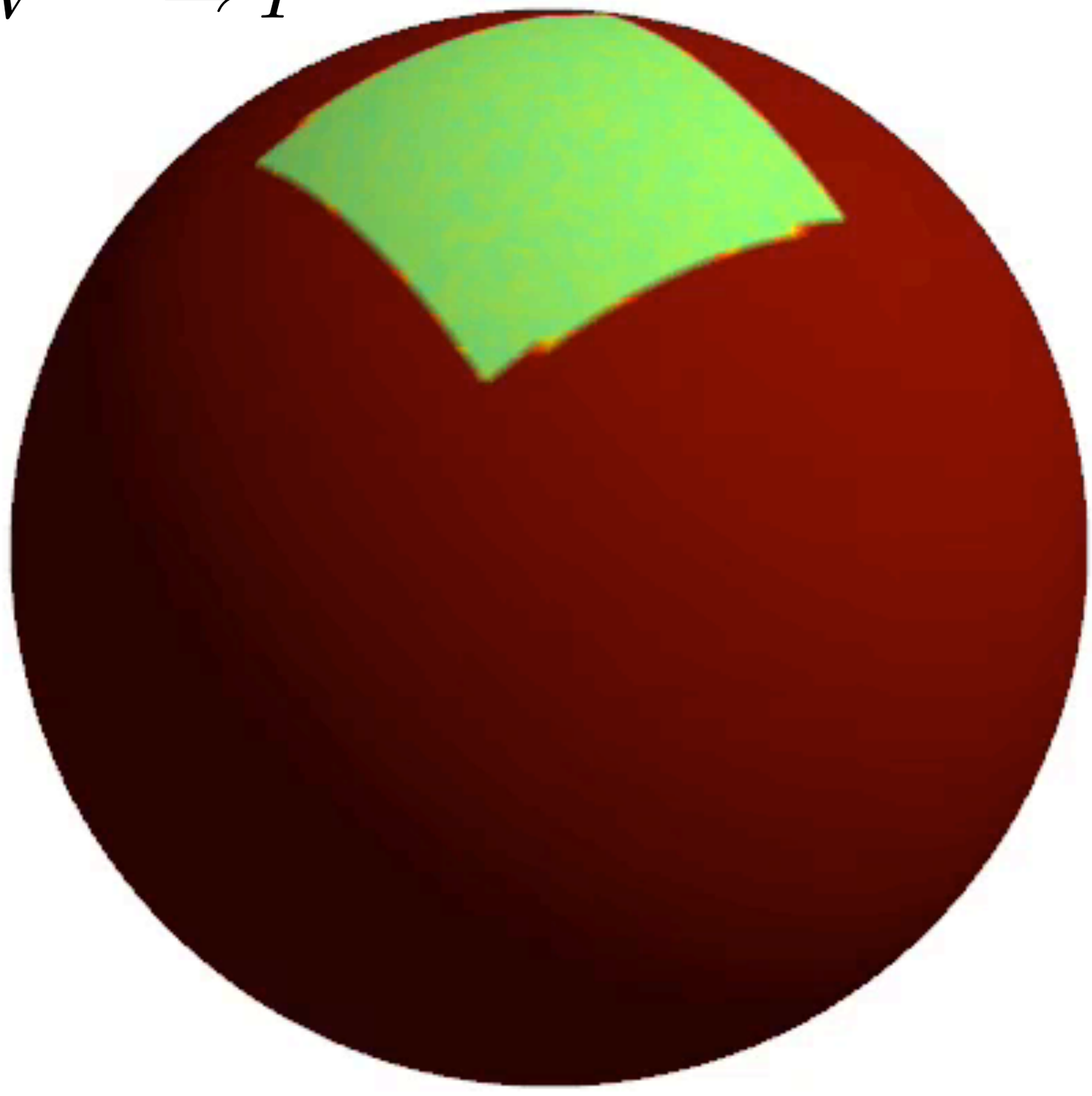
The Lotka-Volterra Model



Reactions on Surfaces

Gray Scott system

$$U + 2V \rightarrow 3V,$$
$$V \rightarrow P$$



Deterministic

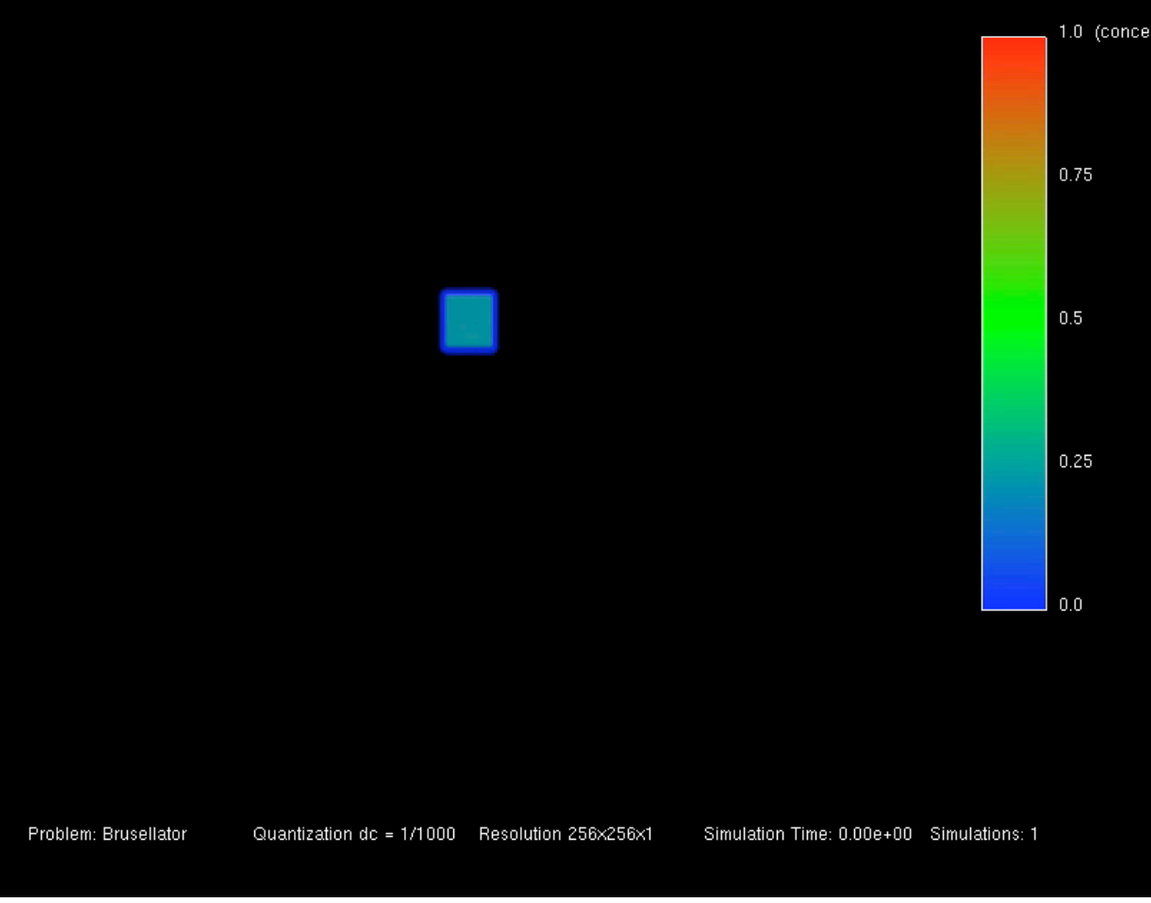
The Spatial Gray-Scott Model

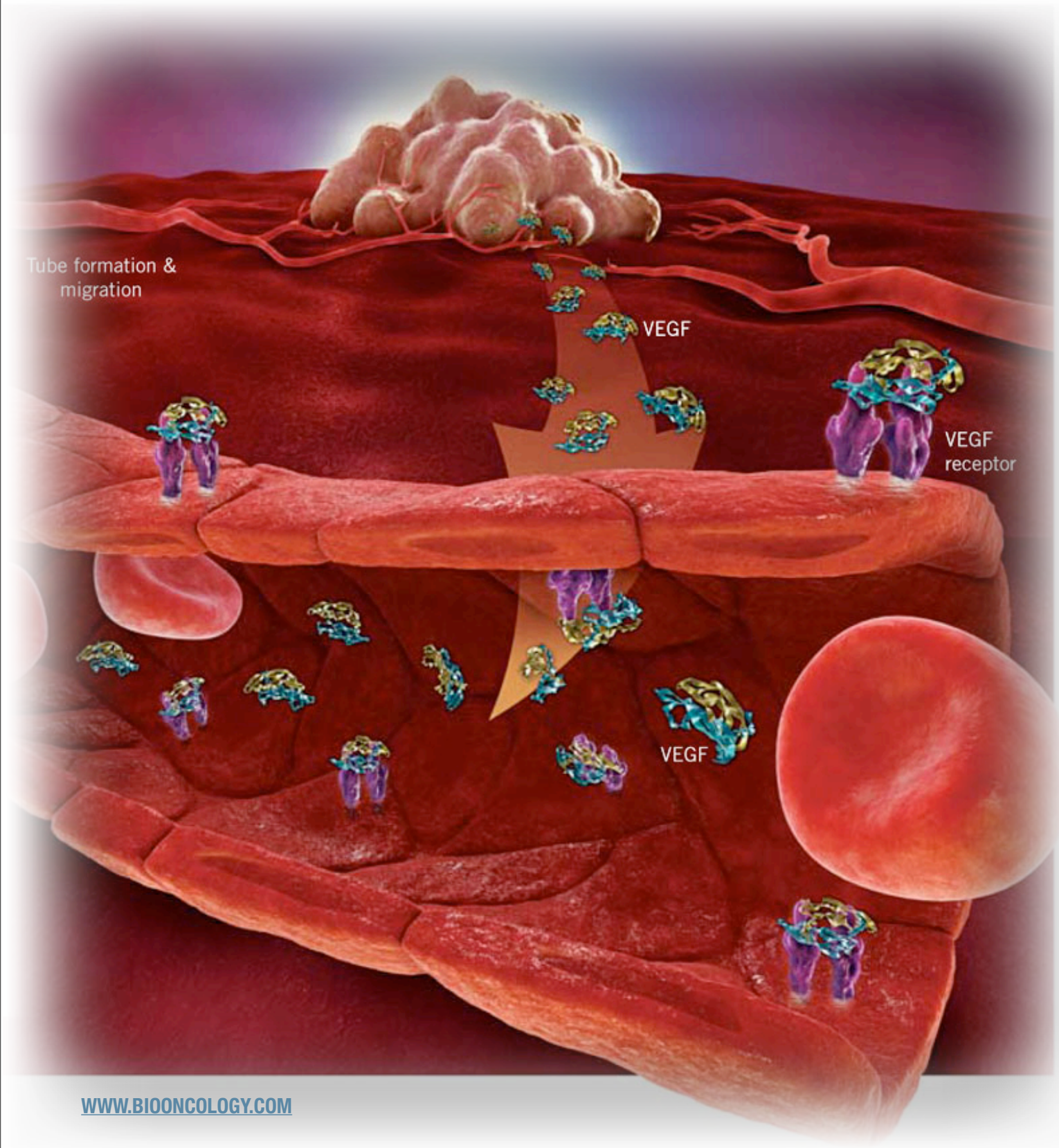
$$U + 2V \quad \rightarrow 3V,$$

Deterministic

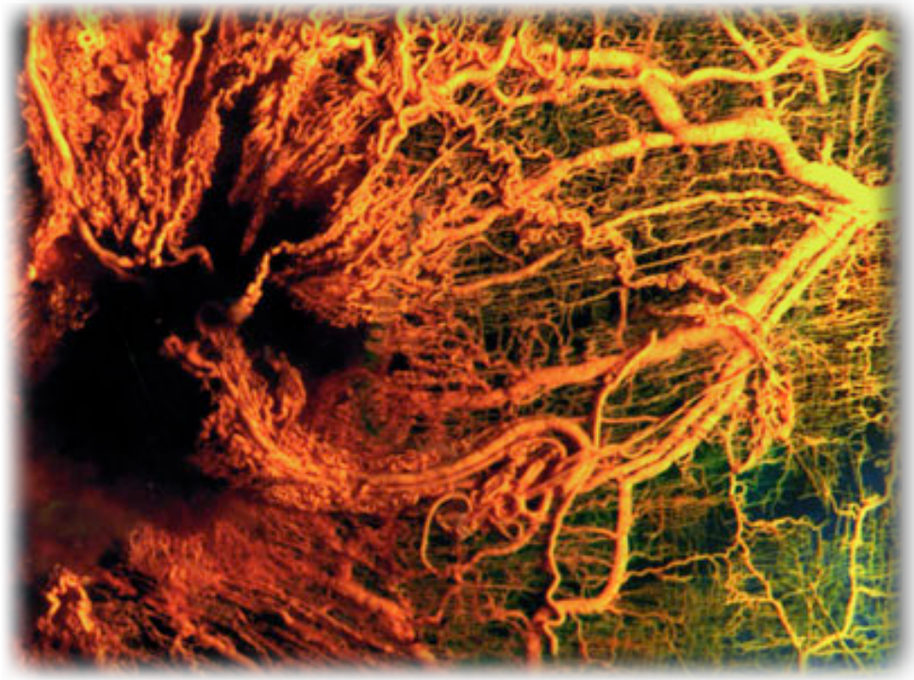
$$V \quad \rightarrow P$$

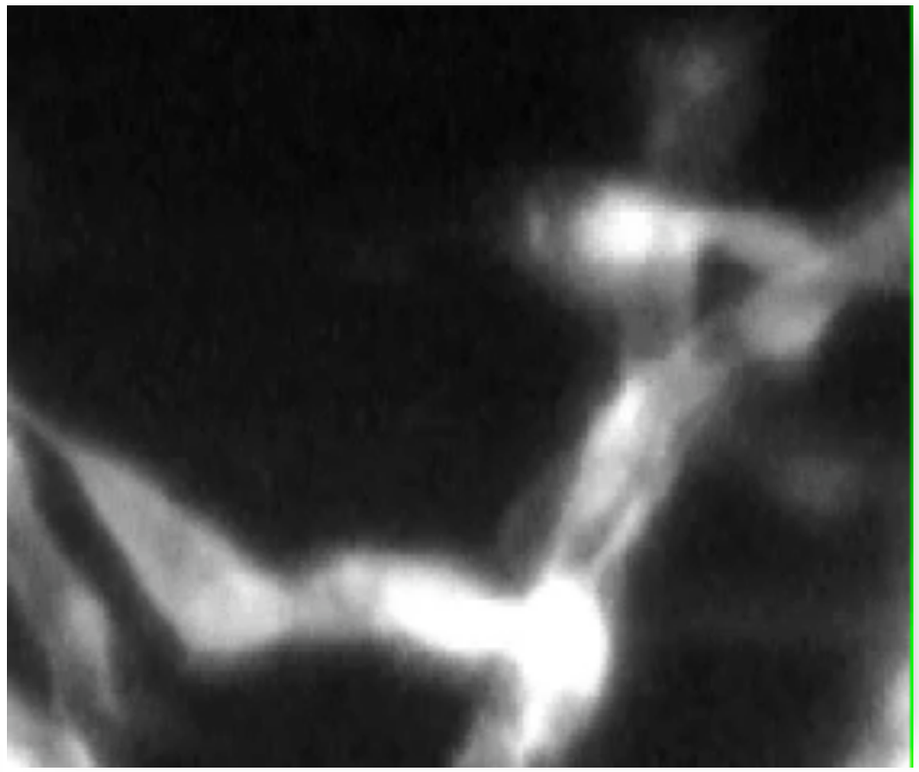
Stochastic





WWW.BIOONCOLOGY.COM

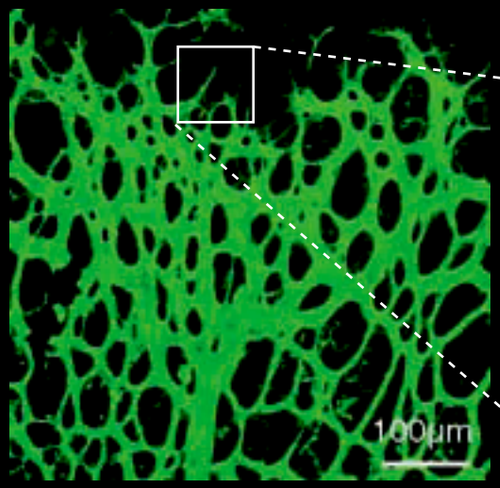




CRANIAL VESSEL ANGIOGENESIS IN ZEBRAFISH
HTTP://ZFISH.NICHD.NIH.GOV/ZFATLAS/FLI-GFP/FLI_MOVIES.HTML

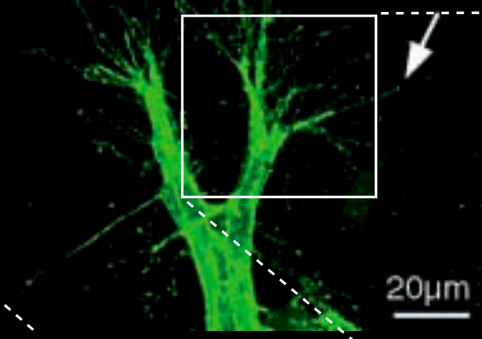
Example of Deterministic Models : **Angiogenesis**

Tumor-Induced Angiogenesis



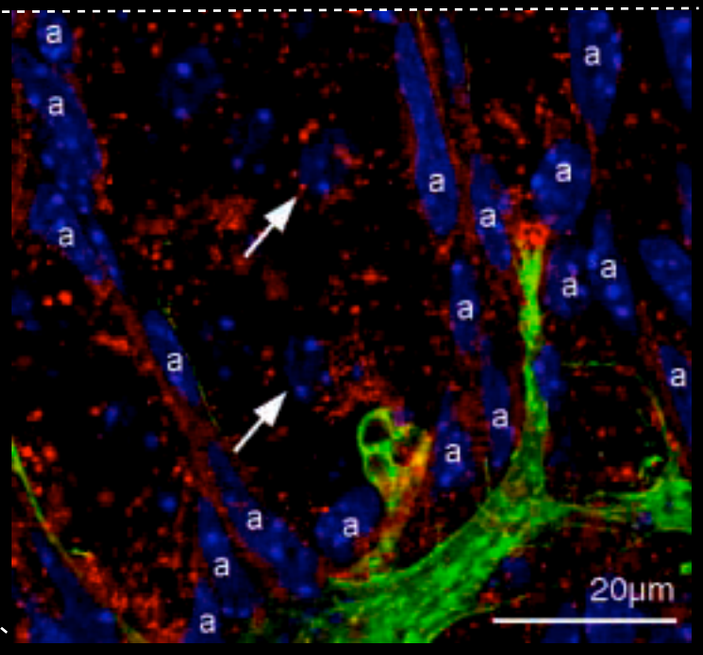
Tissue

Vessel Network



Cellular

Filopodia



Molecular

Growth factors

A Model of Sprouting Angiogenesis

Mechanism:

endothelial cells migrate towards source of growth factors

- form cords
- proliferate
- branch / fuse

Growth factor: VEGF

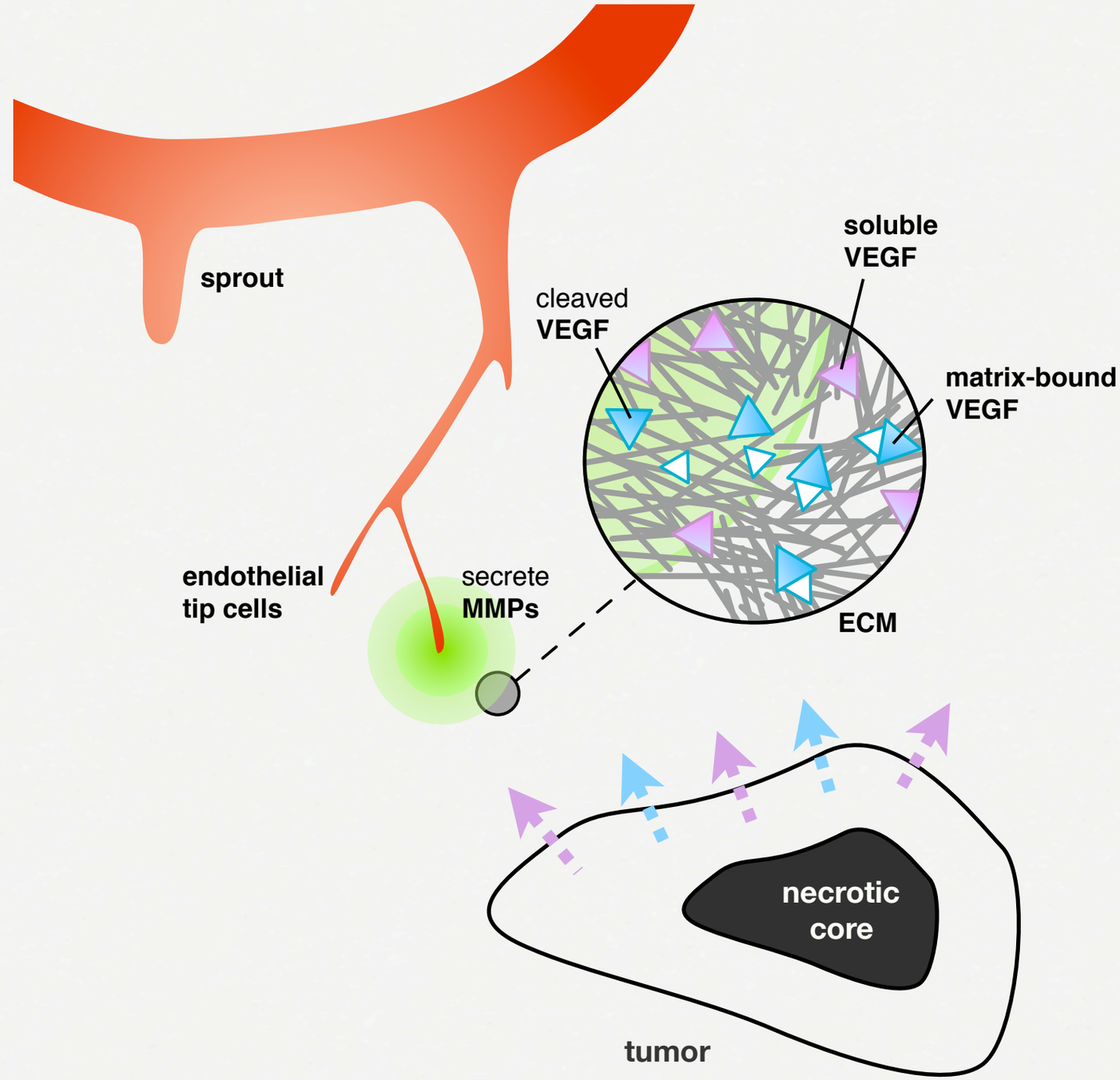
exists in two forms:

- soluble
- bound to the matrix (bVEGF)

Release of bVEGF

endothelial cells secrete proteinases

proteinases cleave bVEGF → soluble



Advantages of an explicit ECM: Angiogenesis

Sprouting angiogenesis:

formation of new blood vessels from existing ones
initiated by tumors with low nutrient supply

Mechanism:

endothelial cells migrate towards source of growth factors

- form cords
- branch
- proliferate

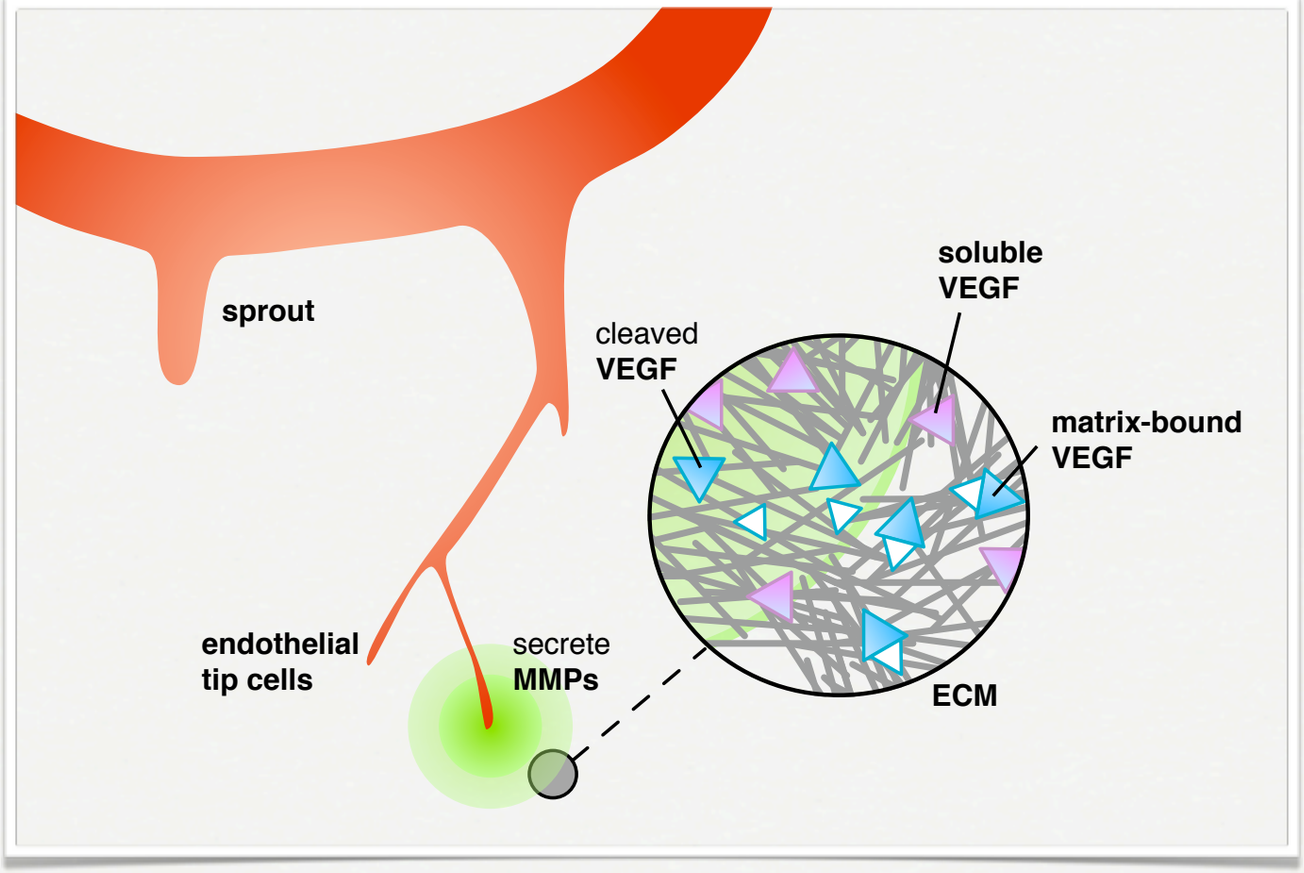
Growth factor: VEGF

exists in two forms:

- soluble
- bound to the matrix (bVEGF)

Release of bVEGF

endothelial cells secrete proteinases
proteinases cleave bVEGF → soluble



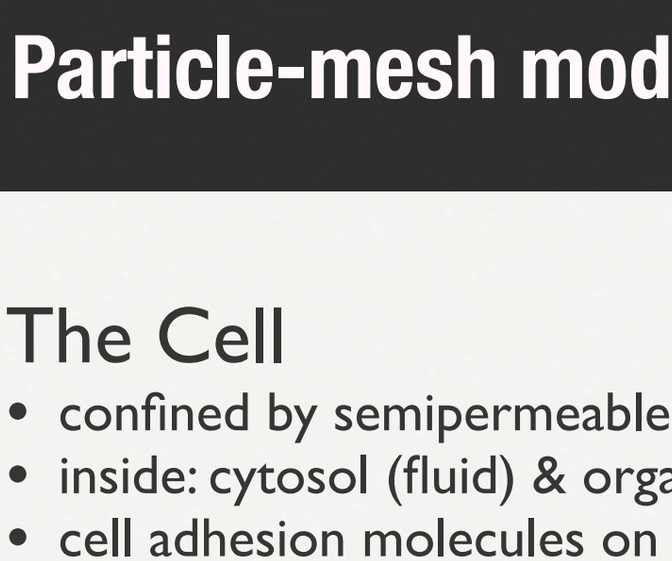
Particle-mesh models for mesenchymal motion / PM4

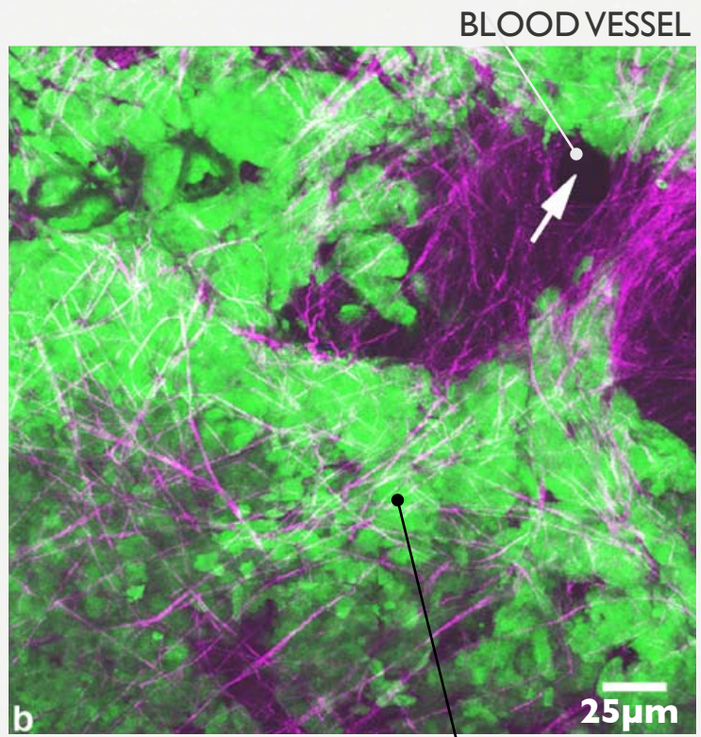
The Cell

- confined by semipermeable membrane
- inside: cytosol (fluid) & organelles
- cell adhesion molecules on the membrane
- extends filopodia for sensing

Extracellular Matrix

- fibrous proteins
- gels of polysaccharides
- sticky scaffolding
- structural support



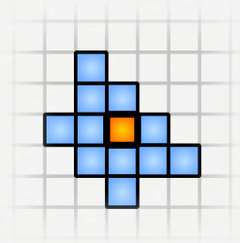


[1] M. SIDANI, J. WYCKOFF, C. XUE, J. E. SEGALL, AND J. CONDEELIS. PROBING THE MICROENVIRONMENT OF MAMMARY TUMORS USING MULTIPHOTON MICROSCOPY. *JOURNAL OF MAMMARY GLAND BIOLOGY AND NEOPLASIA*, V11(2):151–163, 2006.

