Efficient data transmission over dynamical complex networks

Giacomo Baggio

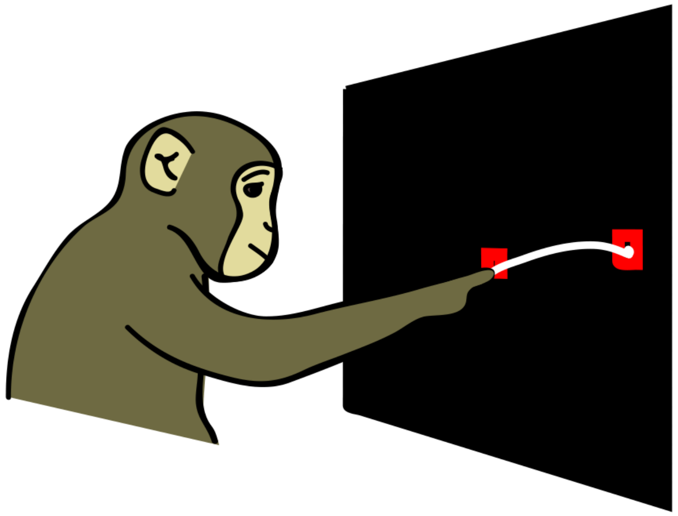
Virginia Rutten

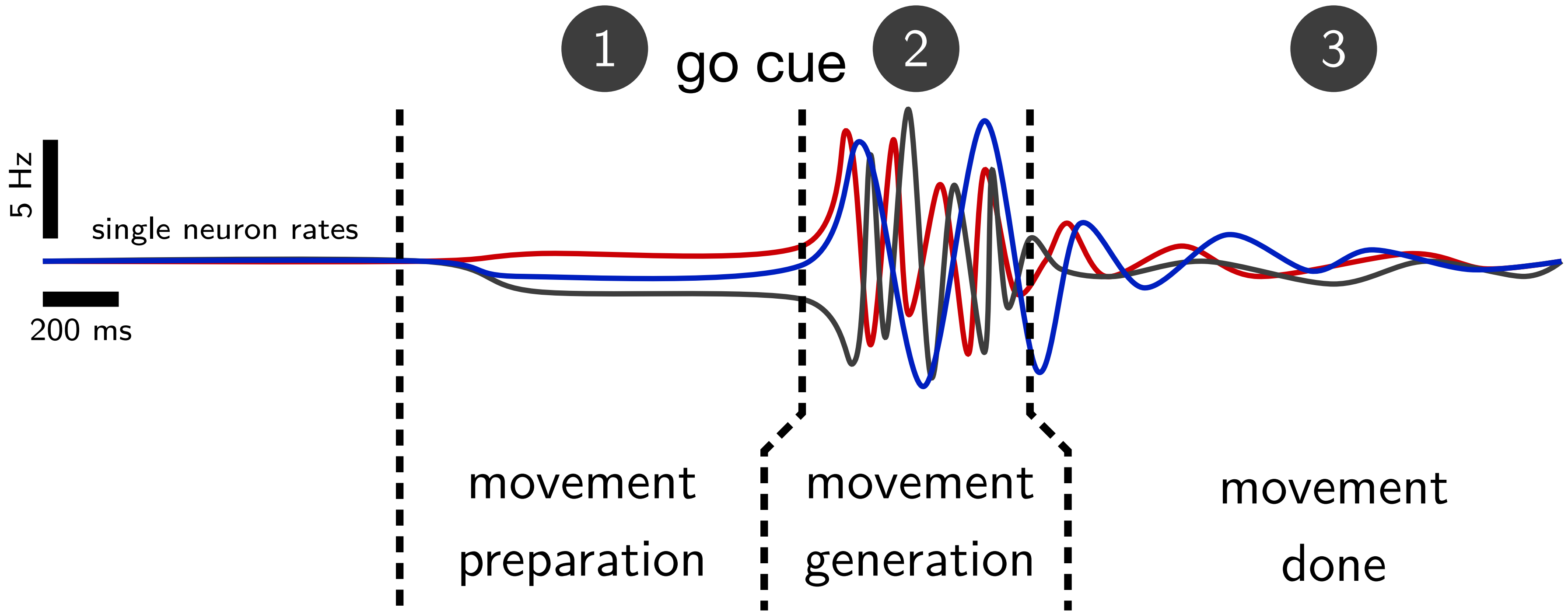
Guillaume Hennequin

Sandro Zampieri

Motor cortex dynamics: from experiments to dynamical models

[Churchland et al., Nature (2012)]





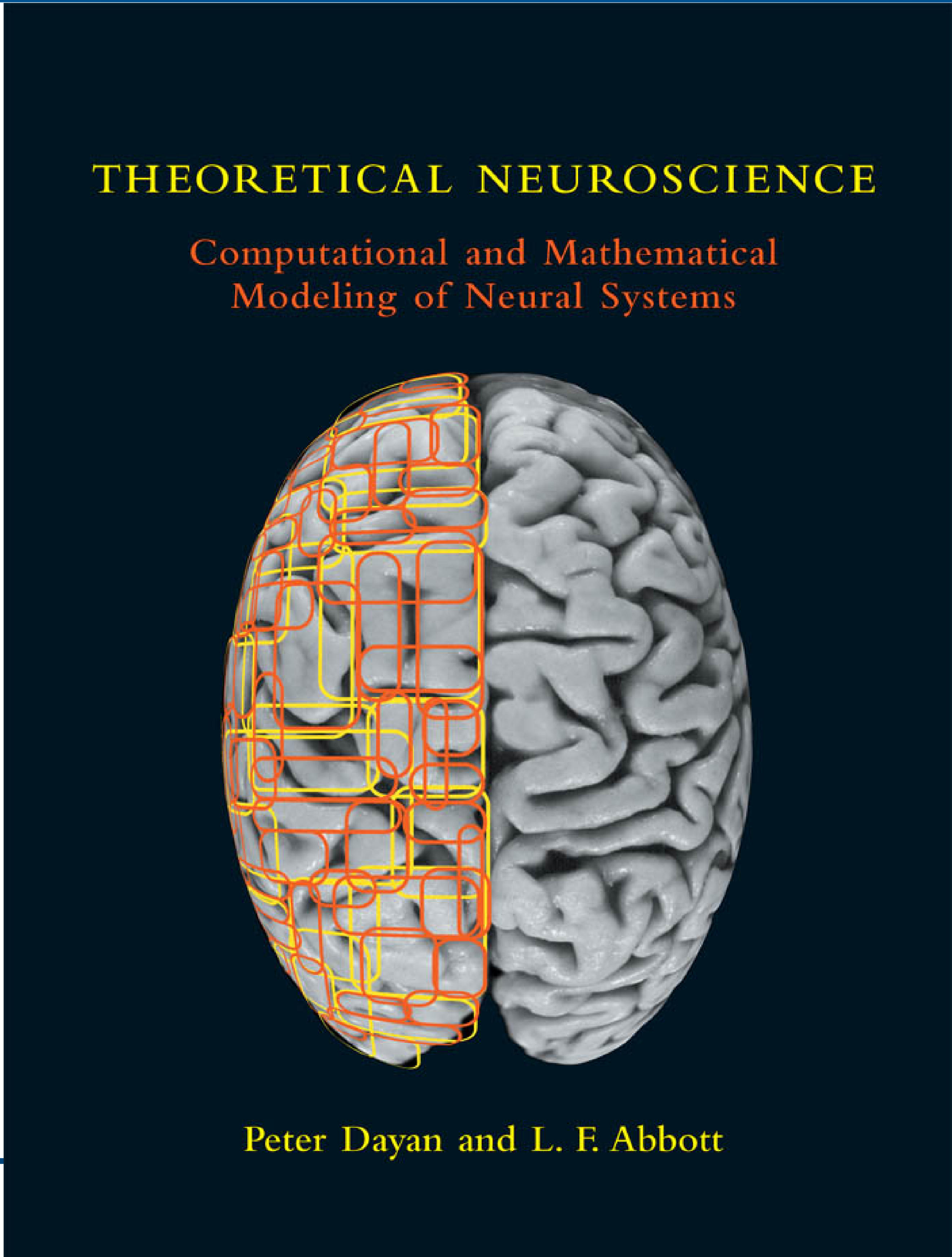
The signal generated is characterized by a “rich” transient and by asymptotic stability target go cue

Introduction to neuronal networks

Peter Dayan and L.F. Abbott
 "Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems" The MIT Press, 2001

This is based on a **linearized model** of the interaction dynamics between neurons

The signals are the **spiking rates**



A model for information transmission

The question is how populations of neurons produce **large amplitude transient signals** for information transmission and storage.

In various papers it is emphasized the role the system **non-normality** for the generation of the “best” signals.

1. S. Ganguli, D. Huh, and H. Sompolinsky “Memory traces in dynamical systems.” Proceedings of the National Academy of Sciences, vol. 105.48, pp. 18970-18975, 2008.S.
2. Ganguli, and P. Latham, “Feedforward to the Past: The Relation between Neuronal Connectivity, Amplification, and Short-Term Memory”, Neuron, Vol. 41, pp. 499-501, 2009.
3. M. S. Goldman, “Memory without Feedback in a Neural Network”, Neuron, Vol. 61, pp. 621–634, 2009.
4. G. Hennequin, T. P. Vogels, and W. Gerstner, “Non-normal amplification in random balanced neuronal networks”, Physical Review E, Vol. 86, pp. 011909, 2012.
5. G. Hennequin, et al. “Optimal control of transient dynamics in balanced networks supports generation of complex movements.” Neuron, 82.6 (2014): 1394-1406.

The key point is to have large transients while keeping the dynamics stable

A model for information transmission

Problems:

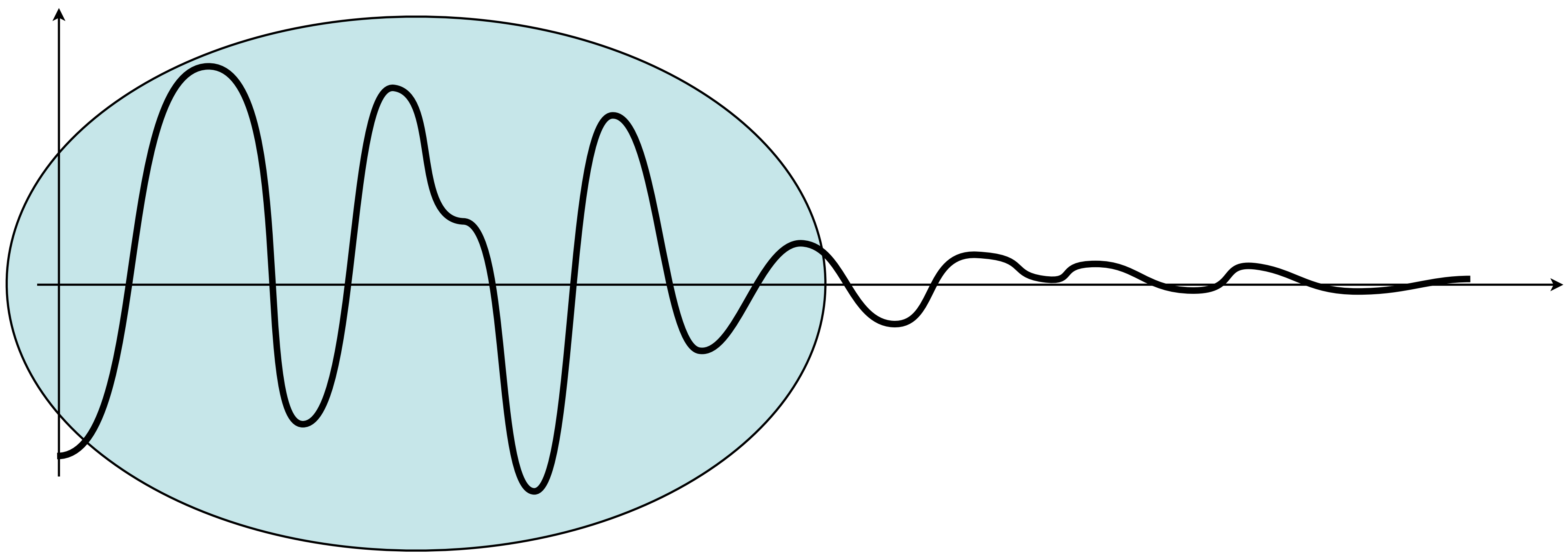
1. Propose a consistent model for understanding why non-normality plays a role in making information transmission more efficient.

