Singular Value Decomposition

Aiichiro Nakano

Collaboratory for Advanced Computing & Simulations
Department of Computer Science
Department of Physics & Astronomy
Department of Quantitative & Computational Biology
University of Southern California

Email: anakano@usc.edu

Goal: Another matrix decomposition (SVD)
for low-rank matrix approximation

cf. Eigen decomposition
$A = Q[\ ]Q^T$
QR decomposition
$A = Q[\ ]$

See note on "least square fit" & *Numerical Recipes* Sec. 2.6

Rank of a Matrix

- $N \times M$ **matrix** A **as a mapping:** $R^M \to R^N$

$$\begin{array}{cc} & 1 \\ M & \begin{bmatrix} \\ x \\ \\ \end{bmatrix} \end{array} \quad x\left(\in R^M\right) \xrightarrow[A]{} b = Ax\left(\in R^N\right) \quad \begin{array}{cc} 1 & \\ \begin{bmatrix} \\ b \\ \\ \\ \\ \end{bmatrix} & N \end{array}$$

- **Range of** A**: Vector space** $\{b = Ax | \forall x\}$
- **Rank of** A**: Number,** m**, of linearly-independent vectors in the range,** *i.e.***, how many linearly-independent** N**-element vectors are there in the range, such that**

$$b = A^{\forall}x = \sum_{\nu=1}^{m} c_\nu |\nu\rangle$$

Low Rank Approximations of a Matrix

- **Rank-1 approximation: $NM \to N + M$**

$$\begin{array}{cc} & M \\ N & \begin{bmatrix} \\ \psi \\ \\ \end{bmatrix} \end{array} \cong \begin{bmatrix} \\ u \\ \\ \end{bmatrix} \begin{bmatrix} \quad v \quad \end{bmatrix}$$

$|u\rangle\langle v|\forall x\rangle \propto |u\rangle$

- **Rank-2 approximation: $NM \to 2(N + M)$**

$$\begin{bmatrix} \\ \psi \\ \\ \end{bmatrix} \cong \begin{bmatrix} \\ u_1 \\ \\ \end{bmatrix} w_1 \begin{bmatrix} \quad v_1 \quad \end{bmatrix} + \begin{bmatrix} \\ u_2 \\ \\ \end{bmatrix} w_2 \begin{bmatrix} \quad v_2 \quad \end{bmatrix}$$

- **Rank-m ($m << N, M$) approximation: $NM \to m(N + M)$**

$$\begin{bmatrix} \\ \psi \\ \\ \end{bmatrix} \cong \sum_{\nu=1}^{m} \begin{bmatrix} \\ u_\nu \\ \\ \end{bmatrix} w_\nu \begin{bmatrix} \quad v_\nu \quad \end{bmatrix}$$

Singular Value Decomposition

- **Problem: Optimal approximation of an $N \times M$ matrix ψ of rank-m ($m << N$)?**
- **Theorem: An $N \times M$ matrix ψ (assume $N \geq M$) can be decomposed as**

$$\psi = UDV^T = \sum_{\nu=1}^{M} U_{i\nu} d_\nu V_{j\nu} = \sum_{\nu=1}^{M} u_i^{(\nu)} d_\nu v_j^{(\nu)}$$

where $U \in R^N \times R^M$ & $V \in R^M \times R^M$ are column orthogonal & D is diagonal

$$U^T U = V^T V = I_M$$

$$\begin{array}{cc} & M \\ N & \begin{bmatrix} \\ \psi \\ \\ \end{bmatrix} \end{array} = \underset{N \times M}{\begin{bmatrix} \\ U \\ \\ \end{bmatrix}} \underset{M \times M}{\begin{bmatrix} d_1 & & \\ & \ddots & \\ & & d_M \end{bmatrix}} \underset{M \times M}{\begin{bmatrix} \\ V^T \\ \\ \end{bmatrix}}$$

See appendix on polar & singular decompositions

- **Theorem: Sort the SVD diagonal elements in descending order, $d_1 \geq d_2 \geq \ldots \geq d_M \geq 0$, & retain the first m terms**

$$\psi^{(m)} \equiv \sum_{\nu=1}^{m} u^{(\nu)} d_\nu v^{(\nu)T}$$

which is optimal among $\forall$rank-m matrices in the 2-norm sense with the error

$$\min_{rank(A)=m} \left\| A - \psi \right\|_2 = \left\| \psi^{(m)} - \psi \right\|_2 = d_{m+1}$$

cf. singular.c & svdcmp.c   cc -o singular singular.c svdcmp.c -lm

Use the program!

