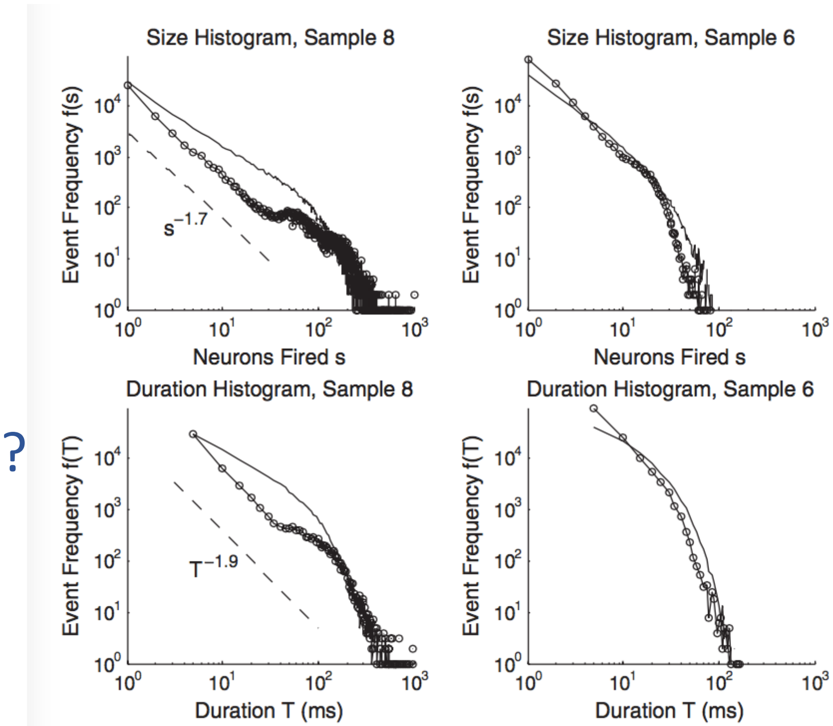
Brain Dynamics and Statistics: Simulation versus Data
BIRS, Feb 27 – March 3, 2017

Power law miscellany and variability/regularity in neurotransmitter systems

Janet Best
Department of Mathematics
The Ohio State University

Preamble, some remarks about power laws in data

- Power law + “bump”
- Short power laws (< 2 decades)
- Where does the power law end??

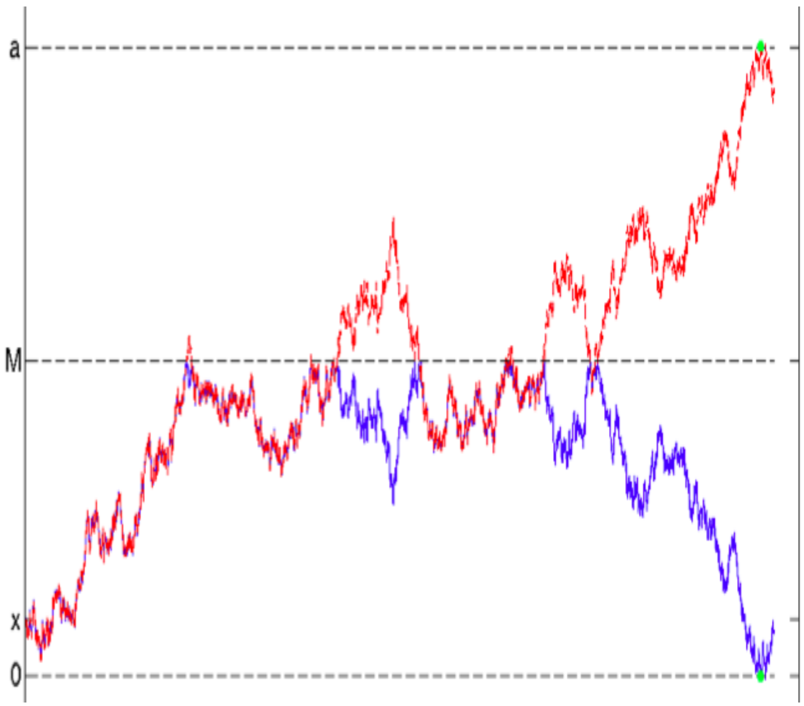


PRL **108,** 208102 (2012) PHYSICAL REVIEW LETTERS week ending 18 MAY 2012

Universal Critical Dynamics in High Resolution Neuronal Avalanche Data

Nir Friedman,[1] Shinya Ito,[2] Braden A. W. Brinkman,[1] Masanori Shimono,[2,5] R. E. Lee DeVille,[3] Karin A. Dahmen,[1]
John M. Beggs,[2] and Thomas C. Butler[4,*]

Reflected Brownian Motion (RBM)



$$\tau_{Pos \to Neg} = \inf\{t : B_t = -1 \text{ or } B_t = 2M+1 \mid B(0) = 1\}$$

For x=2, a=2M+2x, let $f^x_{T_0 \wedge T_a}$ be the p.d.f of

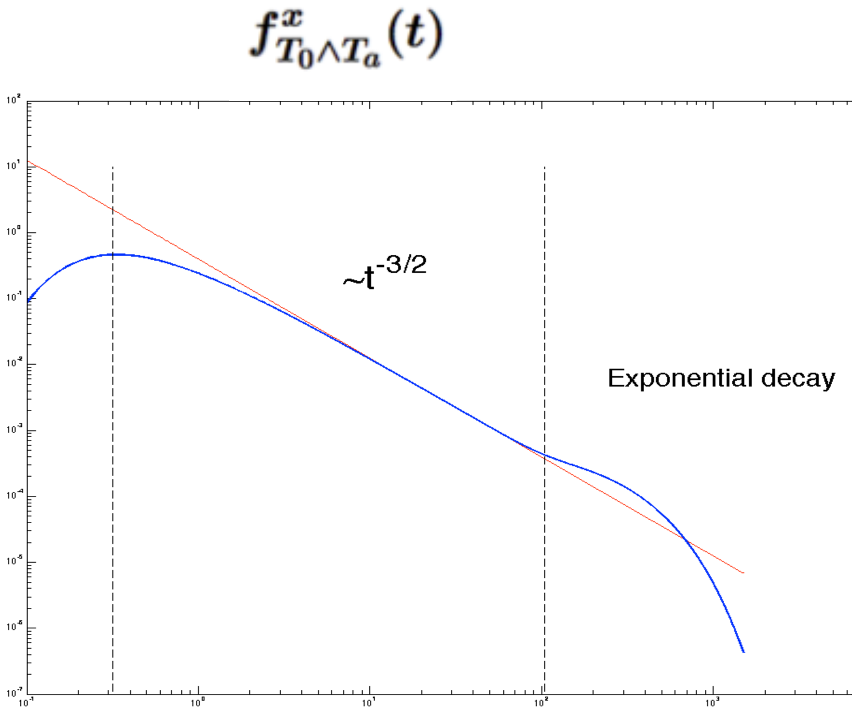
$$T_0 \wedge T_a = \min\{t : W(t) = 0 \text{ or } W(t) = a \mid W(0) = x\}.$$

$$f^x_{T_0\wedge T_a}(t) = \frac{1}{\sqrt{2\pi t^3}} \sum_{n=-\infty}^{\infty} \Big[(2na+x)\exp(-\frac{(2na+x)^2}{2t})$$

$$+ (2na+a-x)\exp(-\frac{(2na+a-x)^2}{2t}) \Big]$$

$$= \sum_{n=1}^{\infty} \frac{n\pi(1+(-1)^{(n+1)})}{a^2} \exp\left(-\frac{(n\pi)^2}{2a^2}t\right) \sin\left(\frac{n\pi x}{a}\right)$$