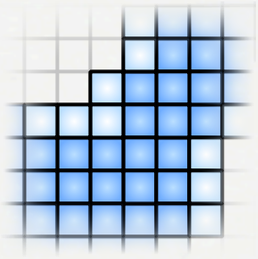
Representing Cells:

About scale:

Cellular Potts
- shape optimization
- interaction energies



Cellular automaton
- intuitive
- behavioral rules
- one “cell” = one cell



Continuum
- cell density (= no individuals)
- PDEs

Continuum modeling of cells

Primary implications:

Cell density: $\rho(\boldsymbol{x}, t)$

$$\frac{\partial \rho}{\partial t} = -\nabla(\boldsymbol{u}\,\rho) + k\,\rho$$

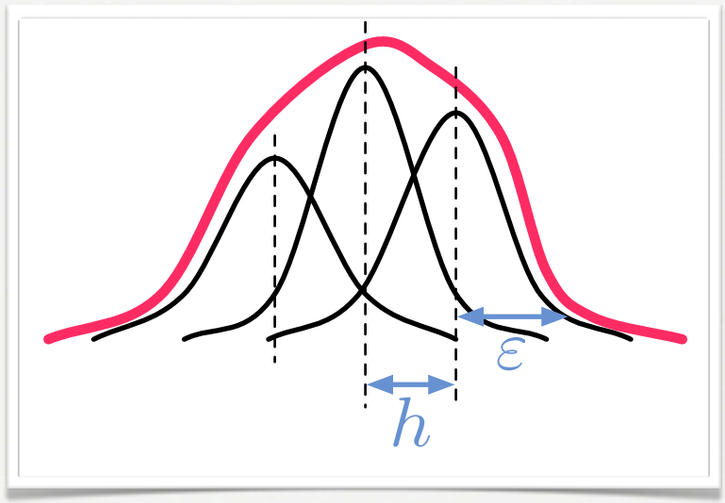
MIGRATION PROLIFERATION

Simulations using Particles

Function approximation:

$$q(\boldsymbol{x},t) = \sum_p Q_p(t)\,\zeta^h\big(\boldsymbol{x} - \boldsymbol{x}_p(t)\big)$$

PARTICLE WEIGHTS — PARTICLE KERNEL — PARTICLE POSITIONS



$$\int \zeta\, \boldsymbol{x}^\alpha\, dx = \mathbf{0}^\alpha \qquad 0 \le \alpha < r \qquad \int \zeta\, |\boldsymbol{x}|^r\, dx < \infty$$

Discretization of Lagrangian form: $\dfrac{Dq}{Dt} = \mathcal{L}(q,\boldsymbol{x},t)$ $\left(\dfrac{\partial q}{\partial t} + \nabla\cdot(\boldsymbol{u}\,q) = \mathcal{L}(q,\boldsymbol{x},t)\right)$

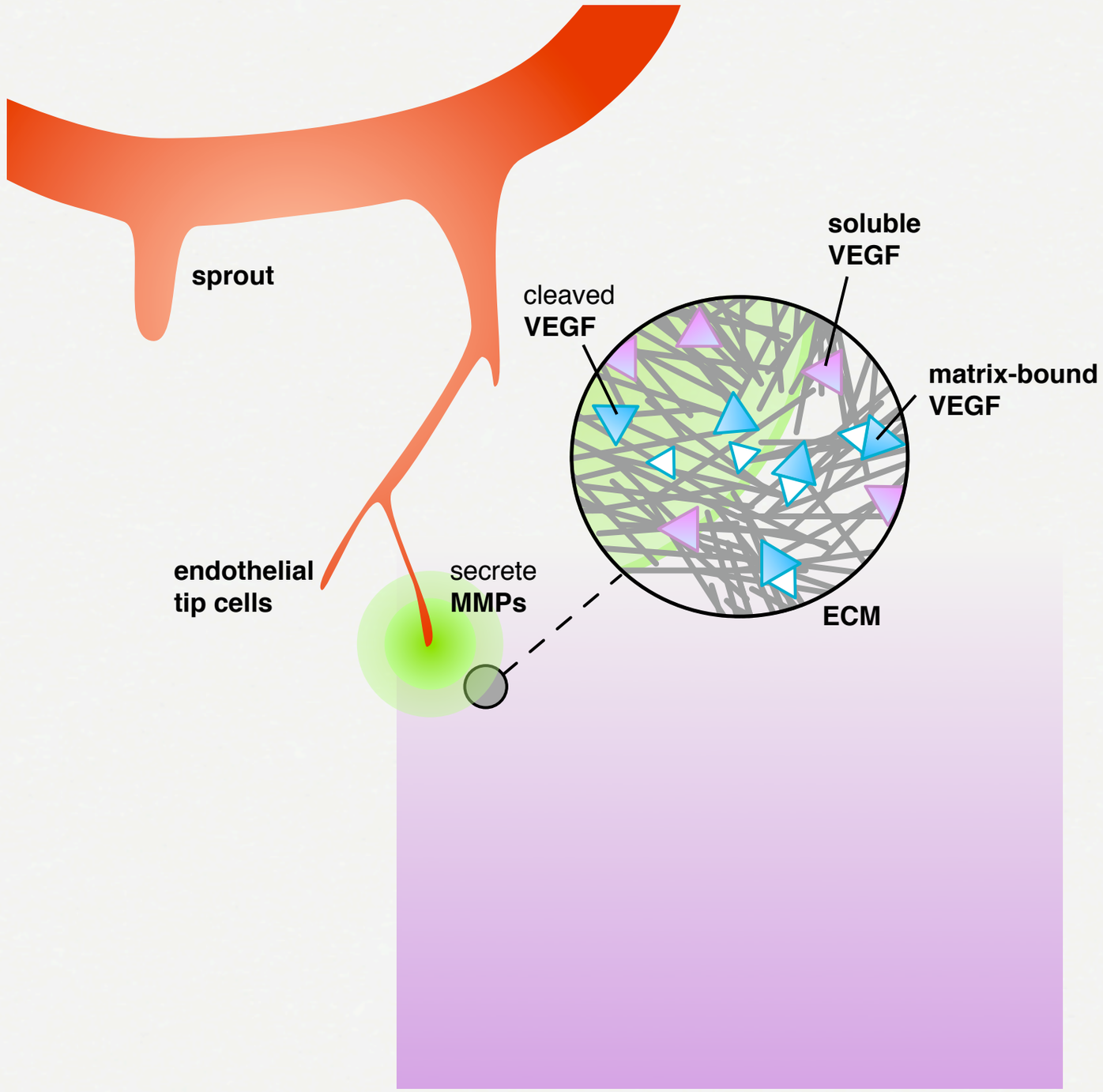
$\dfrac{d\mathbf{x}_p}{dt} = \mathbf{u}(\mathbf{x}_p,t)\,,$	positions	initial values on lattice
$\dfrac{dv_p}{dt} = v_p\,(\nabla\cdot\mathbf{u})\,(\mathbf{x}_p,t)\,,$	volumes	$v_p = h^d$
$\dfrac{dQ_p}{dt} = v_p\,\mathcal{L}^{\varepsilon,h}(q,\mathbf{x}_p,t)\,.$	weights	$Q_p = q(\boldsymbol{x}_p,0)\,v_p$

→ no linear stability constraints = no CFL (dt<dx/u) condition

Model System

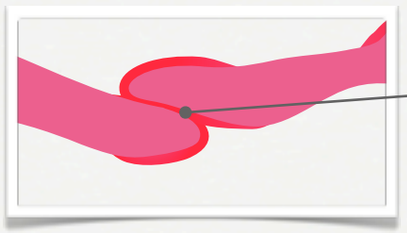
Model Components:

- Endothelial Cells (ECs)
- Extra-cellular Matrix (ECM)
- Growth Factors (VEGF)
- Matrix Metalloproteases (MMPs)



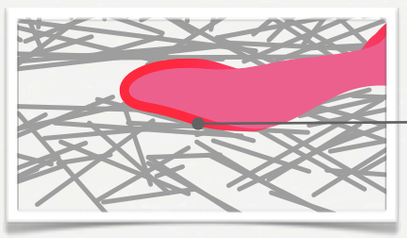
Elements of PM4

The elements of migration



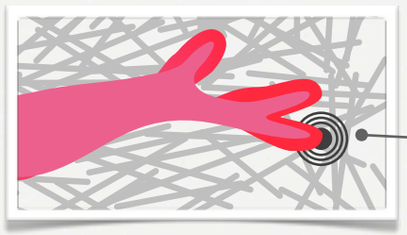
cells stick to cells

transmembrane CAMs: cadherin, ICAM-1, ...
formation of **clusters**, cords



cells guided by the extracellular matrix

transmembrane CAMs: integrins,...
facilitates **migration**

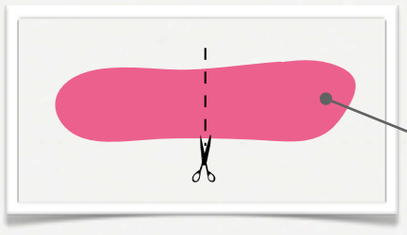


cells secrete proteinases

Matrix metalloproteinases: degrade matrix,
free matrix-bound growth factors

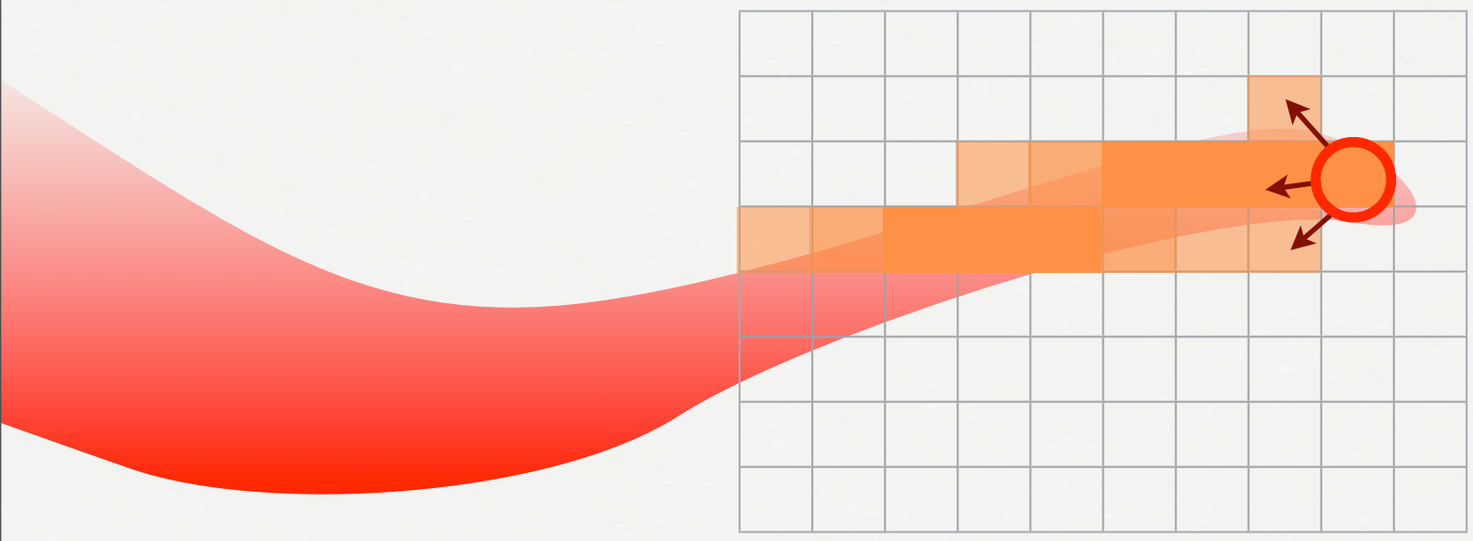
cells sense chemical gradients

gradients of "chemoattractant" serve as migratory cues

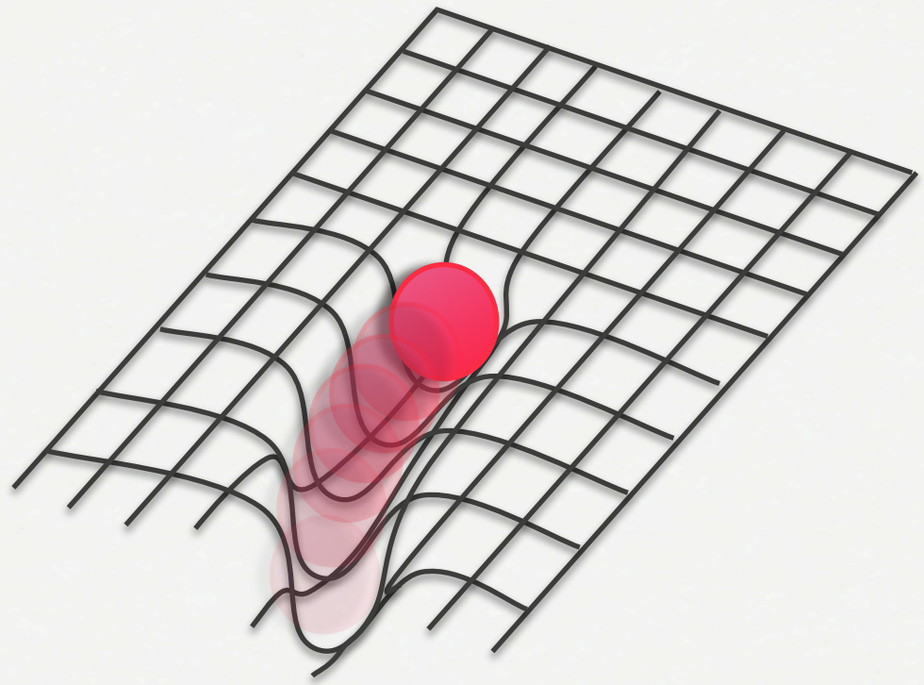


cells proliferate

Endothelial Cell representation



Tip Cell "deposes" endothelial cells



Hybrid representation of ECs:

Tip cell particles Q_p:

- Discrete particle representation
- Particle location: $\boldsymbol{x}_p$
- Migration acceleration: $\boldsymbol{u}_p$
- Drag coefficient: λ

$$\frac{\boldsymbol{x}_p}{\partial t} = \boldsymbol{u}_p, \;\; \frac{\boldsymbol{u}_p}{\partial t} = \boldsymbol{a}_p - \lambda \boldsymbol{u}_p$$

Stalk cell density ρ:

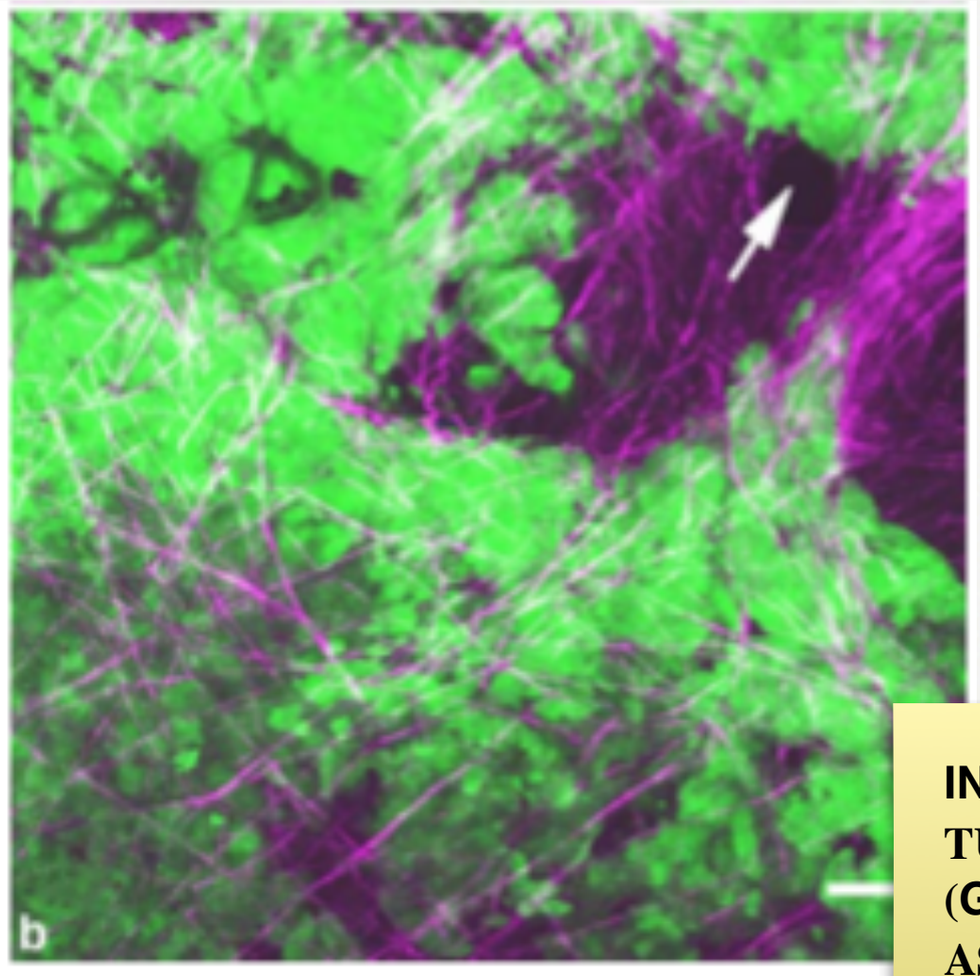
- Continuum vessel representation
- Tip and stalk communicate through Particle-Mesh, Mesh-Particle interpolations

$$\rho_{\boldsymbol{i}}^{n+1} = max\left(\rho_{\boldsymbol{i}}^{n}, \sum_p B(\boldsymbol{i}\,h - \boldsymbol{x}_p)\,Q_p\right)$$

$$Q_p = \sum_{\boldsymbol{i}} h^3 q_{\boldsymbol{i}} M_4'\left(\boldsymbol{x}_p - \boldsymbol{i}h\right)$$

Extracellular Matrix : Structure

- Material occupying the space between cells
- Fibers of structural glycoproteins
  (collagen, laminin and fibrillin are distributed throughout the ECM, occupying ~30% of the ECM)
- Collagens (the main component of the ECM cross-link with neighbouring collagens to form bundles)

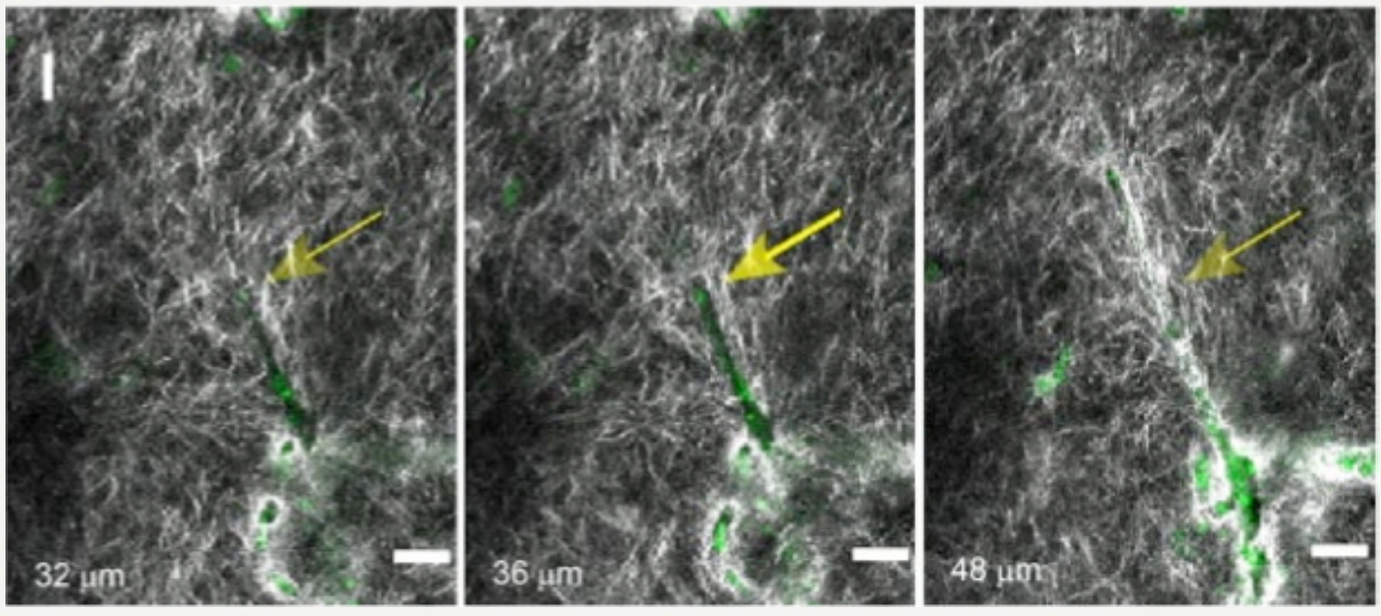


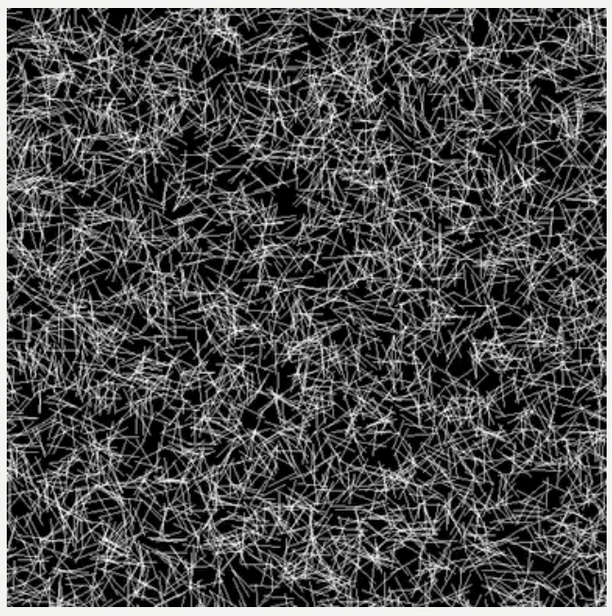
IN M
TUM
(GRE
ASS
WIT
FIBE

[3] M. SIDANI, J. WYCKOFF, C. XUE, J. E. SEALL, AND J. CONDEELIS. PROBING THE MICROENVIRONMENT OF MAMMARY TUMORS USING MULTIPHOTON MICROSCOPY. *J. MAMMARY GLAND BIOL. NEOPLASIA, V11(2):151-163, 2006*

Extracellular Matrix (ECM)

- Fibrous structures in ECM provide a guiding structure for migrating endothelial cells
- ECM fibers are subject of remodeling by migrating EC's
- The ECM expresses binding sites for various growth factors and integrins





[4] N. D. KIRKPATRICK, S. ANDREOU, J. B. HOYING, AND U. UTZINGER. LIVE IMAGING OF COLLAGEN REMODELING DURING ANGIOGENESIS. *AJP HEART.*. PAGES 0124.2006-,2007

Modeling the Matrix:

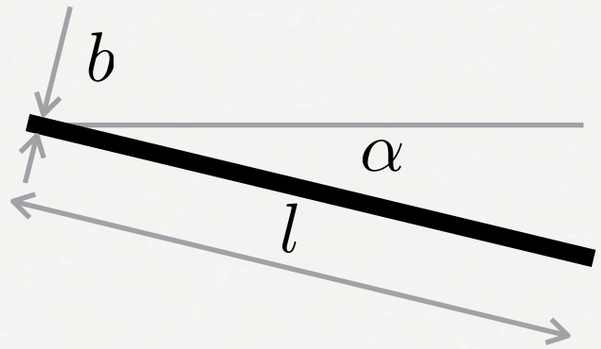
Model matrix explicitly:

- structure: collection of fiber bundles
- function: cell-matrix adhesion sites

Fibers:

- straight
- random direction
- distribution of lengths

$$l = l_0\, 2^{m\, z}$$
$$\alpha \in \mathcal{U}([0, \pi])$$
$$z \in \mathcal{N}(0, 1)$$



Indicator field $\mathcal{E}$

- unity where fibers present
- smoothed (**implicit** filopodia)