2. Quantify the information transmission efficiency.

3. Verify whether the model complexity influences information transmission efficiency.

Generation of “rich” signals

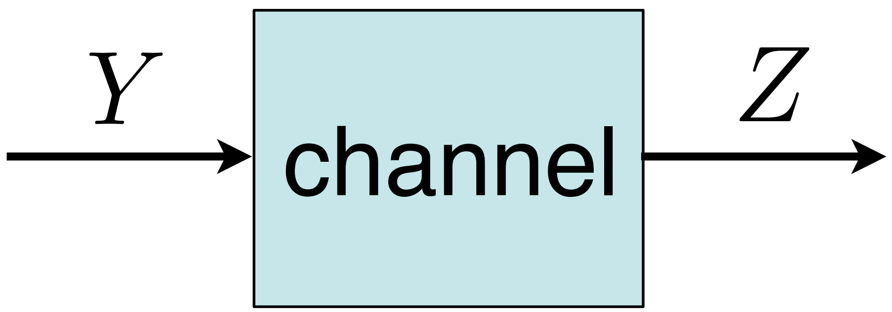
It is important to produce transients that have high energy initially and then converge to zero fast



Channel capacity

Channel model: conditional probability

$$p(Z|Y)$$

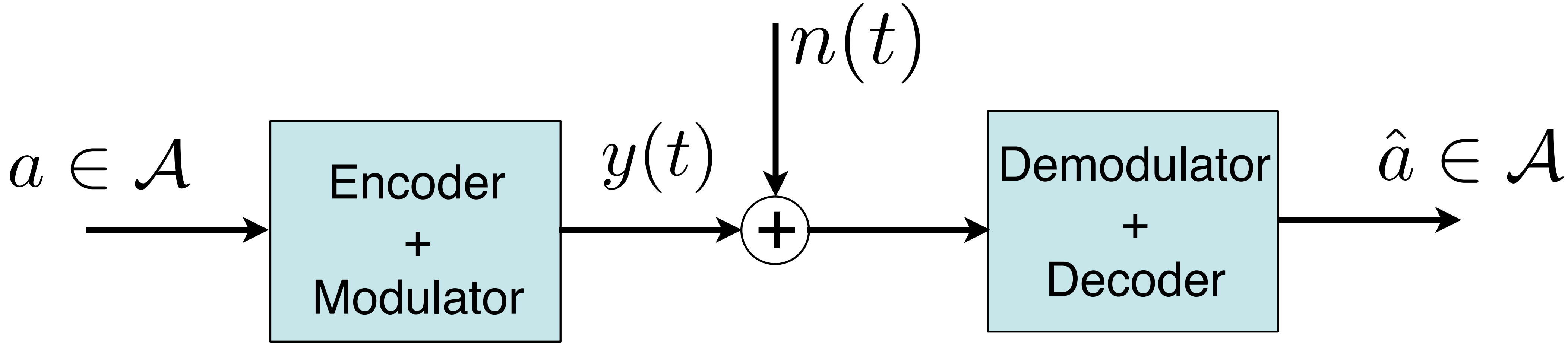


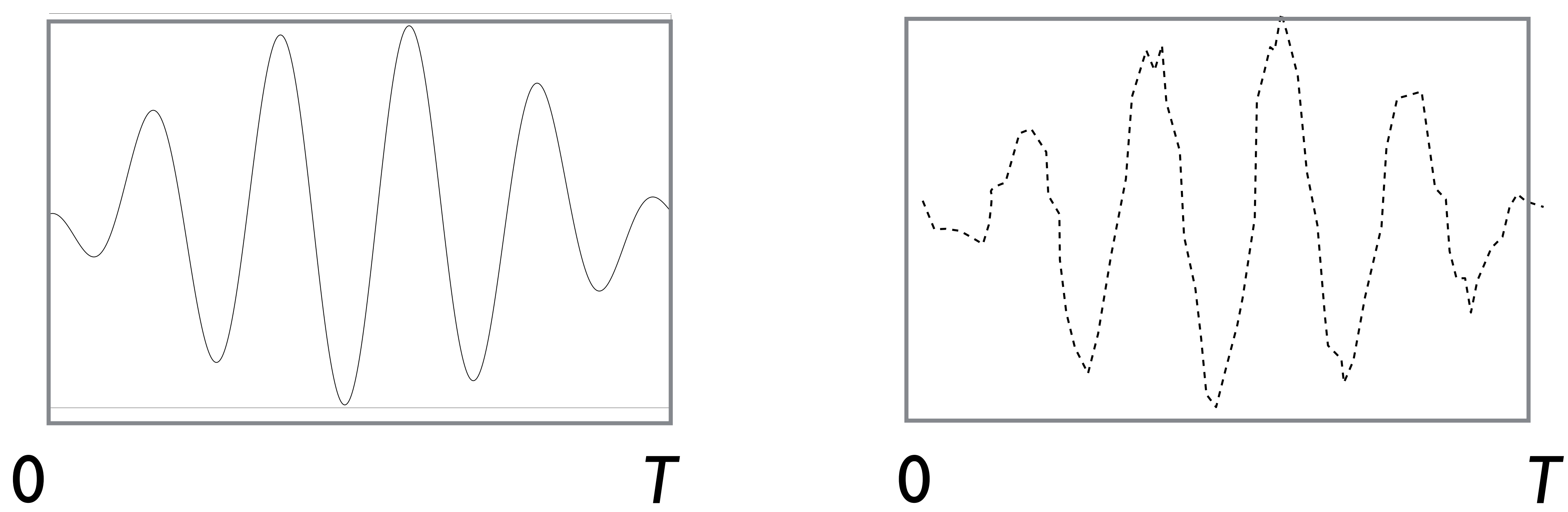
Shannon channel capacity

$$C = \max_{p(Y)} I(Y;Z) = \max_{p(Y)} h(Z) - h(Z|Y)$$

where $I(\cdot)$ is the mutual information and $h(\cdot)$ is the differential entropy

Continuous time channels

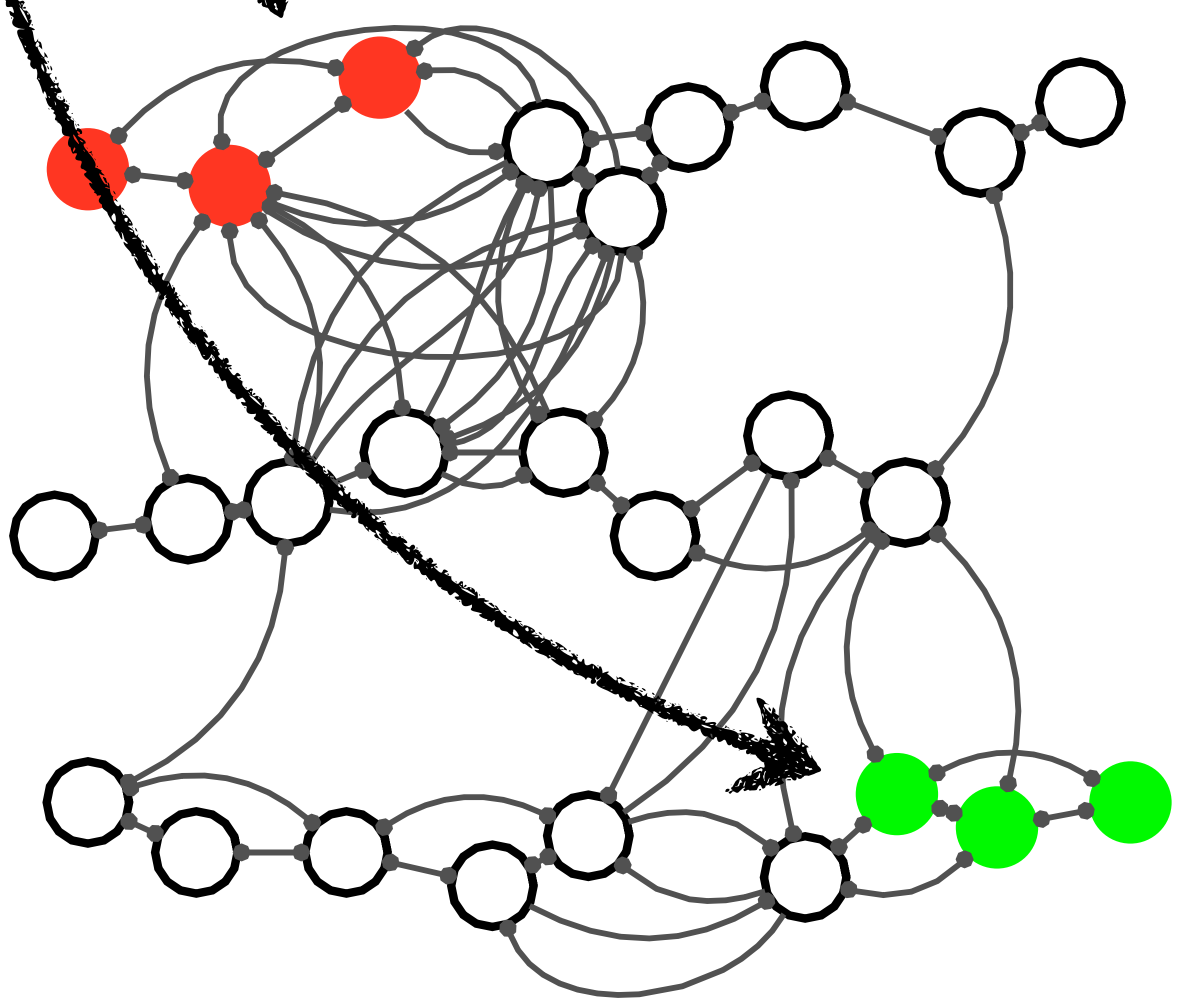




The capacity depends on the signal to noise ratio SNR

Modulation by a system transient

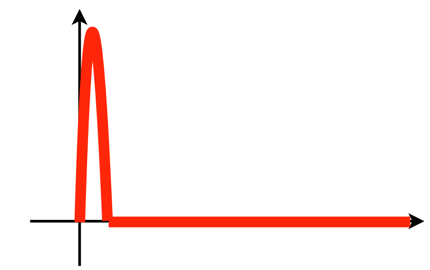
$$\dot{x}(t) = Ax(t) + Bu(t)$$
$$y(t) = Cx(t)$$



Modulation by a system transient

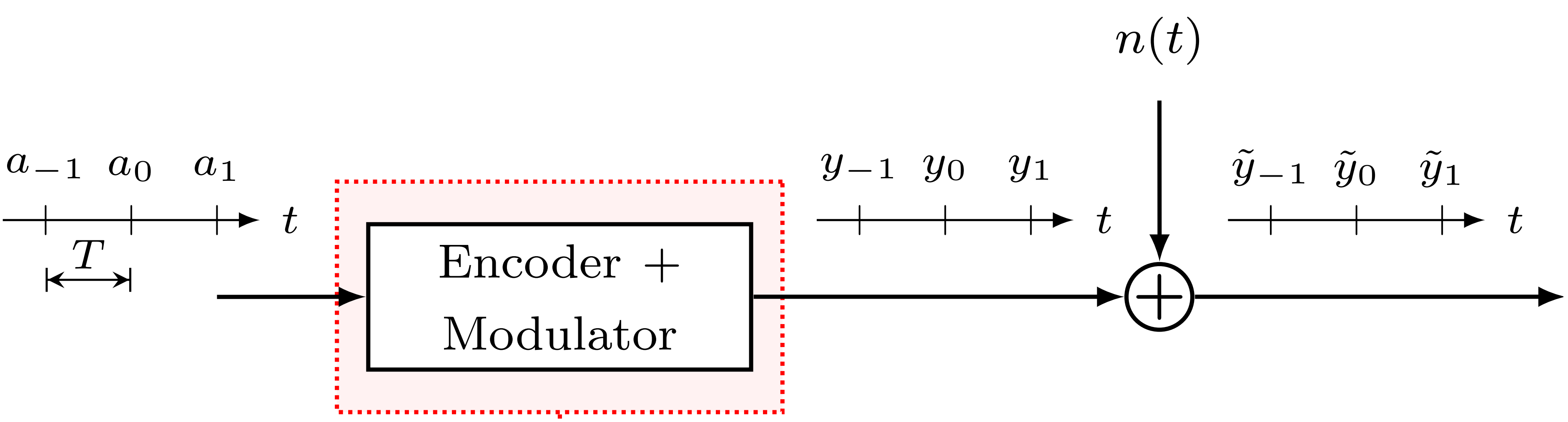
At time $t = 0$ with a symbol $a \in \mathcal{A}$ we associate an impulsive input