Intermediate power law interval for pdf of exit times

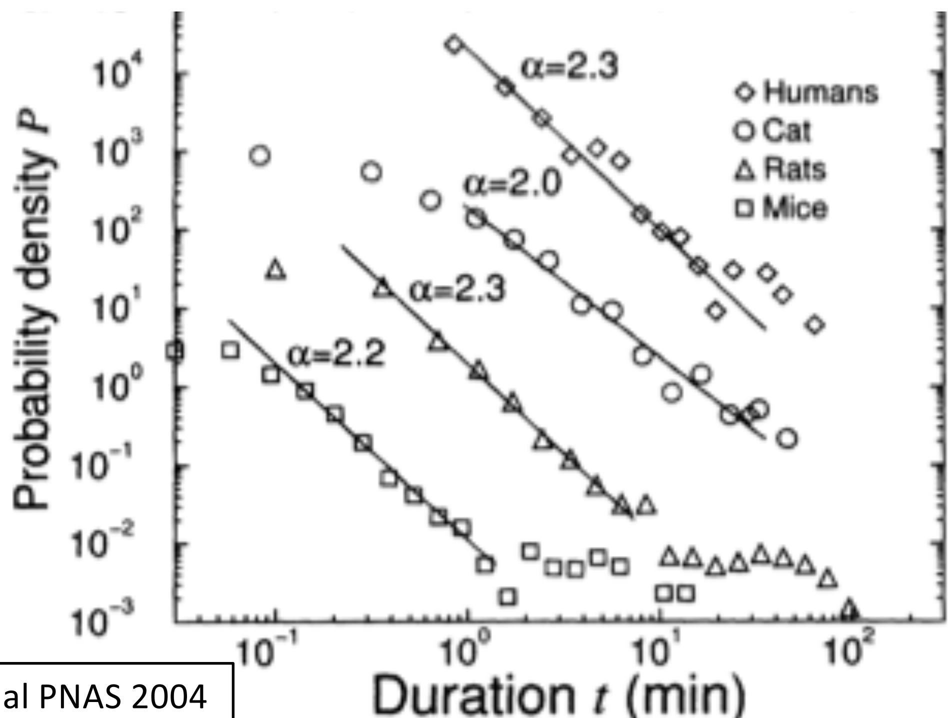


$$
\begin{aligned}
&f^x_{T_0\wedge T_a}(t)\\
&\sim \frac{2\pi}{a^2}\exp\left(-\frac{(\pi)^2}{2a^2}t\right)\sin\left(\frac{\pi x}{a}\right), \text{ if } t > C_1\frac{a^2}{\pi^2}\\
&\sim \frac{1}{\sqrt{2\pi t^3}}x\exp(-\frac{x^2}{2t}), \text{ if } t < C_2\frac{a^2}{\pi^2}\\
&\sim \frac{1}{\sqrt{2\pi t^3}}x, \text{ if } C_3x^2 < t < C_2\frac{a^2}{\pi^2}
\end{aligned}
$$

for some C_1, C_2 and C_3.

Intermediate power law interval for wake durations

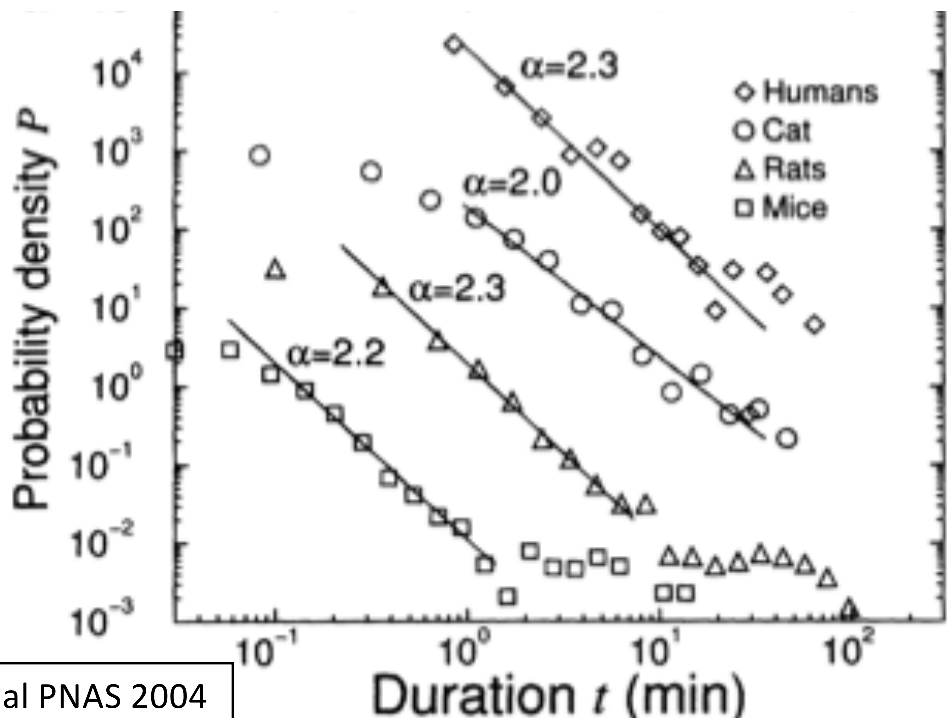
data



Lo et al PNAS 2004

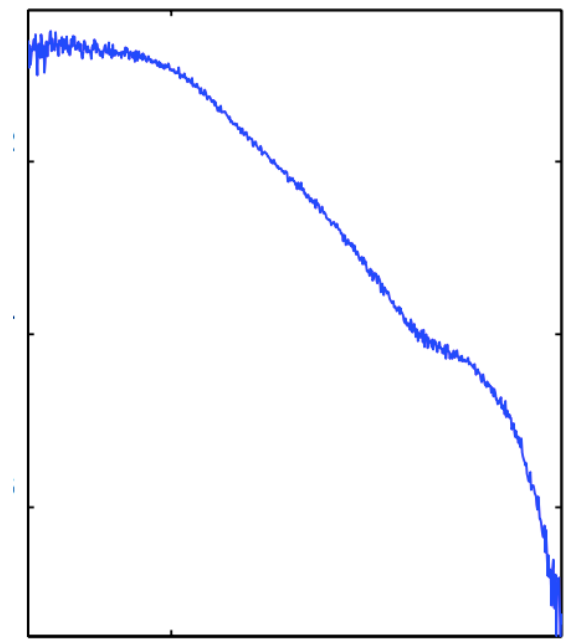
Intermediate power law interval for wake durations

data

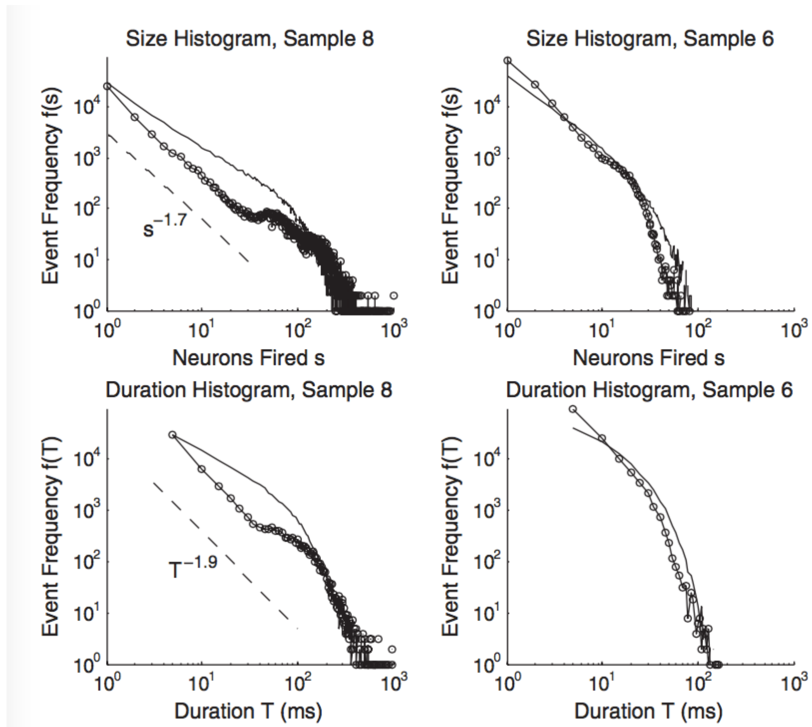


Lo et al PNAS 2004

simulation



How to responsibly fit the data *or* simulations?



For power law tails:

A very popular approach (over 1600 citations) in applied science introduced by Clauset, Shalizi, Newman 2009:

- Loop over candidates for $x_{\min}$
  - do MLE power law fit on $x \geq x_{\min}$
- Choose $x_{\min}$ and corresponding MLE $\hat{\alpha}$ with best model fit
  - measured by Kolmogorov-Smirnov (KS) distance between theoretical and empirical CDFs
- Validate power law fit
  - p-value from semiparametric bootstrap $p > 0.1$
  - likelihood ratios against other candidate models

KS method has a rather obvious extension to intermediate power laws

- Loop over both $x_{\min}$ and $x_{\max}$ candidates
- Choose interval with best power law fit in terms of KS distance
- Same validation procedures for proposed power law fit

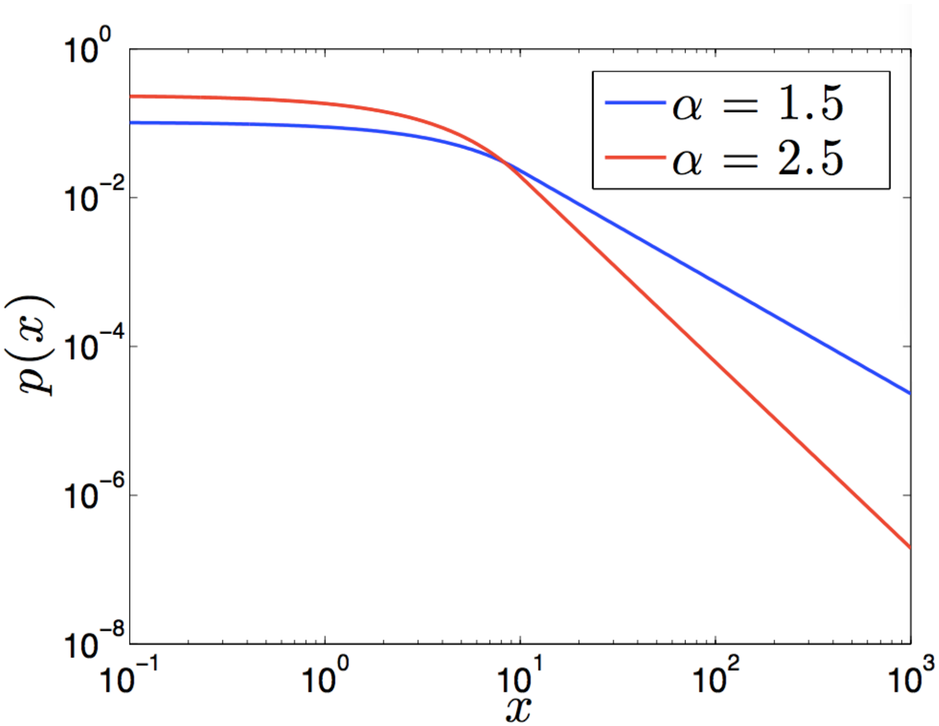
On synthetic trials, extended KS method

- succeeds in estimating the power law exponent α
- gives rather unreliable estimates for $x_{\min}$ and $x_{\max}$

Who cares whether the bounds $x_{\min}$ and $x_{\max}$ are well estimated?