Chemotaxis & cell-matrix adhesion

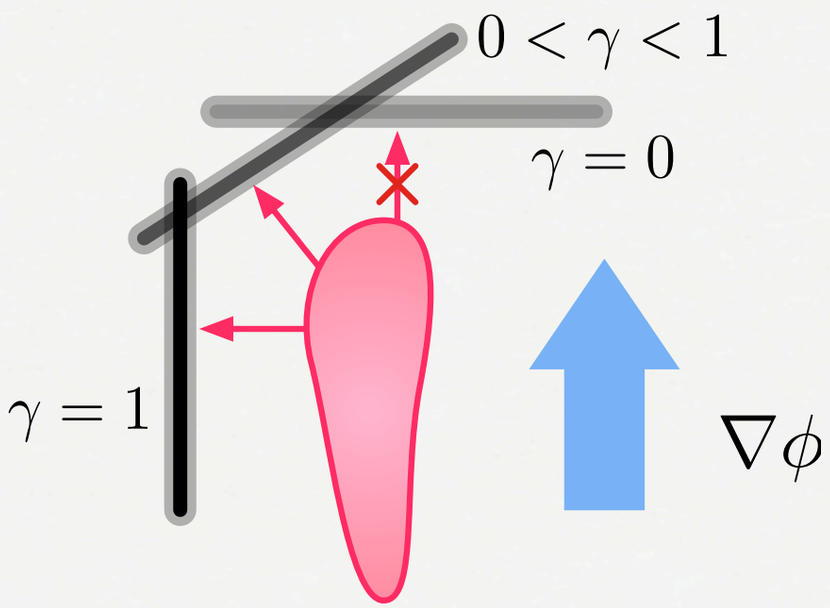
Opportunistic: get to growth factor (GF) source

Existing models: $\boldsymbol{a}_\phi = \nabla\phi$

PM4:

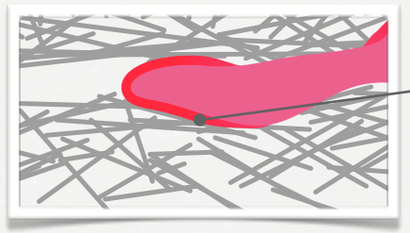
$$\boldsymbol{a}_{\mathrm{ecm},\phi} = \left[\left(1 - \left|\frac{\nabla e}{|\nabla e|} \cdot \frac{\nabla\phi}{|\nabla\phi|}\right|\right)\nabla e + \nabla\phi\right](e + e_o)(\rho_{\mathrm{cpd}} - e)$$

- γ CLING TO FIBER AN ADVANTAGE?
- WHERE IS THE FIBER?
- WHERE IS THE GF SOURCE?
- FIBERS FACILITATE MIGRATION
- TOO MANY FIBERS BLOCK MIG. PATH



Tip Cell Migration

The elements of migration



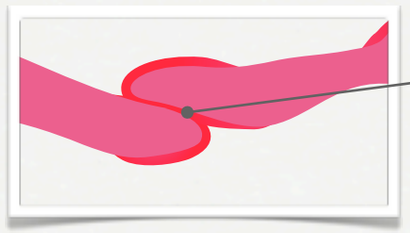
cells are guided by extracellular matrix

transmembrane CAMs: integrins,...)

facilitates migration

cells sense chemical gradients

gradients of "chemoattractant" serve as migratory cues



cells stick to cells

gradient of "haptotactic" molecules serve as migration cues

Migration Speed

$$\mathbf{a} = \overline{\alpha\left(E_\rho\right)\underline{\mathbf{T}}}\left(\overline{w_V \nabla\Psi} + \overline{w_F \nabla\Phi_b}\right)$$

Growth Factors: Assumptions

- We model only one representative growth factor (VEGF)
- VEGF exists in a soluble and a matrix bound isoform
- Soluble VEGF is released from a tumor source
- Unbound VEGF diffuses through the ECM
- VEGF is subject to uptake by endothelial cells
- decays naturally

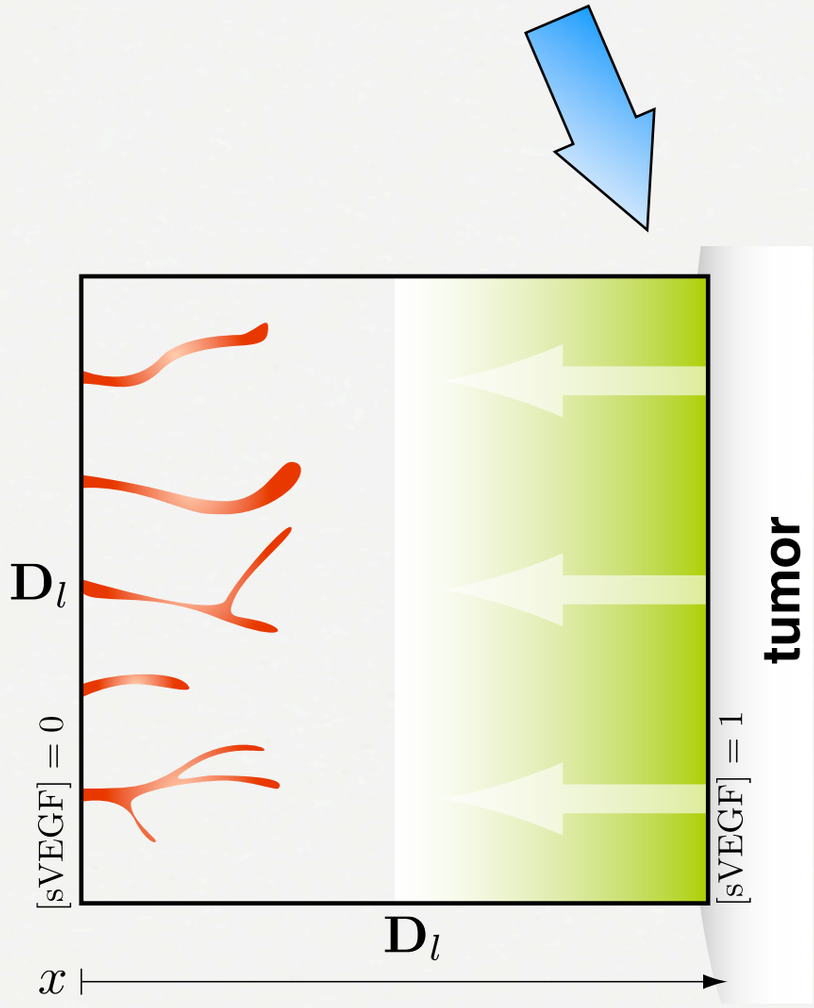
Soluble VEGF (sVEGF) - Assumptions

- Model : One VEGF isoform in soluble and bound state
- sVEGF establishes global chemotactic gradient

  - Tumor source modeled by boundary conditions
  - sVEGF diffuses through ECM
  - Uptake of sVEGF by endothelial cells ρ
  - Subject of natural decay

$$\frac{\partial[\mathrm{sVEGF}]}{\partial t} = k_V \nabla^2[\mathrm{sVEGF}] - U([\mathrm{sVEGF}], \rho) - \delta_V[\mathrm{sVEGF}]$$

$$U([\mathrm{sVEGF}], \rho) = min([\mathrm{sVEGF}], \upsilon_V \rho)$$

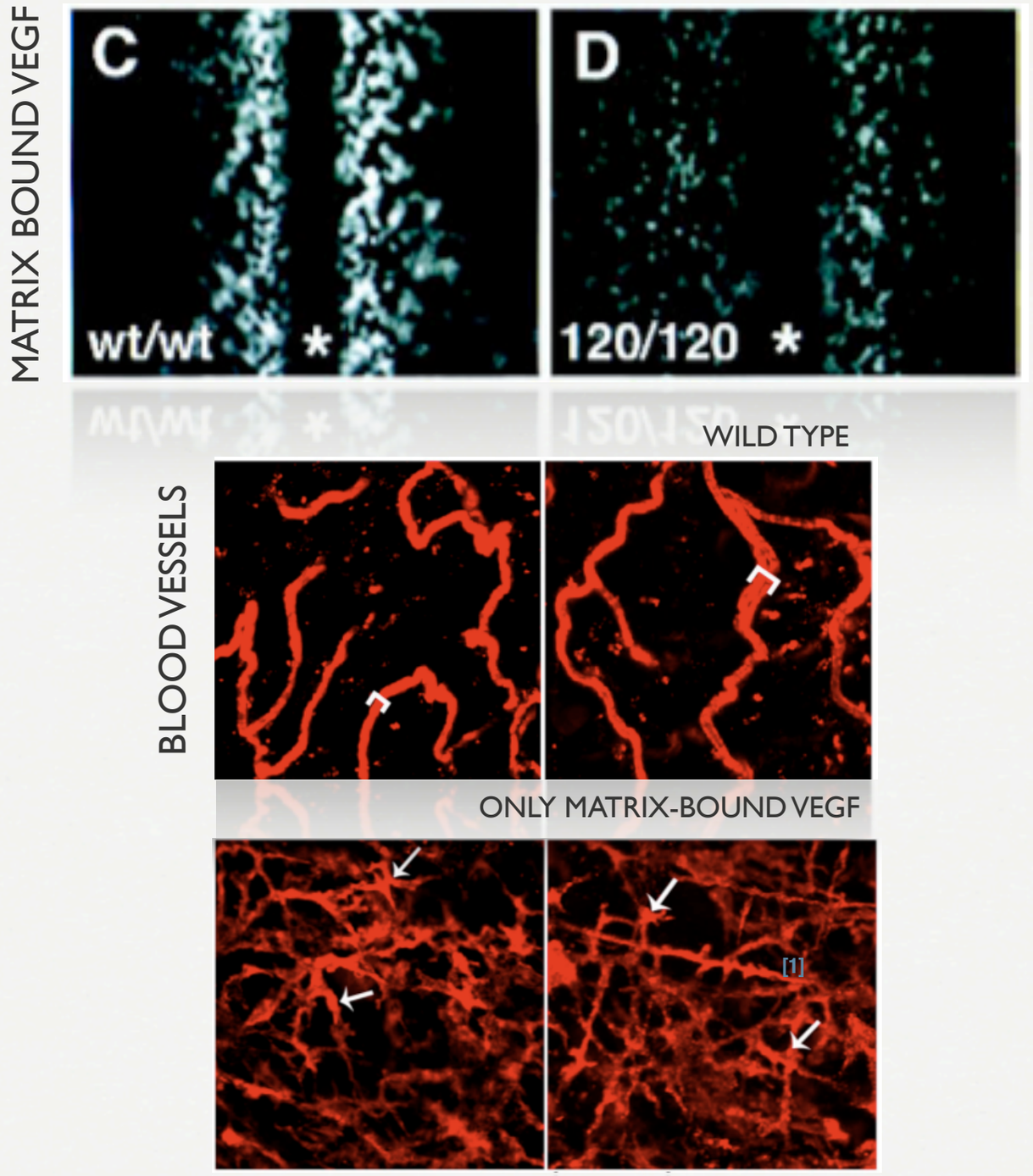


Matrix-bound VEGF (bVEGF)

- Some VEGF isoforms express heparin-binding sites binding to domains in the ECM

- Local gradients of matrix bound VEGF influence sprout morphology

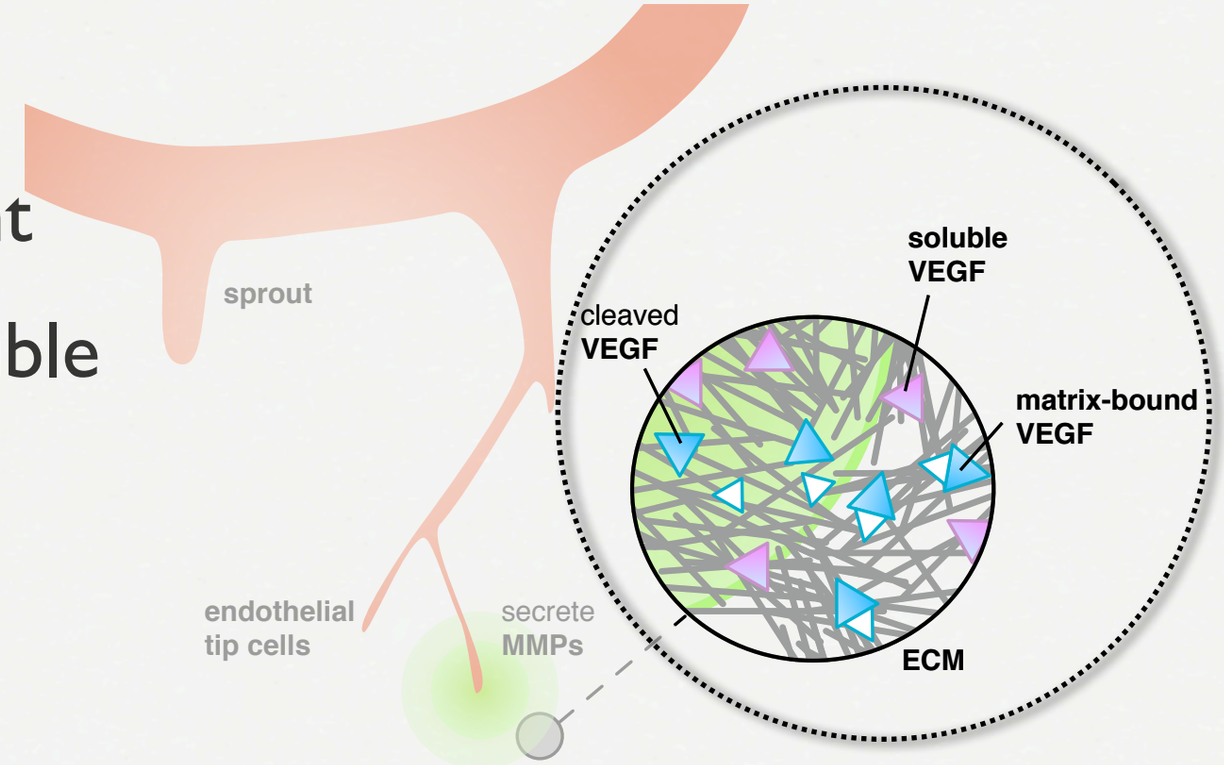
- Matrix bound VEGF is cleaved by MMPs released at endothelial sprout tips



[1] C. RUHRBERG, H. GERHARDT, M. GOLDING, R. WATSON, S. IOANNIDOU, H. FUJISAWA, C. BETSHOLTZ AND D. T. SHIMA. SPATIALLY RESTRICTED PATTERNING CUES PROVIDED BY HEPARIN-BINDING VEGF-A CONTROL BLOOD VESSEL BRANCHING MORPHOGENESIS. *GENES DEV.*, 16(20):2684-2698, 2002.
[2] S. LEE, S. M. JILAI, G. V. NIKOLOVA, D. CARPIZO, AND M. L. IRUELA-ARISPE. PROCESSING OF VEGF-A BY MATRIX METALLOPROTEINASES REGULATES BIOAVAILABILITY AND VASCULAR PATTERNING IN TUMORS. *J. CELL BIOL.*, V42(3):195-238, 2001

Matrix-bound VEGF - Assumptions

- Initially distributed in pockets
- establishes local chemotactic gradient
- cleaved VEGF (cVEGF) becomes soluble
  - bVEGF is cleaved by MMPs
  - Uptake of cVEGF by ECs ρ
  - cVEGF diffuses through ECM
  - cVEGF is subject to natural decay



$$\frac{\partial[\mathrm{bVEGF}]}{\partial t} = -C\left([\mathrm{bVEGF}],[\mathrm{MMP}]\right) - U\left([\mathrm{bVEGF}],\rho\right)$$

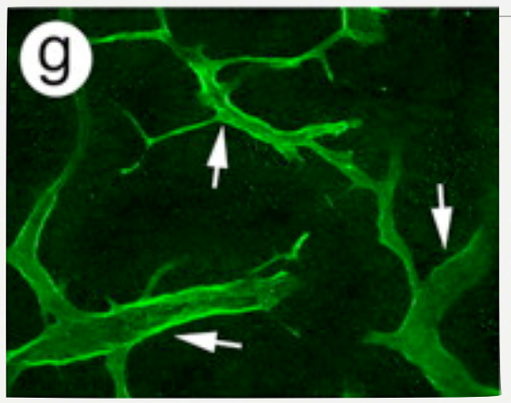
$$C\left([\mathrm{bVEGF}],[\mathrm{MMP}]\right) = min\left([\mathrm{bVEGF}], \upsilon_{bV}[\mathrm{MMP}][\mathrm{bVEGF}]\right)$$

$$\frac{\partial[\mathrm{cVEGF}]}{\partial t} = k_V \nabla^2[\mathrm{cVEGF}] + C\left([\mathrm{bVEGF}],[\mathrm{MMP}]\right) - U\left([\mathrm{cVEGF}],\rho\right) - \delta_V[\mathrm{cVEGF}]$$

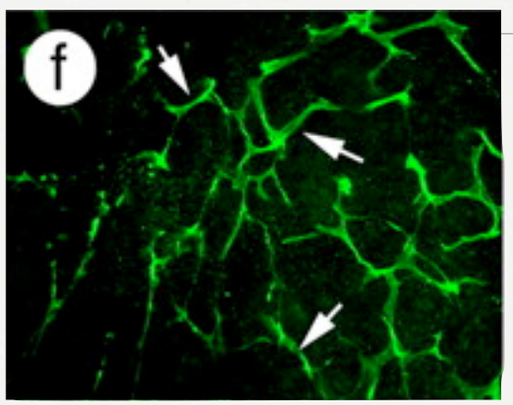
Angiogenesis: Post-dicting Experiments

Matrix-bound VEGF leads to increased branching.
vessel branching ↔ capillary function

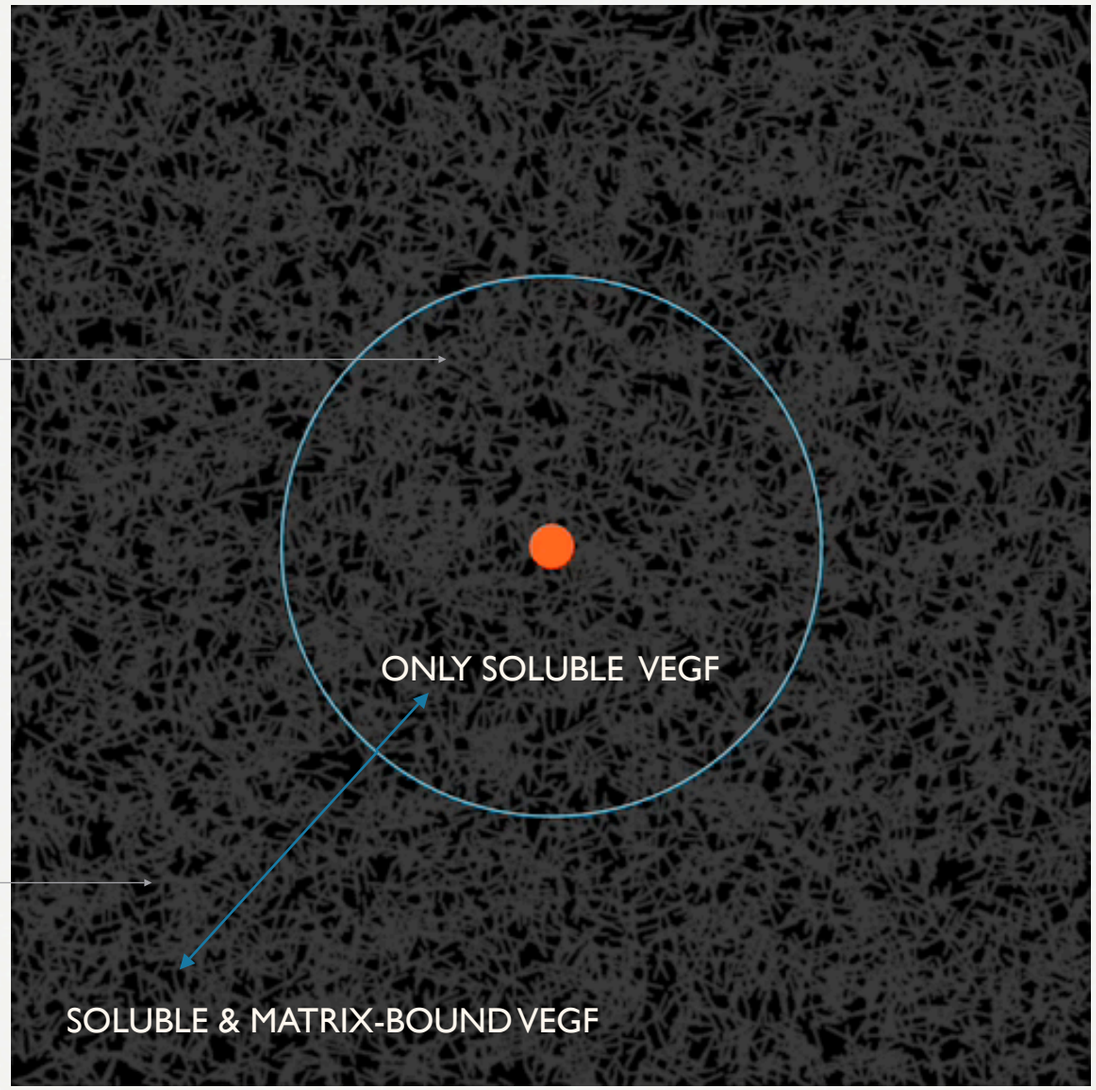
BLOOD VESSEL FORMATION IN A MOUSE MODEL



ONLY SOLUBLE VEGF
> THICKER VESSELS



SOLUBLE + MATRIX-BOUND VEGF
> INCREASED BRANCHING



RADIAL SOLUBLE VEGF GRADIENT AND
LOCALIZED MATRIX-BOUND VEGF

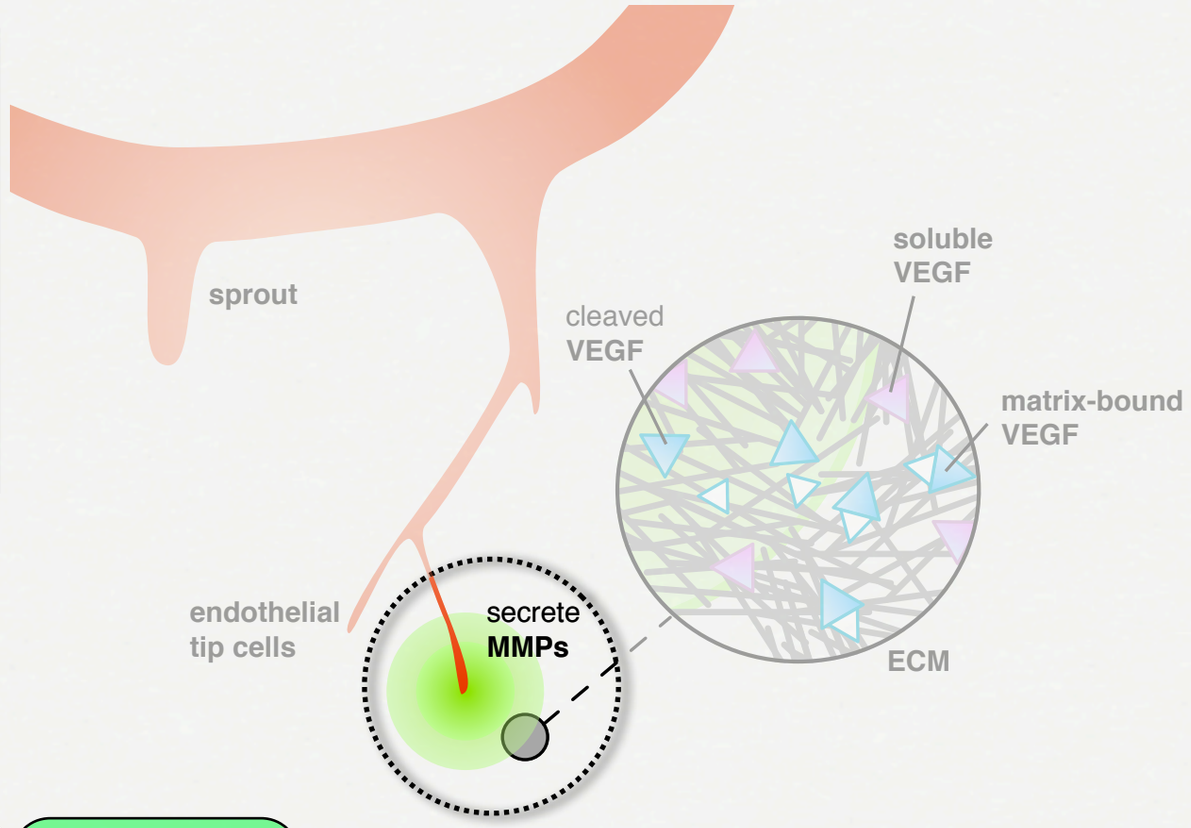
new: **branching is an output of the simulation**

[1] S. LEE, S. M. JILANI, G. V. NIKOLOVA, D. CARPIZO, AND M. L. IRUELA-ARISPE. PROCESSING OF VEGF-A BY MATRIX METALLOPROTEINASES REGULATES BIOAVAILABILITY AND VASCULAR PATTERNING IN TUMORS. *J. CELL BIOL.*, 169(4):681–691, 2005.

