Quantum Change-point and Anomaly Detection

John Calsamiglia

GIQ, Universitat Autònoma de Barcelona

Budapest, September 2017

M. Skotiniotis, R. Hotz, JC, R. Muñoz-Tapia, arXiv:1808.02729

G. Sentís, E. Bagan, JC, G. Chiribella, R. Muñoz-Tapia, PRL 117 150502 (2016)

G. Sentís, JC, R. Muñoz-Tapia, PRL 119, 140506 (2017)

Grup d'Informació Quàntica

Universitat Autònoma de Barcelona

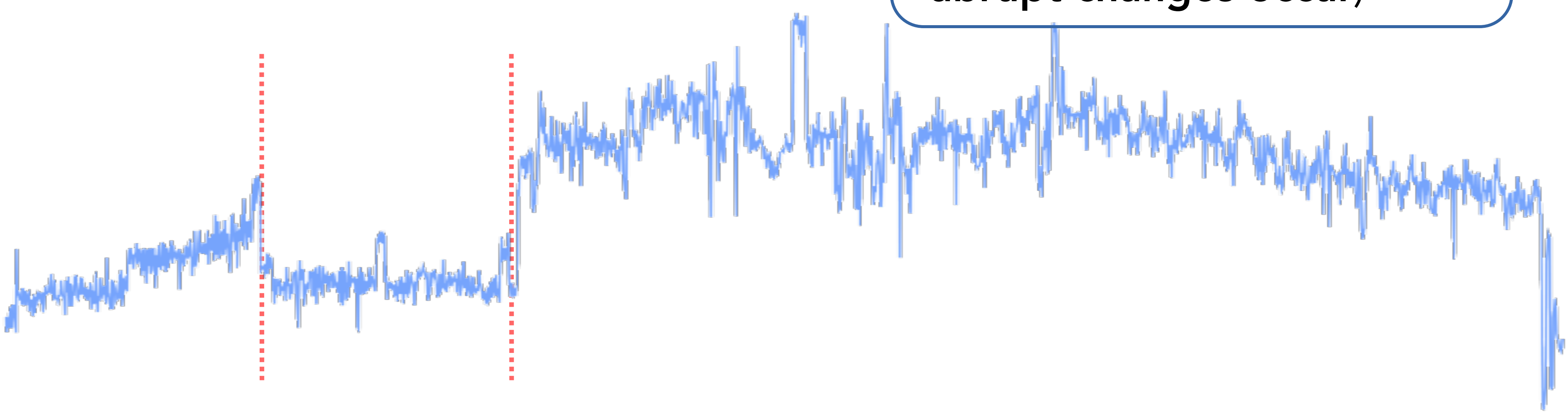
Classical change-point problem

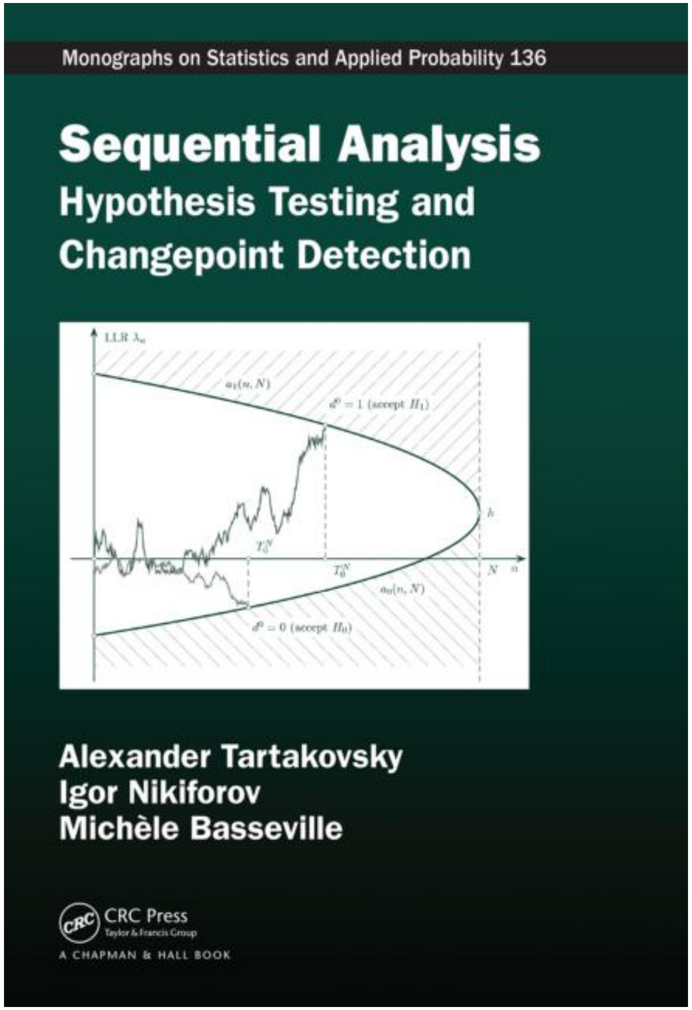
In statistical analysis, **change-point detection** tries to identify times when the probability distribution of a stochastic process or time series changes.