- Short power law intervals generally not considered convincing
  - Two decade criterion (Stumpf and Porter 2012)
- The power law bounds themselves often reflect a meaningful cutoff length scale in theory

Illustrative example with exact power law tail



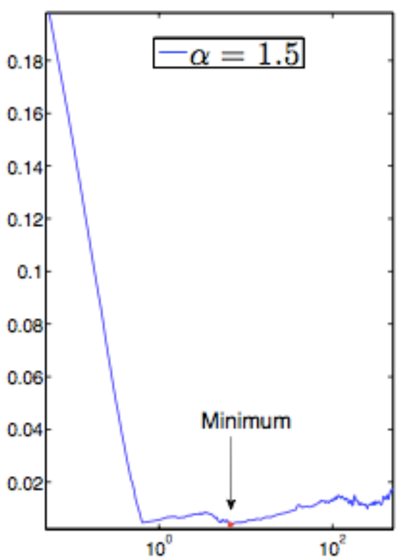
$$p(x) = \begin{cases} Ae^{-\beta x} & : 0 \leq x < x_{\min} \\ Cx^{-\alpha} & : x_{\min} \leq x \end{cases}$$

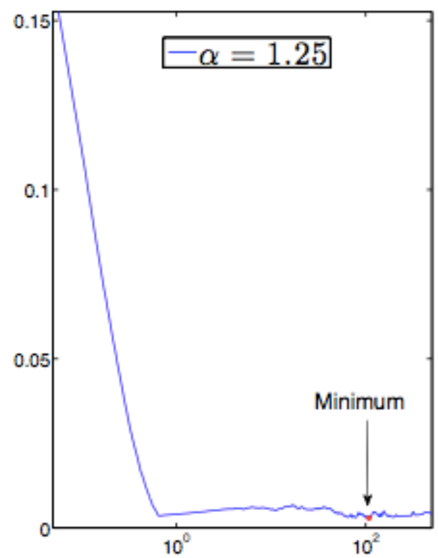
True value: $x_{\min} = 10$

Deficiency of bound estimation by KS method

The KS method exhibits some unnecessary variability because it seeks to globally optimize over a flat region with small bumps

$$\rho_{\mathrm{KS}}(x) = \sup_{y \geq x} |\hat{F}^{x}(y) - F^{x,\hat{\alpha}(x)}(y)|$$





Why not choose $x_{\min}$ as smallest value in the mostly flat region?

Adaptively penalized KS (apKS) method

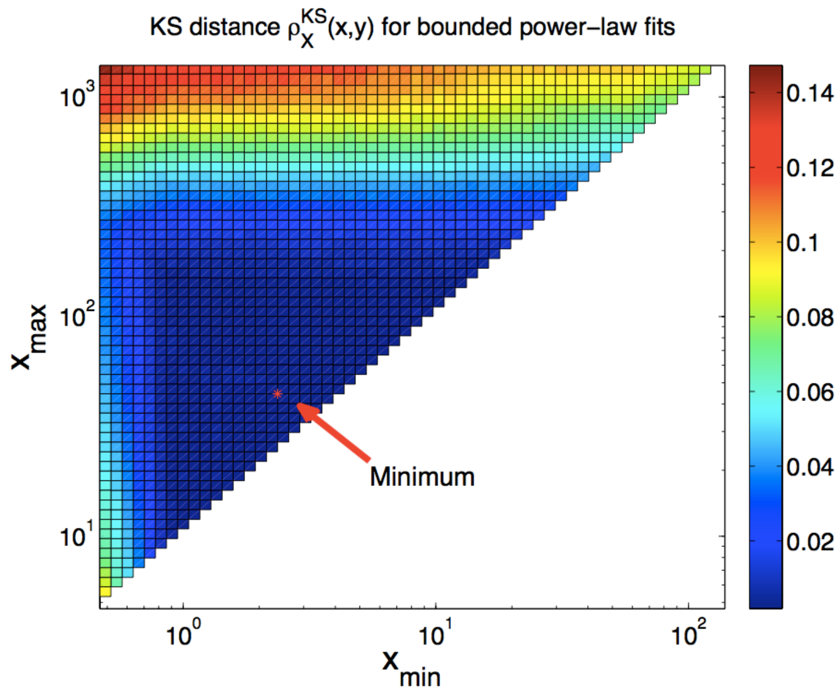
Optimize instead the penalized KS distance

$$\rho_X^{pKS}(x) = \rho_X^{KS}(x) + d \log\left(\frac{x}{x_c}\right),$$

How choose penalty coefficient d? Adaptive iteration

- increase if the $x_{\min}$ produced passes validation step
- decrease otherwise

Flatness of minimum KS distance makes selection of the interval [x_{min}, x_{max}] highly variable between samples from the same probability distribution



An adaptive penalization process finds a *balance* between small KS distance and large interval for validated power law fit.

As data set size N increases, $\hat{\alpha}$ improves as $\sqrt{N}$

$\boxed{\textbf{but } \hat{x}_{min} \textbf{ and } \hat{x}_{max} \textbf{ do not.}}$

Batching: possible solution when data is plentiful (e.g., from simulating a model):

(1) Split a data set into b disjoint subdata sets (i.e. *batches*) whose union is the original data set.

(2) Run the KS (or apKS) method on each batch and validate each bounded power law fit by estimating a p-value obtained by using semi-parametric bootstrap samples. By collecting the estimates from each batch, we obtain power law exponents $\hat{\alpha}^{(1)}$, $\ldots, \hat{\alpha}^{(b)}$, bounded power law intervals $[\hat{x}_{\text{min}}^{(1)}, \hat{x}_{\text{max}}^{(1)}], \ldots, [\hat{x}_{\text{min}}^{(b)}, \hat{x}_{\text{max}}^{(b)}]$, along with the $p-$values.

(3) If all of the bounded power law fits are validated (sufficiently large p-value), report the average of the estimated exponents and bounds as the bounded power law parameters.The bounded power law hypothesis is deemed not valid otherwise.

Intermediate asymptotic power law

Synthetic probability distribution

$$p(x) = C(x + x_{\min})^{-\alpha} e^{-\beta x}, \; x \geq x_0.$$

has intermediate asymptotic power law (IAPL) region, in the sense that

$$p(x) \sim C x^{-\alpha} \text{ for } x_{\min} \ll x \ll x_{\max}$$

with $x_{\max} = \alpha/\beta - x_{\min}$ rather than

$$p(x) = C x^{-\alpha} \text{ for } x_{\min} \leq x \leq x_{\max}$$

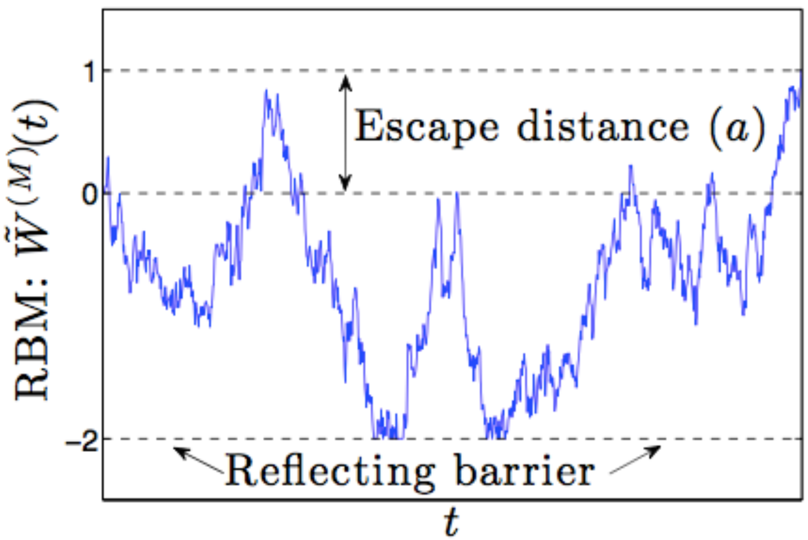
$x_{\min}$, $x_{\max}$ appear in terms of model parameters but not strict boundary