MATRIX METALLOPROTEINASES

- decreases local chemotactic gradients

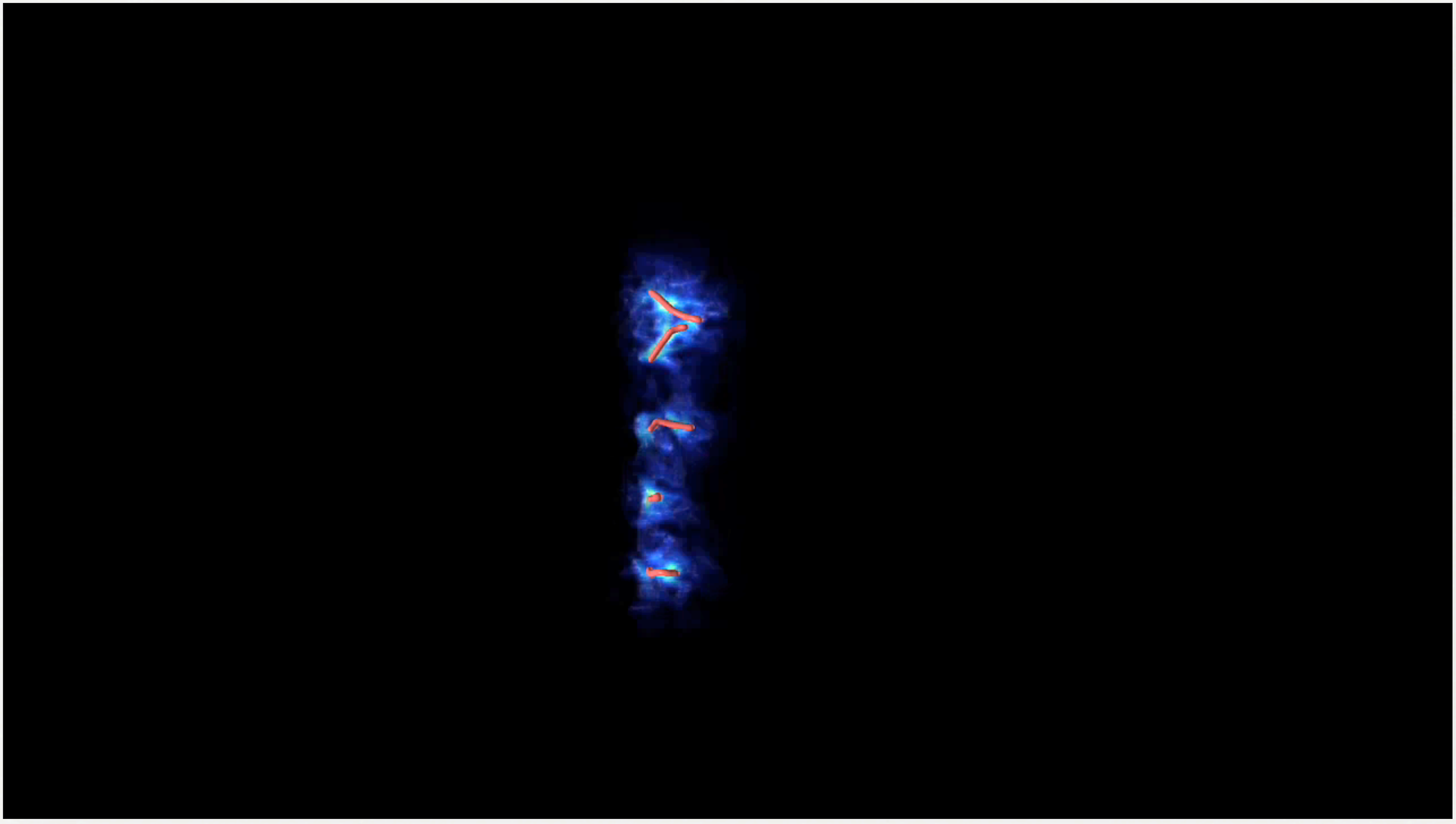
- RELEASED BY MIGRATING TIP-CELLS
- RELEASE BOUND BY THRESHOLD LEVEL
- DIFFUSE THROUGH ECM
- SUBJECT TO NATURAL DECAY



$$\frac{\partial[\mathrm{MMP}]}{\partial t} = k_M \nabla^2[\mathrm{MMP}] + \gamma_M G\left(M_{th}, [\mathrm{MMP}]\right)[\mathrm{EC}] - \delta_M[\mathrm{MMP}]$$

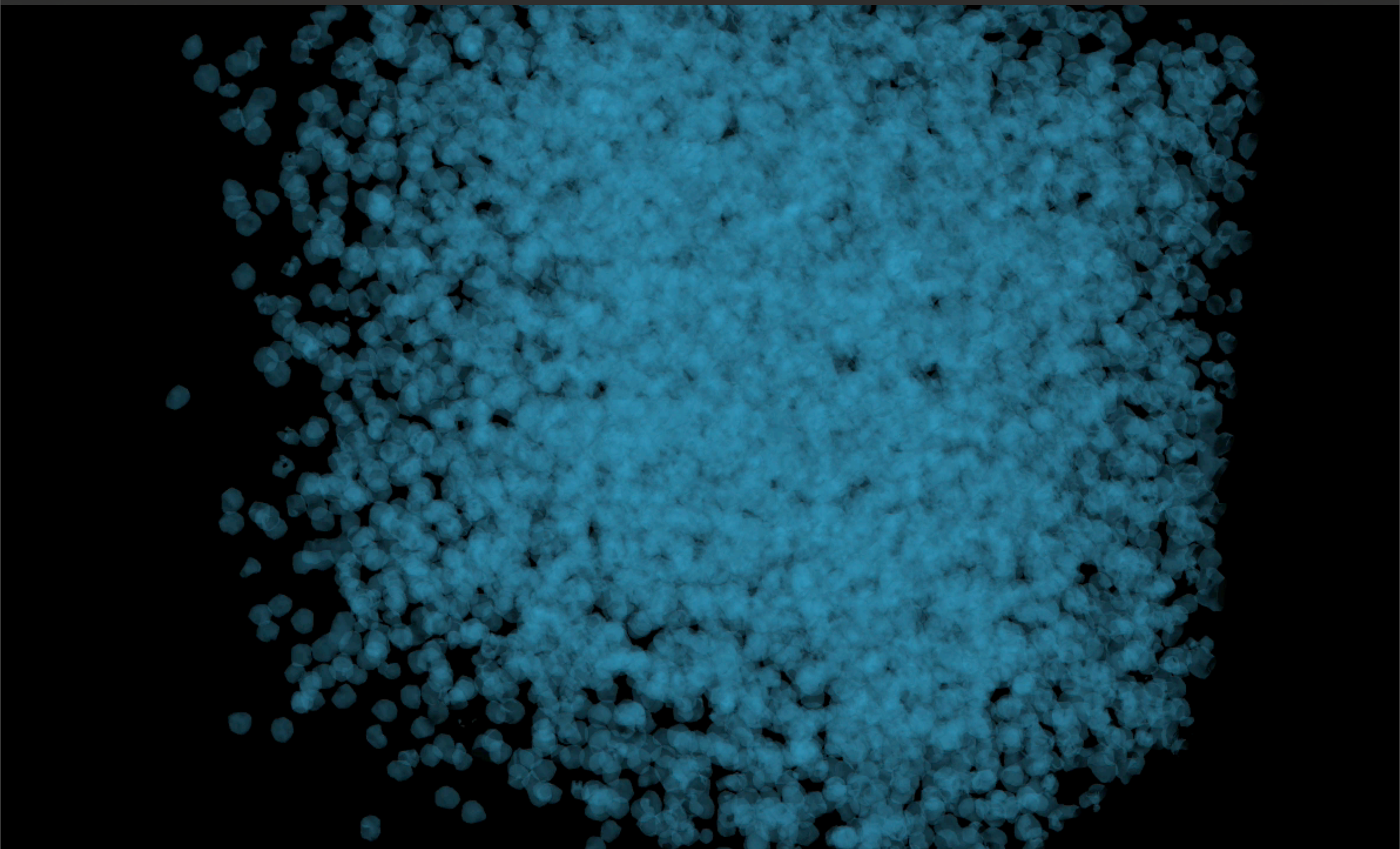
$$G\left(M_{th}, [\mathrm{MMP}]\right) = \frac{M_{th} - [\mathrm{MMP}]}{M_{th}}$$

TUMOR INDUCED ANGIOGENESIS

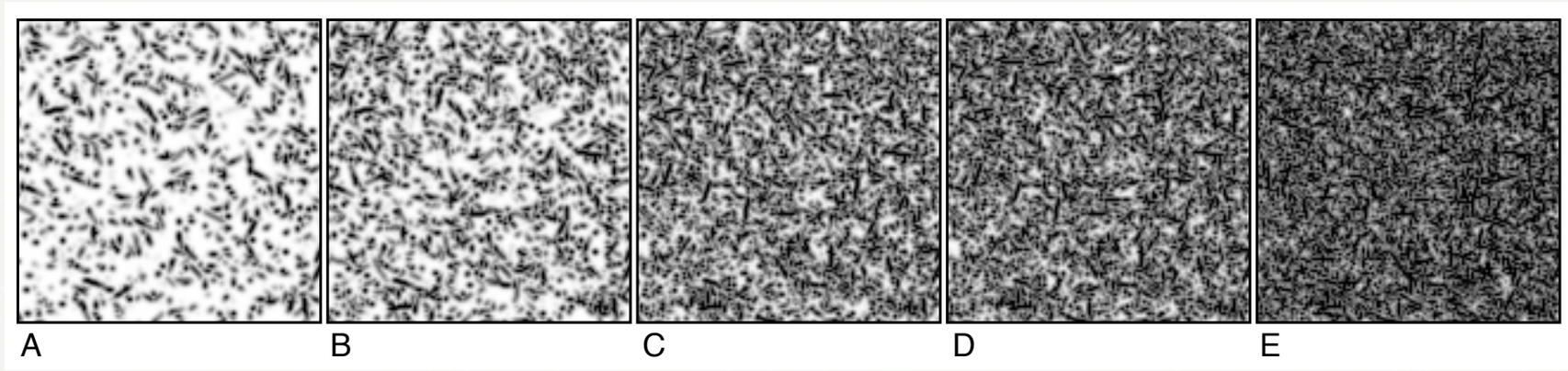


MILDE F., BERGDORF M. AND KOUMOUTSAKOS P., A HYBRID MODEL OF SPROUTING ANGIOGENESIS, BIOPHYSICAL J., 2008

Simulation II (bVEGF)



Extra-cellular Matrix density



FIBER DENSITY:

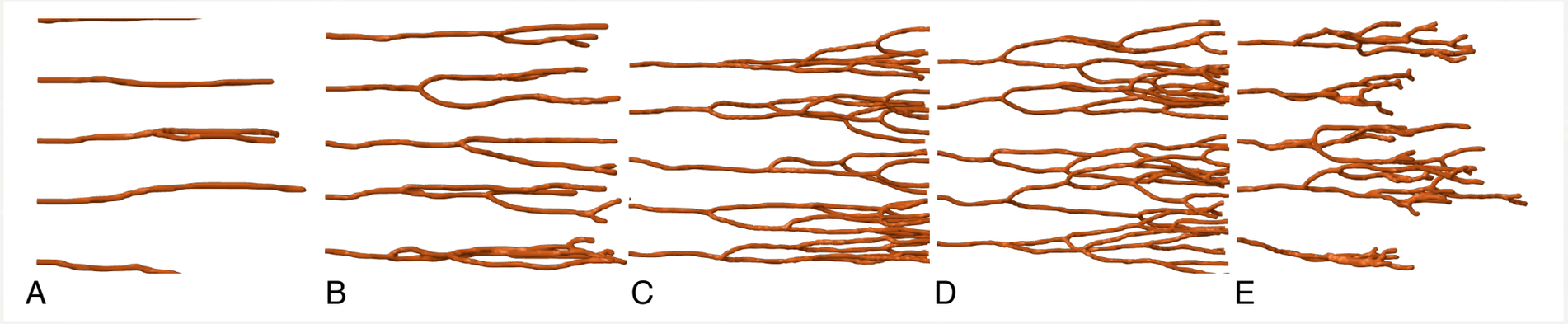
A: 6%,

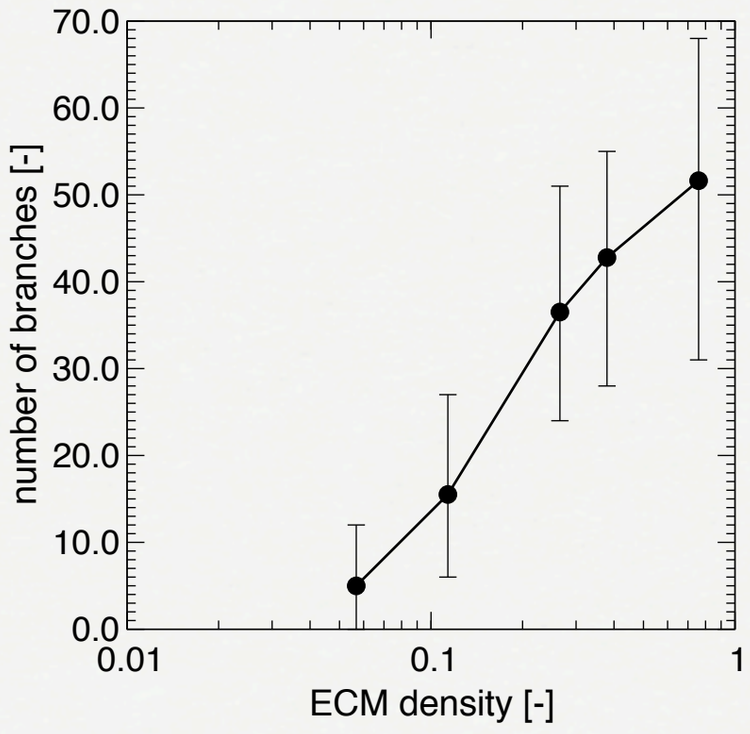
B: 11%,

C: 26%,

D: 48%

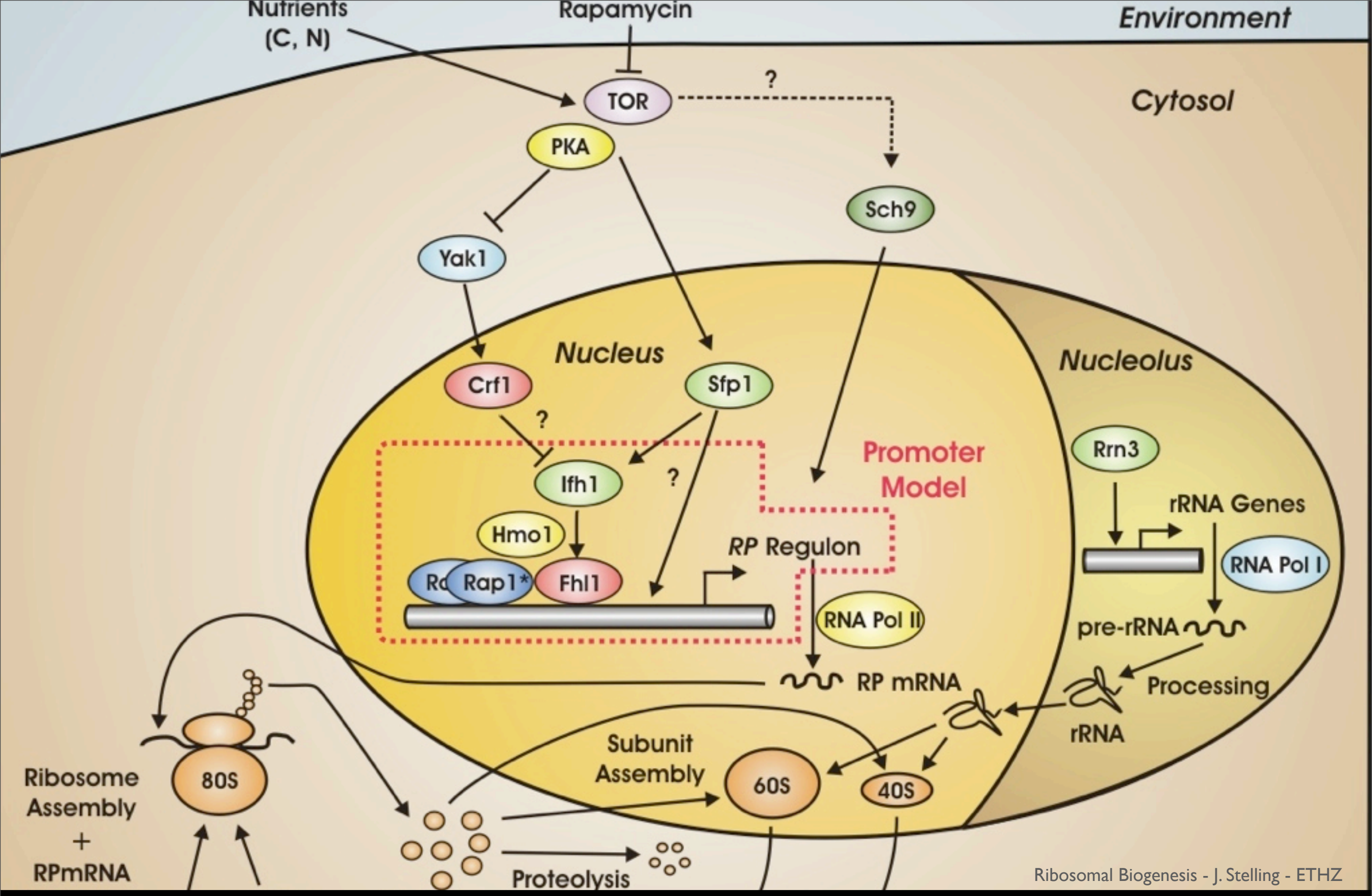
E:75%





NEXT STEPS

- flow through vessel network
- tip cell/stalk cell differentiation
- combine with tumor growth model
- signaling
- validation



Ribosomal Biogenesis - J. Stelling - ETHZ

Stochastic Simulation Algorithms

$$\sigma \leftrightarrow \sigma'$$

$$\sum A_i^j \leftrightarrow \sum B_k^j \qquad j = 1, \cdots, M$$

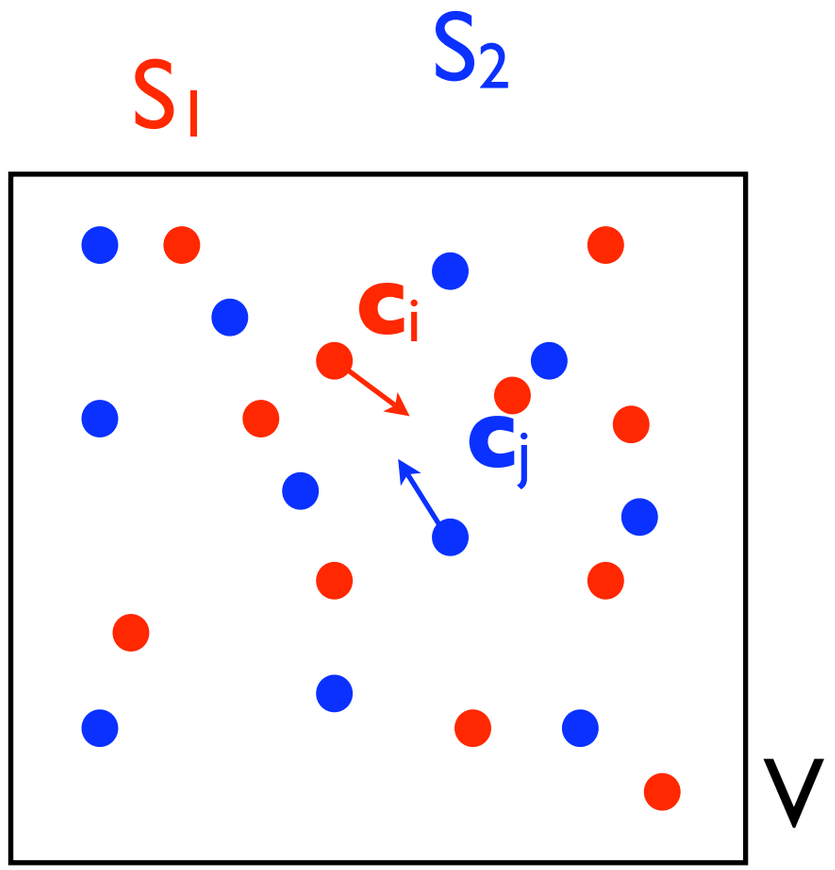
The Chemical Master Equation

$$\frac{dP(\sigma)}{dt} = \sum_{\sigma'} P(\sigma')\, G(\sigma' \to \sigma) - \sum_{\sigma'} P(\sigma)\, G(\sigma \to \sigma')$$

Accelerating Stochastic Simulation Algorithms

Chemical kinetics : Set-up

- Well stirred reaction volume V
- N different species $S_1, S_2, \ldots, S_N$ in numbers $X_1, X_2, \ldots, X_N$
- random collisions and reactions through M channels $R_1, R_2, \ldots, R_M$
- Experiment length T



Chemical Kinetics

- Macroscopic/Continuous approach
  - Species concentrations
  - Reaction rates ~ reactant concentrations product
  - ODEs for x_i...
  - Really result of a limiting process

- **Discrete approach**
  - Species concentrations are random variables
    - macroscopic approach only gives their expected value
    - with variance ~ V^{-2}
  - For small volumes, and small X_i, variance blows up

Stochastic Simulation Algorithm (SSA)

- **Gillespie 1977 -** Well-stirred volume V (example):
  - a single 2nd order reaction $A+B \rightarrow C$
  - Probability of A-B collision within $dt \sim X_A X_B dt$
  - Probability of reaction within dt is $k_{AB} X_A X_B dt$
  - Time until next A-B reaction

$$\tau \sim \mathcal{E}(1/a)$$
$$a = k_{AB} X_A X_B$$

Stochastic Simulation Algorithm

Gillespie,J. Comp. Phys. 1977

- For **M reactions**, time until any reaction

$$\tau \sim \mathcal{E}(1/a_0) \qquad a_0 = \sum_{j=1}^{M} a_j$$

- Reaction index : point-wise distribution $p(j = l) = \frac{a_l}{a_0}$

- One timestep:
  - Sample τ
  - Sample the index j
  - Update the X_i, $t=t+\tau$

exact BUT slow

- The SSA simulates every reaction event !

SSA : acceleration in time

Gillespie,J. Chem. Phys. 2001

- **τ-leaping** : several reaction events over one time step,

- Assumption : reaction propensities a_i remain essentially constant over τ, in spite of several firings

  - Over this given τ, the number of reaction firings K^P_j is governed by a Poisson distribution

$$K_j^{\mathcal{P}} \sim \mathcal{P}(a_j \tau)$$
$$\mathbf{X}(t+\tau) = \mathbf{X}(t) + \sum_{j=1}^{M} K_j^{\mathcal{P}} \boldsymbol{\nu}_j.$$

Cost ~ M Poisson samplings

τ leaping : Fast BUT Inexact

- **τ leaping** : Can generate **negative populations**

- **Binomial τ leaping** : Approximate the unbounded Poisson distributions with Binomial ones

Tian & Burrage, J. Chem. Phys. 2004

Chatterjee et al., J. Chem. Phys. 2005

- **Modified τ leaping**

Cao et al., J. Chem. Phys. 2005

  - Critical reactions, i.e. those likely to drive some populations negative, handled by SSA
  - Other reactions advanced by τ leaping

R-leaping : Accelerate SSA by reaction leaps

Leaps : **prescribe number of firings L across all channels**

- Time increment τ_L is Gamma-distributed $\tau_L \sim \Gamma(L, 1/a_0(\mathbf{x}))$
- In this interval we will have K_m firings of channel R_m
- with : $\sum_{m=1}^{M} K_m = L$
- In R-leaping, (as in SSA), the index j of every firing obeys a point-wise distribution

$$P(j = l) = \frac{a_l(\mathbf{x})}{a_0(\mathbf{x})} \quad \text{for } l = 1, \ldots, M.$$

Auger, Chatelain, Koumoutsakos, **R-leaping: Accelerating the stochastic simulation algorithm by reaction leaps**.
J. Chem. Phys. , 125, 84103, 2006

R-leaping : One step

Auger, Chatelain, Koumoutsakos
J. Chem. Phys. 125, 84103, 2006

- Define L

$$\tau_L \sim \Gamma(L, 1/a_0(\mathbf{x}))$$

- Sample the index j

$$P(j = l) = \frac{a_l(\mathbf{x})}{a_0(\mathbf{x})} \quad \text{for } l = 1, \ldots, L.$$

- Number of reactions for channel m

$$K_m = \sum_{l=1}^{L} \delta_{l,m}$$

- Update species and time :

$$\mathbf{X}(t + \tau_L) = \mathbf{X}(t) + \sum_{j=1}^{M} K_j \boldsymbol{\nu}_j$$

R-leaping : Accelerate SSA by reaction leaps

- L firings distributed across M reaction channels
  - In τ leaping: K^P_j are independent Poisson variables.
  - In R-leaping, K_j are not independent.
- L as a control parameter
  - System can be brought to a desired state X
  - Time is not a-priori specified
  - New approaches to controlling negative species

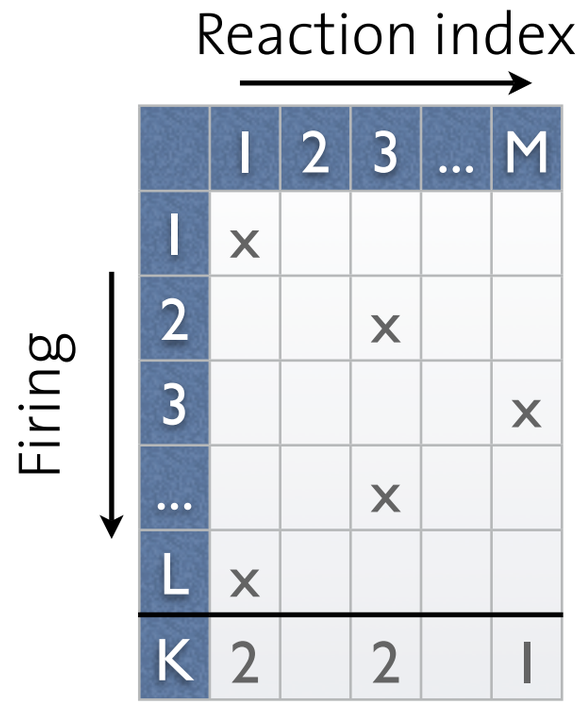
R-leaping : How to Sample the the M K_j

R_0 `Algorithm`

- Pointwise Sampling of **L** *independent* reaction indices

$$p(j = l) = \frac{a_l}{a_0}$$

  - Simple BUT scales with L - close to the work load of SSA!



Ro-sampling scales with L and, in particular when compared with τ -leaping that scales with M, the method is inefficient for large leap sizes, L ≫ M.

R-Leaping Theorem

The distribution of K_1 is a binomial distribution :

$$\mathcal{B}(L, a_1(\mathbf{x})/a_0(\mathbf{x}))$$

and for every $m \in \{2, \ldots, M\}$ the conditional distribution of K_m

given the event $\{(K_1, \ldots, K_{m-1}) = (k_1, \ldots, k_{m-1})\}$ is

$$K_m \sim \mathcal{B}\left(L - \sum_{i=1}^{m-1} k_i, \frac{a_m(\mathbf{x})}{a_0(\mathbf{x}) - \sum_{i=1}^{m-1} a_i(\mathbf{x})}\right).$$

This result is invariant under any permutation of the indices

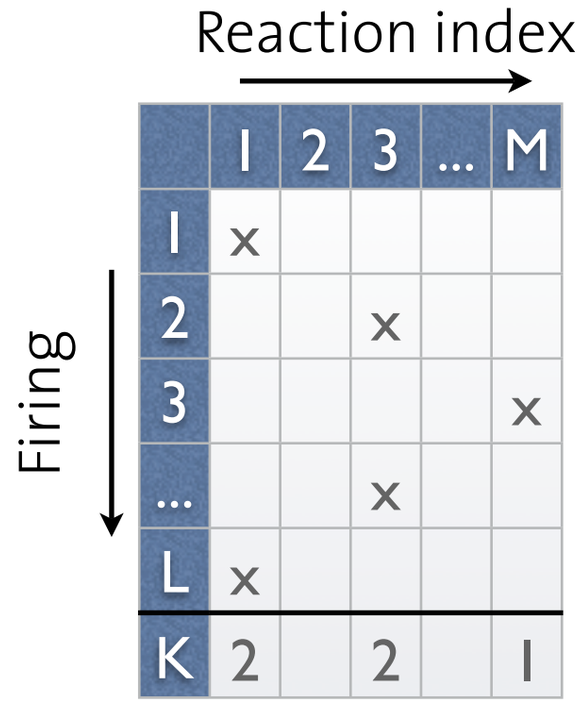
R-leaping : How to Sample the the M K_j

R_0 Algorithm

- Pointwise Sampling of **L** *independent* reaction indices

$$p(j = l) = \frac{a_l}{a_0}$$

  - Simple BUT scales with L - close to the work load of SSA!



Ro-sampling scales with L and, in particular when compared with τ -leaping that scales with M, the method is inefficient for large leap sizes, L ≫ M.