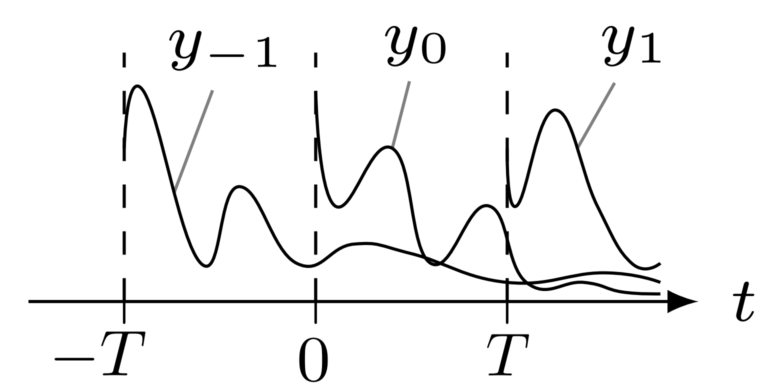
$$u(t) = u_0 \delta(t) \qquad \text{with } \|u_0\| \leq P$$



$$\dot{x}(t) = Ax(t) + Bu(t)$$
$$y(t) = Cx(t)$$

$$\Rightarrow \quad y_0(t) = Ce^{At}Bu_0$$

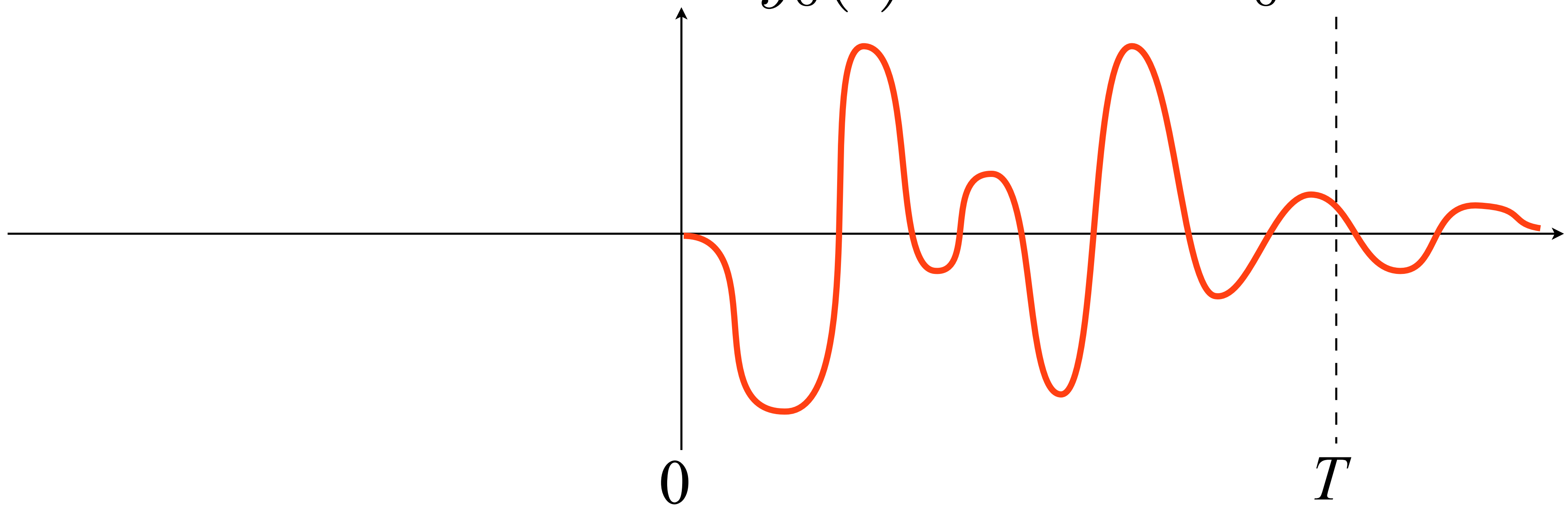




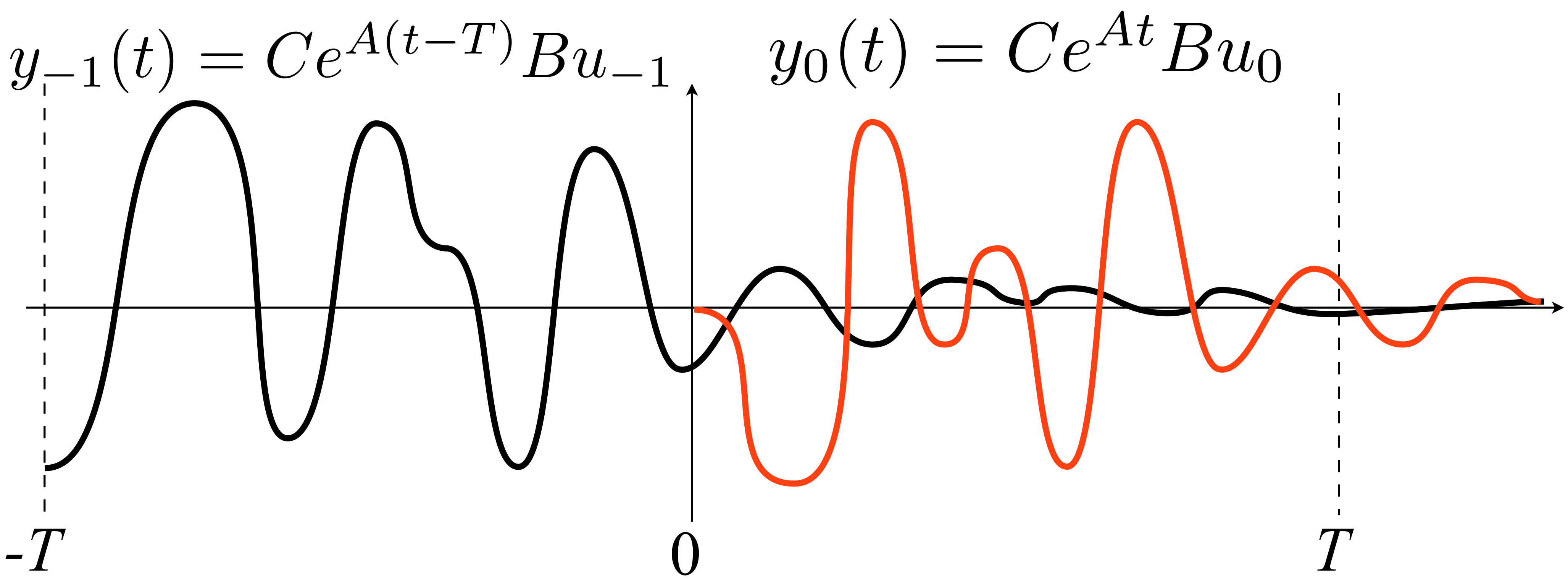
$$y_k(t) = Ce^{A(t-kT)}Bu_k$$

Inter-symbol interference

$$y_0(t) = Ce^{At}Bu_0$$



Inter-symbol interference



Inter-symbol interference: The signal transmitted at time -T, -2T, -3T, ... interfere with the signal transmitted at time 0 and acts as an additional noise.

Expression of the capacity

$$\boxed{R_T = \frac{1}{2T} \max_{\Sigma \geq 0, \mathrm{tr}\Sigma \leq P} \log_2 \frac{\det(\sigma I + \mathbf{O}\mathbf{W})}{\det(\sigma I + \mathbf{O}(\mathbf{W} - B\Sigma B^\top))},}$$

where

$$\mathbf{O} := \int_0^T e^{A^\top t} C^\top C e^{At}\, dt$$

denotes the $[0, T]$-observability Gramian of the pair (A, C) and

$$\mathbf{W} := \sum_{k=0}^{\infty} e^{AkT} B\Sigma B^\top e^{A^\top kT}$$

denotes the discrete-time controllability Gramian of the pair $(e^{AT}, B\Sigma^{1/2})$.

Expression of the capacity

$$\boxed{R_T = \frac{1}{2T} \max_{\Sigma \geq 0, \mathrm{tr}\Sigma \leq P} \log_2 \frac{\det\left(\sigma I + \mathbf{O}\mathbf{W}\right)}{\det\left(\sigma I + \mathbf{O}(\mathbf{W} - B\Sigma B^\top)\right)},}$$

where

$$\mathbf{O} := \int_0^T e^{A^\top t} C^\top C e^{At}\, dt$$

denotes the $[0, T]$-observability Gramian of the pair (A, C) and

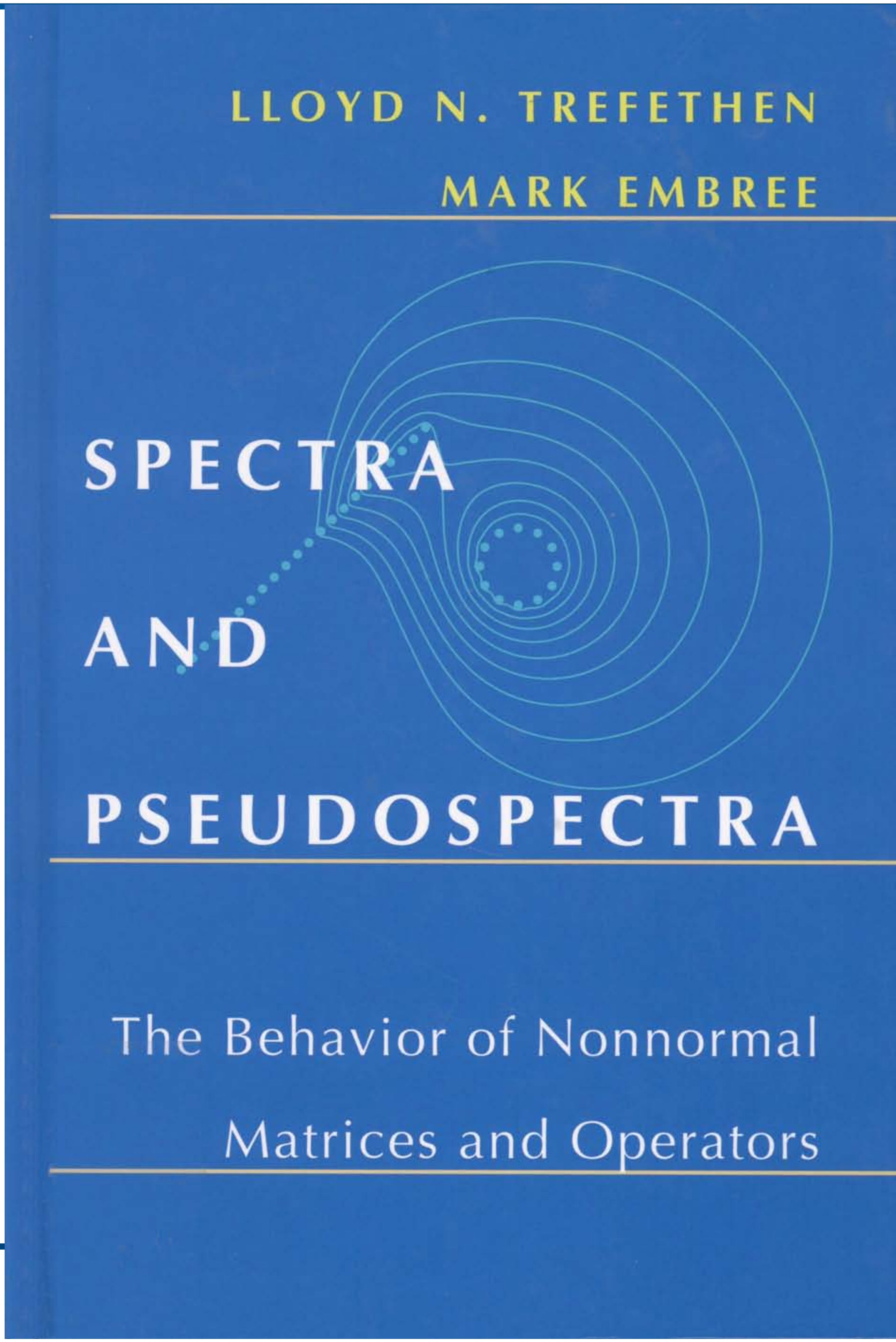
$$\mathbf{W} := \sum_{k=0}^{\infty} e^{AkT} B\Sigma B^\top e^{A^\top kT}$$

denotes the discrete-time controllability Gramian of the pair $(e^{AT}, B\Sigma^{1/2})$.