SVD for Image Compression


Original Image


5 Iterations


10 Iterations

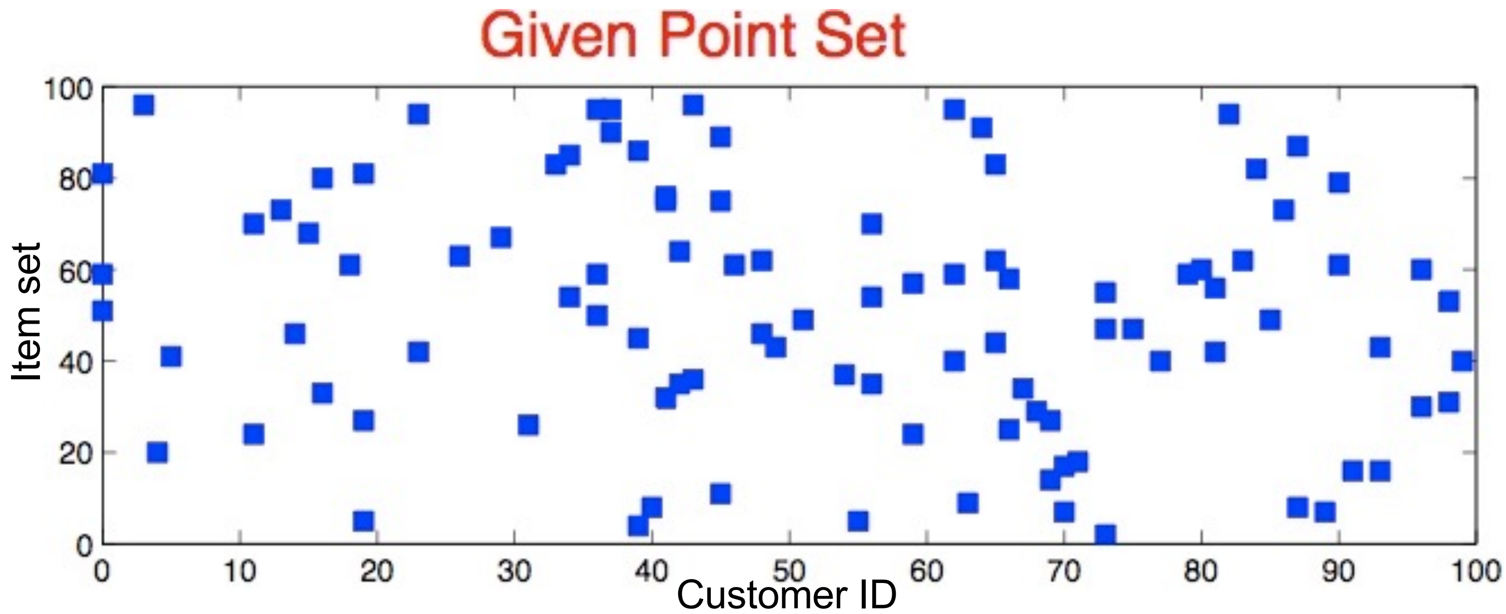
D. Richards & A. Abrahamsen

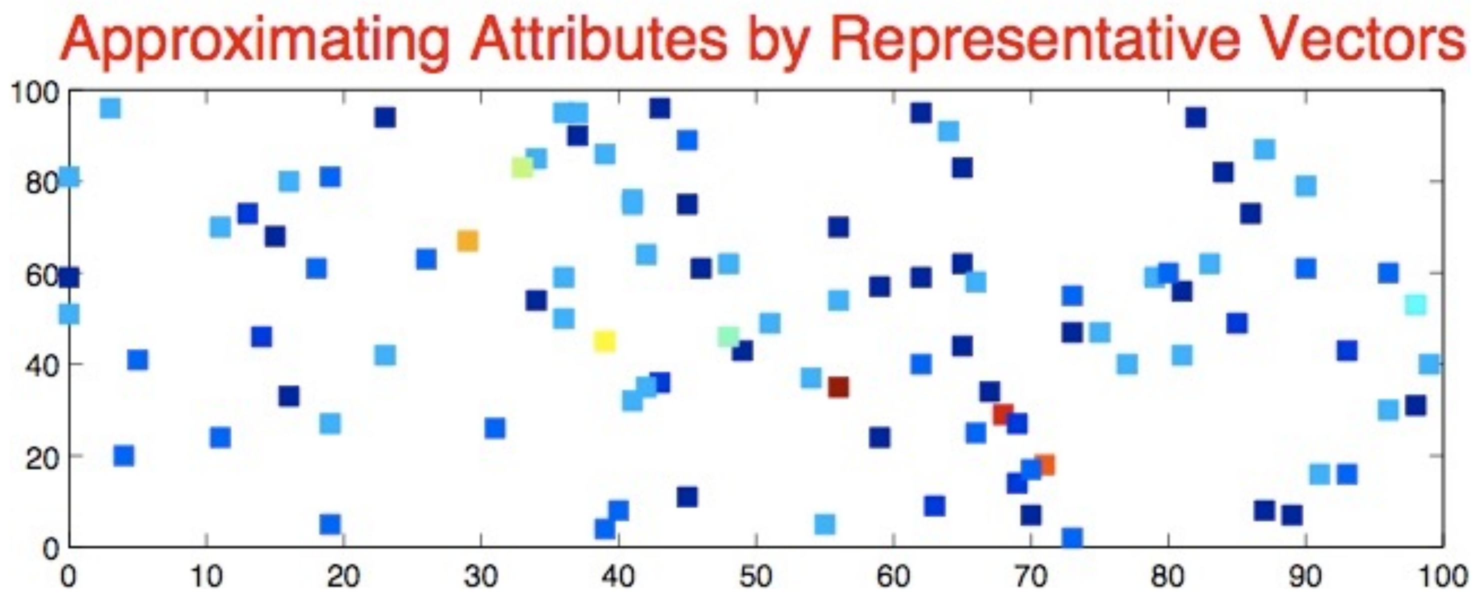

20 Iterations


60 Iterations


100 Iterations

SVD in Data Mining

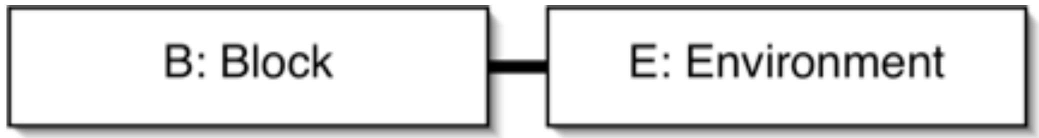




N. Ramakrishnan & A. Y. Grama

Reduced Density Matrix

- **Quantum system coupled to an environment**



$$\{|i\rangle = \psi_i(x) | i = 1, \ldots, N\} \quad \{|j\rangle = \phi_i(X) | j = 1, \ldots, M\}$$

- **∀Quantum state of block + environment**

$$|\psi\rangle = \sum_{i=1}^{N} \sum_{j=1}^{M} \psi_{ij} |i\rangle |j\rangle \quad \text{or} \quad \Psi(x, X) = \sum_{i=1}^{N} \sum_{j=1}^{M} \psi_{ij} \psi_i(x) \phi_j(X)$$

- **Reduced density matrix**

Arbitrary operator in the block

$$\langle {}^{\forall}A \rangle = \sum_i \sum_j \psi_{ij}^* \langle j | \langle i | A \sum_{i'} \sum_{j'} \psi_{i'j'} |i'\rangle |j'\rangle$$

$$= \sum_i \sum_j \sum_{i'} \sum_{j'} \psi_{i'j'} \psi_{ij}^* \langle i | A | i' \rangle \underbrace{\langle j | j' \rangle}_{\delta_{jj'}}$$

$$= \sum_i \sum_{i'} \sum_j \psi_{i'j} \psi_{ij}^* \langle i | A | i' \rangle \equiv \sum_i \sum_{i'} \rho_{i'i} A_{ii'} = \mathrm{tr}_{\mathrm{B}}(\rho A)$$

$$\rho_{i'i} \equiv \sum_j \psi_{i'j} \psi_{ij}^* \qquad A_{ii'} \equiv \langle i | A | i' \rangle$$