R_1 Algorithm

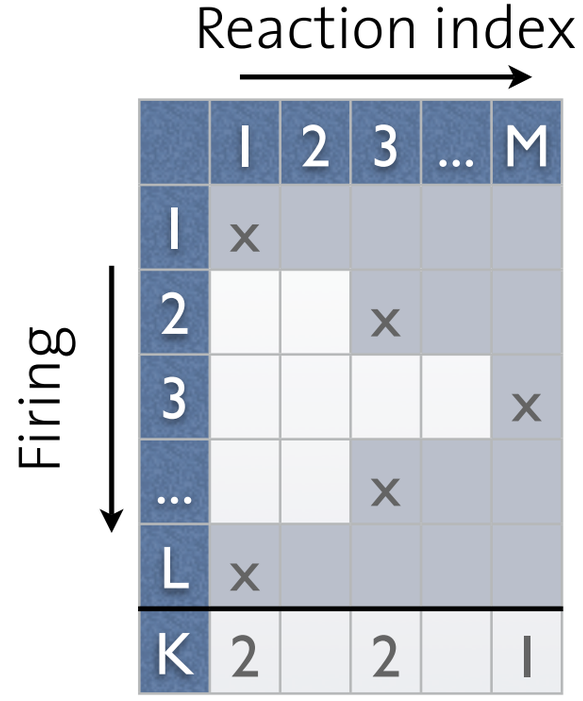
- Sampling **M** *correlated* binomial variables

$$\mathcal{B}(L, a_j/a_0)$$

  - Create correlations with conditional distributions

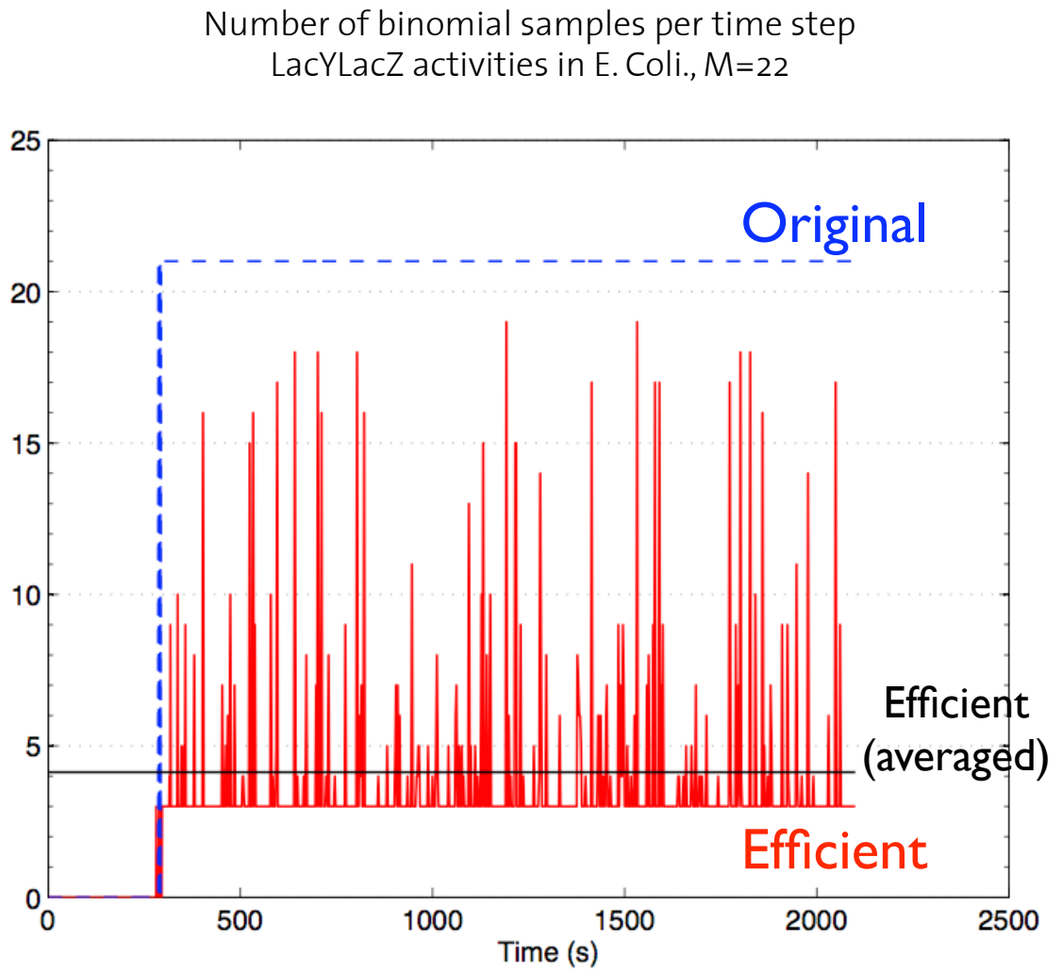
$$\text{If } K_i = k_i,\ \forall i < m,$$

$$K_m \sim \mathcal{B}\left(L - \sum_{i=1}^{m-1} k_i, \frac{a_m}{a_0 - \sum_{i=1}^{m-1} a_i}\right)$$



R-leaping : Efficient Sampling / Sorting

- **Sampling the M K_j efficiently (SORT the reactions)**
  - **M** can be large (~10^2) for bio-chemical systems!
  - Efficient sampling effectively loops over a fraction of **M**.
  - The larger the system, the bigger the payoff.
  - The more disparate the reaction rates are, the smaller the fraction.
  - Price to pay: carry out re-ordering often enough (cheap!)



Stochastic simulation: R-leaping

- **Controlling the leap approximation**
  - All three methods of τ leaping are transposable to R-leaping
    - Absolute change of a_j
    - Relative change of a_j
    - Relative change of a_j but efficiently through the relative changes in populations

Results

- **LacZ/LacY genes expression and enzymatic/ transport activities of LacZ/LacY proteins in E. Coli**

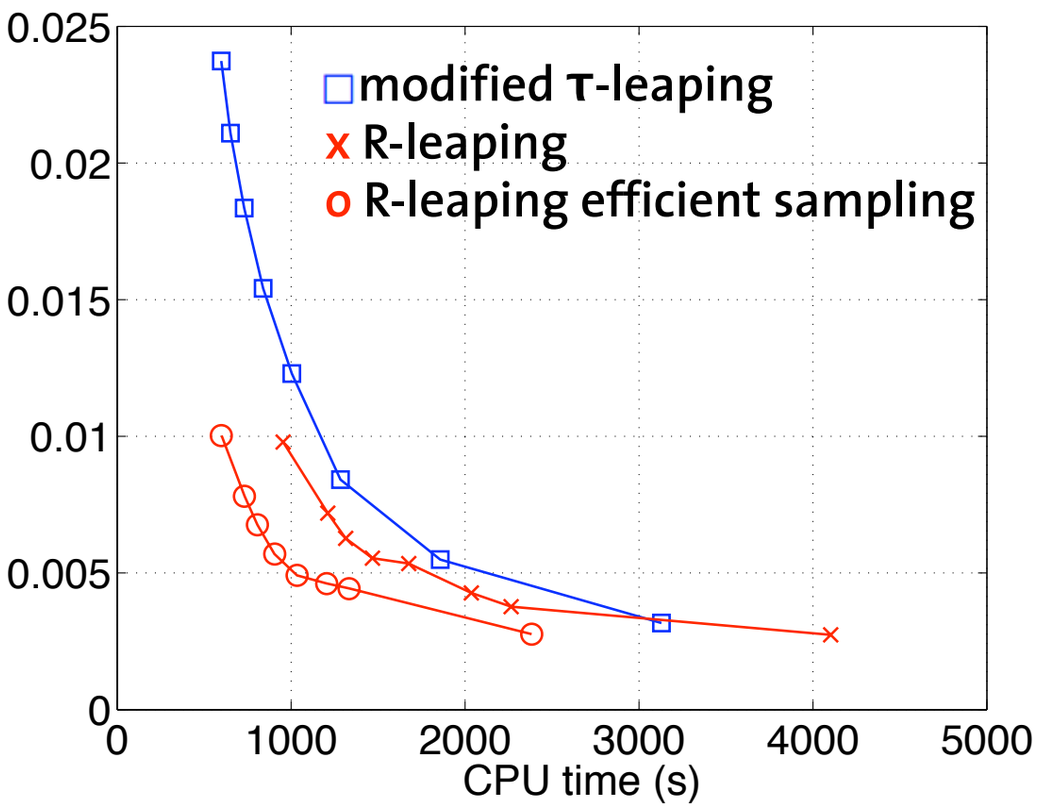
Kierzek, Bioiformatics 2002

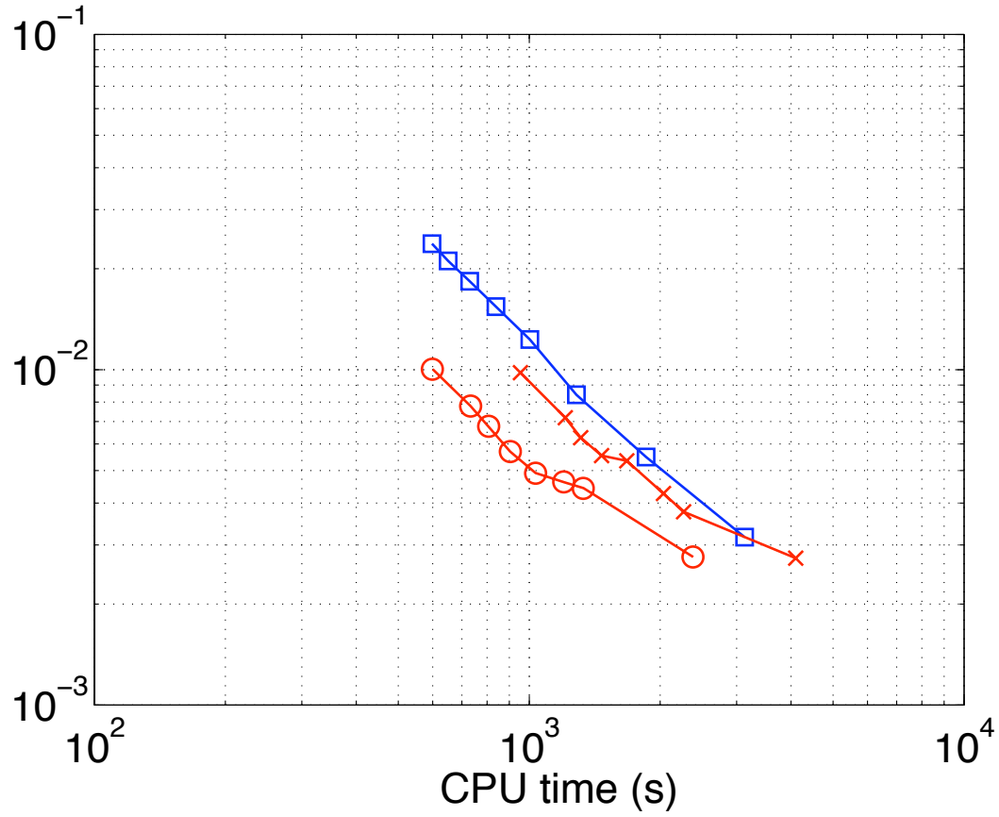
- Moderately large system (M = 22)
- Disparate rates
- Scarce reactants and negative species

	Reaction Channel	Reaction rate
R_1	PLac + RNAP → PLacRNAP	0.17
R_2	PLacRNAP → PLac + RNAP	10
R_3	PLacRNAP → TrLacZ1	1
R_4	TrLacZ1 → RbsLacZ + PLac + TrLacZ2	1
R_5	TrLacZ2 → TrLacY2	0.015
R_6	TrLacY1 → RbsLacY + TrLacY2	1
R_7	TrLacY2 → RNAP	0.36
R_8	Ribosome + RbsLacZ → RbsRibosomeLacZ	0.17
R_9	Ribosome + RbsLacY → RbsRibosomeLacY	0.17
R_{10}	RbsRibosomeLacZ → Ribosome + RbsLacZ	0.45
R_{11}	RbsRibosomeLacY → Ribosome + RbsLacY	0.45
R_{12}	RbsRibosomeLacZ → TrRbsLacZ + RbsLacZ	0.4
R_{13}	RbsRibosomeLacY → TrRbsLacY + RbsLacY	0.4
R_{14}	TrRbsLacZ → LacZ	0.015
R_{15}	TrRbsLacY → LacY	0.036
R_{16}	LacZ → dgrLacZ	$6.42\text{x}10^{-5}$
R_{17}	LacY → dgrLacY	$6.42\text{x}10^{-5}$
R_{18}	RbsLacZ → dgrRbsLacZ	0.3
R_{19}	RbsLacY → dgrRbsLacY	0.3
R_{20}	LacZ + lactose → LacZlactose	$9.52\text{x}10^{-5}$
R_{21}	LacZlactose → product + LacZ	431
R_{22}	LacY → lactose + LacY	14

Results

- **LacZ/LacY genes expression and enzymatic/ transport activities of LacZ/LacY proteins in E. Coli**
  - Histogram errors vs CPU time





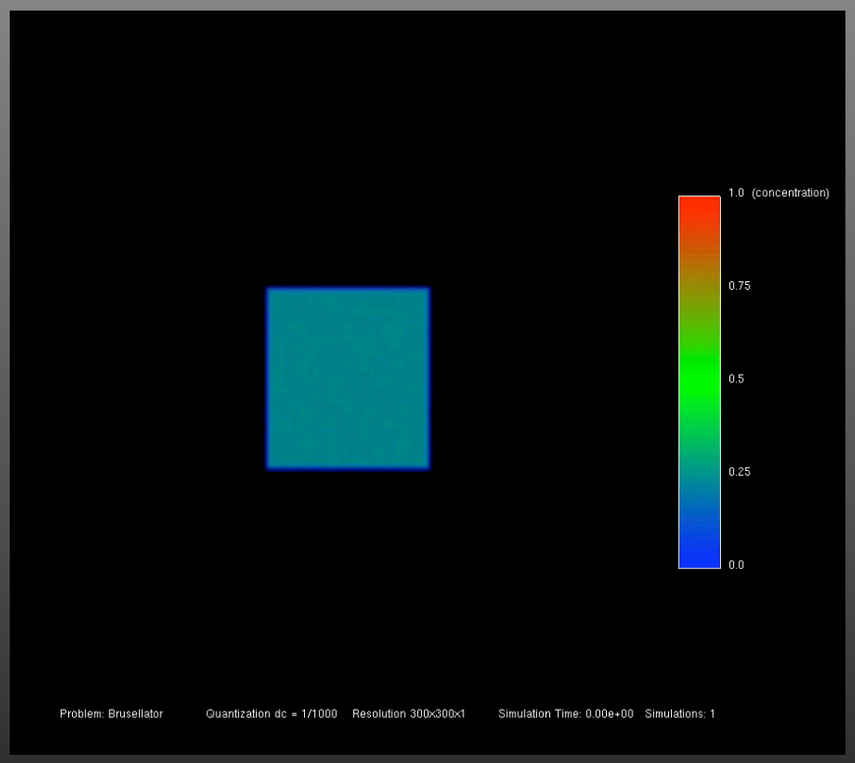
  - Efficient sampling offers factor 2 in speed w.r.t. modified τ-leaping!

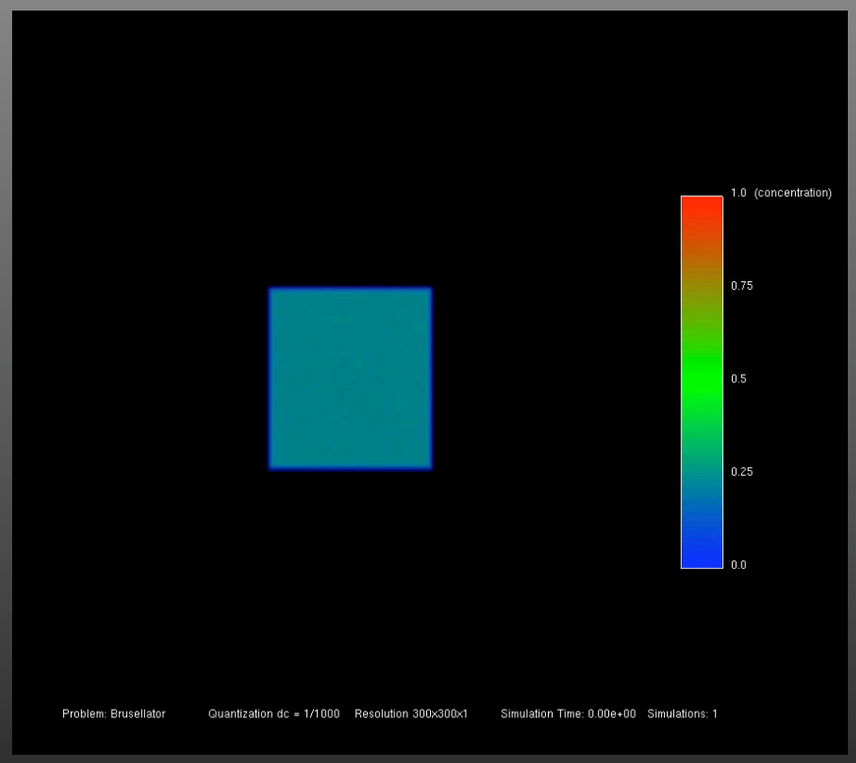
Summary

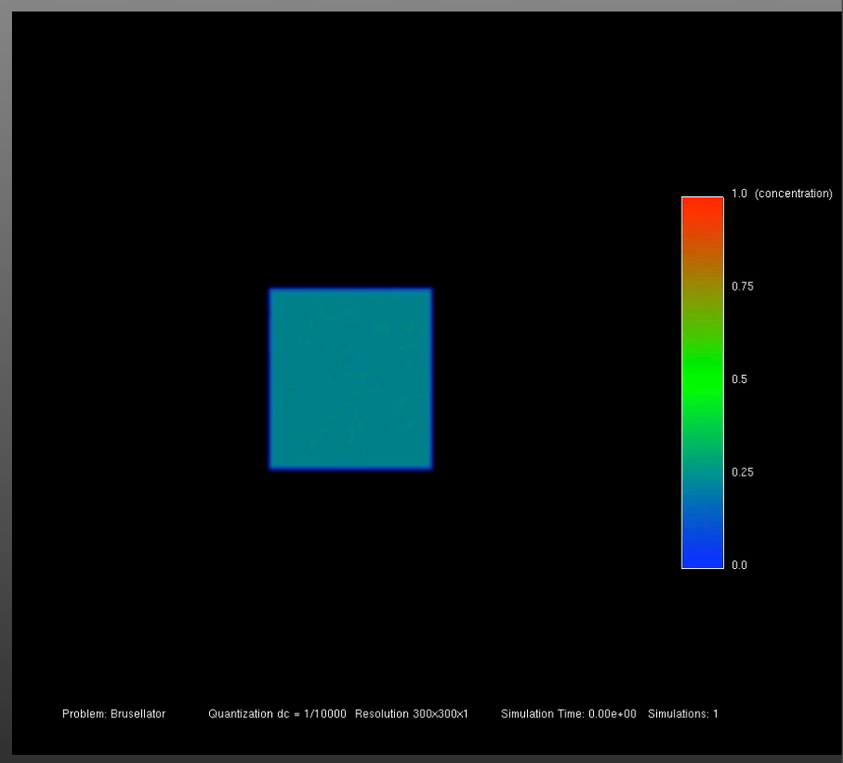
- **R-leaping, an accelerated stochastic algorithm that is complementary to existing τ-leaping algorithms**

- **Efficient binomial sampling offers computational savings for large systems with disparate rates**
  - Efficient sampling exploits size and stiffness of system.
  - Can be transposed to τ-leaping algorithms (!)...

- **Treatment of negative species with a tunable compromise efficiency-accuracy**
  - An alternative to modified τ-leaping, which essentially recurs to SSA when in trouble





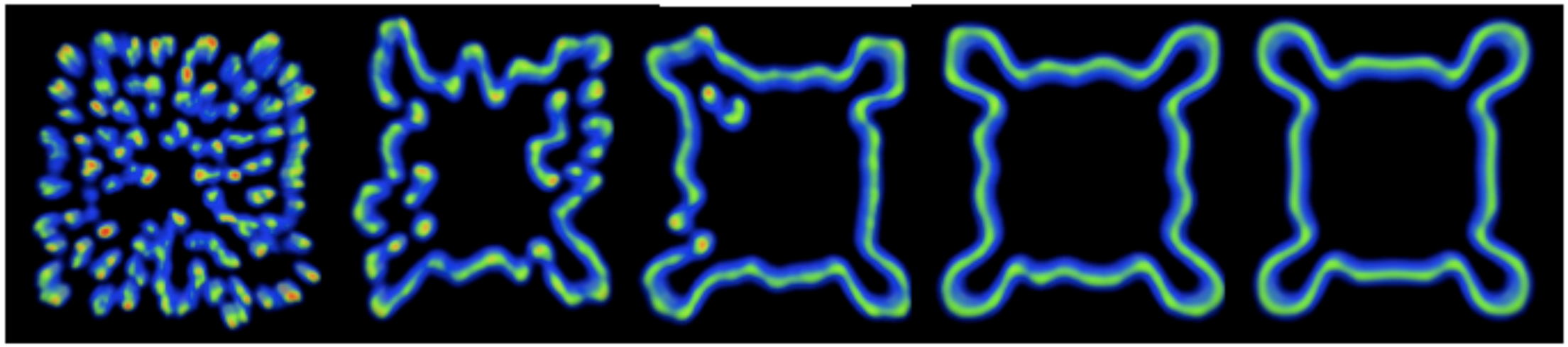


Stochastics in Space: Tau and R-leaping

Simulations of Gray-Scott Equations (2D)

Molecules per grid cell for a 300 x 300 grid

500 1000 5000 10000



Microscopic scale Macroscopic scale

$$F = 0.04, \kappa = 0.06, t = 1000$$

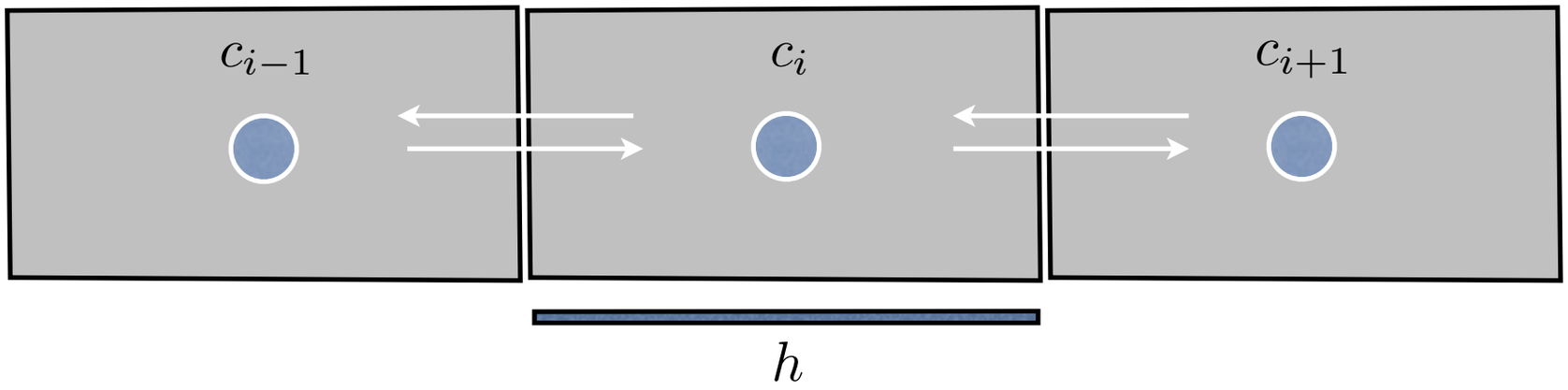
$$\begin{aligned} U + 2V &\rightarrow 3V, \\ V &\rightarrow P \end{aligned}$$

$$\begin{aligned} \frac{\partial u}{\partial t} &= d_u \Delta u - uv^2 + F(1-u), \\ \frac{\partial v}{\partial t} &= d_v \Delta v + uv^2 - (F+\kappa)v. \end{aligned}$$

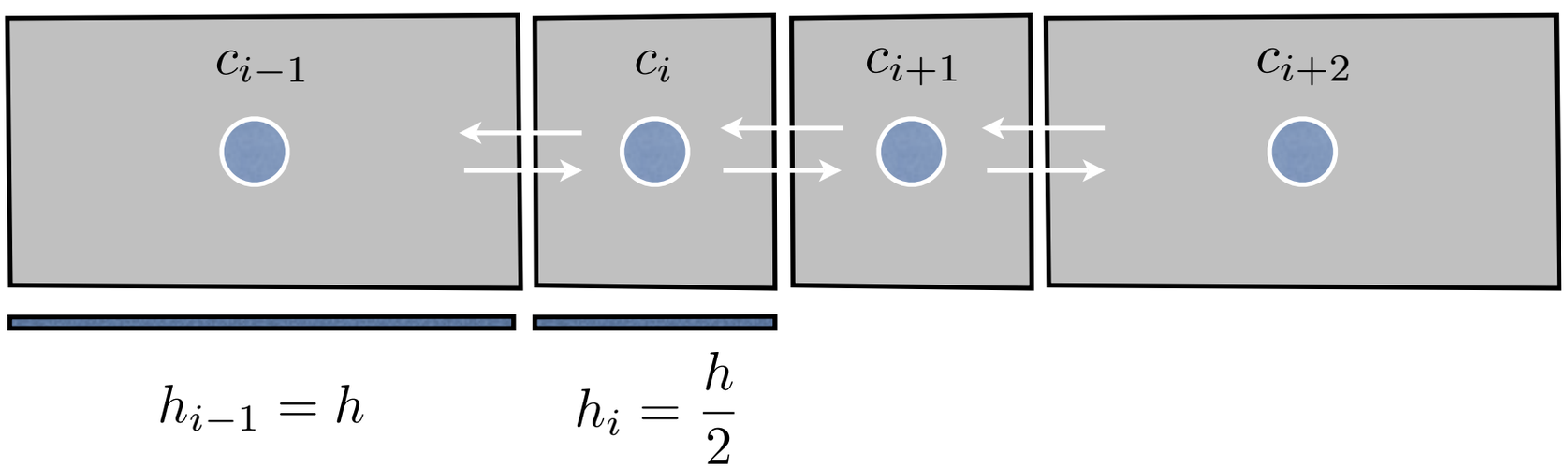
R-LEAP for Stochastic Diffusion on Non-uniform Discretizations

Diffusion events between cells, i.e. propensity for diffusion from cell i to cell j: $a_{i,j}(\mathbf{x}) = X_i \cdot k_{i,j}$

Uniform Cells: $k_{i,j} = \dfrac{D}{h^2}$



Non-uniform Cells: $k_{i,j} = \ ?$



Stochastic Diffusion on Non-Uniform Mesh Using a Finite Volume [1]

Continuum

$$\frac{\partial u}{\partial t} = -\nabla \cdot J$$

$$J = -D(x)\nabla u$$

Diffusion Process

$$\frac{dU_i}{dt} = -(k_{i,i+1} + k_{i,i-1})U_i + k_{i+1,i}U_{i+1} + k_{i-1,i}U_{i-1}$$

$$\frac{\partial U_i}{\partial t} = -\int_i \nabla \cdot J \; dx$$

Using the Divergence Theorem

$$\frac{\partial U_i}{\partial t} = J(c_i - \frac{h_i}{2}) - J(c_i + \frac{h_i}{2})$$

Approximating the Gradient in Fick's Law

$$\nabla u(c_i - \frac{h_i}{2}) \approx \frac{u(c_i) - u(c_{i-1})}{c_i - c_{i-1}} = \frac{1}{c_i - c_{i-1}} \left(\frac{U_i}{h_i} - \frac{U_{i-1}}{h_{i-1}} \right)$$

$$\frac{dU_i}{dt} = -\left(\frac{D_{i,i+1}}{h_i|c_i - c_{i+1}|} + \frac{D_{i,i-1}}{h_i|c_i - c_{i-1}|} \right) U_i + \left(\frac{D_{i+1,i}}{h_{i+1}|c_i - c_{i+1}|} \right) U_{i+1} + \left(\frac{D_{i-1,i}}{h_{i-1}|c_i - c_{i-1}|} \right) U_{i-1}$$

Reaction Rates for Diffusion Events:

$$k_{i,j} = \begin{cases} \frac{D_{i,j}}{h_i|c_i - c_j|} & \text{if } |i-j| = 1 \\ 0 & \text{otherwise} \end{cases}$$

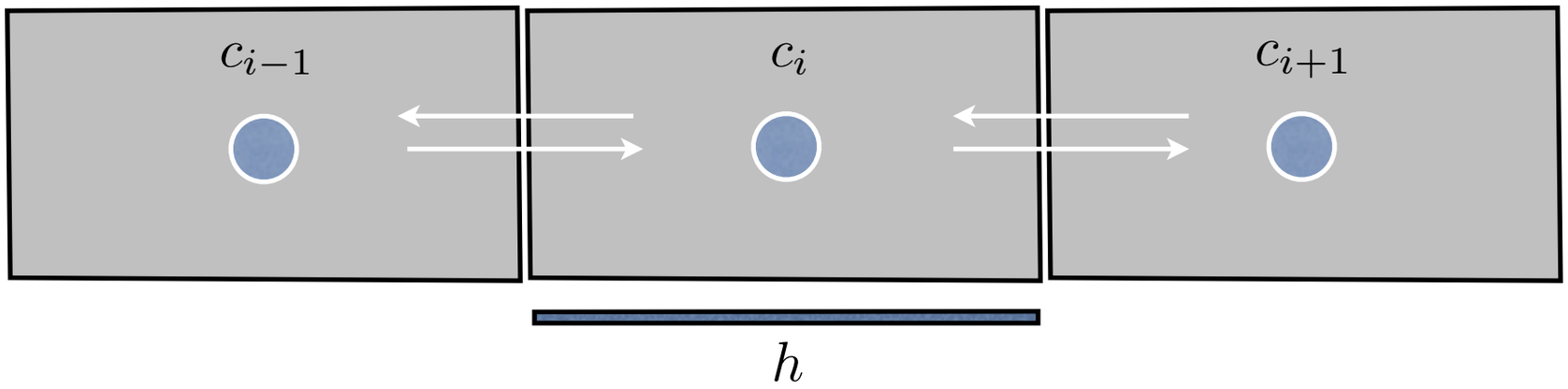
[1] D. Bernstein. Simulating mesoscopic reaction-diffusion systems using the gillespie algorithm. *Phys. Rev. E*, 2005.