$$\boxed{R_{\max} := \max_{T \geq 0} R_T}$$

Matrix non-normality

A matrix is normal iff $AA^T = A^T A$



Matrix non-normality

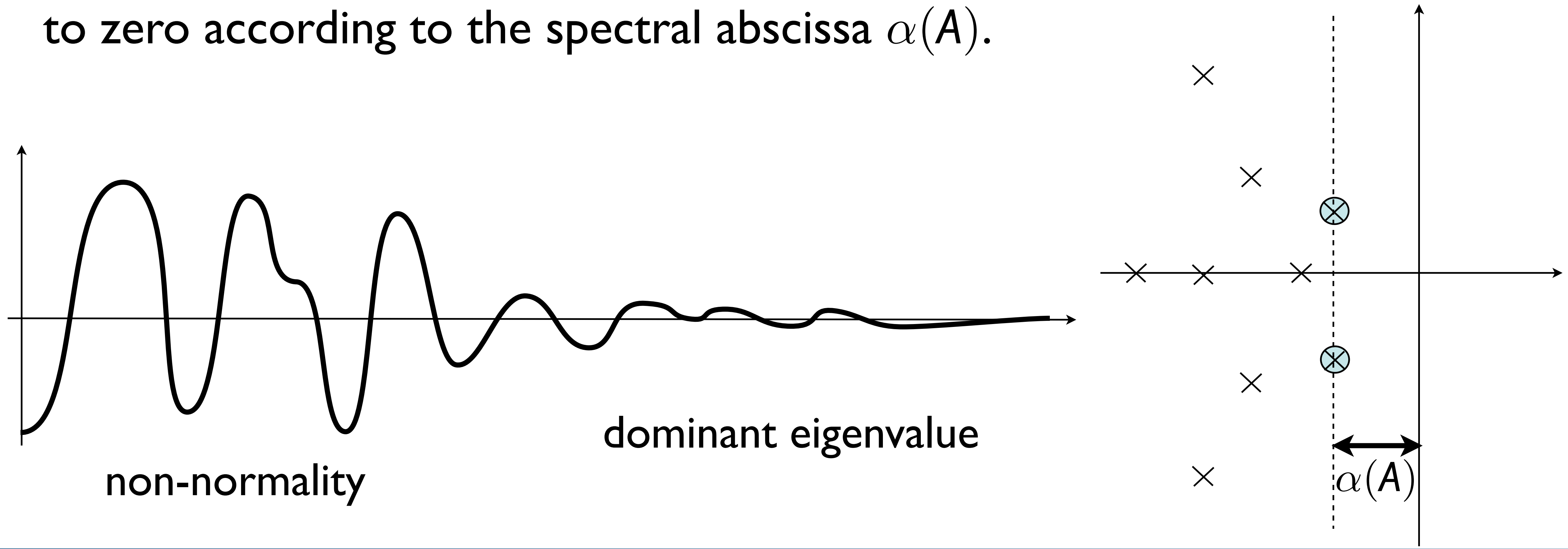
Non-normality has two important features:

1. The eigenvalues of a highly non-normal matrix A are very sensitive to matrix entries variations.

2. The exponential

$$e^{At}$$

of a stable highly non-normal matrix A is large for small t and then decays to zero according to the spectral abscissa $\alpha(A)$.

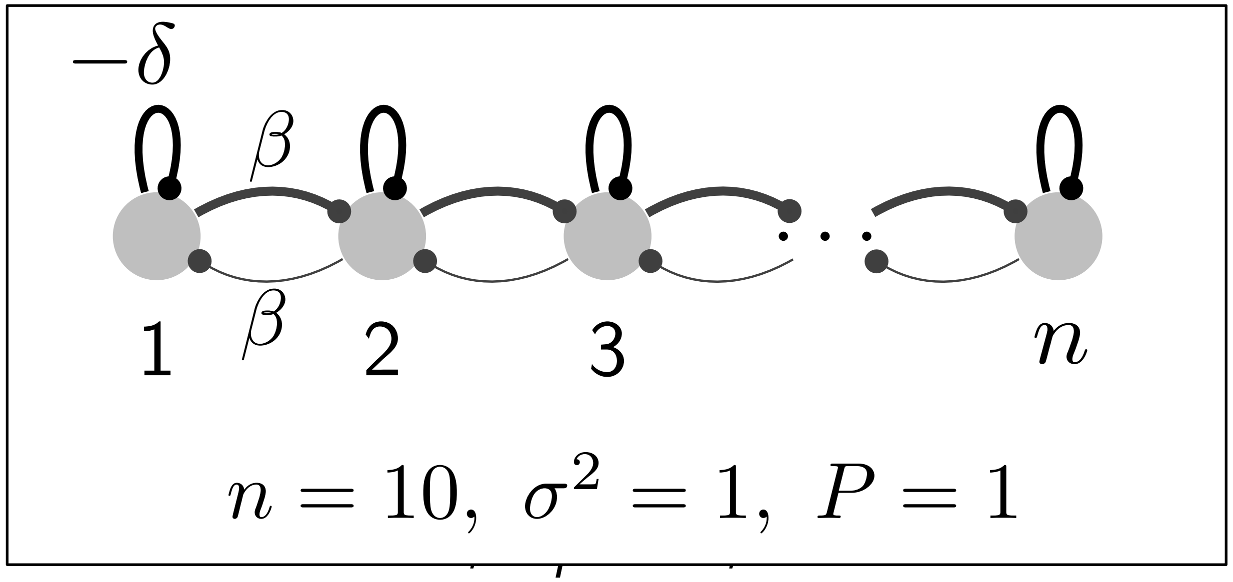


For normal networks R_T is decreasing in T and so the smaller is T the better the performance we obtain.

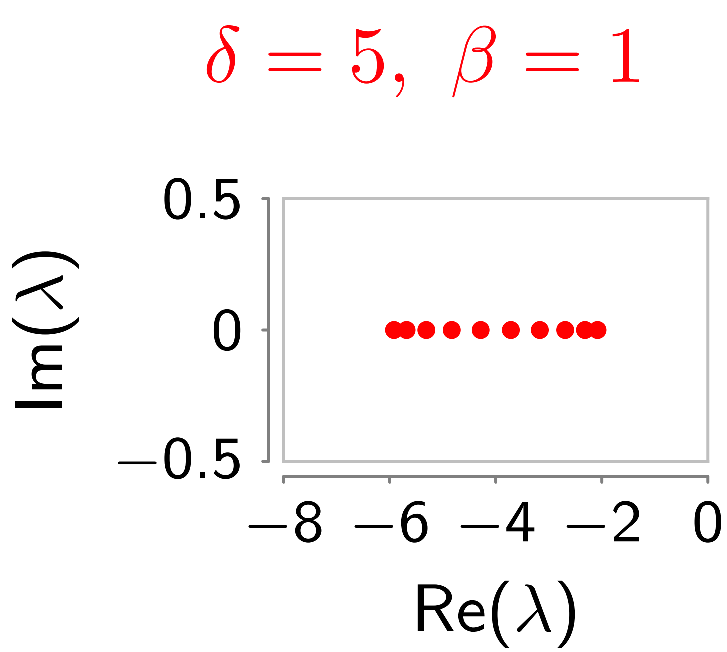
Corollary
If $\mathcal{V}_{in} = \mathcal{V}_{out} = \mathcal{V}$ and $A \in \mathbb{R}^{n\times n}$ is a normal and stable matrix, then

$$R_{\max} := \max_{T\geq 0} R_T = R_0 = \frac{1}{\ln 2} \frac{-\mathrm{tr}(A)\, SNR}{SNR - 2\mathrm{tr}(A)}.$$

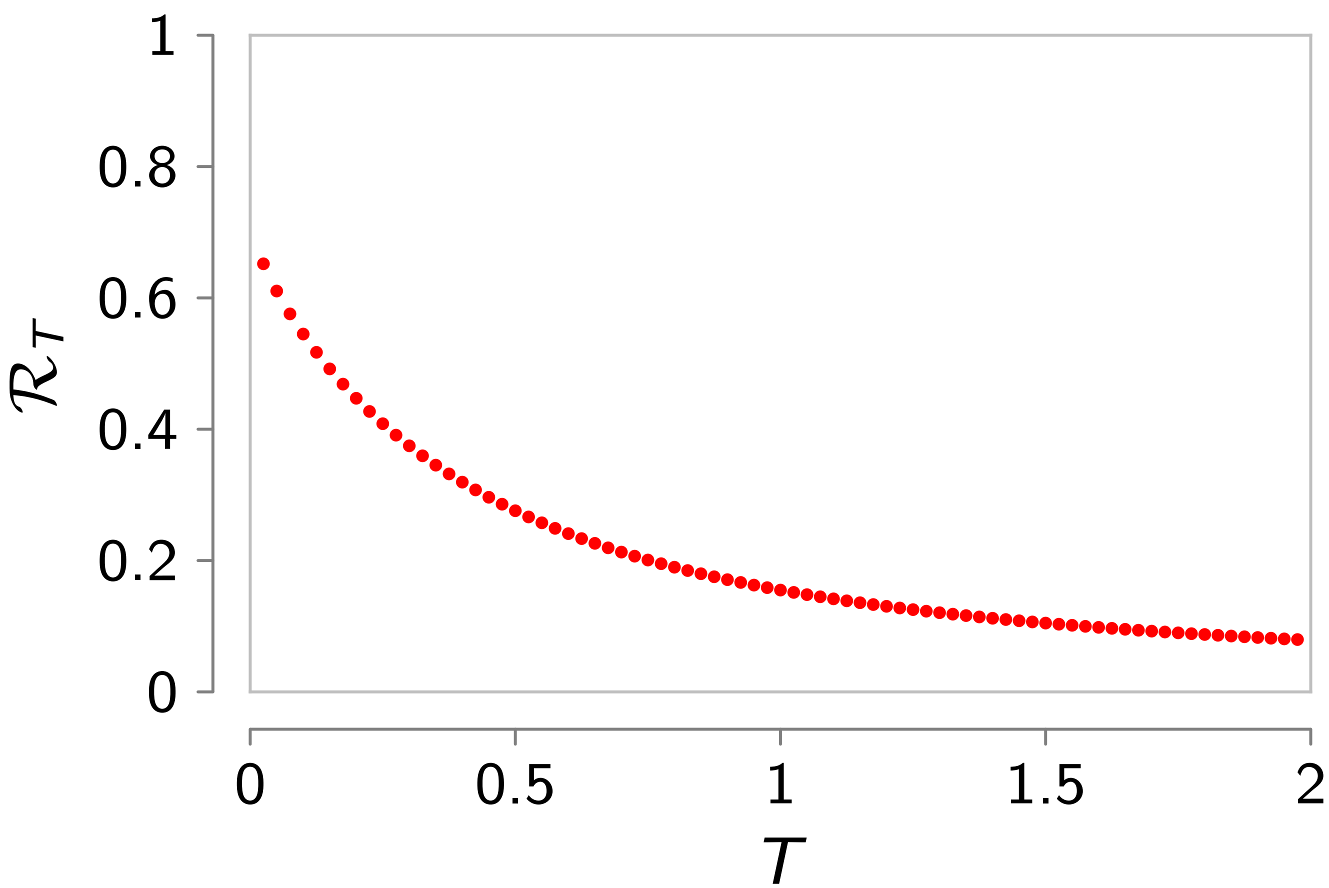
Example



$-\delta$
β
β
1
2
3
n
$n = 10,\ \sigma^2 = 1,\ P = 1$

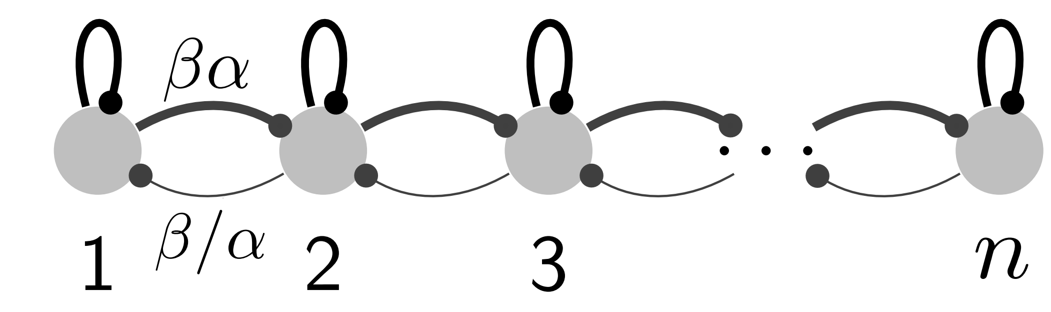


$\delta = 5,\ \beta = 1$
0.5
0
−0.5
−8
−6
−4
−2
0
Re(λ)
Im(λ)



