Singular Value Decomposition

Aiichiro Nakano

Collaboratory for Advanced Computing & Simulations
Dept. of Computer Science, Dept. of Physics & Astronomy,
Dept. of Chemical Engineering & Materials Science,
Dept. of Biological Sciences
University of Southern California

Email: anakano@usc.edu

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

Rank of a Matrix

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

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

- **Range of** A**: Vector space** $\{b = Ax | \forall x\}$
- **Rank of** A**: Number 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:** $NM \rightarrow N + M$

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

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

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

- **Rank-m (m $\ll N, M$) approximation:** $NM \rightarrow m(N + M)$

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

Singular Value Decomposition

- **Problem:** Optimal approximation of an $N \times M$ matrix ψ of rank-m ($m \ll N$)?
- **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

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

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

- **Theorem:** Sort the SVD diagonal elements in descending order $d_1 \geq d_2 \geq \ldots$ & 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

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

cf. `singular.c` & `svdcmp.c`

SVD for Image Compression

Original Image

5 Iterations

10 Iterations

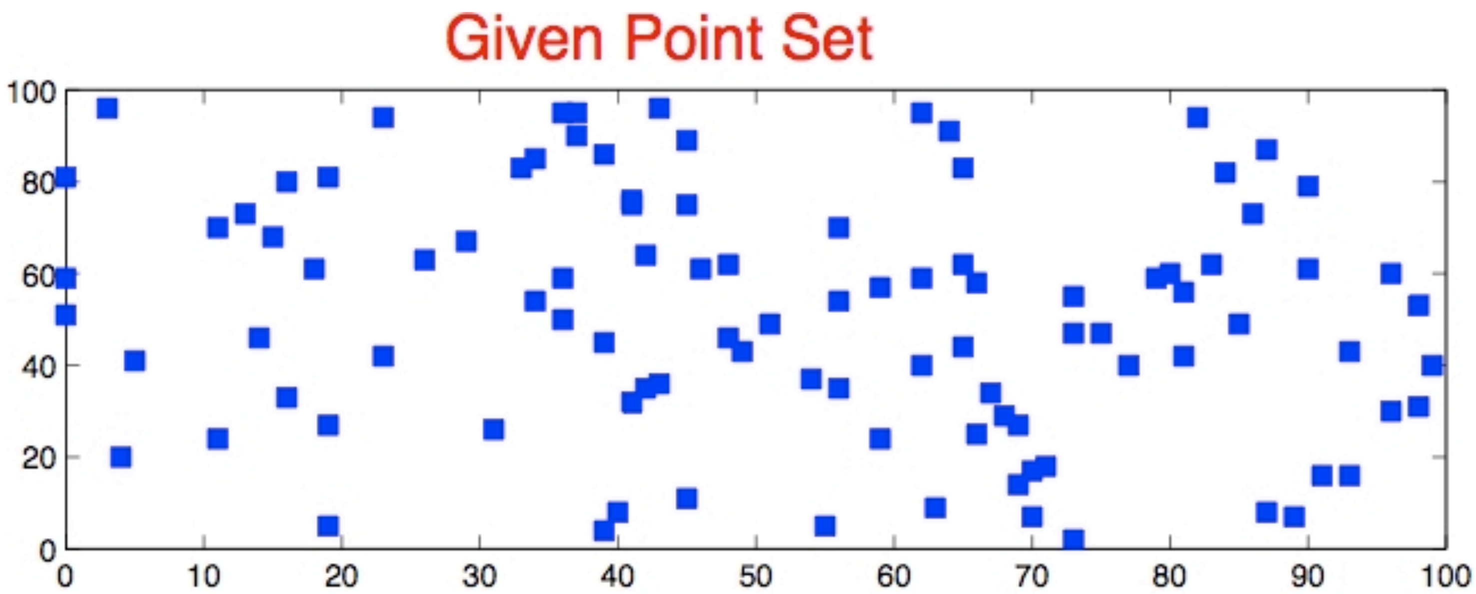
D. Richards & A. Abrahamsen

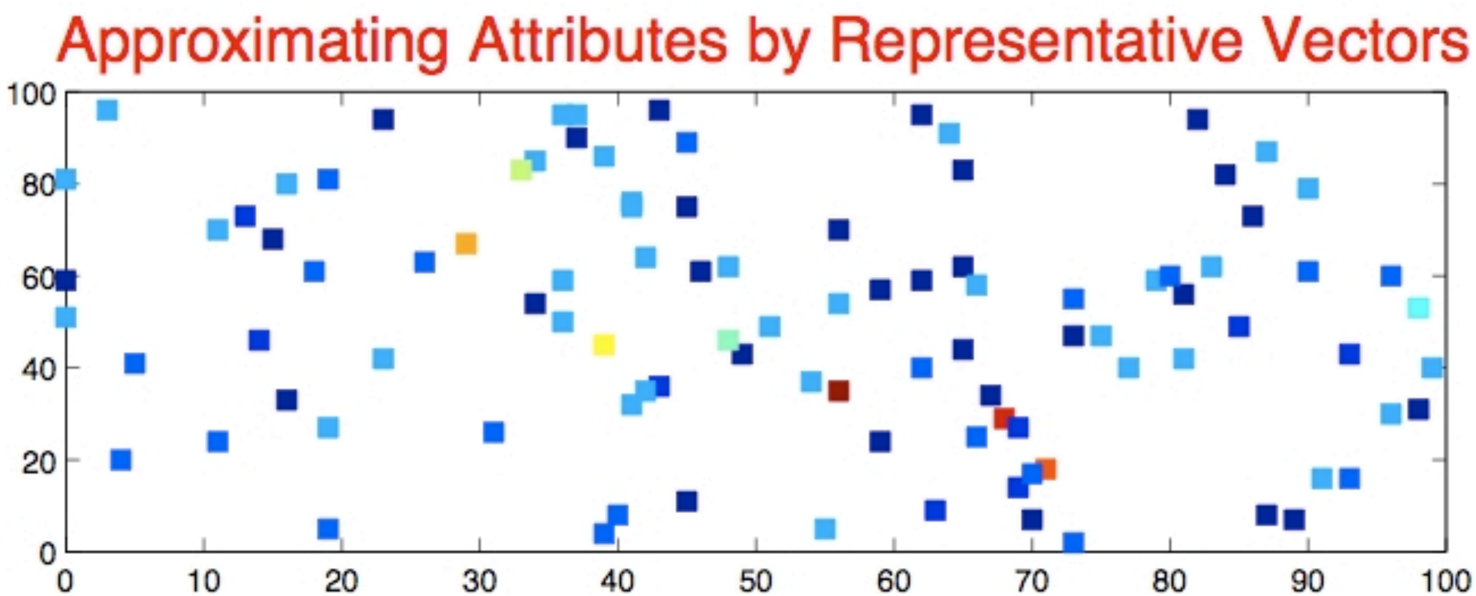
20 Iterations

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**

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

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

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

Arbitrary operator
in the block

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

Low-Rank Approx. to Reduced Density Matrix