Low-Rank Approx. to Reduced Density Matrix

$$\psi \cong \psi^{(m)} = \sum_{\nu=1}^{m} u^{(\nu)} d_\nu v^{(\nu)T} \qquad \psi_{ij}^{(m)} = \sum_{\nu=1}^{m} u_i{}^{(\nu)} d_\nu v_j{}^{(\nu)}$$

$$\rho = \psi\psi^T \cong \psi^{(m)}\psi^{(m)T} = \sum_{\nu=1}^{m}\sum_{\nu'=1}^{m} u^{(\nu)} d_\nu \left(v^{(\nu)T} v^{(\nu')}\right) d_{\nu'} u^{(\nu')T}$$

$$= \sum_{\nu=1}^{m}\sum_{\nu'=1}^{m} u^{(\nu)} d_\nu (\delta_{\nu\nu'}) d_{\nu'} u^{(\nu')T} = \sum_{\nu=1}^{m} u^{(\nu)} d_\nu^2 u^{(\nu)T} \equiv \rho^{(m)}$$

$$\rho_{ii'}^{(m)} = \sum_{\nu=1}^{m} u_i{}^{(\nu)} d_\nu^2 u_{i'}{}^{(\nu)}$$

- **Density matrix renormalization group = systematic procedure to accurately obtain a quantum ground state:**
  1. **Incrementally add environment to a block**
  2. **Solve the global (= block + environment) ground state**
  3. **Construct a low-rank approx. to represent the block with reduced d.o.f.**

S. R. White, *Phys. Rev. B* **48**, 10345 ('93);
G. K.-L. Chan & S. Sharma, *Annu. Rev. Phys. Chem.* **62**, 465 ('11)

Entanglement Entropy

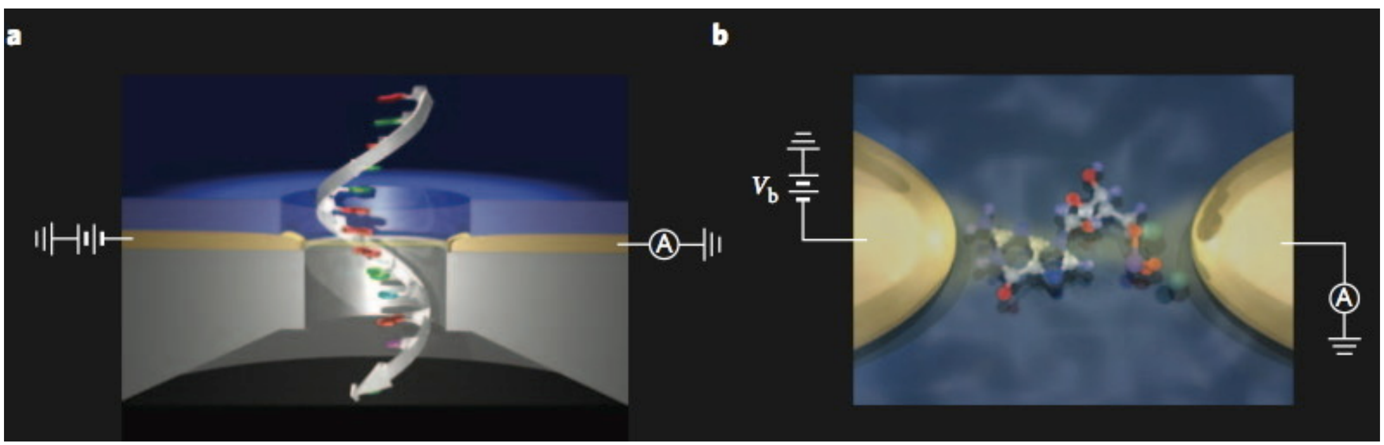
- **Entanglement entropy:** A measure of the degree of quantum entanglement between two subsystems. If a state describing two subsystems A and B is a ***separable*** state $|\Psi_{AB}\rangle = |\phi_A\rangle|\phi_B\rangle$, then the reduced density matrix $\rho_A = \mathrm{Tr}_B|\Psi_{AB}\rangle\langle\Psi_{AB}| = |\phi_A\rangle\langle\phi_A|$ is a ***pure state***. Thus, the entropy of the state is zero. A reduced density matrix having a non-zero entropy is therefore a signal of the existence of entanglement in the system.
- **Area law:** A quantum state satisfies an ***area law*** if the leading term of the entanglement entropy grows at most proportionally with the ***boundary*** between the two partitions. Area laws are remarkably common for ground states of local gapped quantum many-body systems. ***It greatly reduces the complexity of quantum many-body systems***. ***The density matrix renormalization group and matrix product states, for example, implicitly rely on such area laws***.