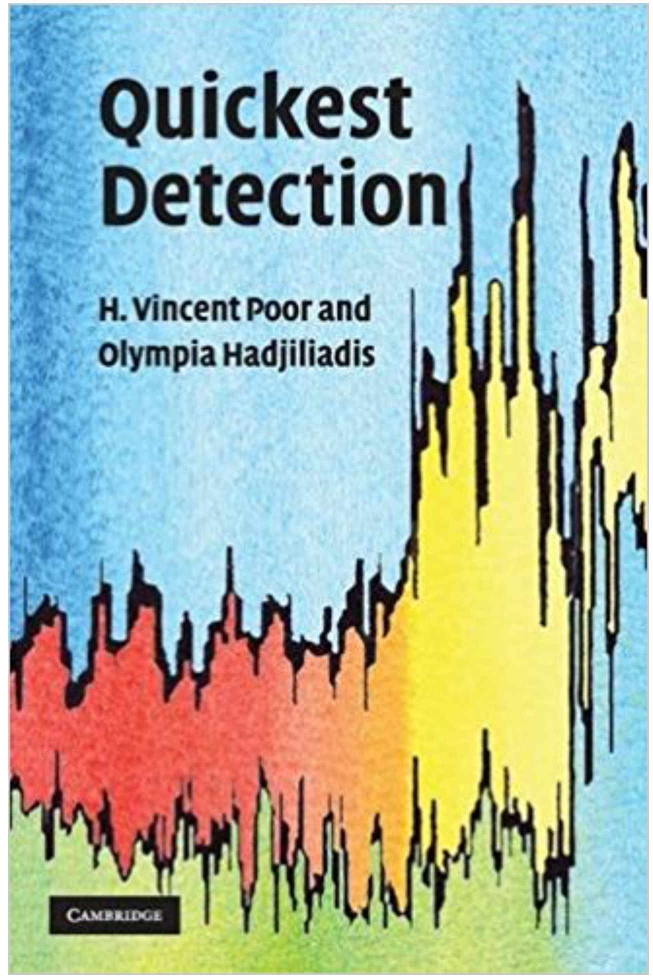
- Landscape changes
- Quality control, fault detection
- Stock market variations
- Network intrusion
- Protein folding