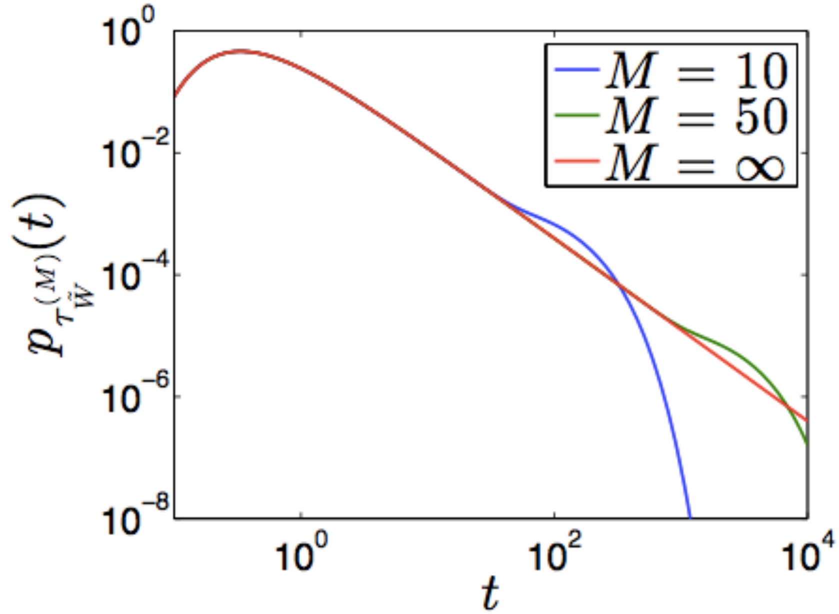
Parametric scaling of bounds

Bounds of IAPL regions may have scaling with respect to meaningful model parameters.

First passage time $\tau_{\tilde{W}}^{(M)} = \inf\{t > 0 : \tilde{W}^{(M)}(t) = a\}$ of reflected Brownian motion $\tilde{W}^{(M)}(t) = |W(t) + M| - M$ has explicit PDF $p_{\tau_{\tilde{W}}^{(M)}}(t)$ expressed as infinite series, with IAPL

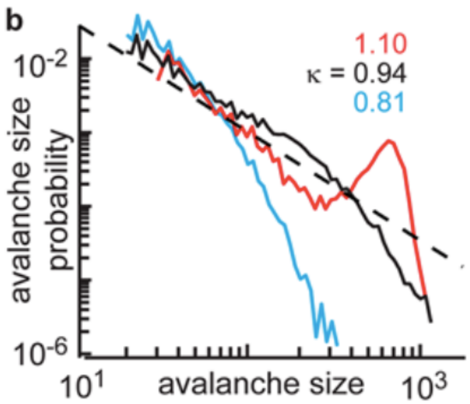
$$p_{\tau_{\tilde{W}}^{(M)}}(t) \sim \frac{a}{\sqrt{2\pi}} t^{-3/2} \text{ for } a^2 \ll t \ll M^2$$



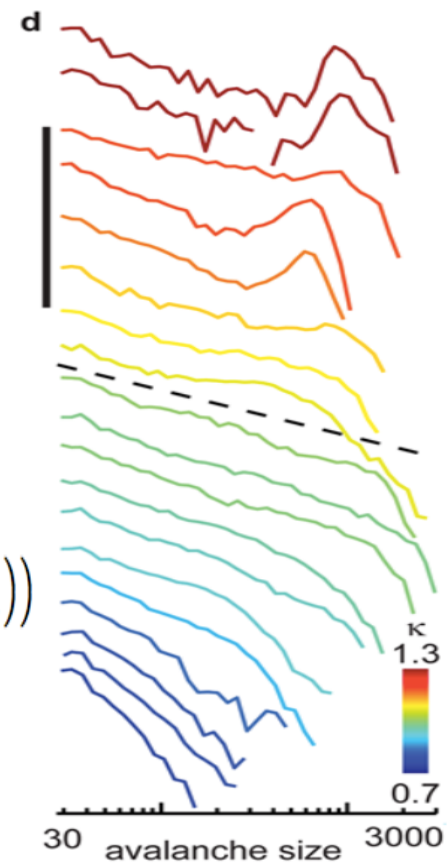


$$p_{\tau^{(M)}_{\tilde{W}}}(t) \sim \frac{a}{\sqrt{2\pi}} t^{-3/2} \text{ for } a^2 \ll t \ll M^2$$

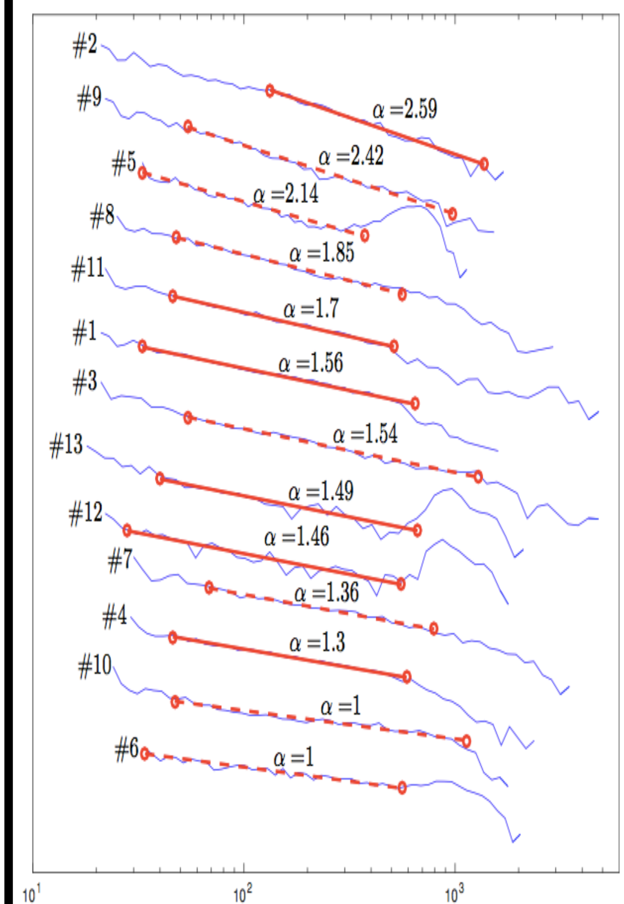
Application



$$\kappa = 1 + \frac{1}{10}\sum_{i=1}^{10}\left(F^{\mathrm{NA}}(\beta_i) - F(\beta_i)\right)$$