https://en.wikipedia.org/wiki/Entropy_of_entanglement

SVD for Rapid Genome Sequencing

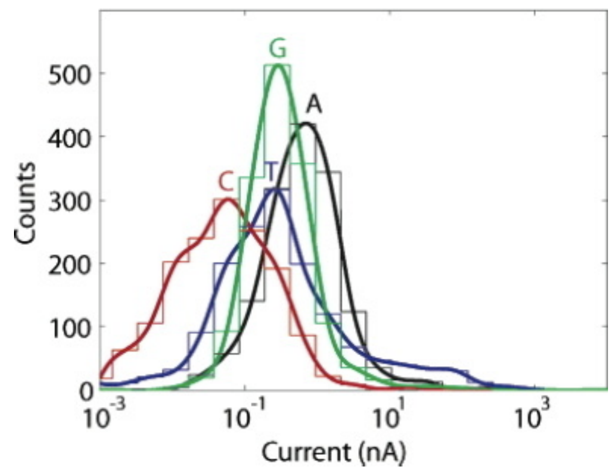
- **$10M Archon X prize for decoding 100 human genomes in 10 days & $10K per genome (**http://genomics.xprize.org**): Preemptive attack on diseases**

- **Quantum tunneling current for rapid DNA sequencing?**



Tsutsui *et al.*, *Nature Nanotechnology* ('10)

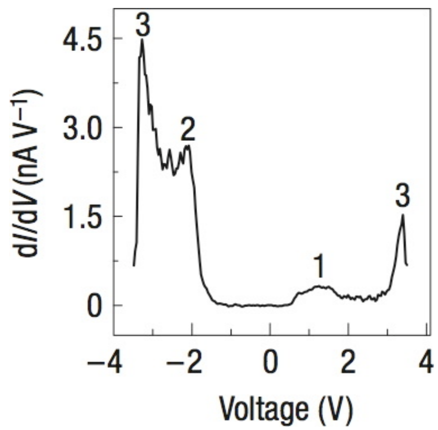
Lagerqvist *et al.*, *Nano Letters* ('06)



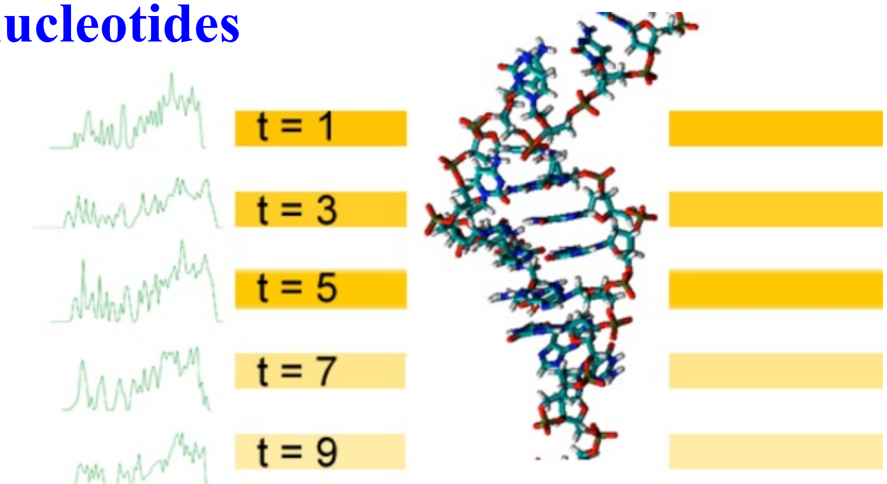
- **Tunneling current alone cannot distinguish the 4 nucleotides (A, C, G, T)**