Stochastic Diffusion on Non-uniform Discretizations: Propensity Disparities

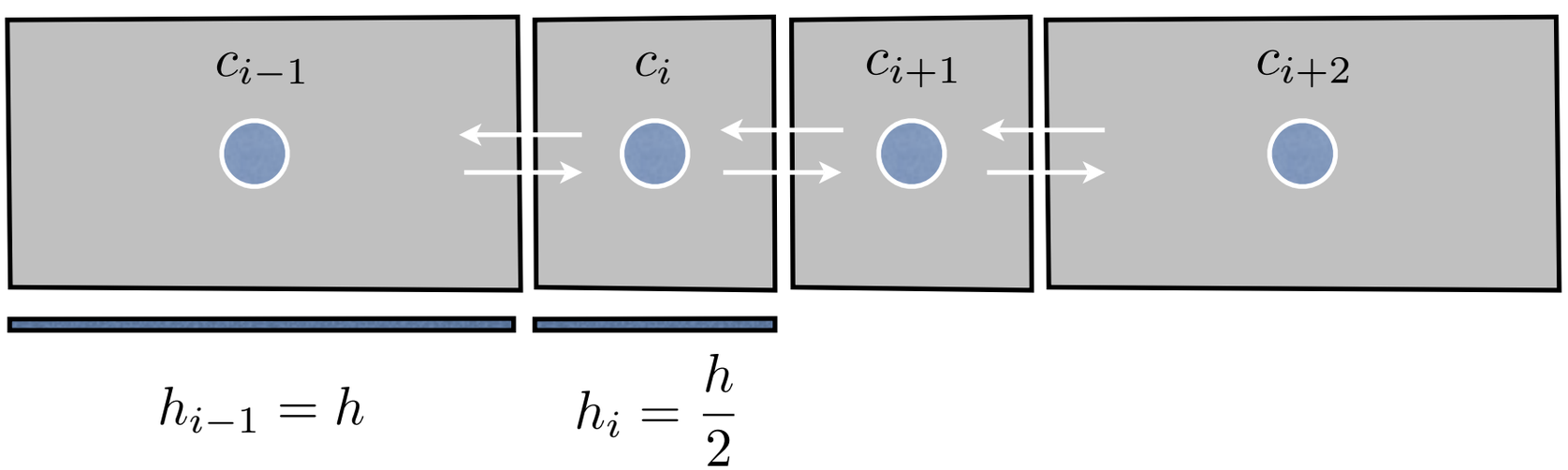
Disparity in diffusion propensities arise from two sources:

- Gradients in the concentration, which are problem specific.
- Non-uniform cells. E.g. consider the diffusion of a uniform concentration field as shown below.

Uniform Cells: $a_{i,i-1}(\mathbf{x}) = a_{i,i+1}(\mathbf{x})$

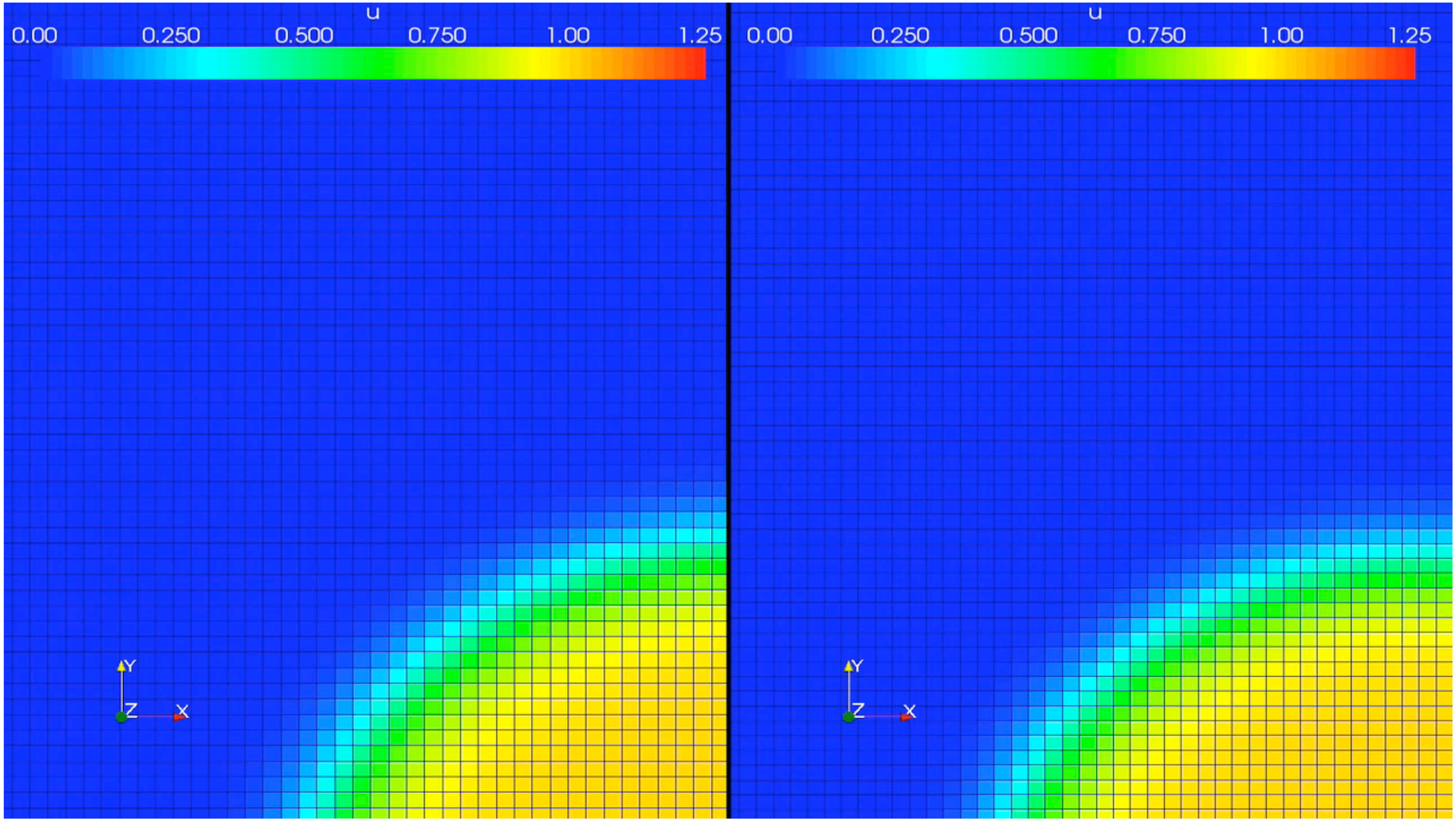


Non-uniform Cells: $a_{i,i-1}(\mathbf{x}) < a_{i,i+1}(\mathbf{x})$



Stochastic Diffusion on Non-uniform Discretizations: Optimization

2D Fisher Equation



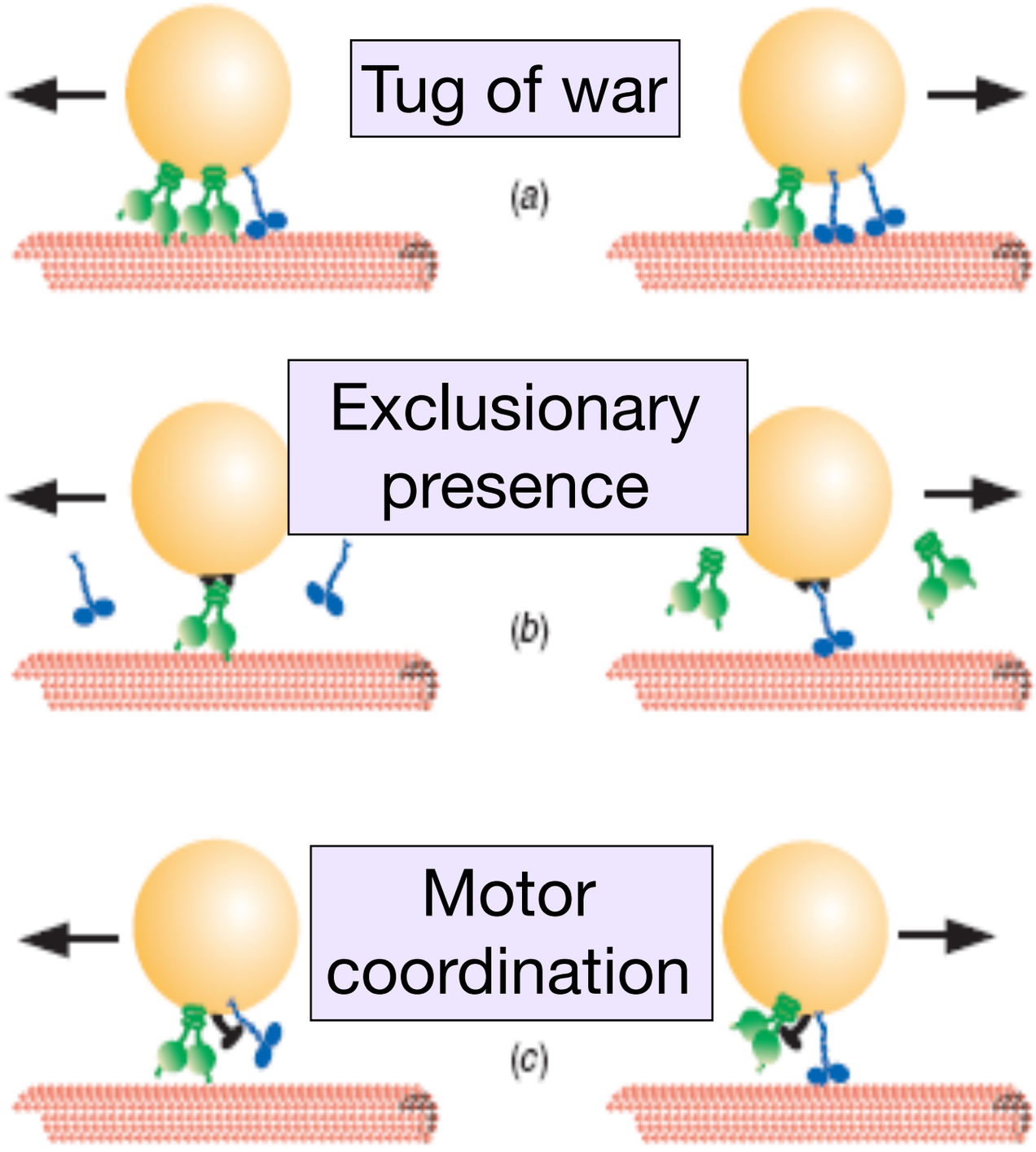
Multiresolution Stochastic Simulations of Reaction-Diffusion Processes, B. Bayati, P. Chatelain, P. Koumoutsakos, Phys. Chem. Chem. Phys., 2008

with Urs Greber, Christoph Burkhardt, Uni ZH

optimizing Stochastic Models : Adenovirus Transport

Proposed models for bi-directional transport

from Gross, Physical Biology (2004)

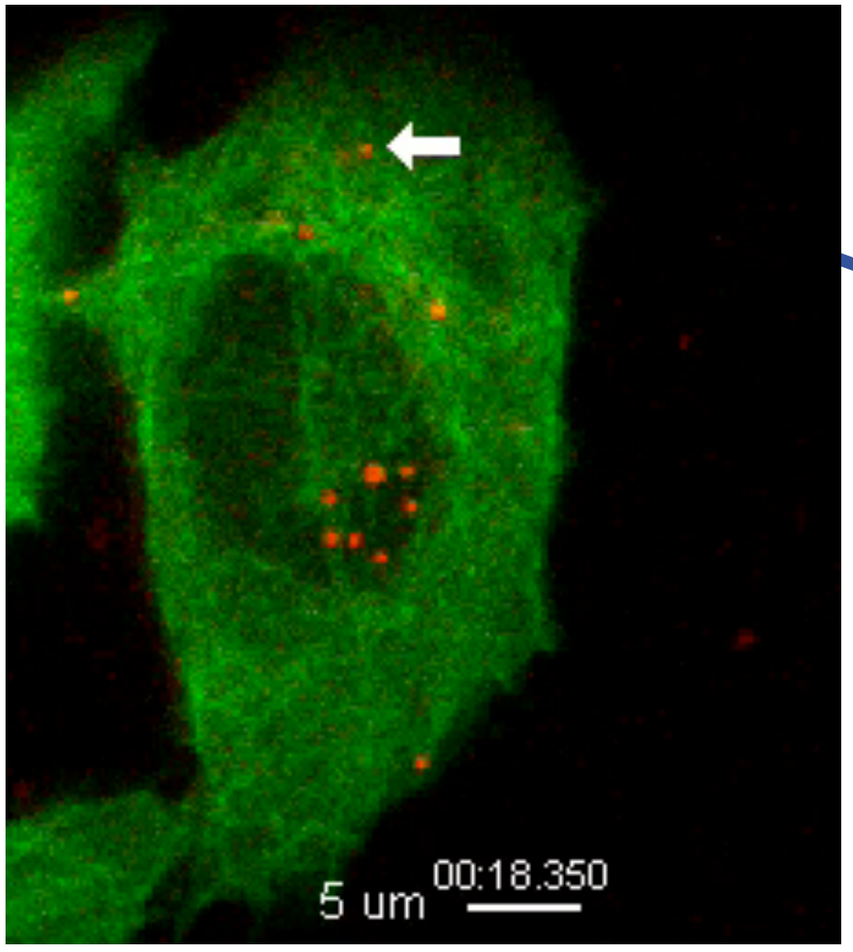


Our modeling approach assumes that ...

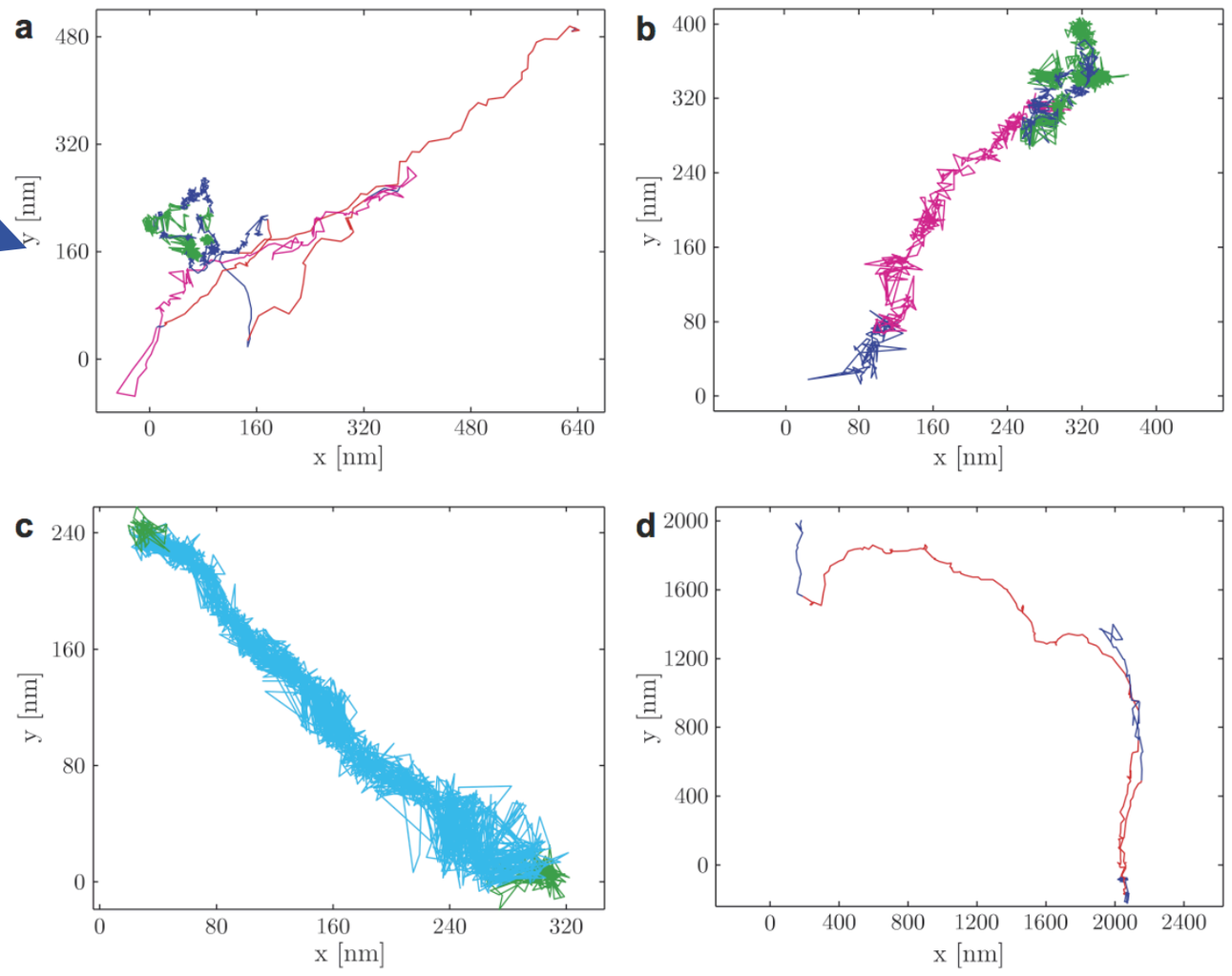
- viruses are cytosolic, and invariant
- motions along a filament are equivalent to 1D runs
- planar virus movements in flat regions of the cell
- homogeneous microtubules without MAPs & dynamics
- individual microtubules independent of each other
- motor step size 8 nm
- no motor inhibitors

From directed motions to segmented 1-D tracks

- 2D virus trajectories are extracted by a single particle tracking algorithm [1]
- Different motion patterns are identified through segmentation and classification by Support Vector Machine [2]



Images of HeLa cells infected with Adenovirus serotype 2 imaged using a spinning disk confocal microscope

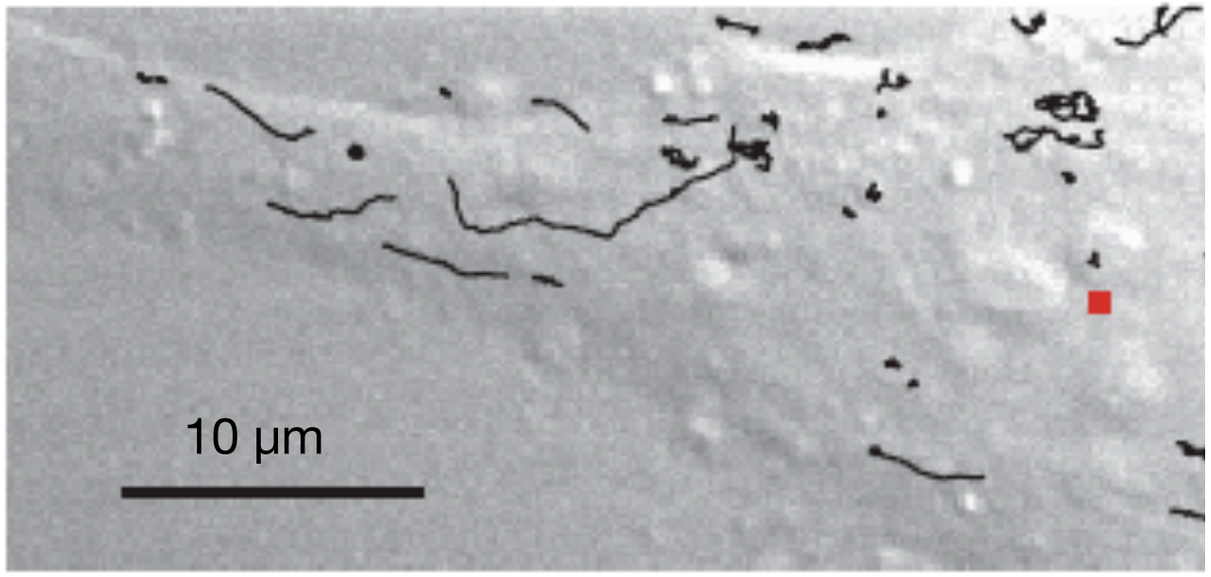


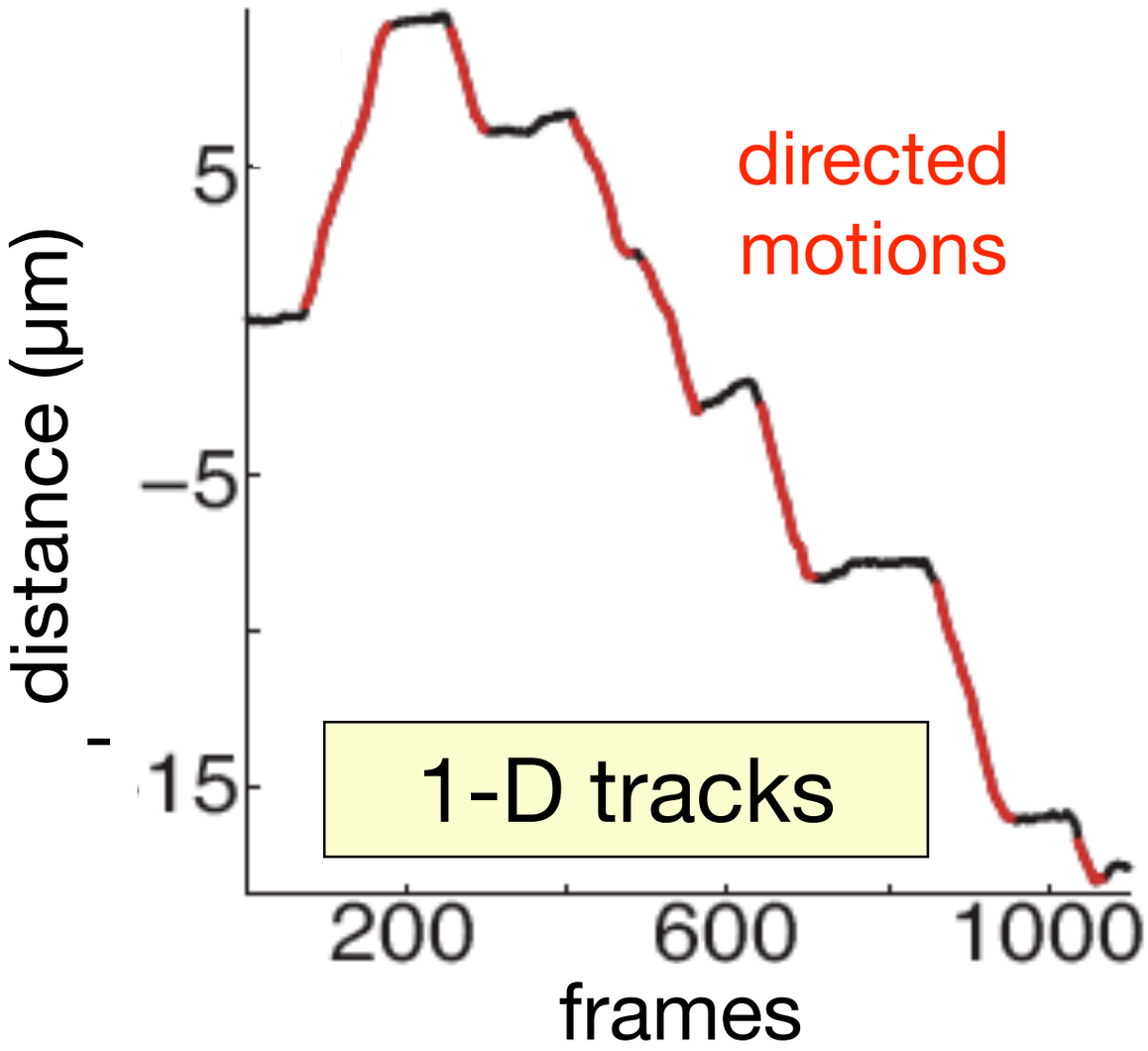
confined motion
slow drifting
fast drifting
directed motion

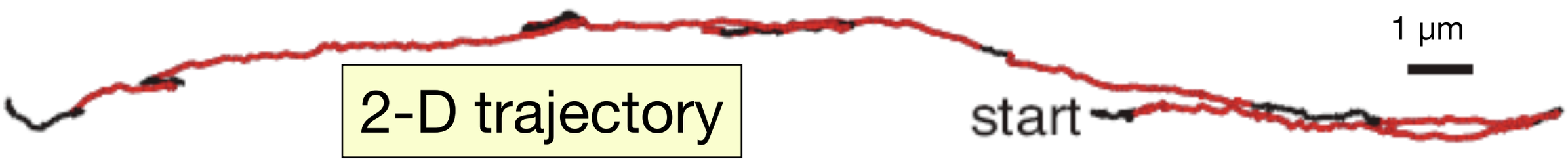
[1] I.F. Sbalzarini and P. Koumoutsakos, Feature point tracking and trajectory analysis for video imaging in cell biology, Journal of Structural Biology, 2005
[2] J.A. Helmuth, C.J. Burckhardt and P. Koumoutsakos and U.F. Greber and I.F. Sbalzarini, Feature point tracking and trajectory analysis for video imaging in cell biology, Journal of Structural Biology, 2007