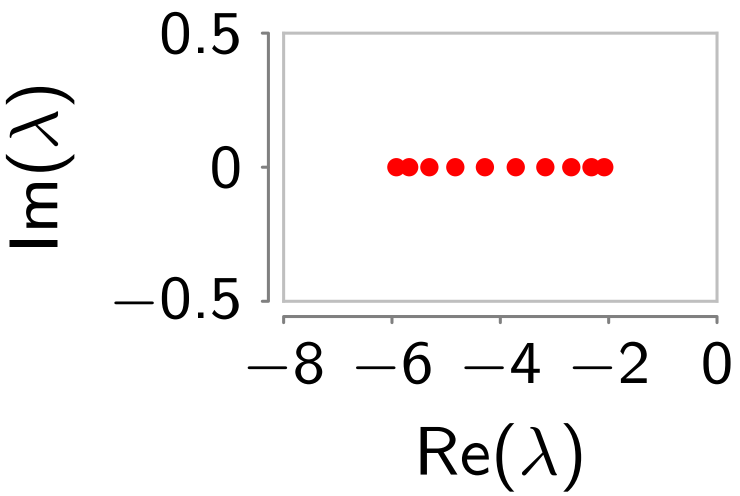
1
0.8
0.6
0.4
0.2
0
$\mathcal{R}_T$
0
0.5
1
1.5
2
T