Rapid DNA Sequencing *via* Data Mining

- **Use tunneling current (I)-voltage (V) characteristic (or electronic density-of-states) as the 'fingerprints' of the 4 nucleotides**

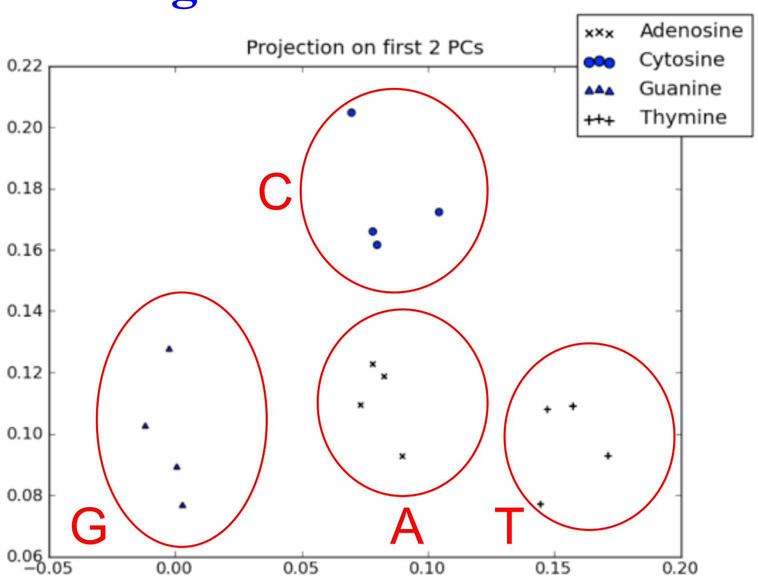


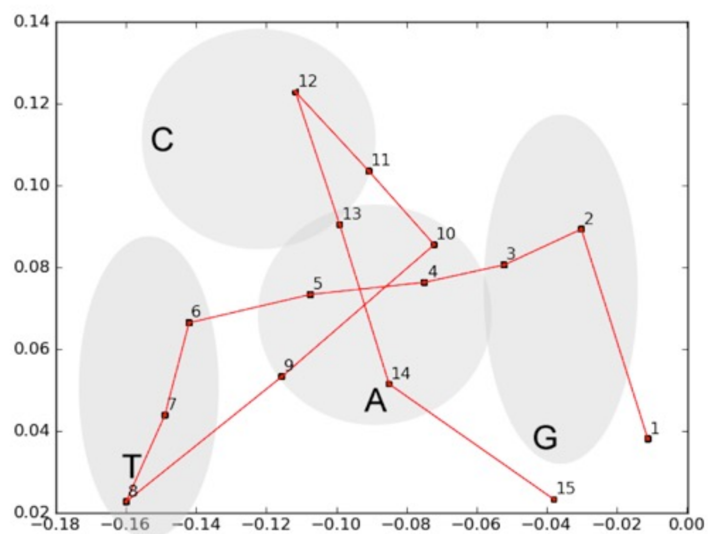
Shapir *et al*.,
Nature Materials ('08)



- **Principal component analysis (PCA) & fuzzy c-means clustering clearly distinguish the 4 nucleotides**

H. Yuen *et al*., *IJCS* **4**, 352 ('10)





http://www.henryyuen.net/

- **Viterbi algorithm for even higher-accuracy sequencing**

SVD *vs.* PCA (in Economics)

- SVD of N (number of companies) × T (number of time points) of stock-price time series

$$\underset{T\times N}{\Xi^{T}} = \underset{T\times N}{\mathbf{U}}\ \underset{N\times N}{\Sigma}\ \underset{N\times N}{\mathbf{V}^{T}}$$