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

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

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

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

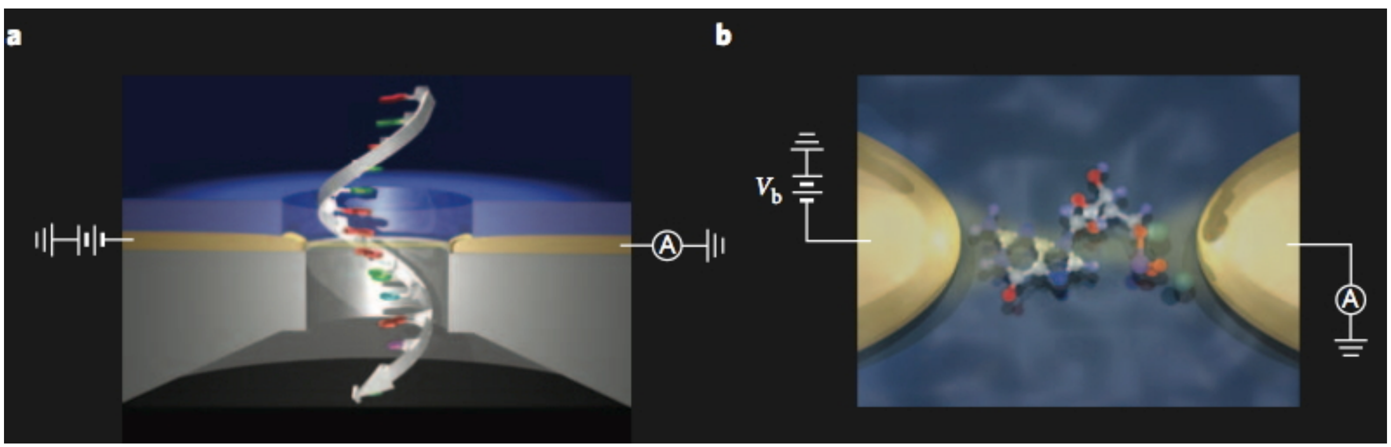
- **Density matrix renormalization group = systematic procedure to accurately obtain a quantum ground state:**
  1. **Incrementally add environment to a block**
  2. **Solve the global (= block + environment) ground state**
  3. **Construct a low-rank approx. to represent the block with reduced d.o.f.**

S. R. White, *Phys. Rev. B* **48**, 10345 ('93)

Rapid Genome Sequencing

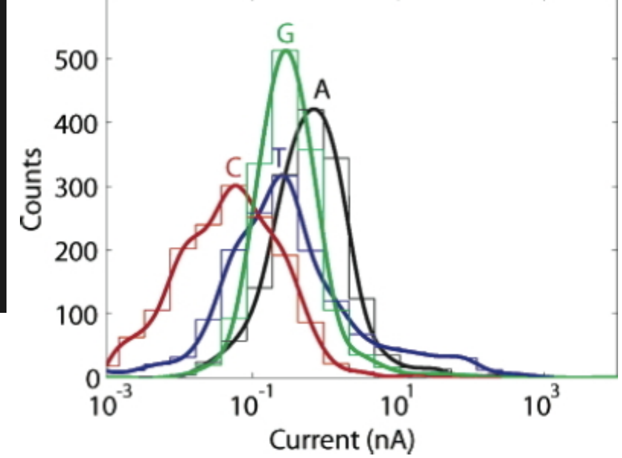
- **$10M Archon X prize for decoding 100 human genomes in 10 days & $10K per genome (**http://genomics.xprize.org**): Preemptive attack on diseases**

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



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

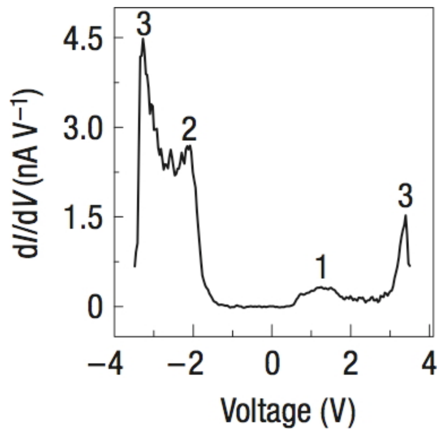
Lagerqvist *et al*., *Nano Letters* ('06)