From directed motions to segmented 1-D tracks

Ad5 trajectories, HeLa cells



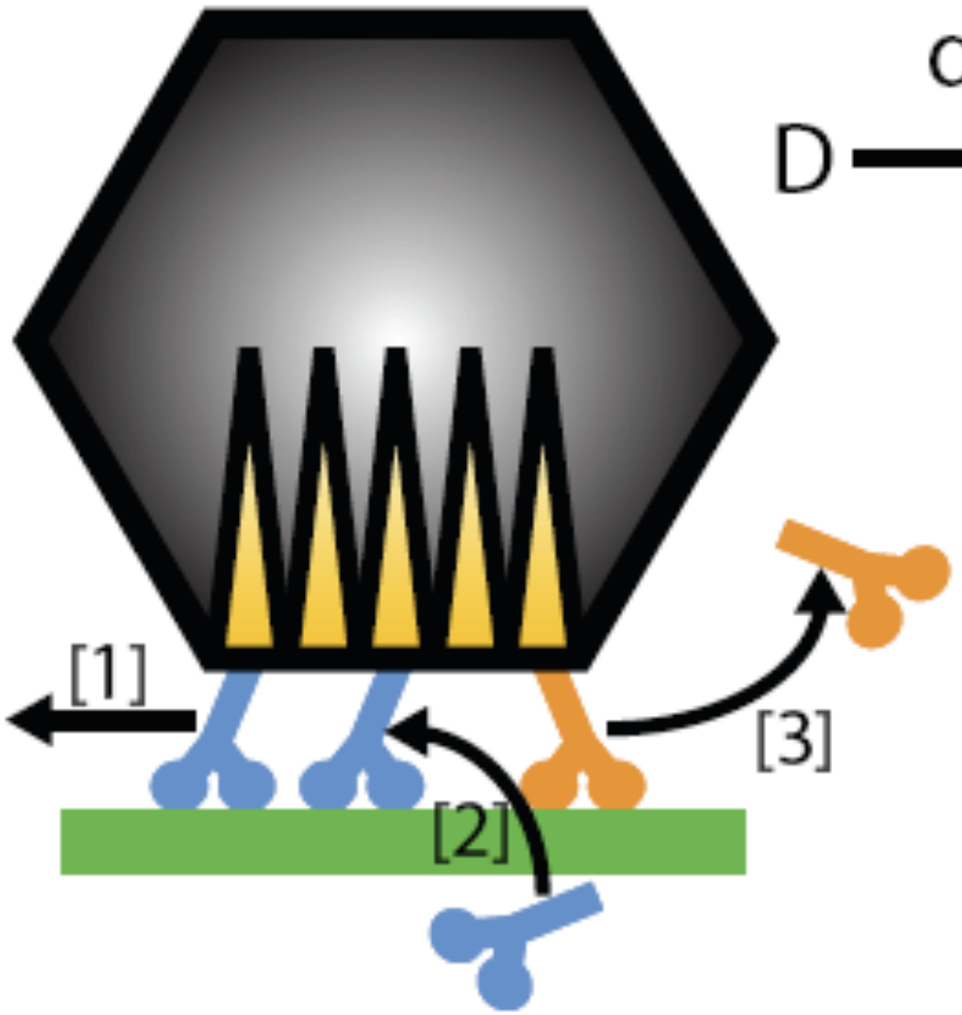




Our input into the model

- Segmented 25Hz virus trajectories
- Motor binding to virus at distinct or overlapping sites
- Step size of the plus & minus end motor: +/- 8nm

Six parameters optimized by the algorithm



$$D \xrightarrow{d_\mu} D + \mu_D, \quad K \xrightarrow{k_\mu} K + \mu_K \quad [1]$$

$$\rho \xrightarrow{d_\alpha} D, \quad D \xrightarrow{d_\delta} \rho \quad [2]$$

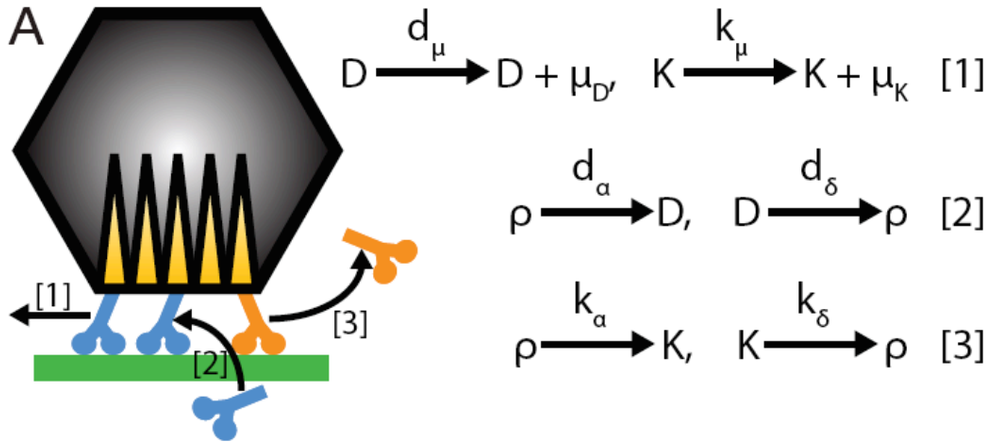
$$\rho \xrightarrow{k_\alpha} K, \quad K \xrightarrow{k_\delta} \rho \quad [3]$$

dynein (D) and kinesin (K) can bind, unbind and move the virus cargo along a microtubule.

Eq. 1 : movement of the motor proteins, where d_μ and k_μ are the displacement rates for dynein and kinesin, respectively.

Eq. 2(3) : binding and unbinding of dynein(kinesin), where ρ is the number of binding sites, d_α and d_δ are the binding and unbinding rates.

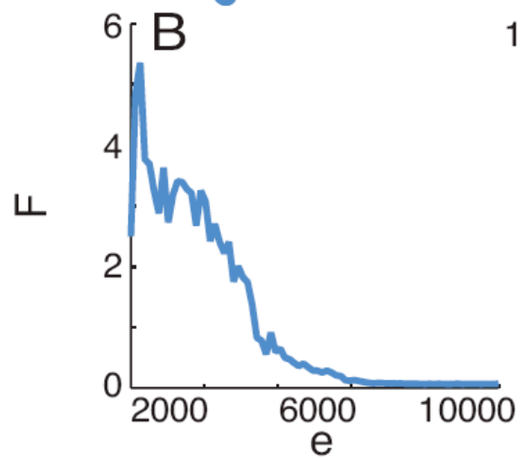
Six parameters optimized by the algorithm



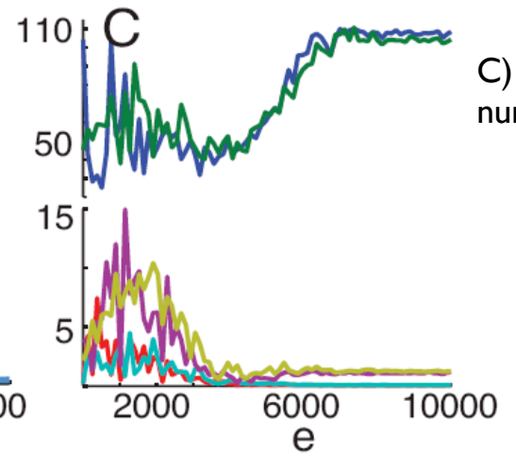
$$D \xrightarrow{d_\mu} D + \mu_{D'} \quad K \xrightarrow{k_\mu} K + \mu_K \quad [1]$$

$$\rho \xrightarrow{d_\alpha} D, \quad D \xrightarrow{d_\delta} \rho \quad [2]$$

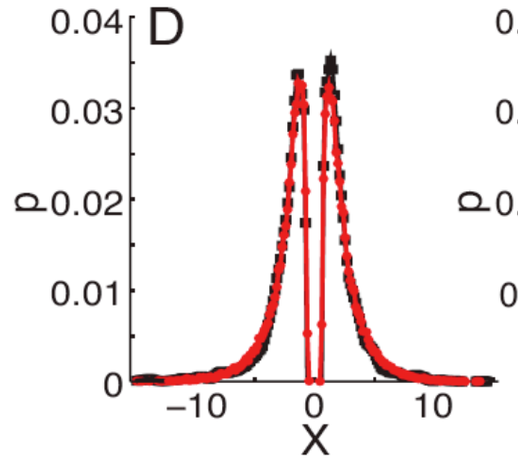
$$\rho \xrightarrow{k_\alpha} K, \quad K \xrightarrow{k_\delta} \rho \quad [3]$$



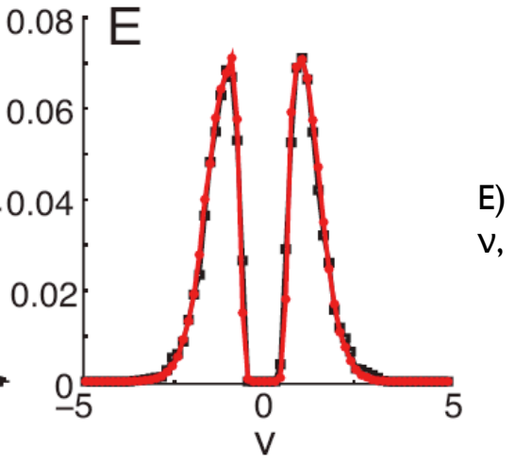
B) Fitness value F from the optimization, plotted against the number of fitness evaluations e, for the competing binding sites model with 14 binding sites.



C) The parameter values plotted against the number of fitness evaluations e.



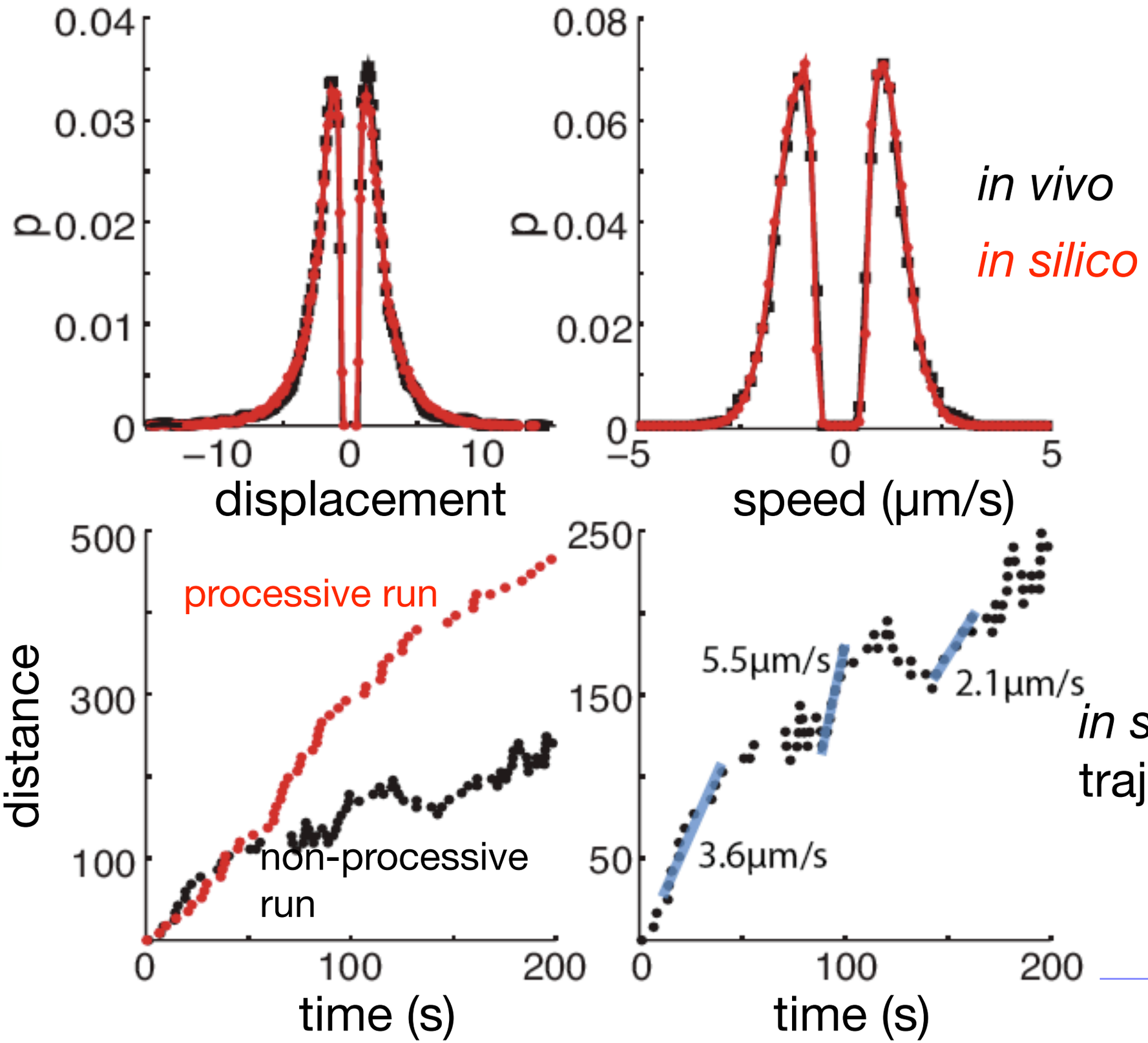
D) Run length distributions (probability p versus X, μm) for the biological (black) and in silico (red) data.



E) Velocity distributions (probability p versus v, μm/s).

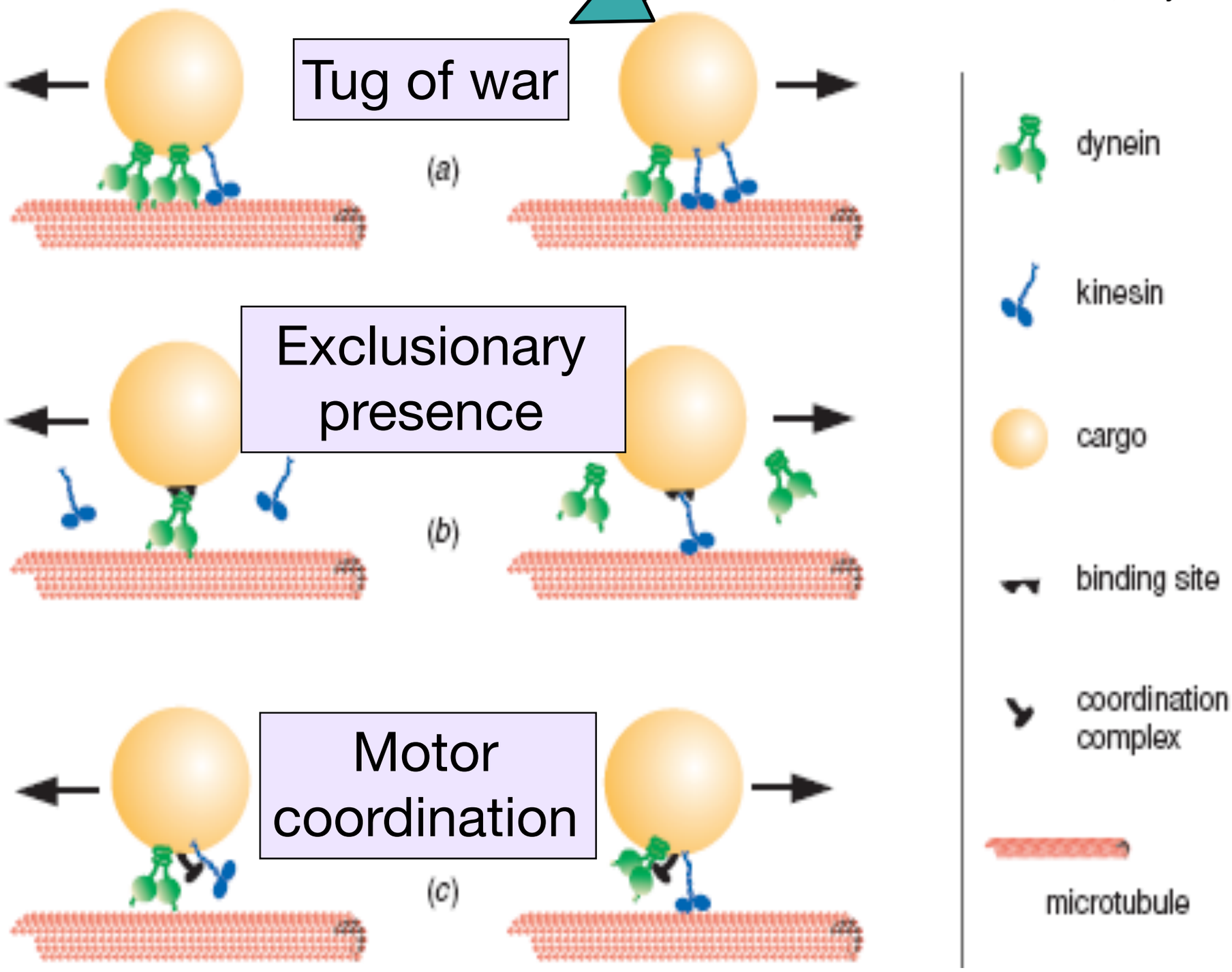
The optimized model reproduces *in vivo* data

> 100’000
CPU hours



and the winner is .....

from Gross, Physical Biology (2004)



Model predicts an non-coordinated ‘tug-of-war’

- Distinct or overlapping binding sites for dynein & kinesin are possible

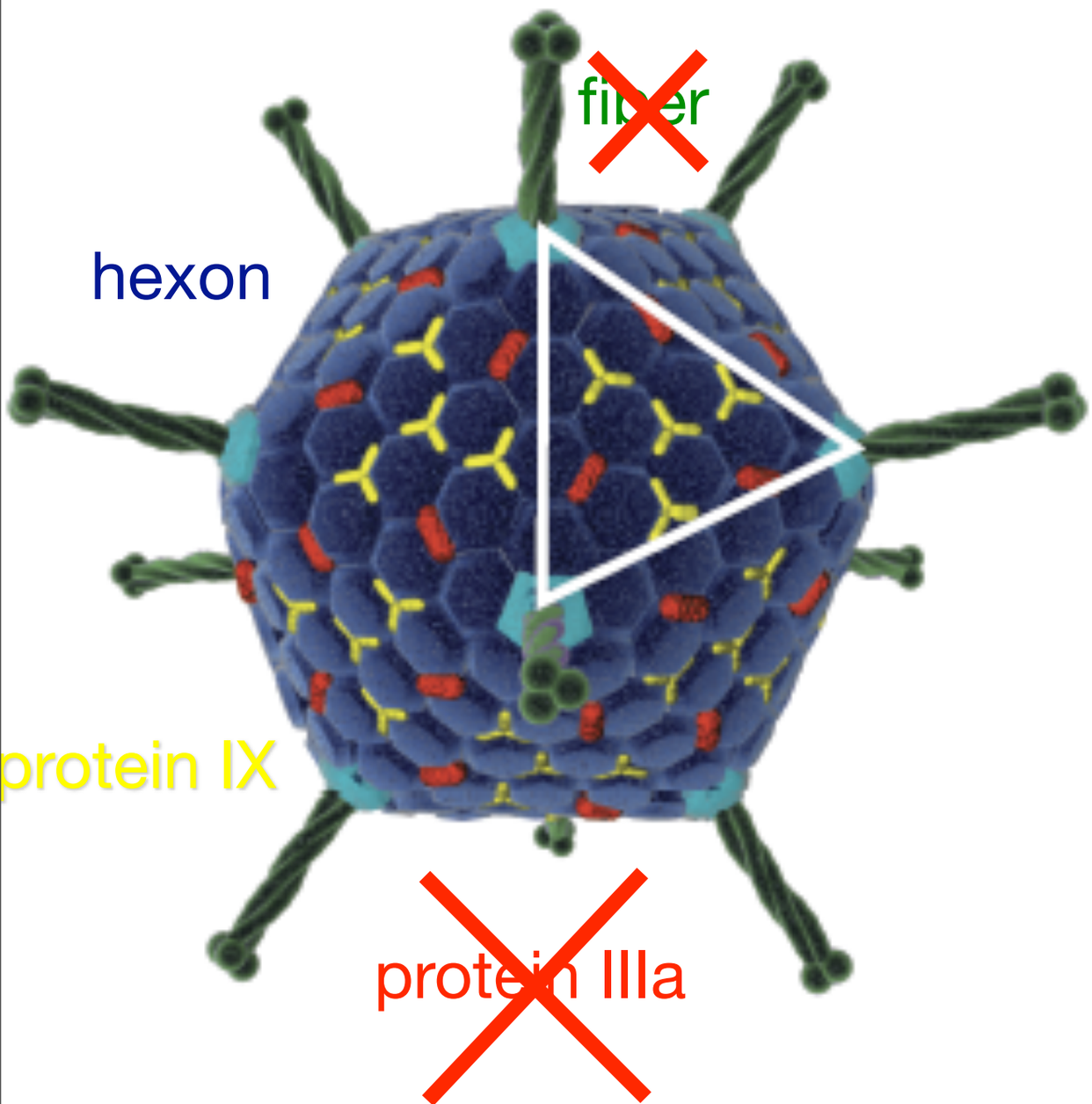
- k_{on} and k_{off} for dynein and kinesin are lower than expected

- 1-3 motors bound per virus during runs

- a total of 8-14 motor binding sites required

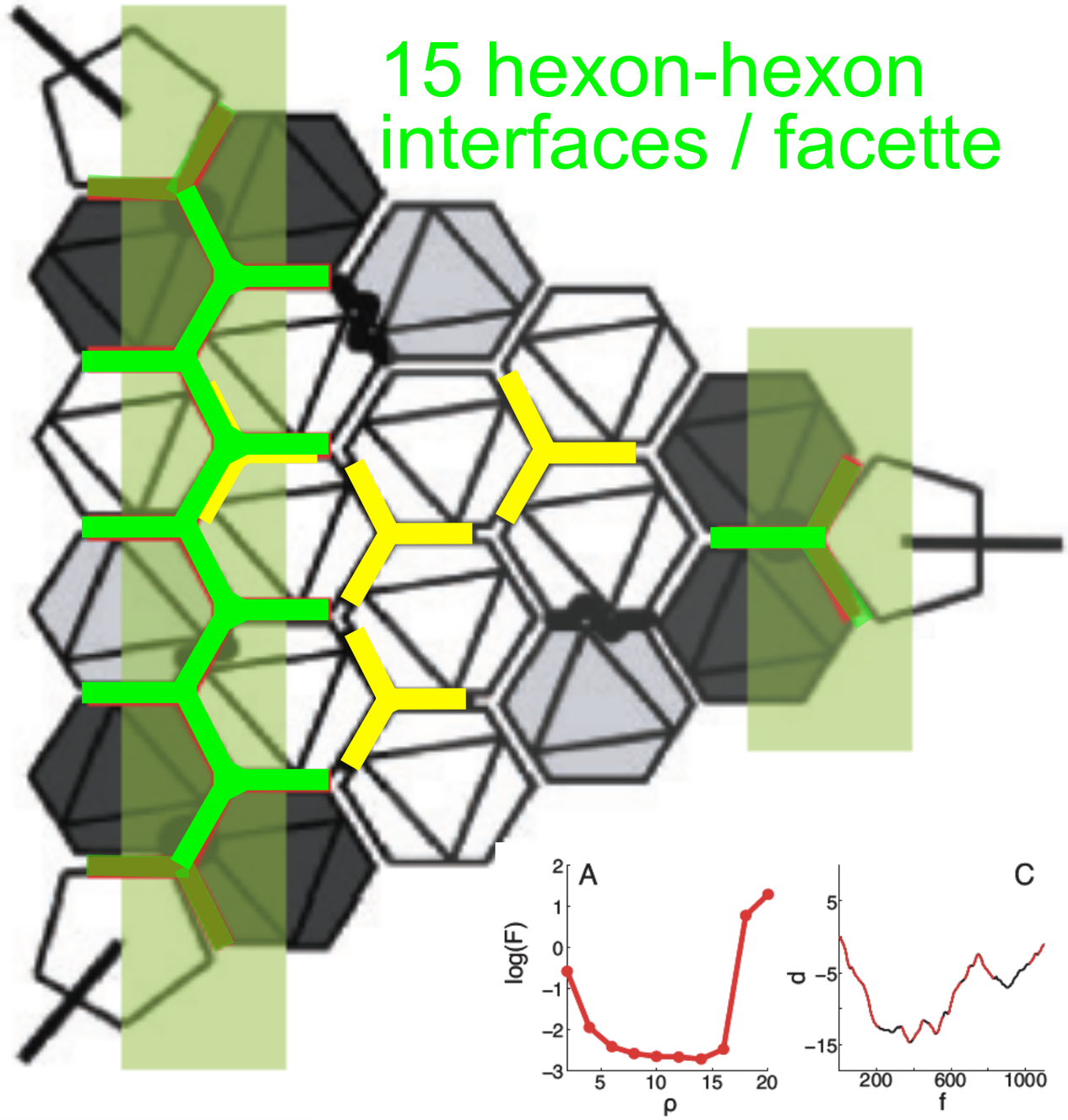
Where is motor binding site, on hexon or pIX?

Adenovirus model

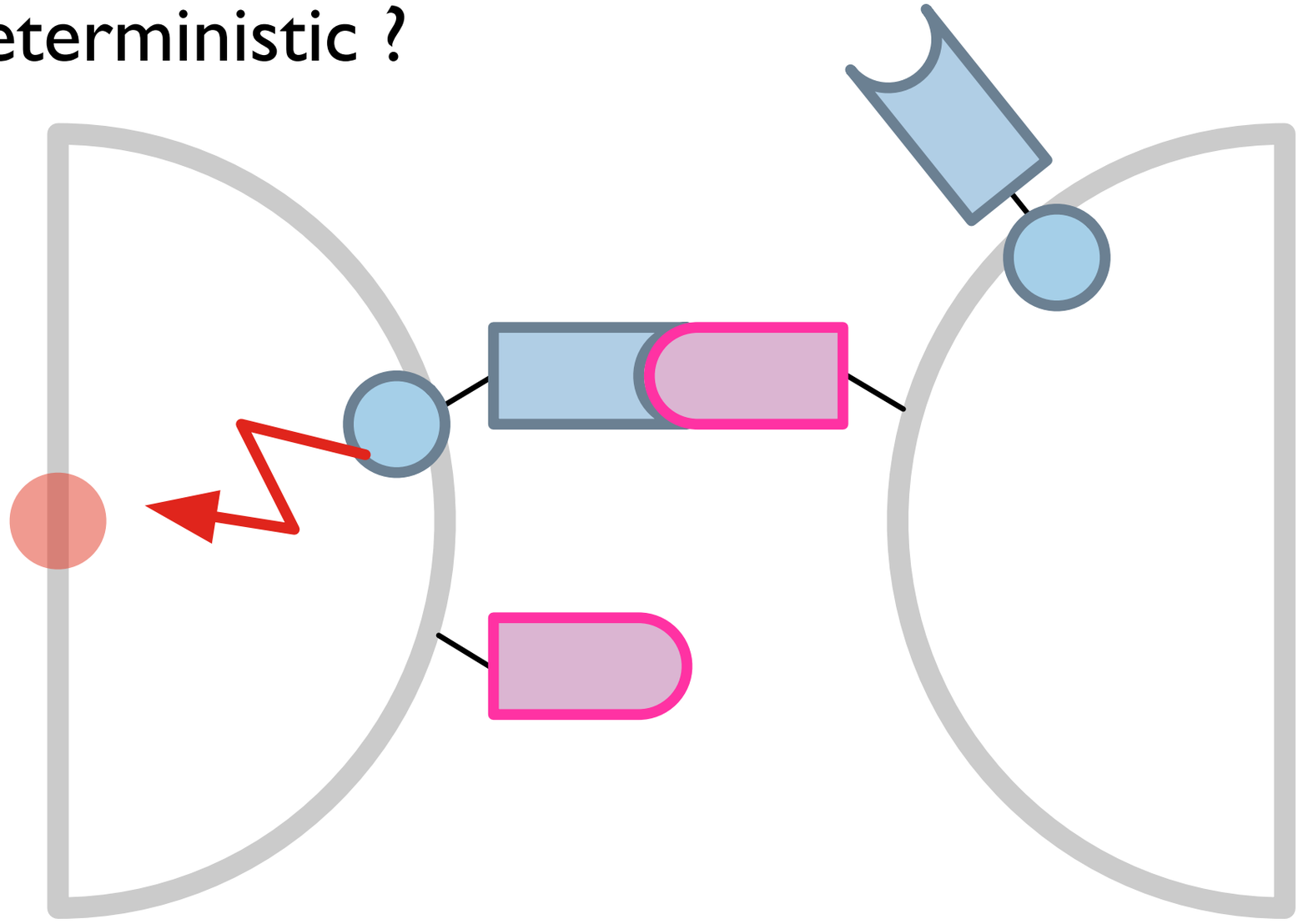


Facette

15 hexon-hexon interfaces / facette



Deterministic ?