Gautam et al 2015



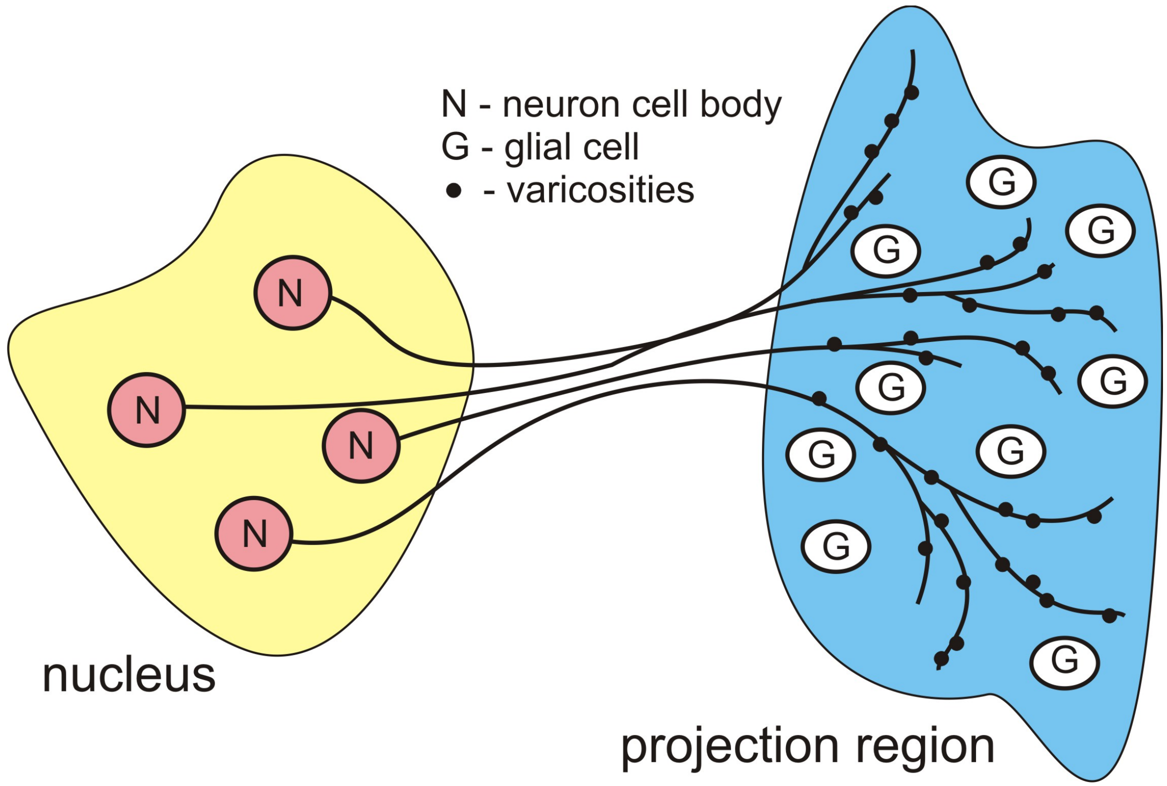
apKS

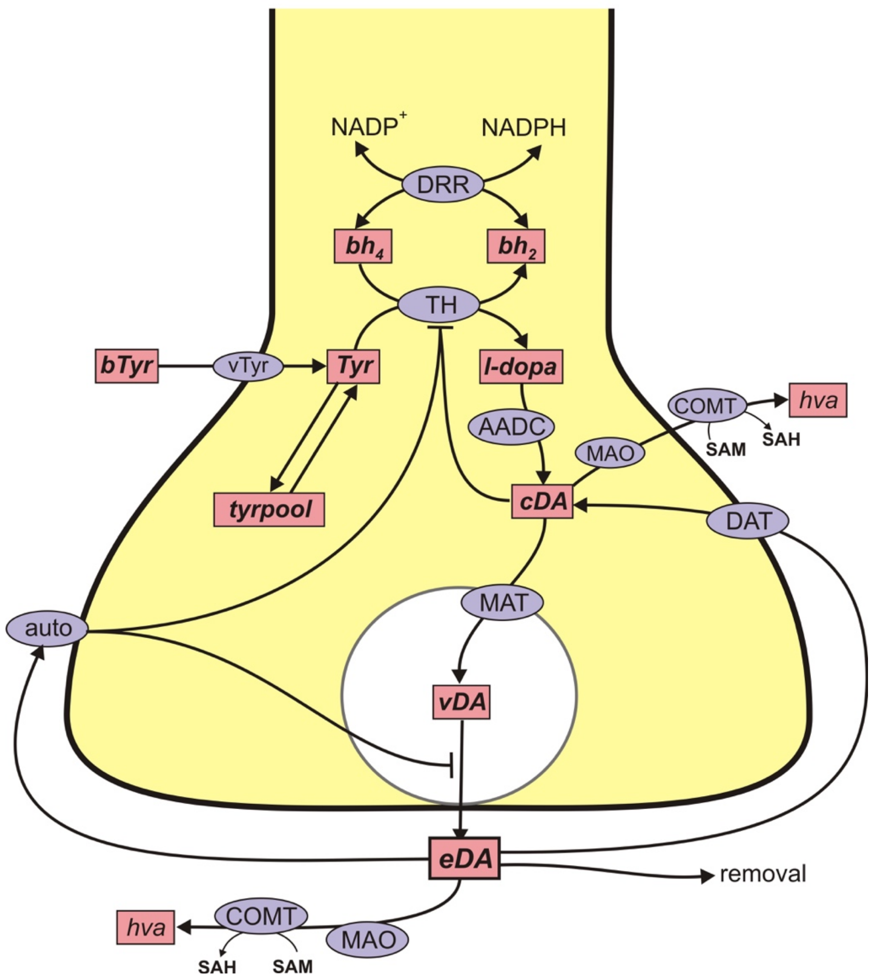
Conclusions – power law interval fitting

An adaptively penalized version of the KS method for inferring power law tails

- gives reasonable quality estimates for bounds of intermediate asymptotic power law regions
- allows inference of parametric scaling of power law region bounds
- performs better on an ensemble of data sets of sizes $\gtrsim 10^4$
  - batch for larger data sets