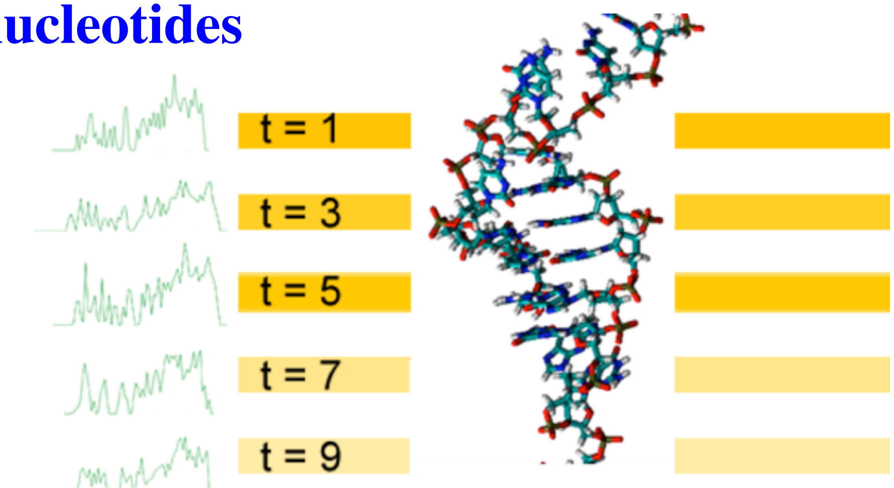
- **Tunneling current alone cannot distinguish the 4 nucleotides (A, C, G, T)**

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**

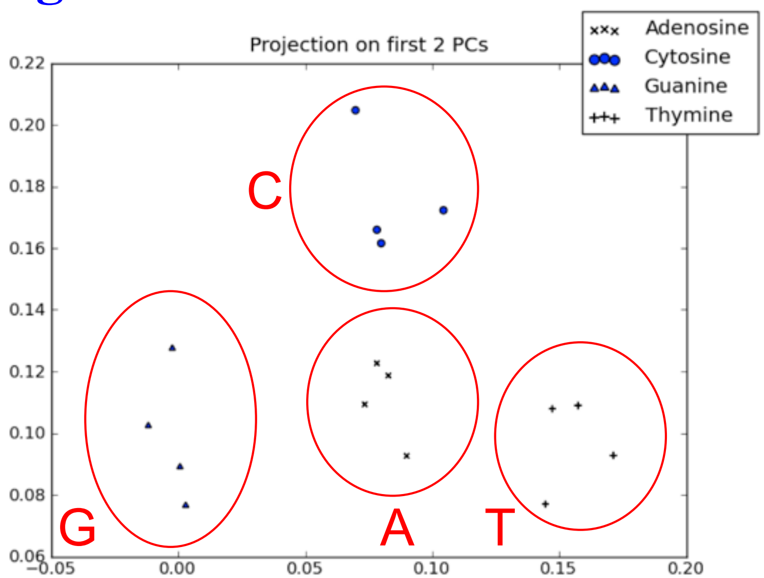


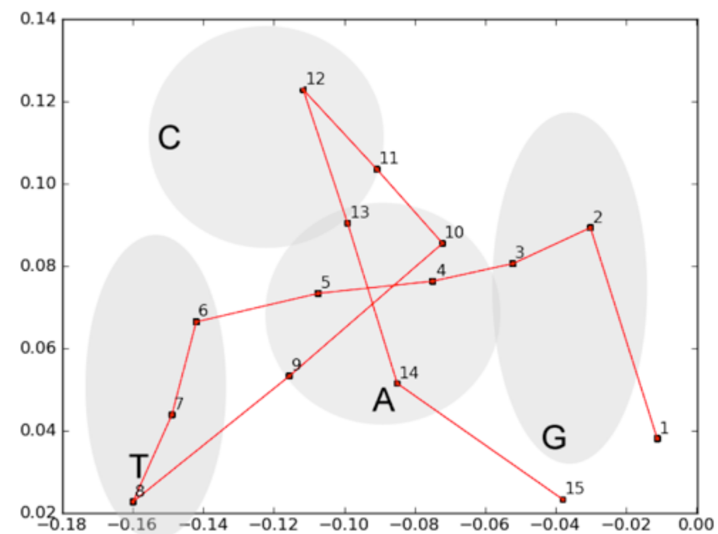
Shapir *et al.*,
Nature Materials ('08)