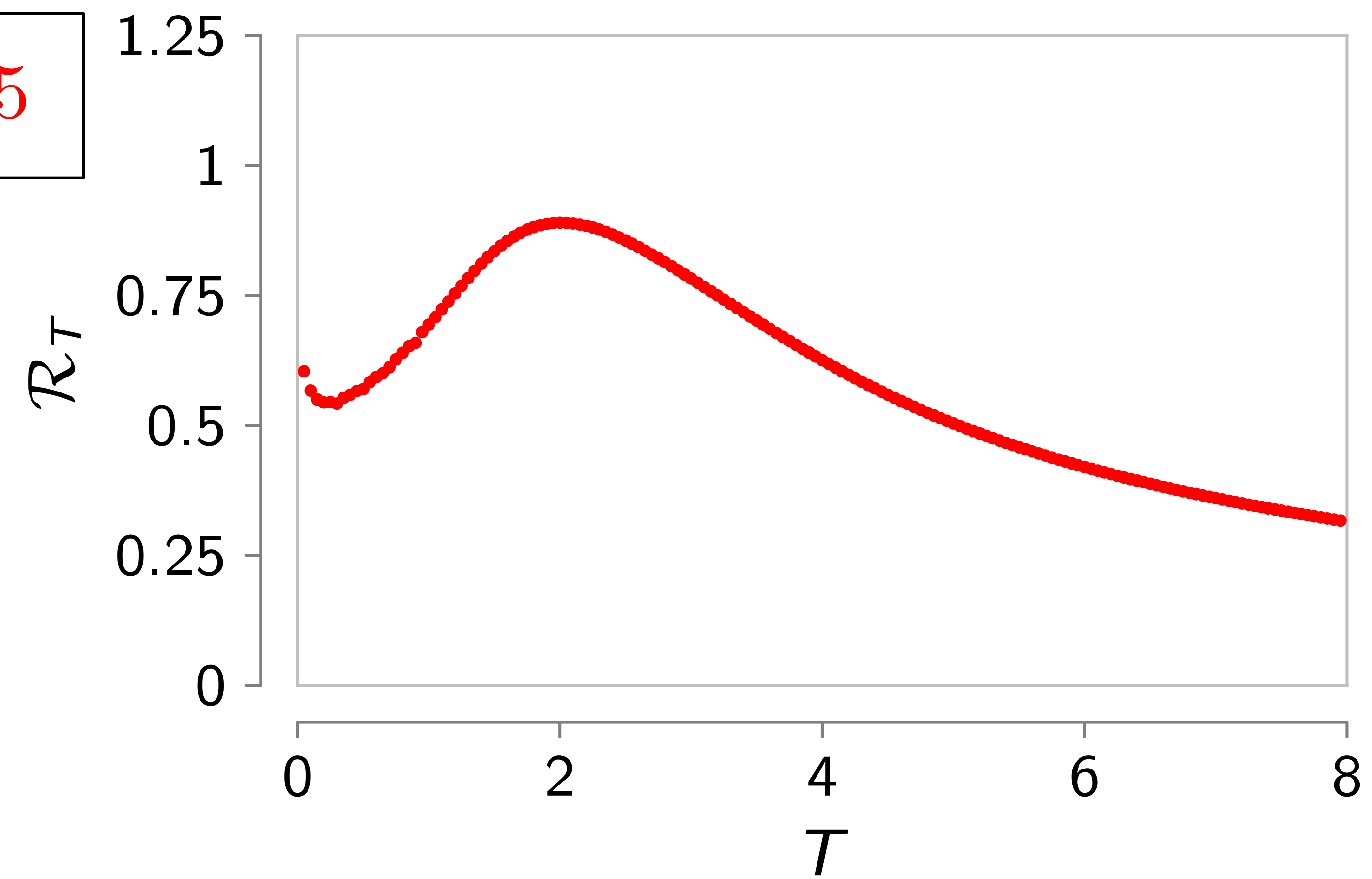
Example: Non normal matrix



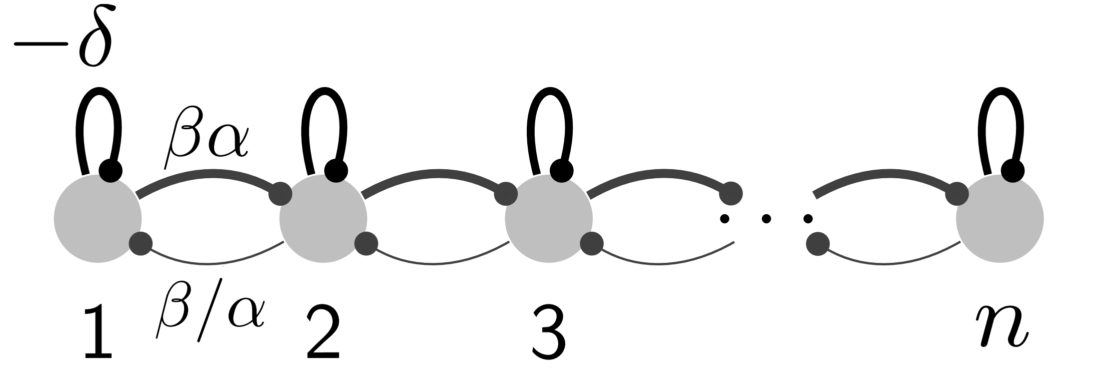
$$n = 10,\ \sigma^2 = 1,\ P = 1$$

$$\delta = 4,\ \beta = 1,\ \alpha = 5$$



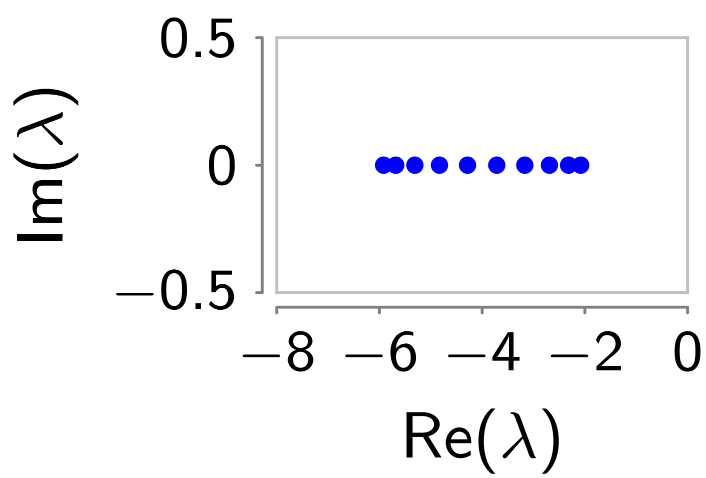


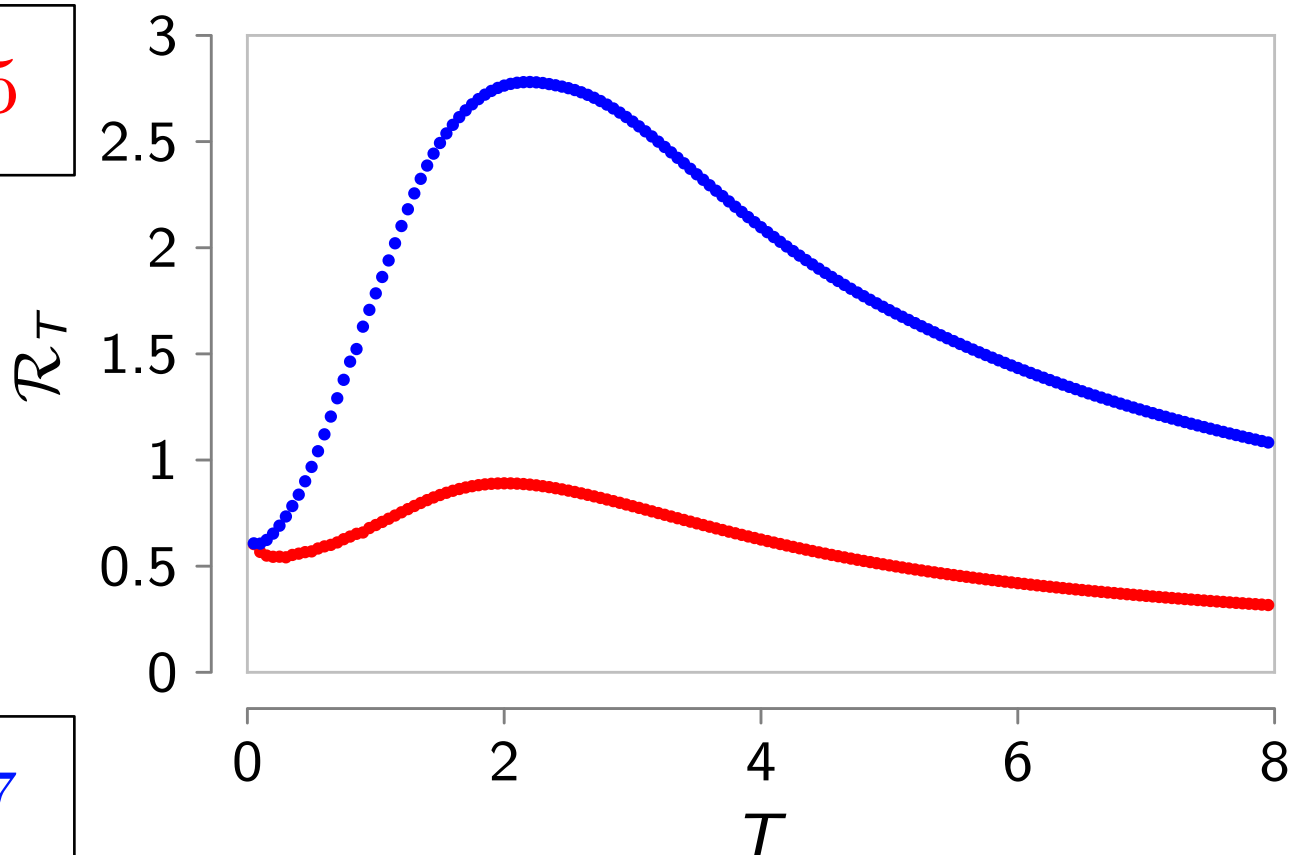
Example: Non normal matrix



$$n = 10,\ \sigma^2 = 1,\ P = 1$$

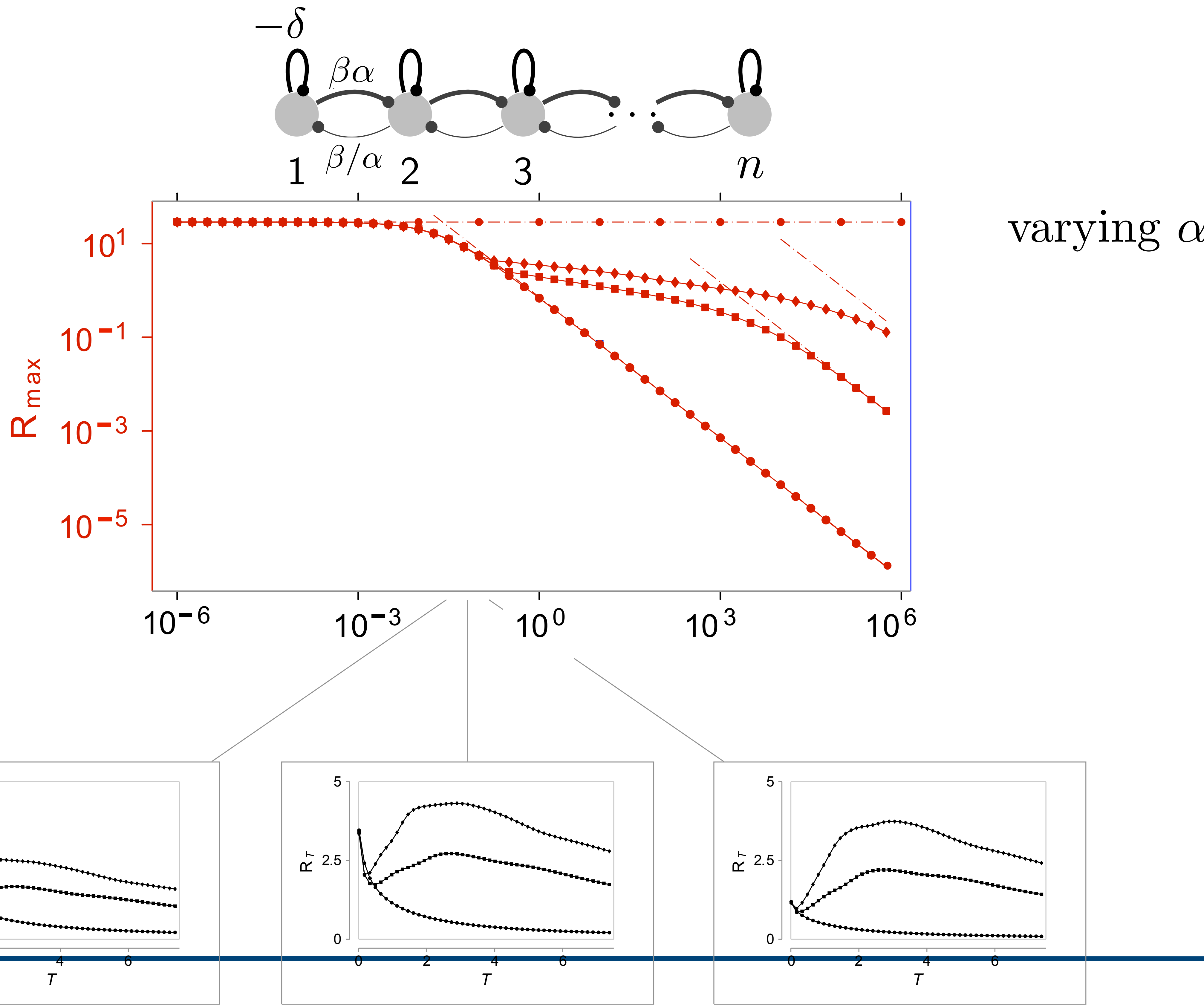
$\delta = 4,\ \beta = 1,\ \alpha = 5$



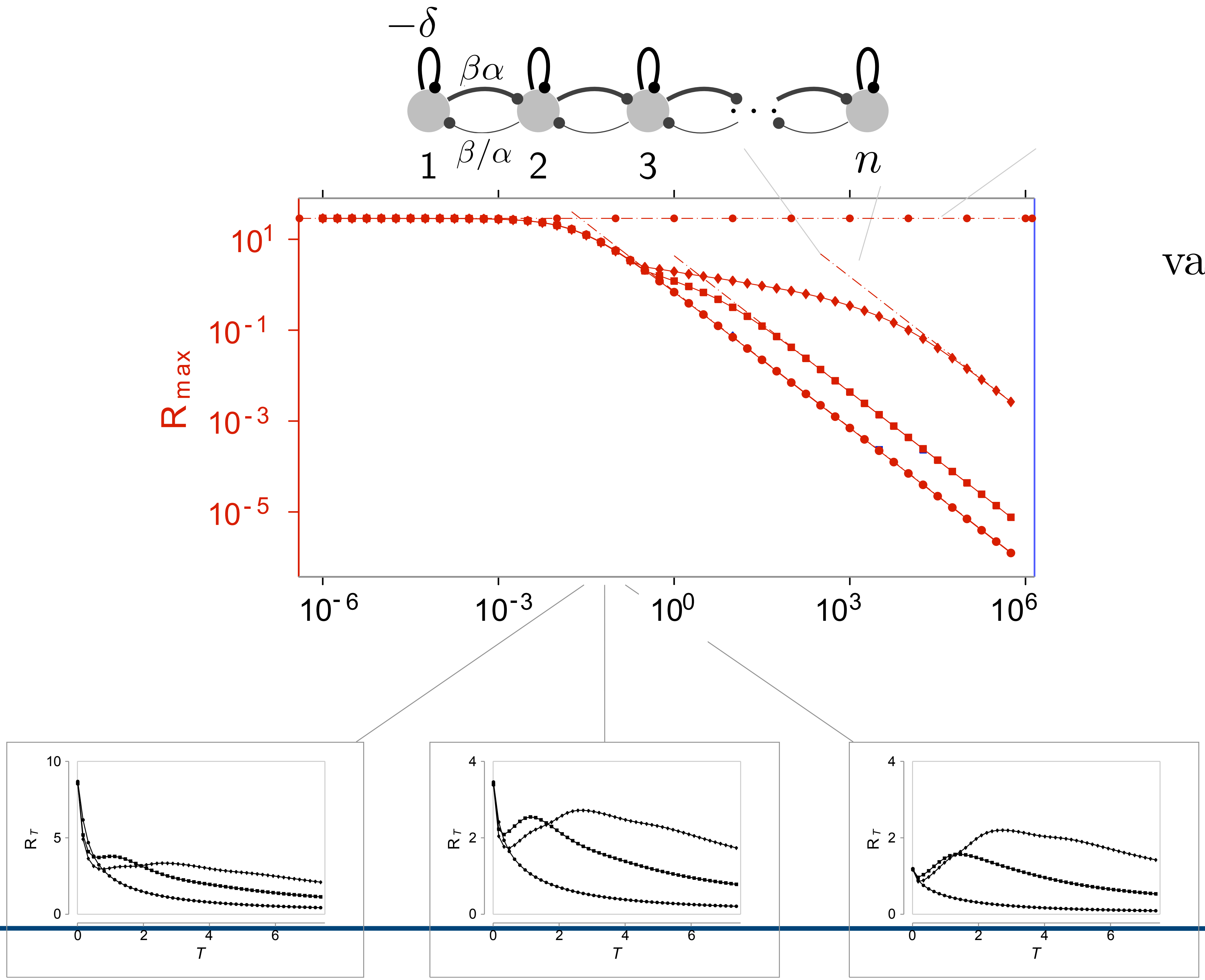


$\delta = 4,\ \beta = 1,\ \alpha = 7$

Line network: varying anysotropy strength



Line network: varying length



varying n

Limit behaviors in the noise

If the noise variance $\sigma^2 \to 0$, then

$$\boxed{R_{\max} \simeq -\frac{1}{\ln 2}\operatorname{tr}(A)}$$

If the noise variance $\sigma^2 \to \infty$, then

$$R_{\max} \simeq \frac{1}{2\ln 2}\frac{\ell}{\sigma^2}$$

where

$$\boxed{\ell := \max_{T\geq 0}\frac{\|B^T O_T B\|}{T}} \qquad O_T = \int_0^T e^{A^T t} C^T C e^{At} dt$$

For normal matrices A we have that

1. A is diagonizable by an orthonormal matrix.
2. The eigenvalues of $\frac{A+A^T}{2}$ are the real values of the eigenvalues of A and so

$$\begin{aligned}\omega(A) \quad &:= \quad \max\{\lambda_i : \ \lambda_i \text{ eigenvalues of } \frac{A+A^T}{2}\} \\ &= \quad \max\{Re[\lambda_i] : \ \lambda_i \text{ eigenvalues of } A\}\end{aligned}$$

Theorem

Let $A \in \mathbb{R}^{n \times n}$ be a stable matrix.