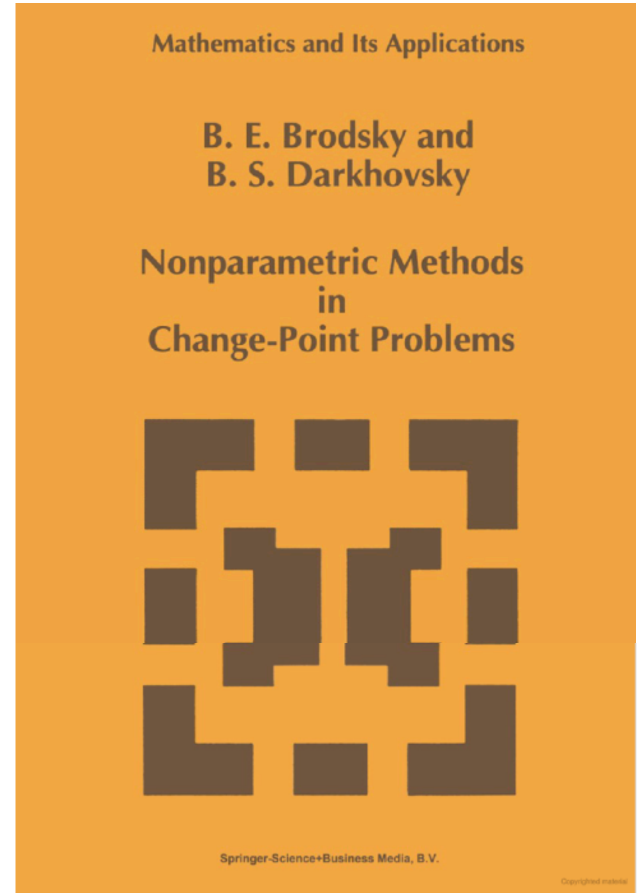
away from i.i.d setting!

statistical analysis requires the stability of the system parameters (**not suitable if abrupt changes occur**)









Econometrica, Vol. 61, No. 4 (July, 1993), 821–856

TESTS FOR PARAMETER INSTABILITY AND STRUCTURAL CHANGE WITH UNKNOWN CHANGE POINT

By Donald W. K. Andrews[1]

A Non-Parametric Approach to the Change-Point Problem
Author(s): A. N. Pettitt
Source: *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, Vol. 28, No. 2 (1979), pp. 126-135

Quantum change point (QCP)

- A source is expected to prepare *N* quantum particles in the same default state $|0\rangle$
- At position k something occurs and the source starts producing altered states $|\phi\rangle$

$$|0\rangle_1 |0\rangle_2 \cdots |0\rangle_{k-1} |\phi\rangle_k |\phi\rangle_{k+1} \cdots |\phi\rangle_N$$

- Assumptions:
  - states and *N* are known
  - change point is equally likely to occur in all positions.

Task: identify k

QCP | Multihypothesis discrimination

- Hypotheses: $\{|\Psi_k\rangle\}_{k=1}^{N}$

$$|\Psi_k\rangle = |0\rangle_1 |0\rangle_2 \cdots |0\rangle_{k-1} |\phi\rangle_k |\phi\rangle_{k+1} \cdots |\phi\rangle_N$$

- Collective measurement with N outcomes (POVM)

$$\{M_k\} \qquad \sum_k M_k = \mathbb{1} \qquad M_k \geq 0$$

...but hypotheses are linearly independent → von Neumann ok

$$M_k = |m_k\rangle\langle m_k| \qquad \langle m_k|m_l\rangle = \delta_{kl} \quad p(k'|k) = |\langle m_{k'}|\Psi_k\rangle|^2$$

QCP | Multihypothesis discrimination

- All relevant information is encoded in the Gram matrix

$\{|\Psi_k\rangle\}_{k=1}^{N}$ linearly independent

$$G_{ij} = \langle\Psi_i|\Psi_j\rangle = \sum_k \langle\Psi_i|m_k\rangle\underbrace{\langle m_k|\Psi_j\rangle}_{B_{kj}} = \sum_k B^\dagger_{ik} B_{kj}$$

$$G = B^\dagger B > 0$$

- Figure of merit: success probability of correct identification

$$P_s = \max_{\{|m_k\rangle\}} \frac{1}{N} \sum_k |\langle m_k|\Psi_k\rangle|^2$$

$$P_s = \max_B \frac{1}{N} \sum_k |B_{kk}|^2 \quad \text{s.t.} \quad B^\dagger B = G$$

QCP | Success probability

- Each choice of B corresponds to a choice of measurement basis:

$$P_s = \max_B \frac{1}{N} \sum_k |B_{kk}|^2 \quad \text{s.t.} \quad B^\dagger B = G$$

- Self-adjoint choice ⟶ Square Root Measurement

$$S = \sqrt{G} = S^\dagger \qquad G = S^2 \qquad P_s \geq \frac{1}{N} \sum_k |S_{kk}|^2 \equiv P_{\text{SRM}}$$

Saturated iff $S_{kk} = S_{k'k'} \quad \forall k, k'$

N. Dalla Pozza N, G. Pierobon, Phys Rev A ;91 042334 (2015)

QCP | Success probability

How well does the SRM approximate the optimal measurement?