Volume transmission



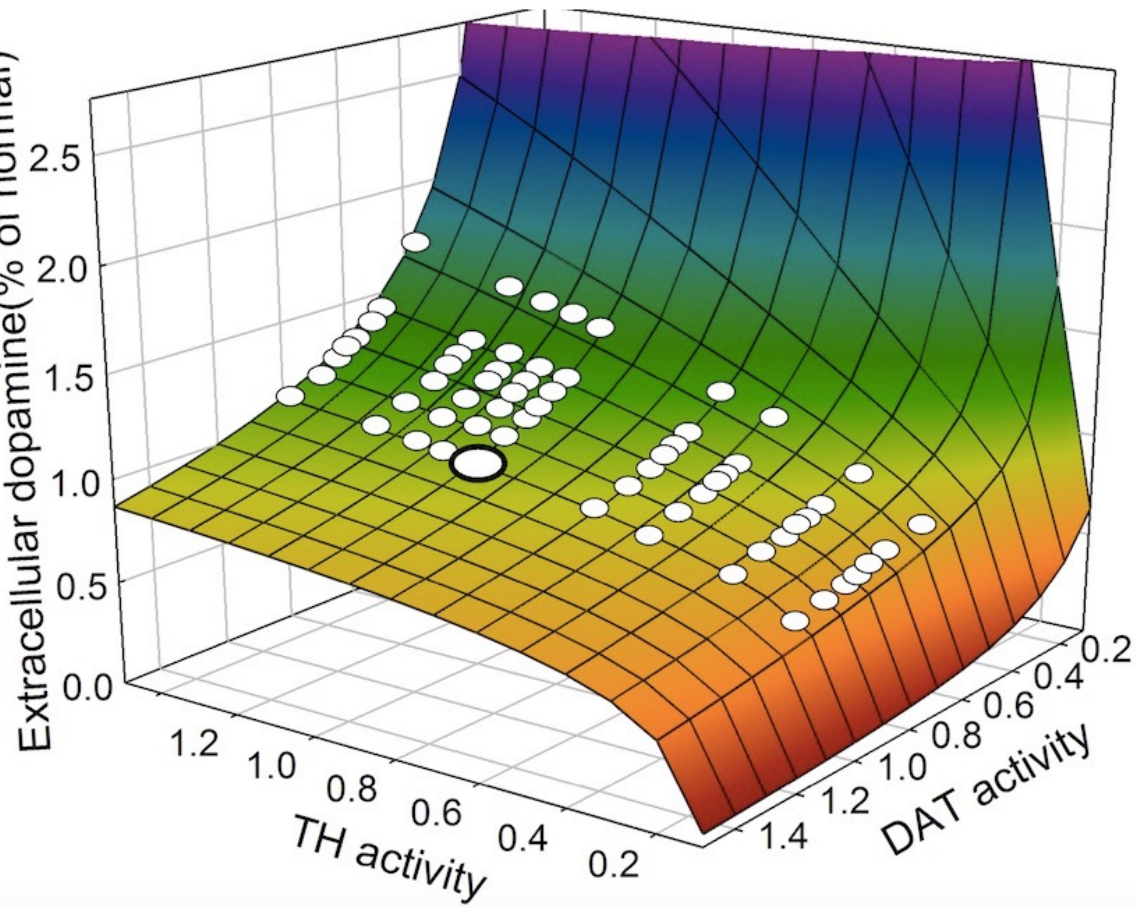


tyr = tyrosine

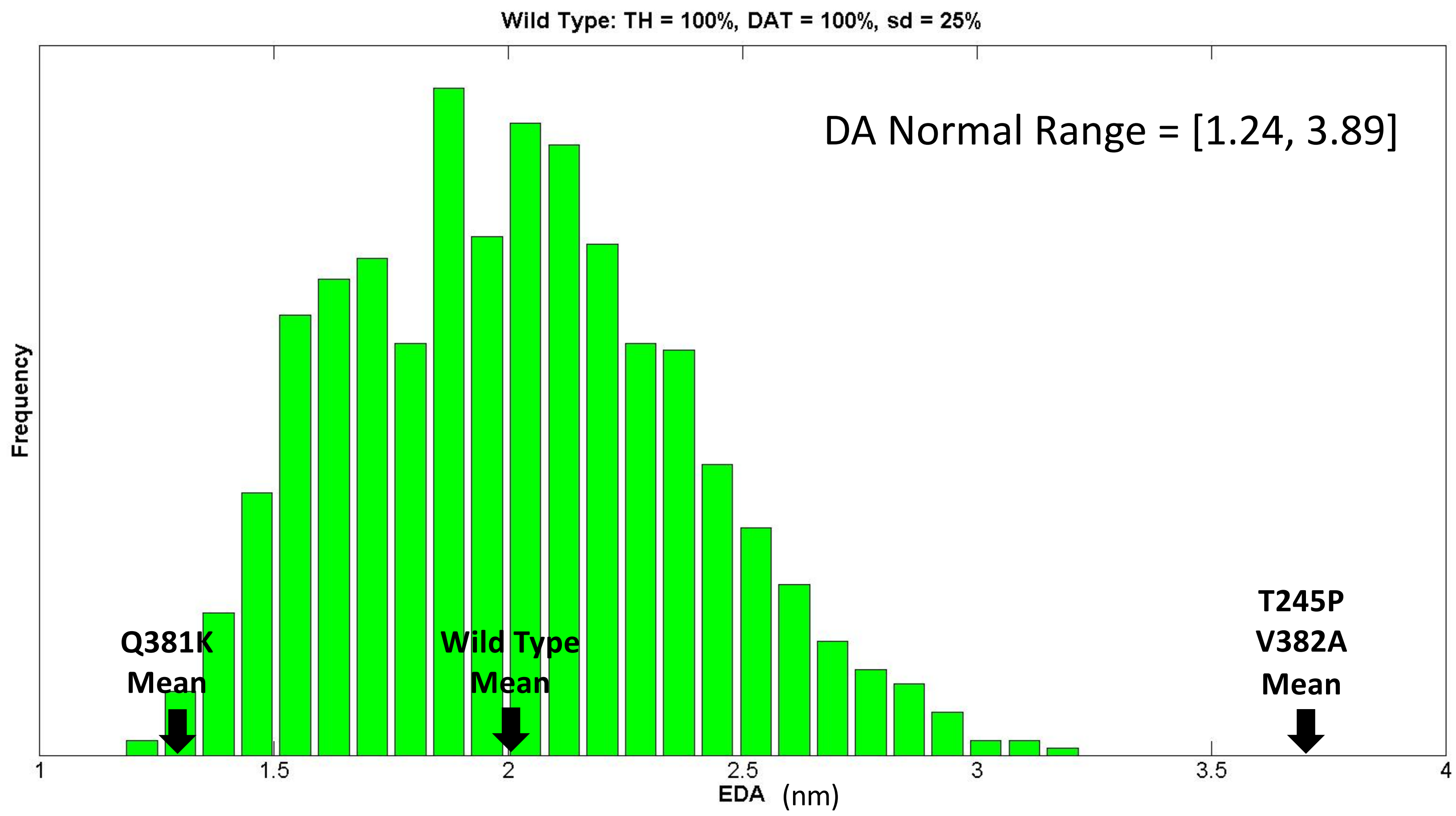
cda = cytosolic dopamine

vda = vesicular dopamine

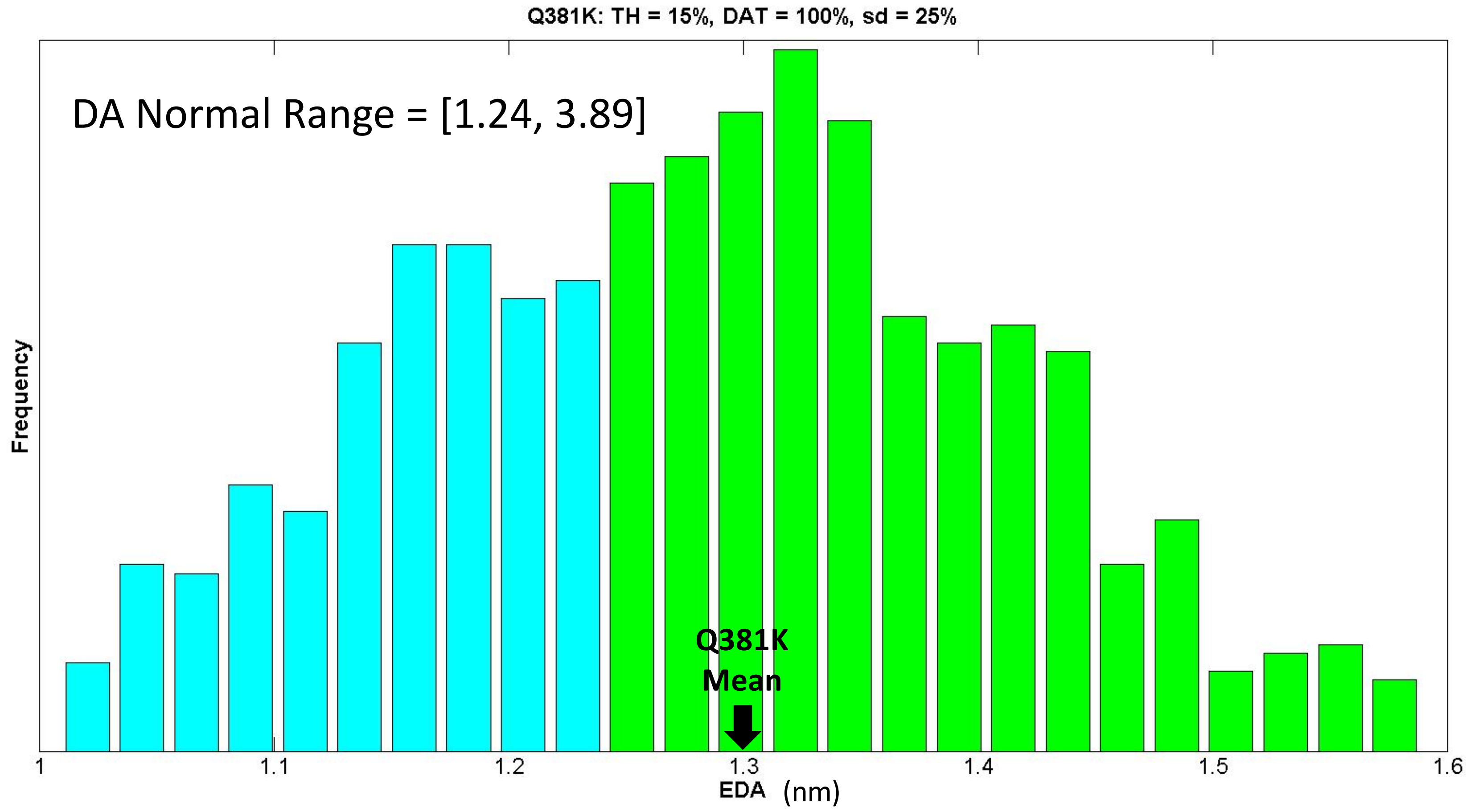
eda = extracellular dopamine



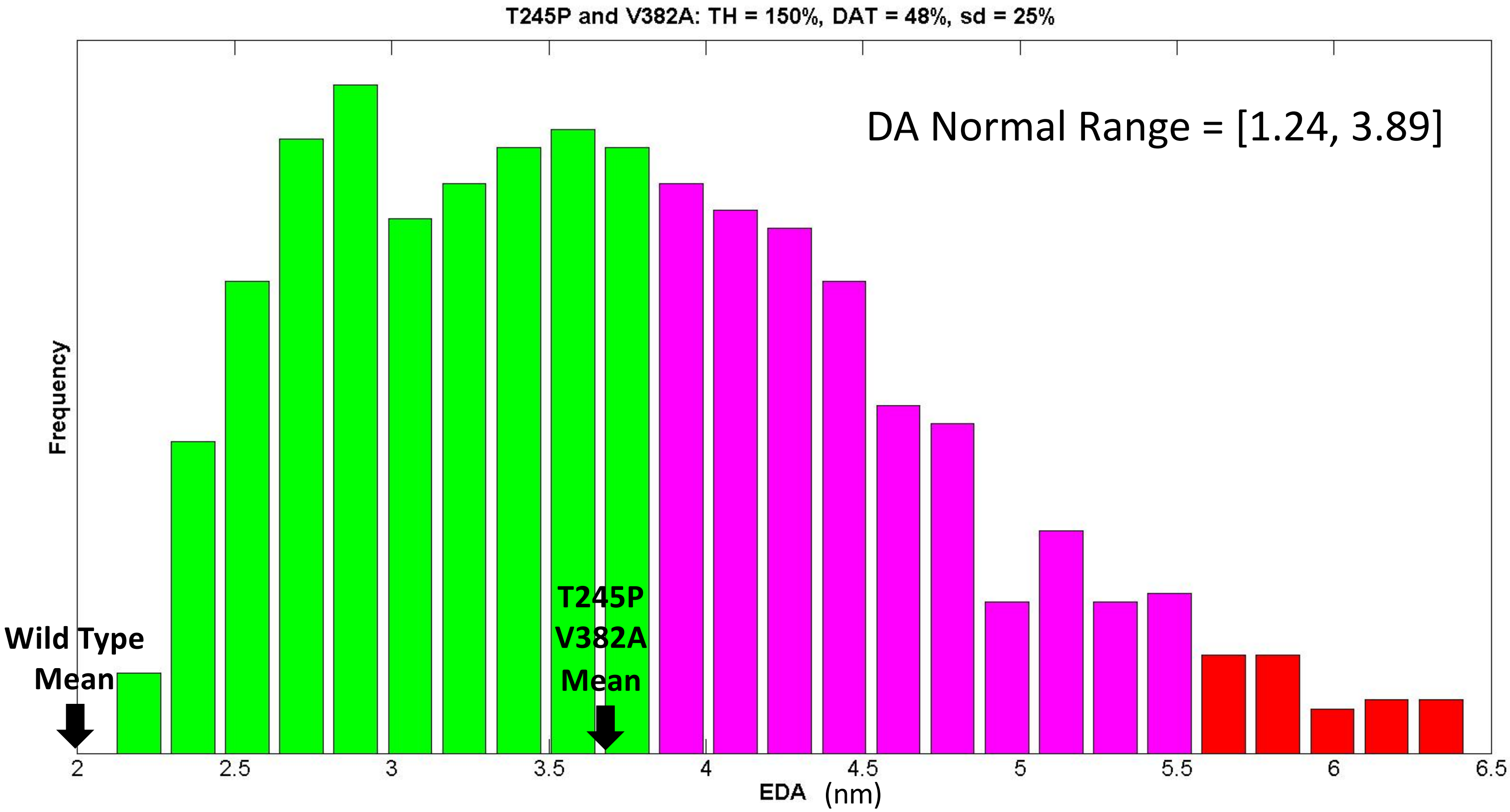
Extracellular dopamine(% of normal)
2.5
2.0
1.5
1.0
0.5
0.0
1.2
1.0
0.8
0.6
0.4
0.2
TH activity
0.2
0.4
0.6
0.8
1.0
1.2
1.4
DAT activity