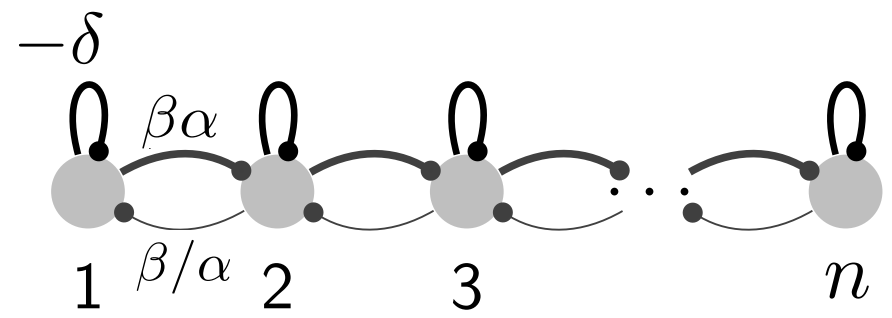
If $\omega(A) \leq 0$, then

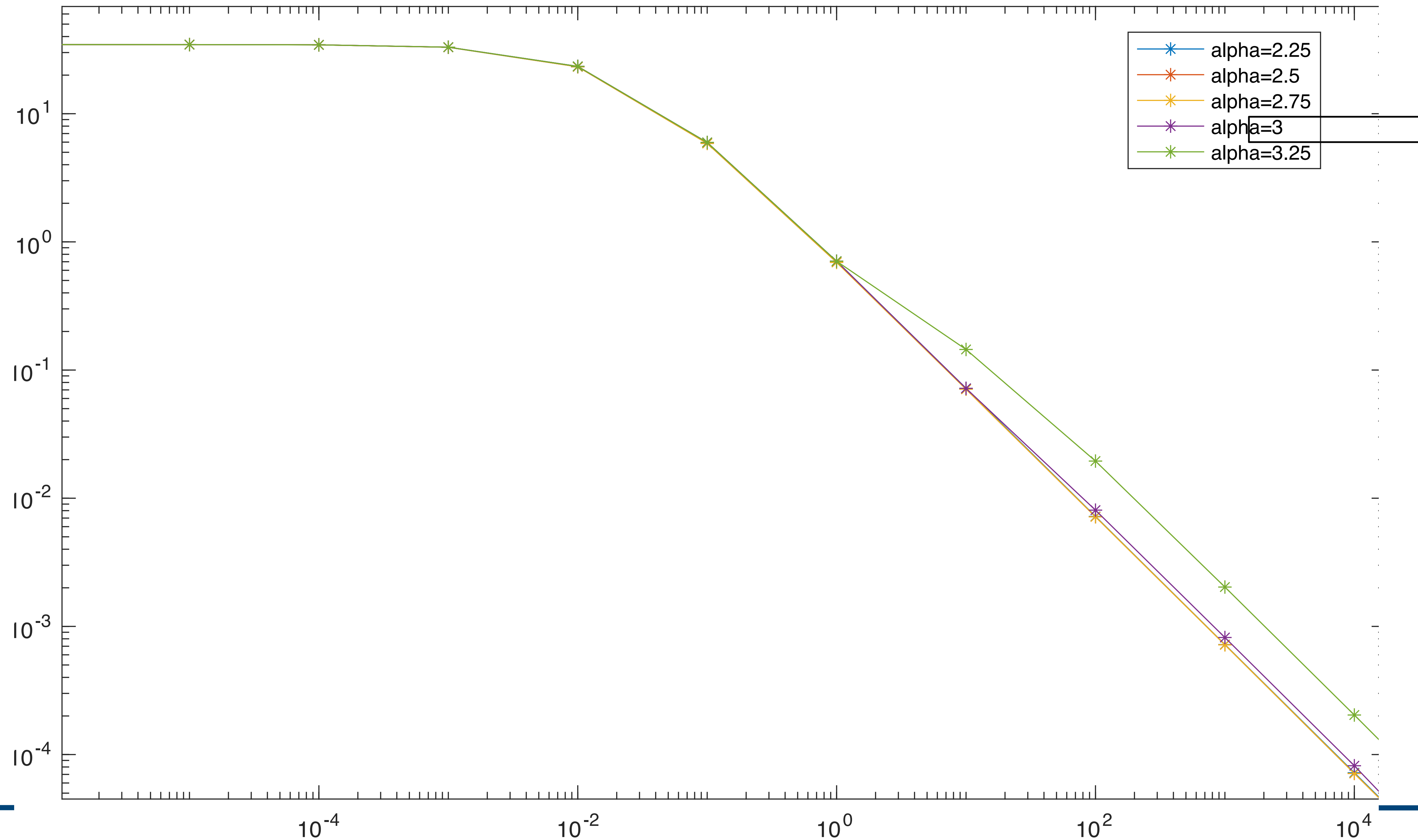
$$\ell(A, B, C) \leq 1$$

If $\omega(A) > 0$ and $B = C = I$, then

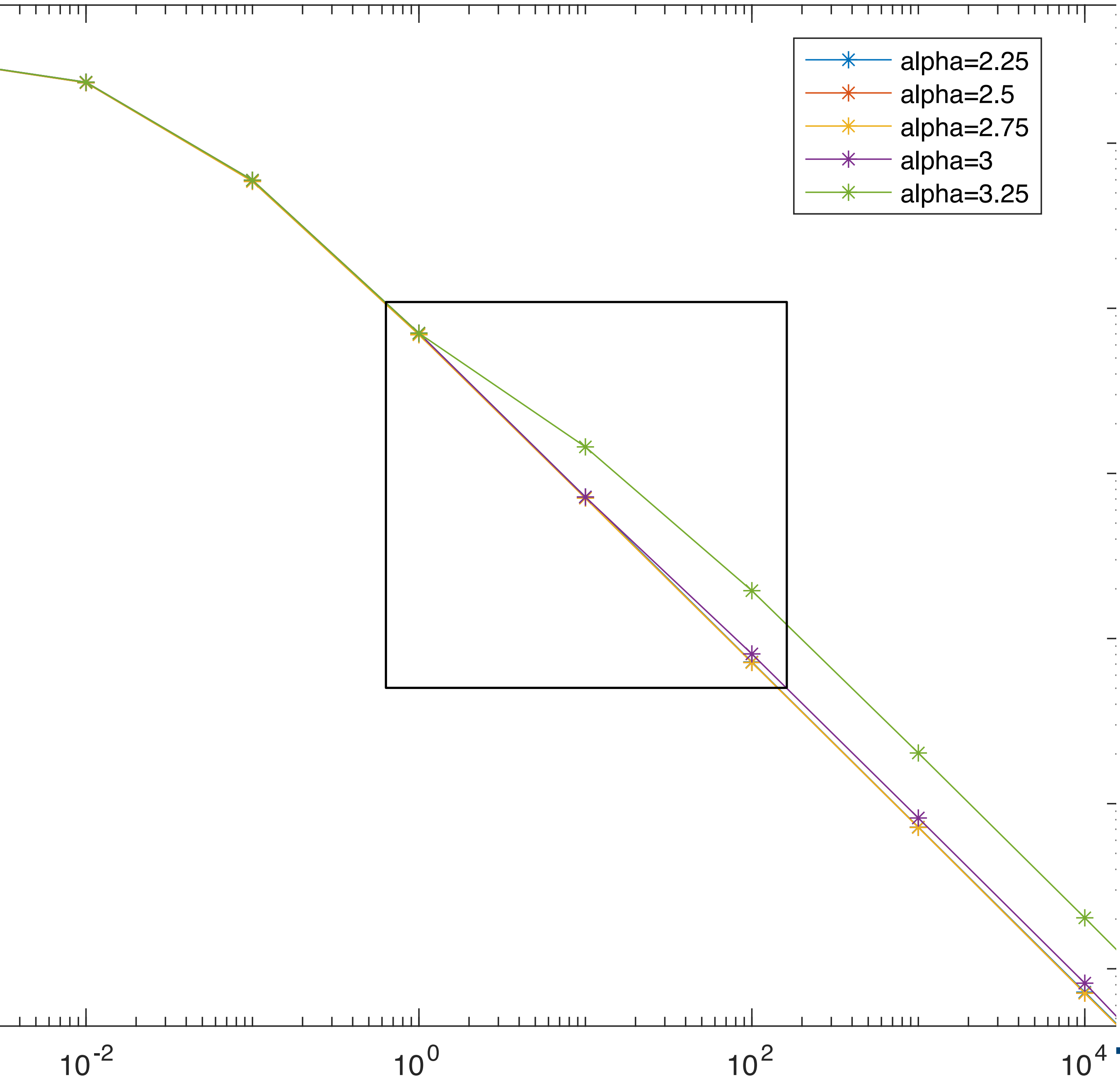
$$\ell(A, B, C) > 1$$

Line network

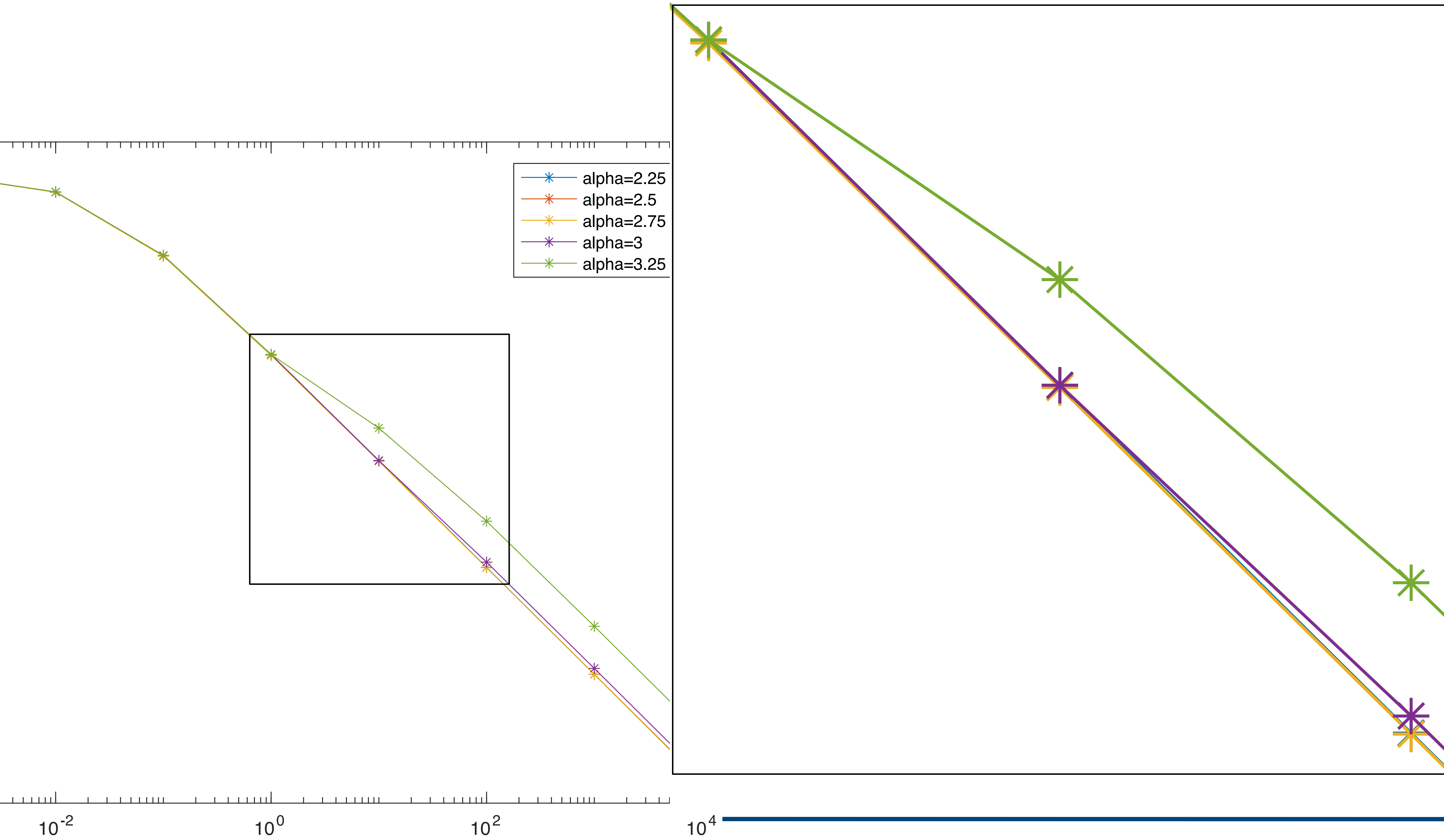




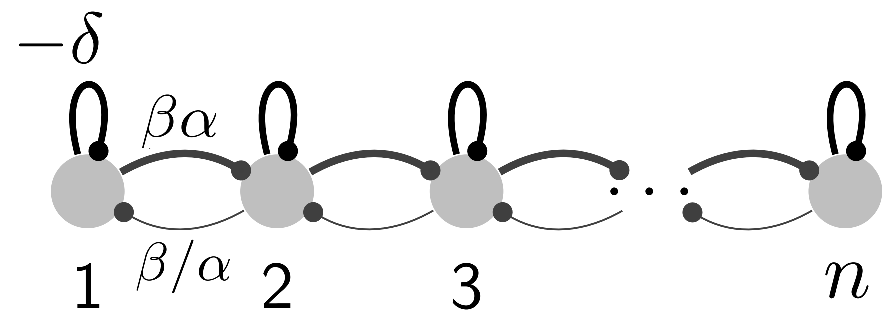
Line network

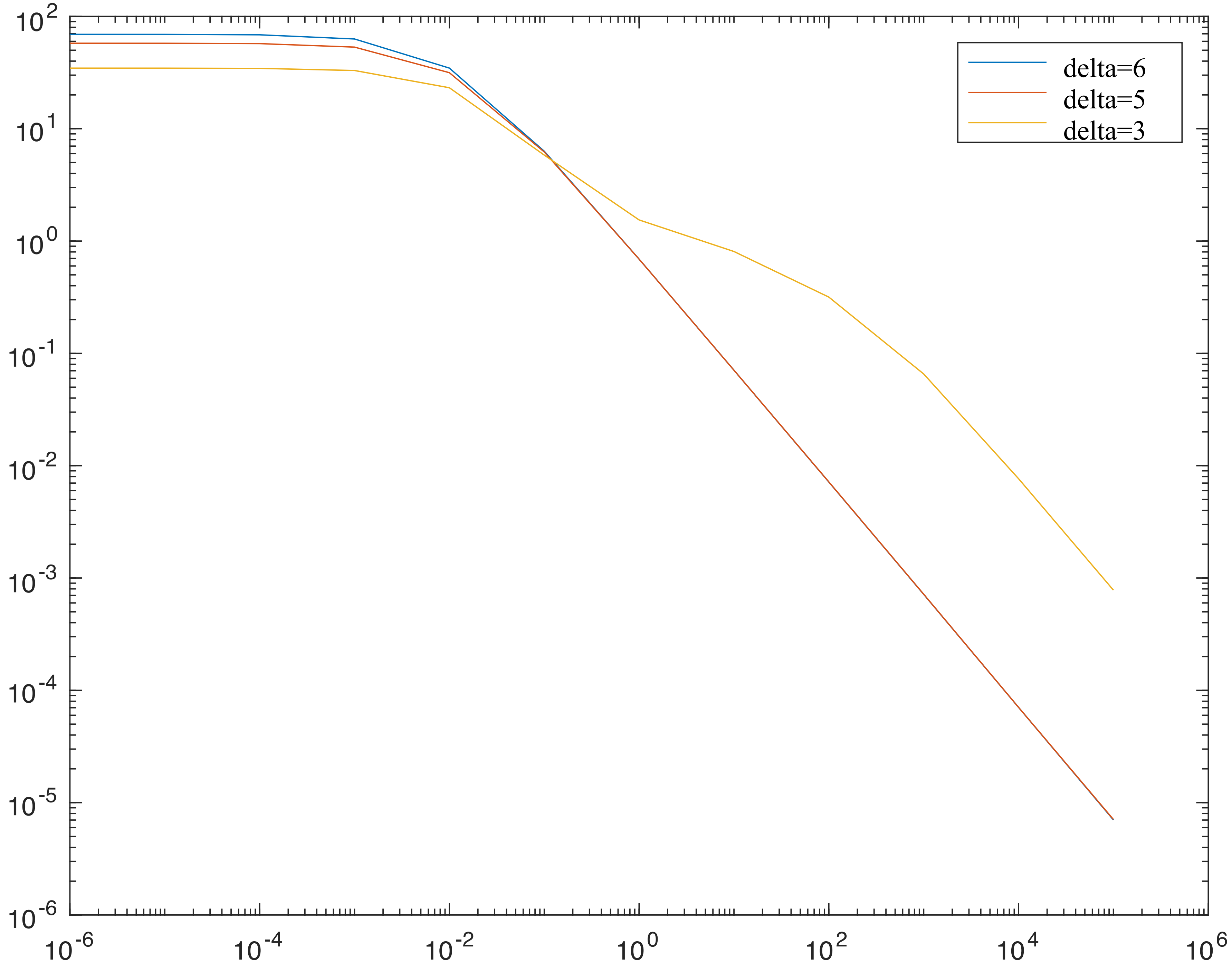


Line network

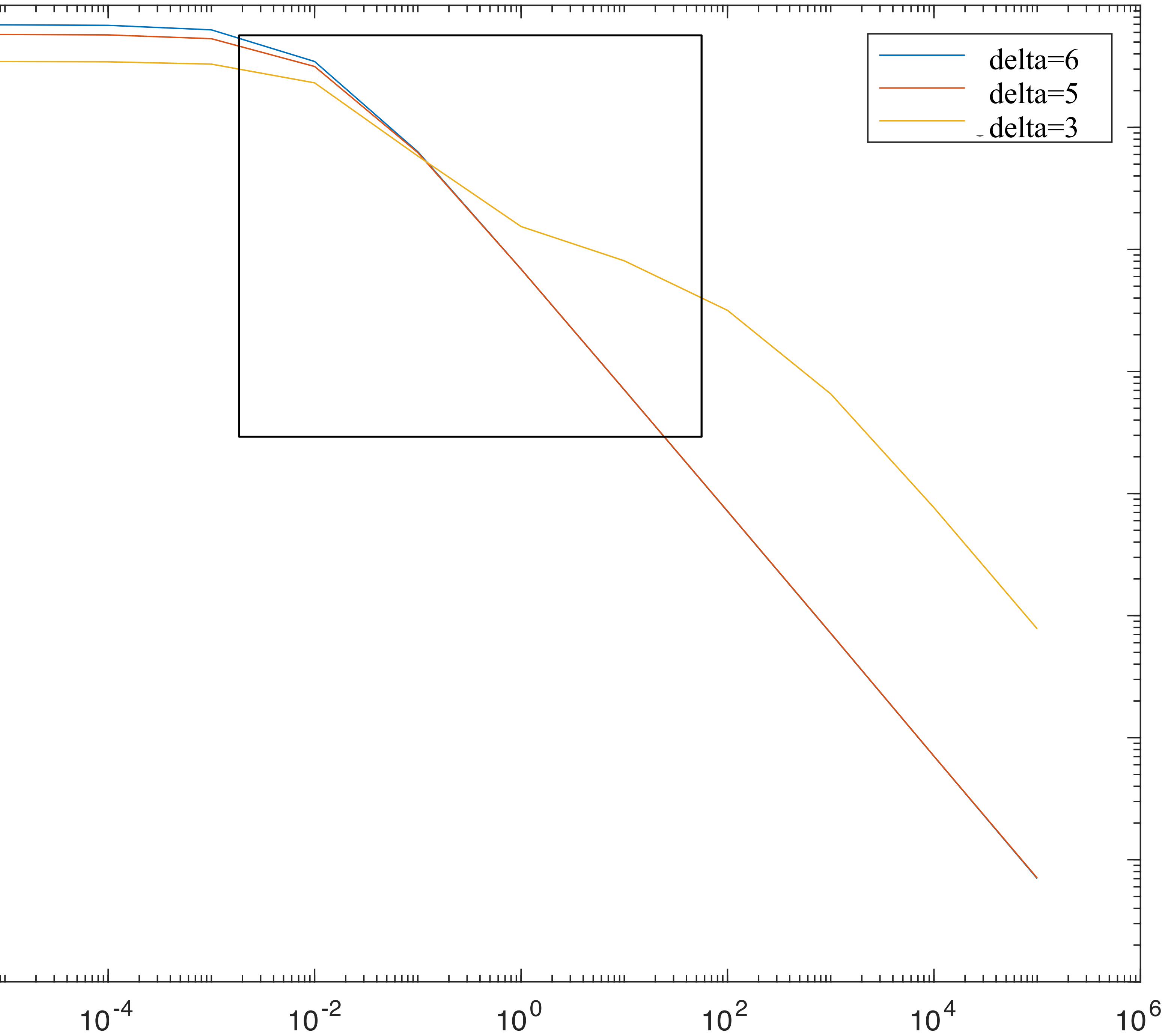


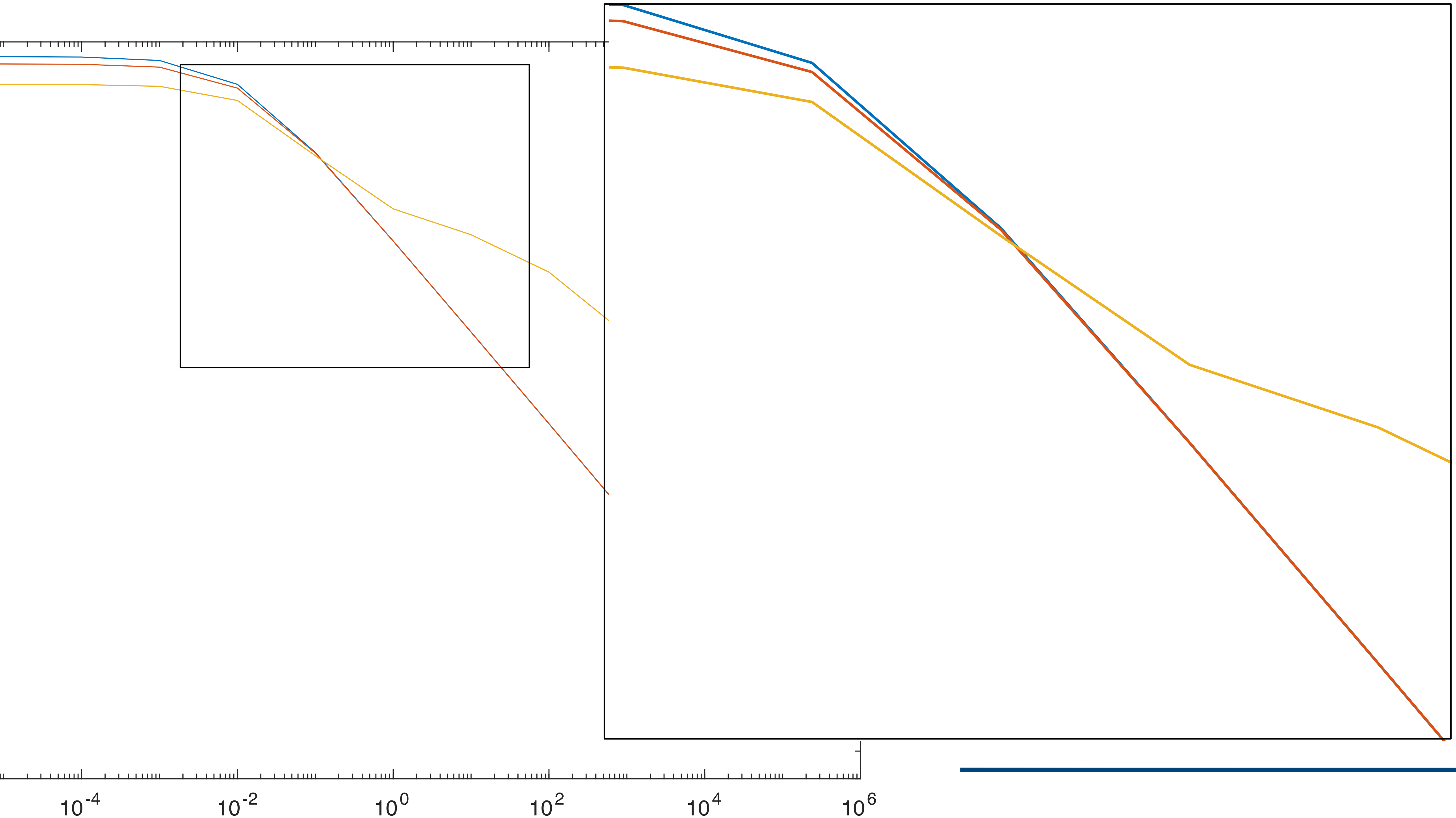
Line network





Line network




10^-4
10^-2
10^0
10^2
10^4
10^6

Example

$$A = \begin{bmatrix} -\delta & \alpha & 0 & & \\ 0 & -\delta & \alpha & \ddots & \\ & \ddots & \ddots & \ddots & \ddots \\ & & \ddots & -\delta & \alpha \\ & & & 0 & -\delta \end{bmatrix}, B = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}, C = \begin{bmatrix} 0 & 0 & \dots & 0 & 1 \end{bmatrix},$$

If the noise is small then

$$R_{\max} \simeq \frac{\delta n}{\ln 2}$$

Notice moreover that $\omega(A) \simeq -\delta + \alpha$ and so $\alpha > \delta$ guarantees that $\omega(A) > 0$.

$$\ell(A, B, C) \geq \frac{1}{4(2n-1)\sqrt{\pi(n-1)}} \left(\frac{\alpha}{\delta}\right)^{2n-2}$$

Conclusions

1. We proposed a model to quantify the information transmission performance in linear dynamic networks.