Theorem

$$\left(\frac{\operatorname{tr} S}{N}\right)^2 \leq P_s^{\text{opt}} \leq \left(\frac{\operatorname{tr} S}{N}\right)^2 + \sqrt{\lambda_{\max}}\, \|\mathbf{q} - \mathbf{u}\|_1 \qquad q_k = \frac{S_{kk}}{\sum S_{kk}} \qquad u_k = \frac{1}{N}$$

For constant S_{kk} : $P_{SRM} = \frac{1}{N}\sum_k |S_{kk}|^2 = |S_{11}|^2 = \left(\frac{\operatorname{tr} S}{N}\right)^2$

Obtaining the explicit form of S_{kk} is not easy, since we need to diagonalize G:

$$G_{ij} = \langle \Psi_i | \Psi_j \rangle = c^{|i-j|} \qquad c := \langle 0 | \phi \rangle$$

$$G = \begin{pmatrix} 1 & c & c^2 & \cdots & c^{N-1} \\ c & 1 & c & \cdots & c^{N-2} \\ c^2 & c & 1 & \cdots & c^{N-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c^{N-1} & c^{N-2} & c^{N-3} & \cdots & 1 \end{pmatrix}$$

Not tridiagonal, not circulant: **Life is hard**

QCP | Asymptotic optimality

With the explicit eigensystem $\{\lambda_l, v_l\}$, we can compute...

$$S = \sum_l \sqrt{\lambda_l}\, |v_l\rangle\langle v_l| \qquad S_{kk} = \sum_{l=1}^{N} \sqrt{\lambda_l} |v_k^l|^2 \qquad P_s = \frac{1}{N} \sum_{k=1}^{N} |S_{kk}|^2$$

...but it's a mess for finite N

Only a vanishing fraction of S_{kk} deviate from the mean for large N:

$$S_{kk} - \lim_{N\to\infty} \frac{\operatorname{tr} S}{N} = O\left(\frac{c^{2k}}{k^{3/2}}\right) \qquad \|\mathbf{q} - \mathbf{u}\|_1 \leq \frac{1+c}{1-c} \frac{4}{N^{1-\epsilon}} + O\left(\frac{1}{N}\right)$$

$$\left(\frac{\operatorname{tr} S}{N}\right)^2 \leq P_s^{\text{opt}} \leq \left(\frac{\operatorname{tr} S}{N}\right)^2 + \sqrt{\lambda_{\max}}\, \|\mathbf{q} - \mathbf{u}\|_1$$

As $N \to \infty$ only a vanishing fraction of S_{kk} deviate from the mean.

QCP | Asymptotic optimality

So, in the asymptotic limit:

$$P_s^{\rm opt} = \lim_{N\to\infty}\left(\frac{\operatorname{tr} S}{N}\right)^2 = \lim_{N\to\infty}\left(\frac{1}{N}\sum_{l=1}^{N}\sqrt{\lambda_l}\right)^2 = \lim_{N\to\infty}\left(\frac{1}{N}\sum_{l=1}^{N}\sqrt{\frac{1-c^2}{1-2c\cos\theta_l+c^2}}\right)^2$$

Almost there!