- Stock correlation matrix

$$\underset{N\times N}{\mathbf{C}} = \underset{N\times T}{\Xi}\ \underset{T\times N}{\Xi^{T}}$$

- Principal component analysis (PCA): Eigen decomposition of the correlation matrix

$$\mathrm{C} = \Xi\Xi^{T}$$

$$= \mathrm{V}\Sigma\,\overbrace{\mathrm{U}^{T}\mathrm{U}}^{\mathrm{I}}\,\Sigma\mathrm{V}^{T}$$

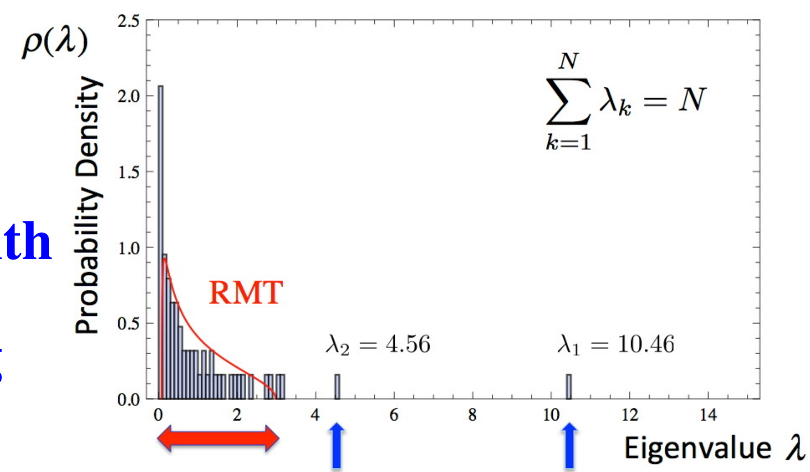
$$= \mathrm{V}\Sigma^{2}\mathrm{V}^{T}$$

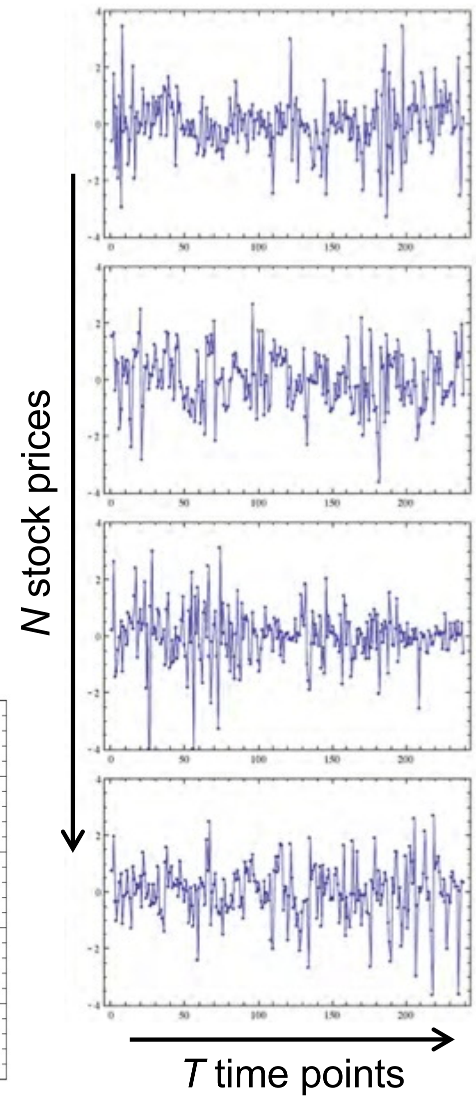
- Compare the spectrum with that of random matrix theory (RMT) for judging statistical significance

$$\rho_{\mathrm{RMT}} = \frac{Q}{2\pi}\frac{\sqrt{(\lambda_{+}-\lambda)(\lambda-\lambda_{-})}}{\lambda}$$

$$\lambda_{\pm} = \left(1 \pm \frac{1}{\sqrt{Q}}\right)^{2}$$

$$Q = T/(2N) \quad N,T \to \infty$$





Y. Kichikawa *et al.*, *Proc. Comp. Sci.* **60**, 1836 ('15)