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

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





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

SVD *vs*. PCA

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

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

- **Stock correlation matrix**

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

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

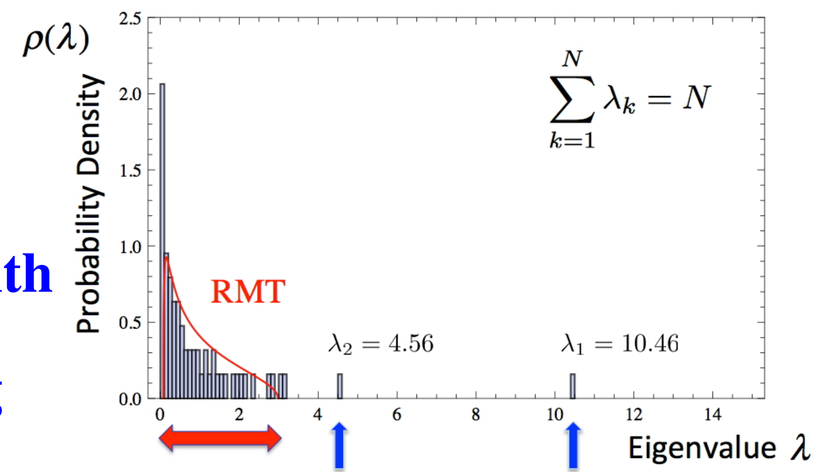
$$\begin{aligned}\mathbf{C} &= \boldsymbol{\Xi}\boldsymbol{\Xi}^T \\ &= \mathbf{V}\boldsymbol{\Sigma}\,\underbrace{\mathbf{U}^T\mathbf{U}}_{\mathbf{I}}\,\boldsymbol{\Sigma}\mathbf{V}^T \\ &= \mathbf{V}\boldsymbol{\Sigma}^2\mathbf{V}^T\end{aligned}$$

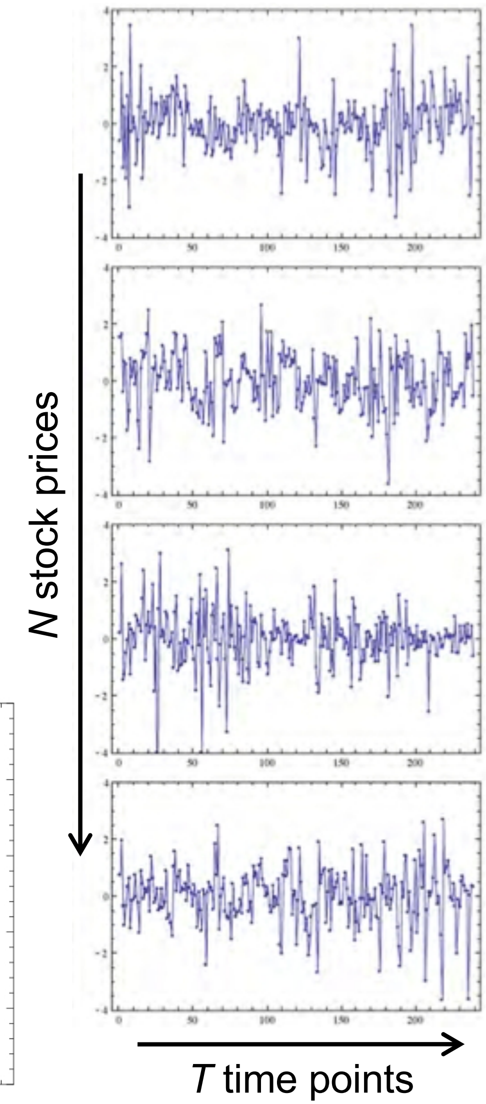
- **Compare the spectrum with that of random matrix theory (RMT) for judging statistical significance**

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

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

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





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