We still have to sum the eigenvalues…

Last small technicality: uniform distribution of zeroes: $\{\theta_l\}$ uniform in $(0,\pi)$

$$P_s^{\rm opt} = \frac{1-c^2}{\pi^2}\left(\int_0^\pi \frac{{\rm d}\theta}{\sqrt{1-2c\cos\theta+c^2}}\right)^2$$

$\sum \to \int$

$$P_s^{\rm opt} = \frac{4(1-c^2)}{\pi^2} K^2\left(c^2\right)$$

It's finite!

complete elliptic function of the first kind

Local QCP strategies

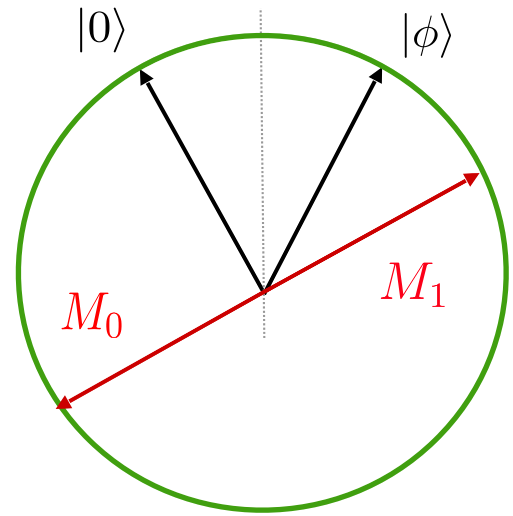
- Does it pay off to use full quantum correlations?

- How far are local strategies from optimality?

Local QCP | Fixed measurements

Measure each particle with a generic Stern-Gerlach measurement.

In Bloch sphere



This is the same as classical CP for coins with different bias:

$$p_A(0) = \langle 0|\, M_0 \,|0\rangle = p$$
$$p_A(1) = \langle 0|\, M_1 \,|0\rangle = 1 - p$$

$$p_B(0) = \langle \phi|\, M_0 \,|\phi\rangle = q$$
$$p_B(1) = \langle \phi|\, M_1 \,|\phi\rangle = 1 - q \quad ,$$

Local QCP | Fixed measurements

$i =$	1	2	3	4	5	6	7	8	9	10	11	12	13	14=N
$r_i =$	0	1	0	1	1	1	0	1	1	1	0	0	1	0

k

$$P(\boldsymbol{r}|k) = \prod_{i=1}^{k-1} p(r_i) \prod_{i=k}^{N} q(r_i) \qquad k = 1, \ldots, N$$

$$P_s = \sum_{k=1}^{N} \sum_{\mathbf{r}} P(\mathbf{r}, k)\delta_{k\,\hat{k}(\mathbf{r})} = \sum_{\mathbf{r}} \max_k P(\mathbf{r}, k) = \frac{1}{N} \sum_{\mathbf{r}} \max_k P(\mathbf{r}|k)$$

Given a sequence of results $\mathbf{r}$ the optimal guess (estimator) is given by

$$\hat{k}(\mathbf{r}) = \underset{k}{\operatorname{argmax}}\, P(\mathbf{r}|k)$$

Local QCP | Fixed measurements

$i =$	1	2	3	4	5	6	7	8	9	10	11	12	13	14=N
$r_i =$	0	1	0	1	1	1	0	1	1	1	0	0	1	0

k

The value of k that optimizes $P_s(\mathbf{r}|k)$ also optimizes the likelihood ratio:

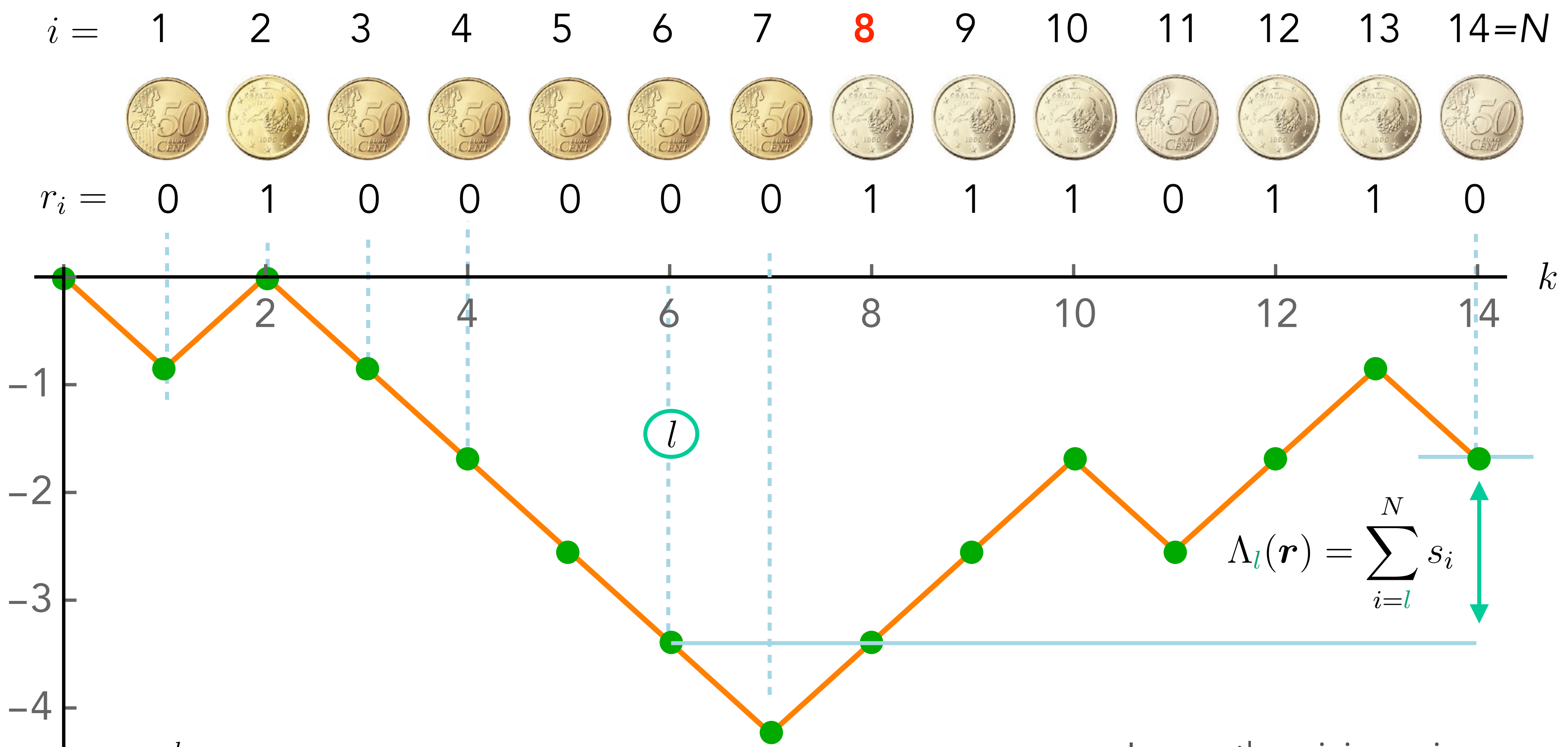
$$\Lambda_k(\boldsymbol{r}) = \log\left(\frac{P(\mathbf{r}|k)}{P(\mathbf{r}|\mathrm{nc})}\right) = \log\prod_{i=k}^{n}\frac{q(r_i)}{p(r_i)} = \sum_{i=k}^{N}\log\frac{q(r_i)}{p(r_i)} = \sum_{i=k}^{N} s_i \quad \text{where} \quad s_i = \log\frac{q(r_i)}{p(r_i)}$$

Optimal guess: $\hat{k}(\boldsymbol{r}) = \underset{k}{\operatorname{argmax}}\,\Lambda_k(\boldsymbol{r})$

Change at k: $P(\boldsymbol{r}|k) \;=\; \prod_{i=1}^{k-1} p(r_i)\prod_{i=k}^{N} q(r_i)$

No change: $P(\boldsymbol{r}|\mathrm{nc}) = \prod_{i=1}^{N} p(r_i) \;=\; \prod_{i=1}^{k-1} p(r_i)\prod_{i=k}^{N} p(r_i)$