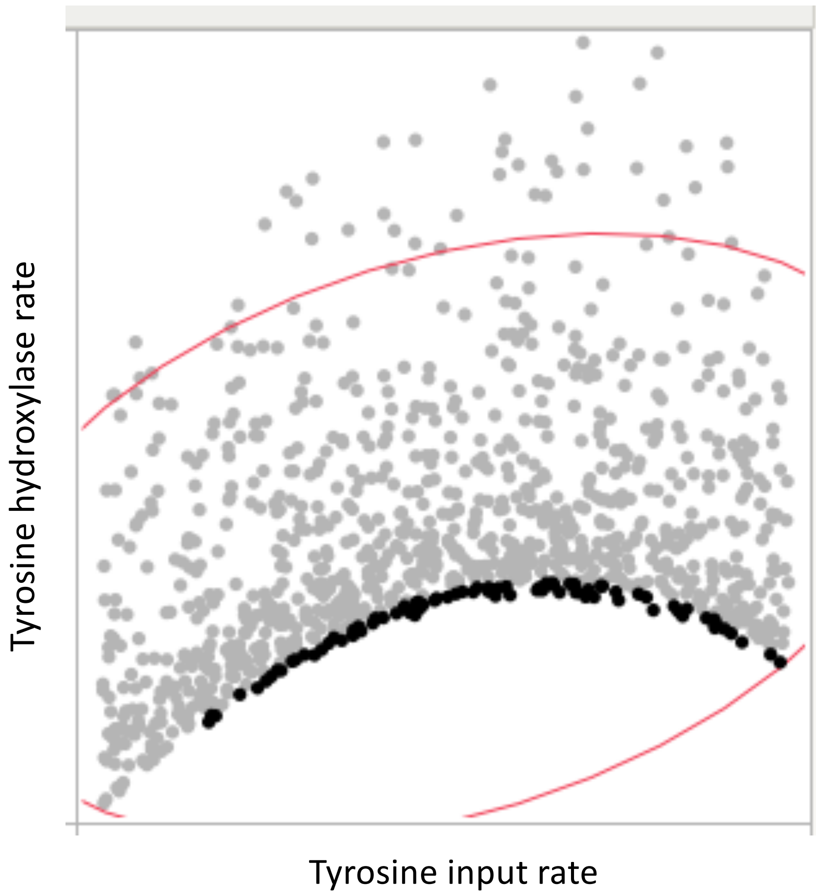
Wild Type: TH = 100%, DAT = 100%, sd = 25%
DA Normal Range = [1.24, 3.89]
Frequency
Q381K
Mean
Wild Type
Mean
T245P
V382A
Mean
1
1.5
2
2.5
3
3.5
4
EDA (nm)



Q381K: TH = 15%, DAT = 100%, sd = 25%
DA Normal Range = [1.24, 3.89]
Frequency
Q381K
Mean
1
1.1
1.2
1.3
1.4
1.5
1.6
EDA (nm)

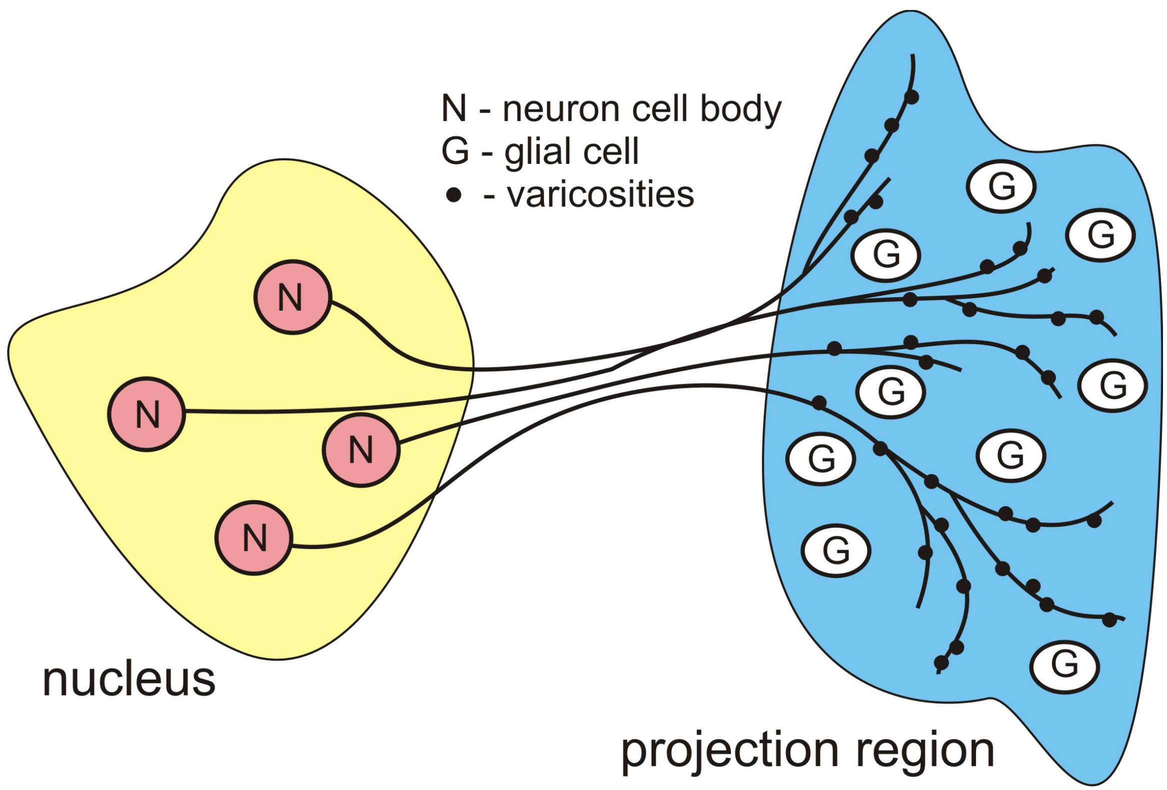


T245P and V382A: TH = 150%, DAT = 48%, sd = 25%
DA Normal Range = [1.24, 3.89]
Frequency
Wild Type Mean
T245P V382A Mean
2
2.5
3
3.5
4
4.5
5
5.5
6
6.5
EDA (nm)



Black dots represent very high extracellular dopamine

Volume transmission



Volume transmission, questions

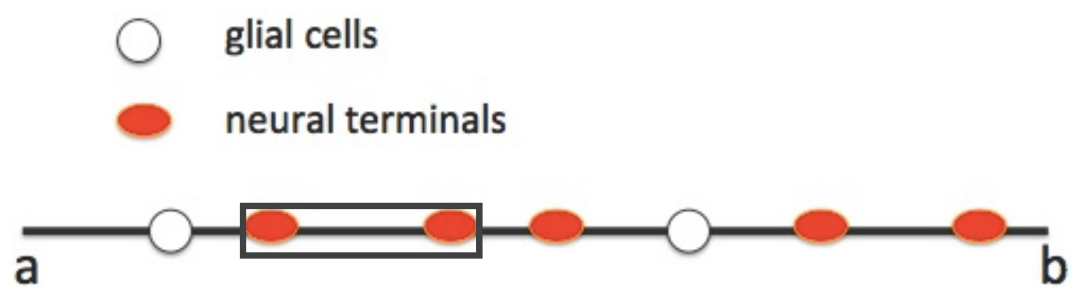
Given the statistics of the stochastic firing of each neuron,

- How to calculate mean neurotransmitter level over whole extracellular space?
- How to calculate the spatial dependence of expected neurotransmitter level?
- How do these answers depend on firing rates, amounts released, distances between terminals, diffusion constants, etc?

1-dimensional extracellular space

1-dimensional extracellular space

1-dimensional extracellular space



1-dimensional extracellular space

$$\partial_t u = D\Delta u \quad \text{in } (0, L)$$

$$(q)\begin{cases}\partial_x u(0,t) = 0\\ u(L,t) = 0\end{cases} \qquad \text{and} \qquad (f)\begin{cases}\partial_x u(0,t) = 0\\ \partial_x u(L,t) = c > 0\end{cases}$$

quiescent

firing

stochastic hybrid system

Continuous-time stochastic process with

- continuous component $(X_t)_{t \geq 0}$
- jump component $(J_t)_{t \geq 0}$: jump process on finite set. For each element of state space, assign some continuous dynamics to X_t.