Stochastic ?

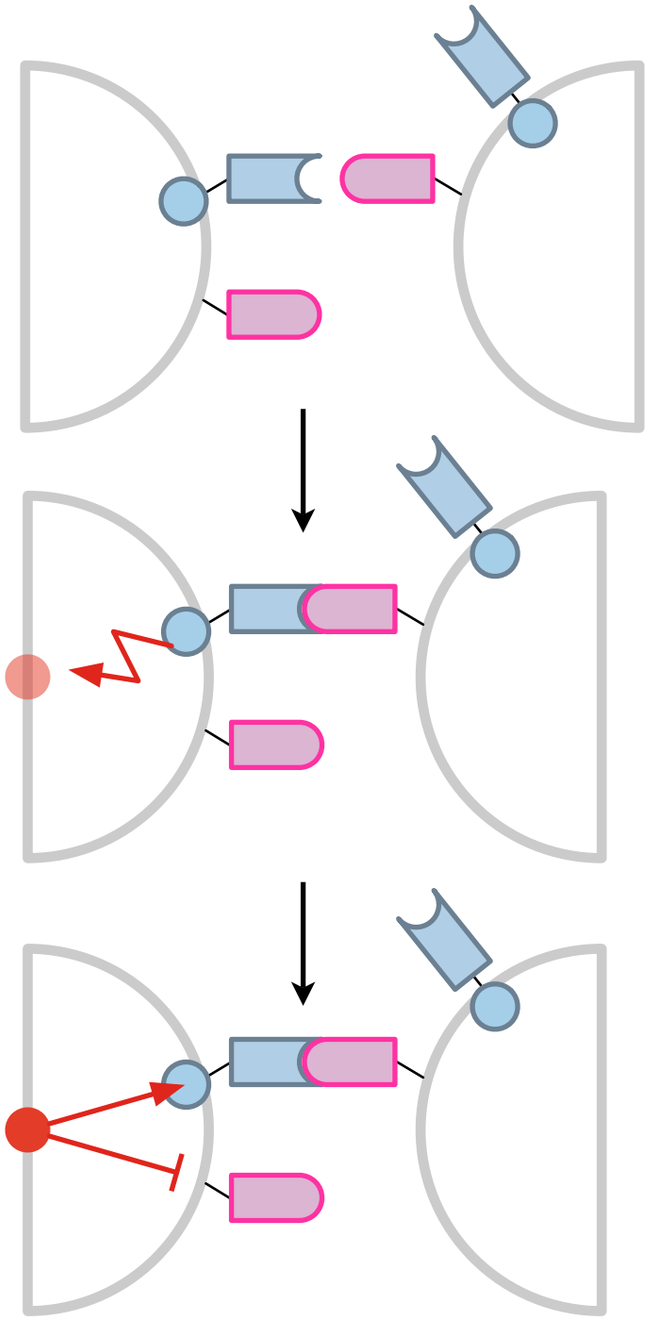
A message from your mother : Chose your models carefully

Juxtacrine Signaling

Definition

Intercellular signaling induced by physical cell contact.
Mainly involved in

- developmental processes,
- wound healing,
- angiogenesis



Examples

Receptor: Notch Ligands: Delta, Jagged, Serrate
Receptor: EGFR Ligand: TGFa

Pattern formation

through positive and negative feedback loops
(autoinduced up-regulation and down-regulation of ligand and receptor expression)

Computational Models (EGFR/TGFa)

3 Species: R Receptor L Ligand B Bound receptor (=R-L complex)

Deterministic

$$\frac{\partial R_i}{\partial t} = \underbrace{-k_a\, R_i\, \langle L \rangle_i}_{\text{BINDING}} + \underbrace{k_d\, B_i}_{\text{DISASSOCIATION}} - \underbrace{d_r\, R_i}_{\text{DECAY}} + \underbrace{\mathcal{P}_r(B_i)}_{\text{FEEDBACK}}$$

$$\frac{\partial L_i}{\partial t} = \underbrace{-k_a\, \langle R \rangle_i\, L_i}_{\text{BINDING}} + \underbrace{k_d\, \langle B \rangle_i}_{\text{DISASSOCIATION}} - \underbrace{d_l\, L}_{\text{DECAY}} + \underbrace{\mathcal{P}_l(B_i)}_{\text{FEEDBACK}}$$

$$\frac{\partial B_i}{\partial t} = \underbrace{k_a\, R_i\, \langle L \rangle_i}_{\text{BINDING}} - \underbrace{k_d\, B_i}_{\text{DISASSOCIATION}} - \underbrace{k_{\text{int}}\, B_i}_{\text{INTERNALIZATION}}$$

2X POSITIVE FEEDBACK:

$$\mathcal{P}_l(B) = \frac{C_1^m B^m}{C_2^m + B^m} \qquad \mathcal{P}_r(B) = C_3 \frac{C_4^n B^n}{C_5^n + B^n}$$

Spatial operator $\langle \cdot \rangle_i$ = CONCENTRATION ON CELLS j TOUCHING CELL i

Stochastic e.g. ID

$$R_i + L_{i+1} \xrightarrow{1/2\, k_a} B_i \qquad \text{BINDING}$$

$$R_i + L_{i-1} \xrightarrow{1/2\, k_a} B_i$$

$$B_i \xrightarrow{1/2\, k_d} R_i + L_{i+1} \qquad \text{DISASSOCIATION}$$

$$B_i \xrightarrow{1/2\, k_d} R_i + L_{i-1}$$

$$L_i \xrightarrow{d_l} \qquad R_i \xrightarrow{d_r} \qquad B_i \xrightarrow{k_i} \qquad \text{DECAY \& INTERNALIZATION}$$

$$\xrightarrow{\mathcal{P}_l(B_i)} L_i \qquad \xrightarrow{\mathcal{P}_r(B_i)} R_i \qquad \text{FEEDBACK}$$

OWEN&SHERRATT, MATHEMATICAL BIOSCIENCES, 1998

Stochastic vs. Det. simulation

Spatiality:

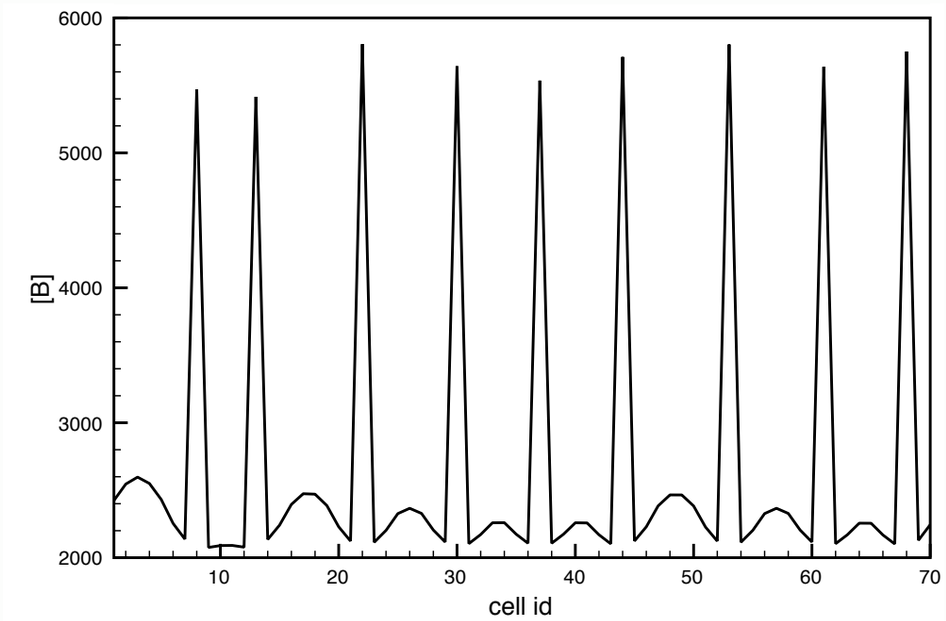
Species of different cells i encoded as different species, i.e.

$$B_1,\ R_1,\ L_1,\ B_2,\ R_2\ \ \ldots,\ L_N$$

For 70 cells: **770** reactions, prohibitively expensive with SSA → **R-leap**

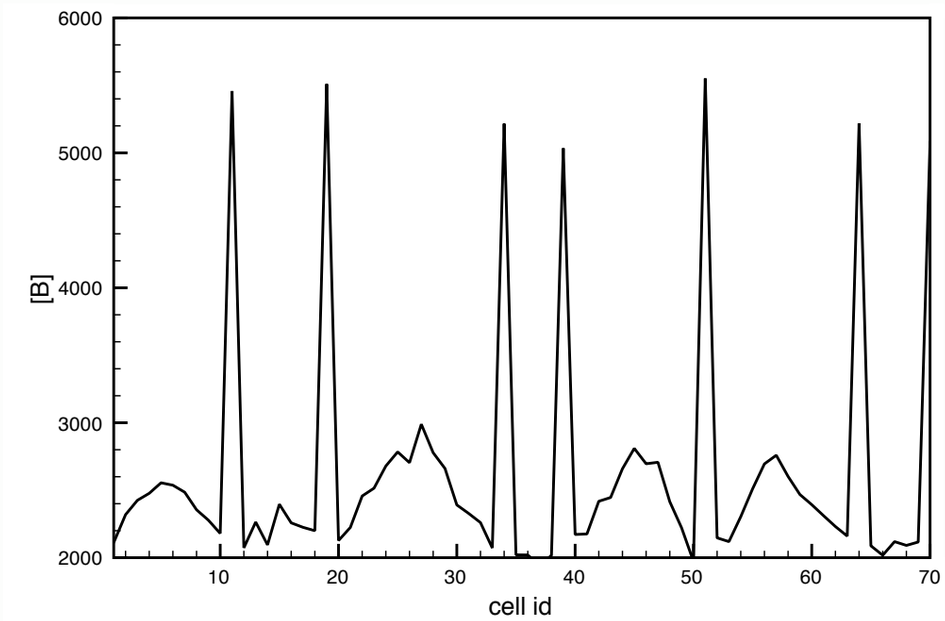
Patterning after 100h:

initial condition: $B_j = R_j = 3000 \ \ L_j = 462$
+ **10%** Uniform noise



Deterministic

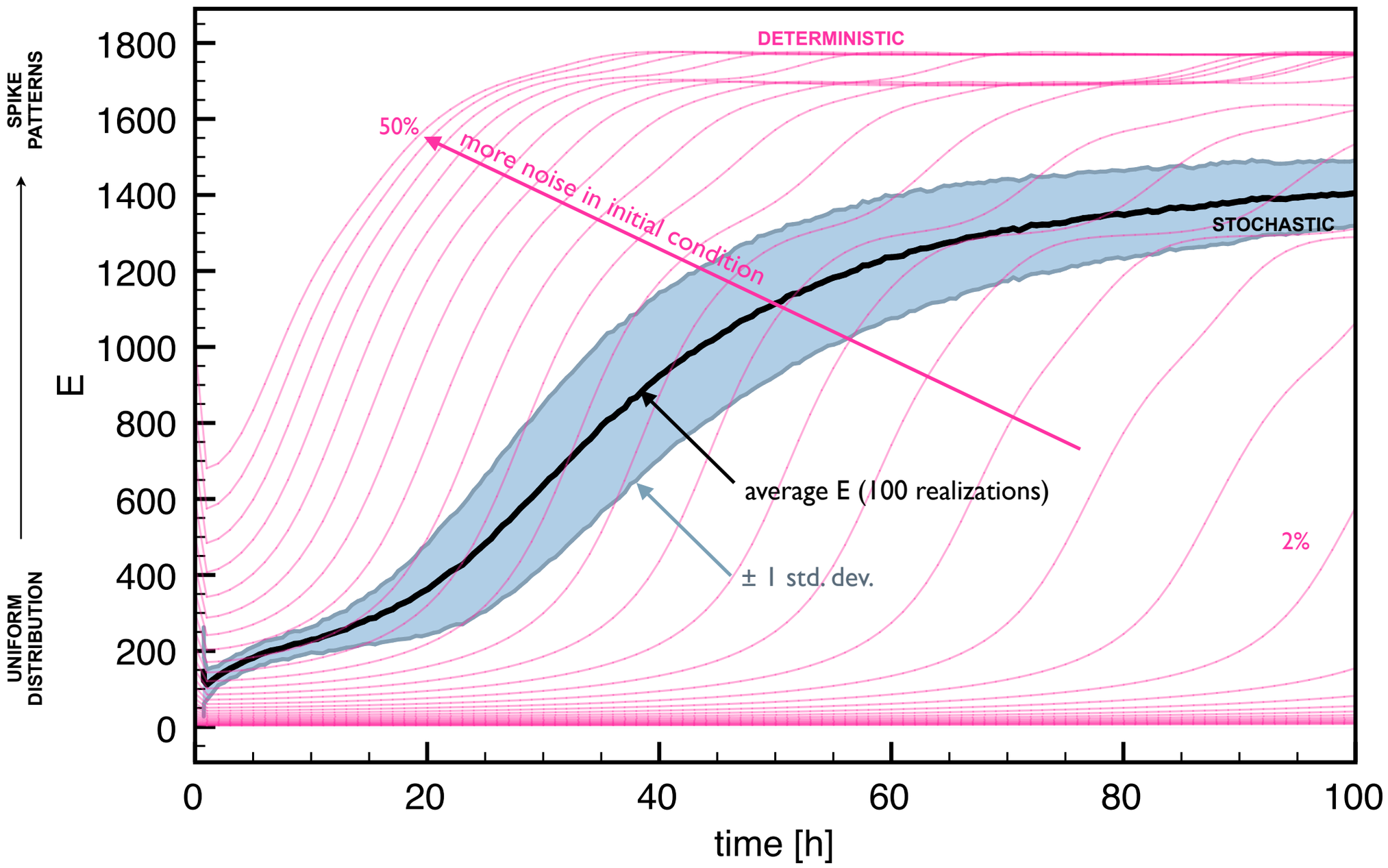
initial condition: $B_j = R_j = 3000$
$L_j = 462$



Stochastic

Dynamics: Stochastic vs. Deterministic

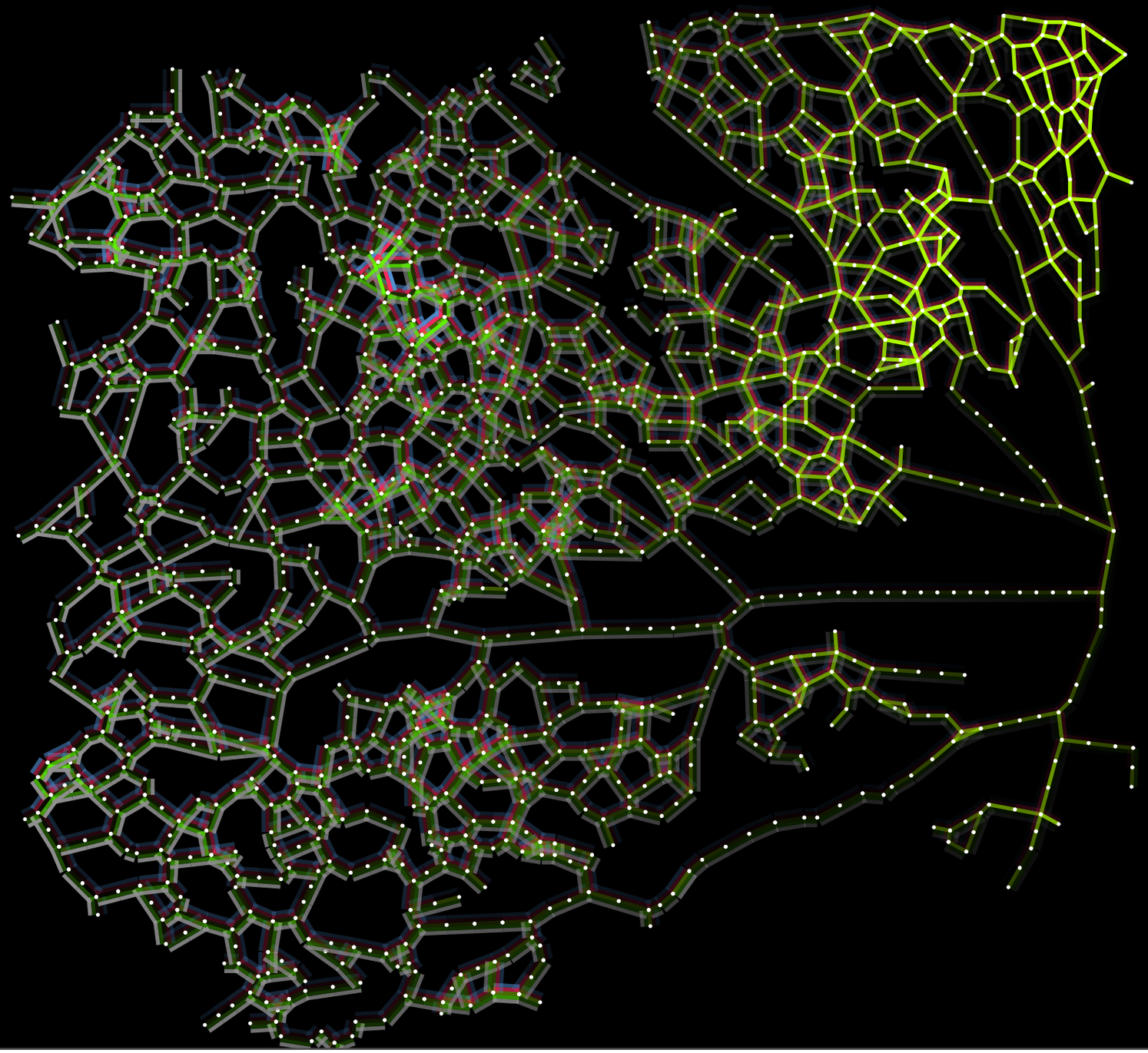
Energy of 1-spike mode $E = \sqrt{\sum_{j \in \text{cells}} |B_j - \frac{1}{2}(B_{j+1} + B_{j-1})|^2}$

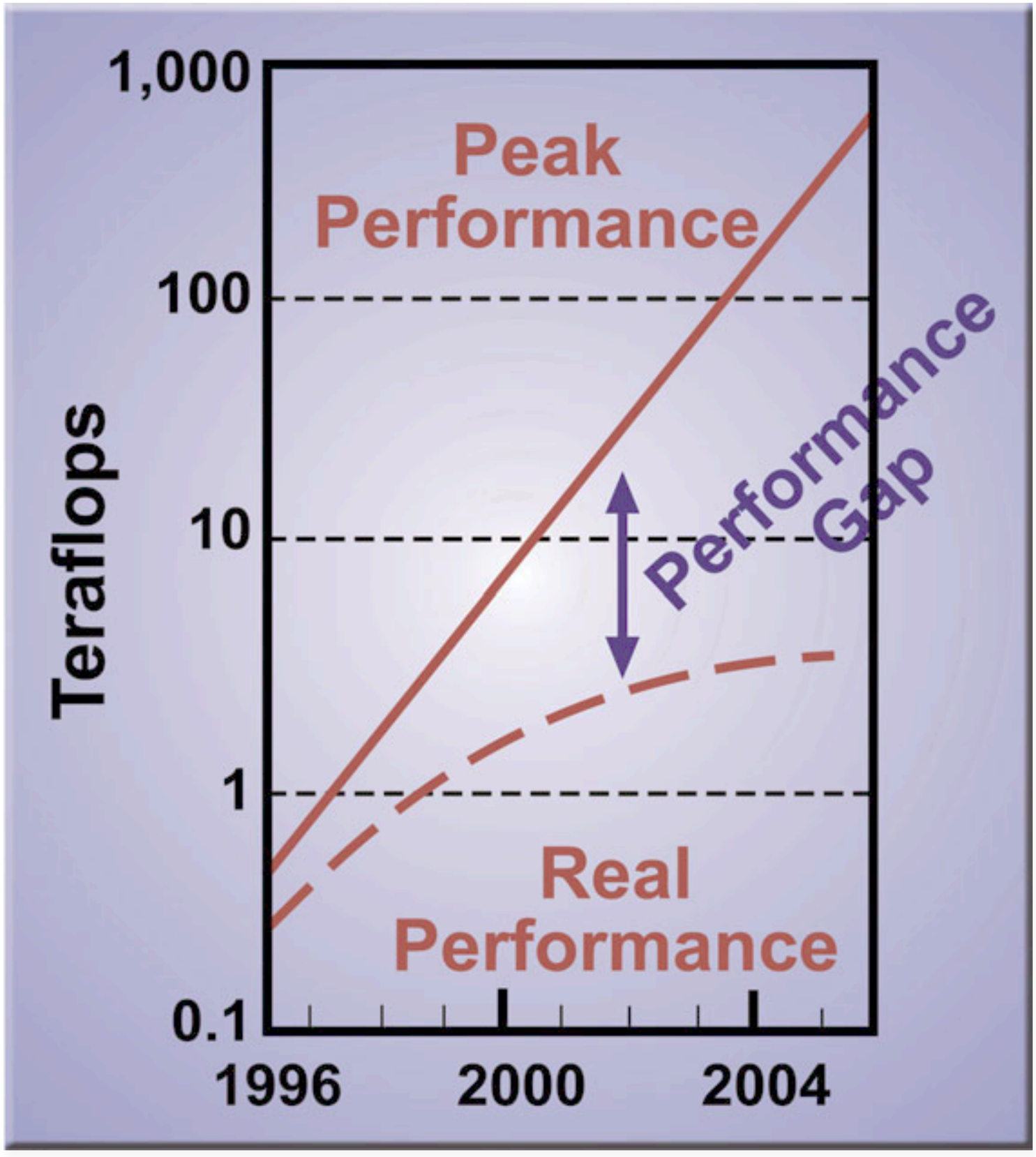


Deterministic simulation is **unable** to assess dynamics

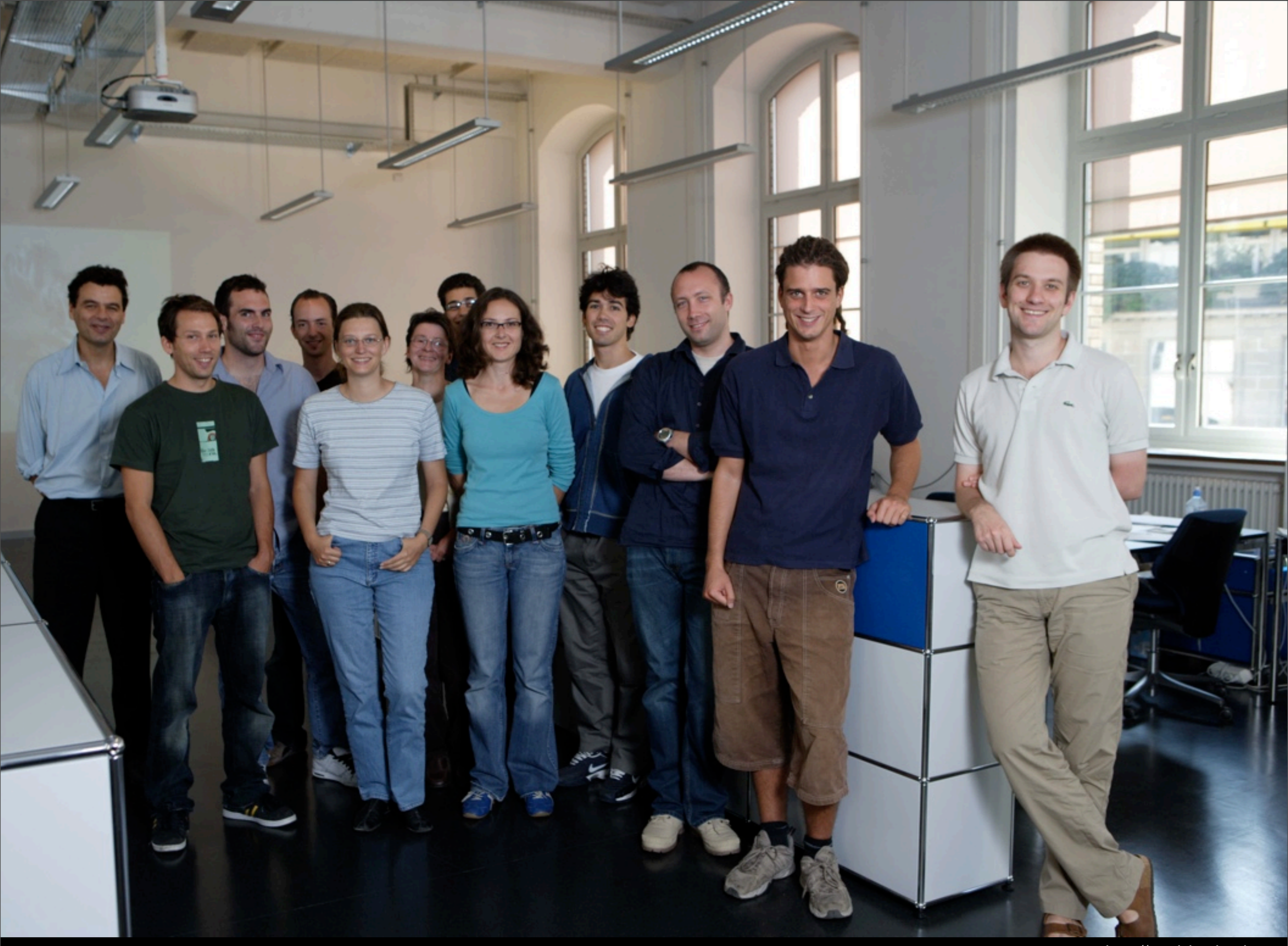
Outlook: Notch/Dll4 system on vessels

- Reverse-Engineering of Notch/Dll4 feedback system
- Coupling with Simulations of sprouting angiogenesis
- Juxtacrine signaling on vessel-graphs reconstructed from experiments




1,000
100
10
1
0.1
Teraflops
Peak Performance
Performance Gap
Real Performance
1996
2000
2004

A WARNING



http://cse-lab.ethz.ch