We define a random walker that upon outcome r_i it makes a step $s_i = \log \frac{q(r_i)}{p(r_i)}$



$$S_k(\boldsymbol{r}) = \sum_{i=1}^{k} s_i$$

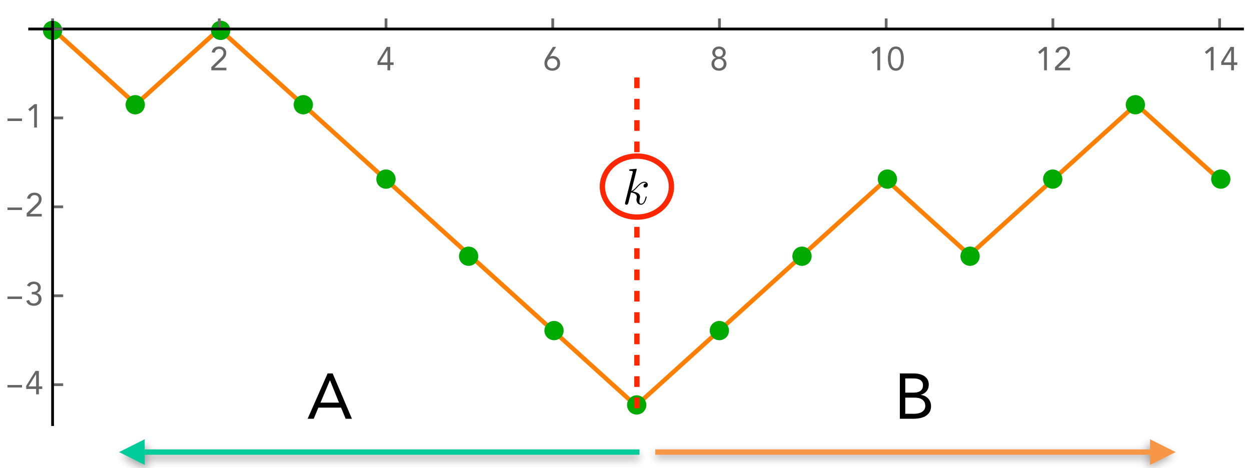
$$\hat{k}(\boldsymbol{r}) = \underset{l}{\operatorname{argmax}}\, \Lambda_l(\boldsymbol{r}) = 7$$

In case the minimum is degenerate, pick any of those, e.g. the first minimum.

Local QCP | Fixed measurements

Since, $P_s^{\text{fix}} = \frac{1}{N}\sum_{\mathbf{r}} \max_k P(\mathbf{r}|k) = \frac{1}{N}\sum_{\mathbf{r}} P(\mathbf{r}|\hat{k}(\boldsymbol{r}))$