In between jumps of J_t, the component X_t evolves according to the dynamics associated with the current state of J_t

E.g., the stochastic process $u(t,x) \in L^2[0,L]$ that solves

$$\partial_t u = D\Delta u \quad \text{in } (0,L)$$

$$u(0,t) = 0 \quad \text{and} \quad J_t u_x(L,t) + (1 - J_t)(u(L,t) - b) = 0$$

Lawley, Mattingly, Reed "Stochastic switching in infinite dimensions with applications to random parabolic PDE." *SIAM J Math Anal* 2015

$$\partial_t u = D\Delta u \quad \text{in } (0, L)$$

$$(q)\begin{cases}\partial_x u(0,t) = 0 \\ u(L,t) = 0\end{cases} \qquad \text{and} \qquad (f)\begin{cases}\partial_x u(0,t) = 0 \\ \partial_x u(L,t) = c > 0\end{cases}$$

Can show: the mean of $u(x,t)$ is constant in x at large time.

(the process converges in distribution to an $L^2[0,L]$-valued random variable $u(x)$ with constant expectation for almost every x in [0,L].)

Lawley, Best, Reed, DCDS-B 2016

$$\partial_t u = D\Delta u \quad \text{in } (0, L)$$

$$(q) \begin{cases} \partial_x u(0,t) = 0 \\ u(L,t) = 0 \end{cases} \qquad \text{and} \qquad (f) \begin{cases} \partial_x u(0,t) = 0 \\ \partial_x u(L,t) = c > 0 \end{cases}$$

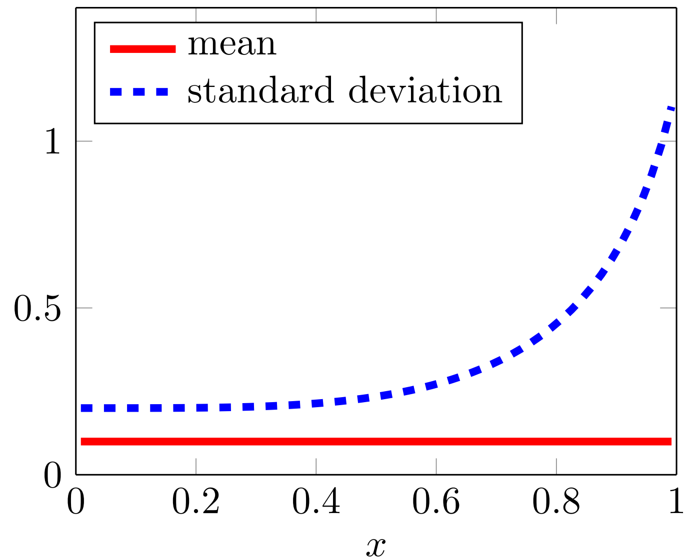
If the switching time distributions, μ_f and μ_q, are exponential with rates r_f and r_q, then the constant value of the expectation is

$$\mathcal{M} \quad = \quad c\frac{\mu}{\eta} \coth L\eta$$

where $\mu := \frac{r_q}{r_f}$ and $\eta := \sqrt{\frac{r_f + r_q}{D}}$

Lawley, Best, Reed, DCDS-B 2016

If the switching time distributions, μ_f and μ_q, are exponential



Large time mean and standard deviation for the process.

$L=D=r_q=1$
$c=r_f=100$

$$\mathcal{M} = c\frac{\mu}{\eta}\coth L\eta$$

where $\mu := \frac{r_q}{r_f}$ and $\eta := \sqrt{\frac{r_f + r_q}{D}}$

- Increase μ: increase *M*
- μ constant, increase r_q, r_f: *M* decreases
- μ constant, decrease r_q, r_f: *M* increases
- Decrease/increase D: decrease/increase *M*
- *M* gets smaller as L increases. *But*, once *L* is large compared to η, *M* is almost independent of *L*:

$$\mathcal{M} \approx c\frac{\mu}{\eta}$$

real neural parameters

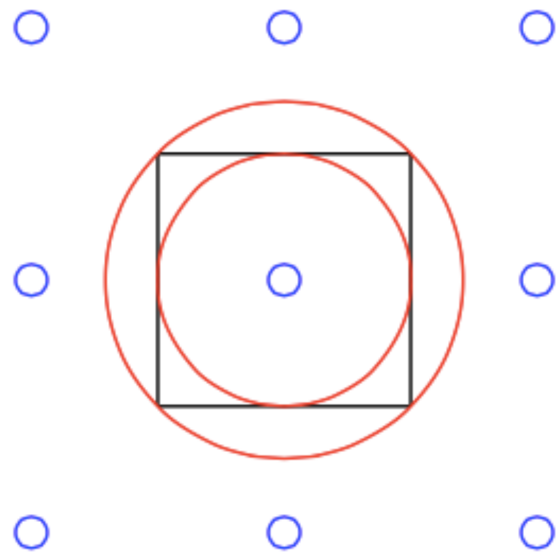
Many dopaminergic and serotonergic neurons fire at a basal rate of about 1 spike/sec
Assume that the release of neurotransmitter lasts about 5 milliseconds
Then reasonable values are $r_q = 1/sec, r_f = 200/sec$
$\mu = \frac{r_q}{r_f} = \frac{1}{200}$.
For dopamine, $D \approx 10^{-6}(cm)^2/sec$, so

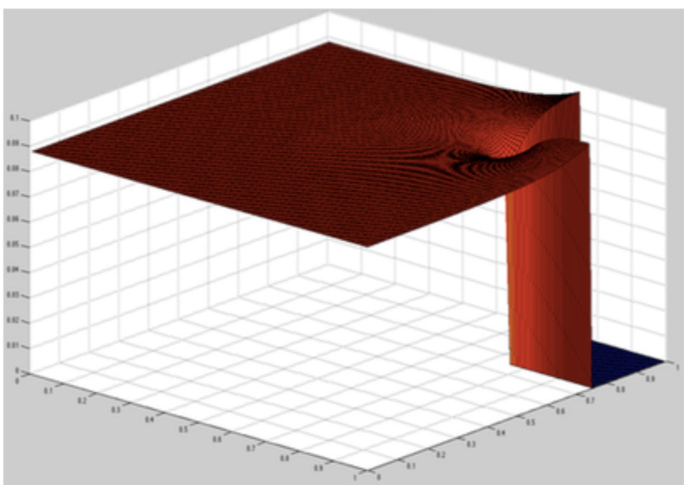
$$\eta \; = \; \sqrt{\frac{r_q + r_f}{D}} \; \approx \; \sqrt{2}\, 10^4/cm \; = \; \sqrt{2}\,/\mu m.$$

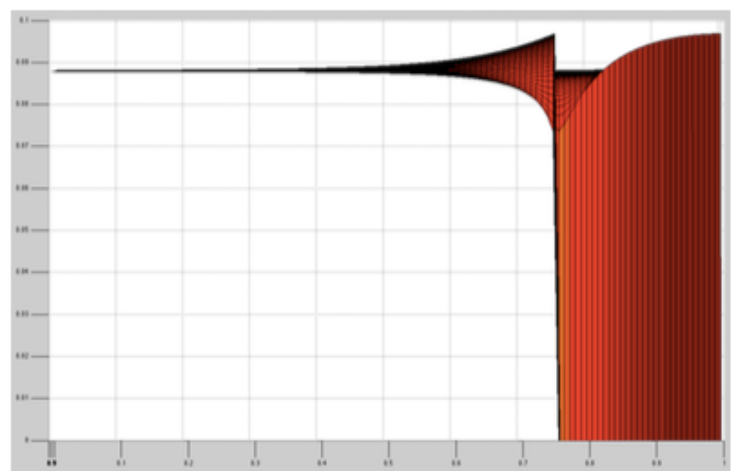
About $(2.6)10^6$ terminals per cubic millimeter or a distance of about 7μm between terminals. If we assume that $7\mu\text{m} \leq L \leq 20\mu\text{m}$, then

$$9.9 \; \leq \; \eta L \; = \; (\frac{\sqrt{2}}{\mu m})(L\mu m) \; \leq \; 28.$$

Thus $\coth{(\eta L)} \approx 1$ and we are well within the range of L where $\mathcal{M}$ is approximately independent of L.







Thanks!

Power laws

Peter Kramer, *Rensselaer Polytechnic Institute*
John Fricks, *Arizona State University*
Deena Schmidt, *U Nevada at Reno*
Jung Eun Kim, *UNIST, Korea*
Fatih Olmez, *Shanghai Center for Mathematical Sciences*
Anthony Trubiano, *Rensselaer Polytechnic Institute*
Woody Shew, *U Arkansas*
Mark Blumberg, *U Iowa*

Volume transmission

Mike Reed, *Duke*
Fred Nijhout, *Duke*
Sean Lawley, *U Utah*
Farrah Sadre-Marandi, *MBI, OSU*

mbi

Funding:

National Science Foundation
WHERE DISCOVERIES BEGIN

National Institutes of Health