2. By introducing the intersymbol interference, we can highlight the role of non-normality of the dynamics.

3. For normal dynamics a theoretical analysis is possible. However this is the less interesting case.

4. Non-normal dynamics is more interesting and harder to study.

Thank you

Matrix non-normality

$$\boldsymbol{A} \in \mathbb{R}^{N\times N}$$

normal matrices

$$\boldsymbol{A}\boldsymbol{A}^\top = \boldsymbol{A}^\top\boldsymbol{A}$$

$$\boldsymbol{A} = \boldsymbol{U}^*\boldsymbol{D}\boldsymbol{U},$$

$\boldsymbol{U}$ unitary, $\boldsymbol{D}$ diagonal

(spectral decomposition)

$$\boldsymbol{D} = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_N \end{bmatrix}$$

$$\sigma(\boldsymbol{A}) = \{\lambda_i\}_{i=1}^N$$

non-normal matrices

$$\boldsymbol{A}\boldsymbol{A}^\top \neq \boldsymbol{A}^\top\boldsymbol{A}$$

$$\boldsymbol{A} = \boldsymbol{U}^*\boldsymbol{T}\boldsymbol{U},$$

$\boldsymbol{U}$ unitary, $\boldsymbol{T}$ triangular

(Schur decomposition)

$$\boldsymbol{T} = \begin{bmatrix} \lambda_1 & & & \\ \star & \lambda_2 & & \\ \star & \star & \ddots & \\ \star & \star & \star & \lambda_N \end{bmatrix}$$

$$\boldsymbol{T} = \boldsymbol{D} + \boldsymbol{N}$$