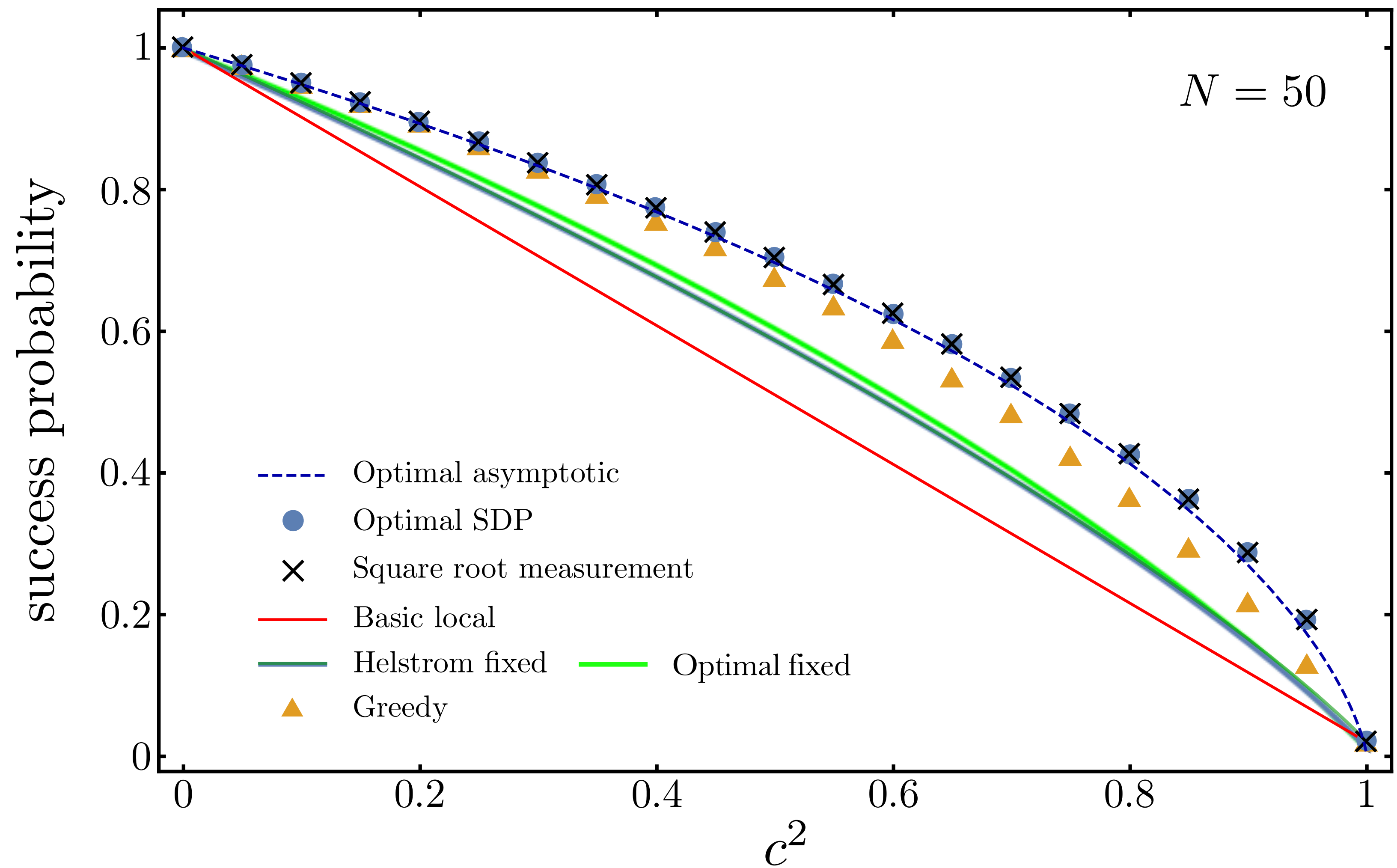
we just need to compute the probability that given that a change occurred in position *k*, the random walk exhibits a maximum precisely at position *k*.



$P_s^{\text{fix}} =$ Prob(walker A never returns to origin & walker B never crosses origin.)

HARD to compute!

Local QCP | Fixed Helstrom



Unambiguous QCP

Task: identify exactly the change point, no errors allowed, by allowing for an inconclusive outcomes:

$$\{M_1 \, , \ldots , M_N , M_? \,\}$$

M_k "clicks" if and only if input state is $|\Psi_k\rangle$, i.e. change occurred at position k:

$$p(k' \neq k|k) = \langle\Psi_k| \, M_{k'} \, |\Psi_k\rangle = 0$$

Unambiguous QCP

Task: identify exactly the change point, no errors allowed.

We can do this because $\{|\Psi_k\rangle\}$ are linearly independent.

$$G = B^\dagger B \qquad B = \sum_k |\Psi_k\rangle\langle k| \qquad \Omega = BB^\dagger = \sum_k |\Psi_k\rangle\langle\Psi_k|$$

The inverse of Ω provides the unambiguous POVM

$$\Omega^{-1} = \sum_k |\Phi_k\rangle\langle\Phi_k| \ , \quad \langle\Phi_j|\Psi_k\rangle = \delta_{jk}$$

$$M_k^{\mathrm{U}} = \gamma_k \, |\Phi_k\rangle\langle\Phi_k|$$

$$M_?^{\mathrm{U}} = \mathbb{1} - \sum_k \gamma_k \, |\Phi_k\rangle\langle\Phi_k| \geq 0$$

efficiencies

$$\gamma_k = p(\hat{k} = k|k) = \langle\Psi_k| \, M_k^U \, |\Psi_k\rangle$$

$B^\dagger(\,\cdot\,)B$

SDP formulation

$$P_s^U = \max_{\{\gamma_k\}} \frac{1}{N} \sum_k \gamma_k = \frac{1}{N} \max_\Gamma \operatorname{tr}(\Gamma)$$

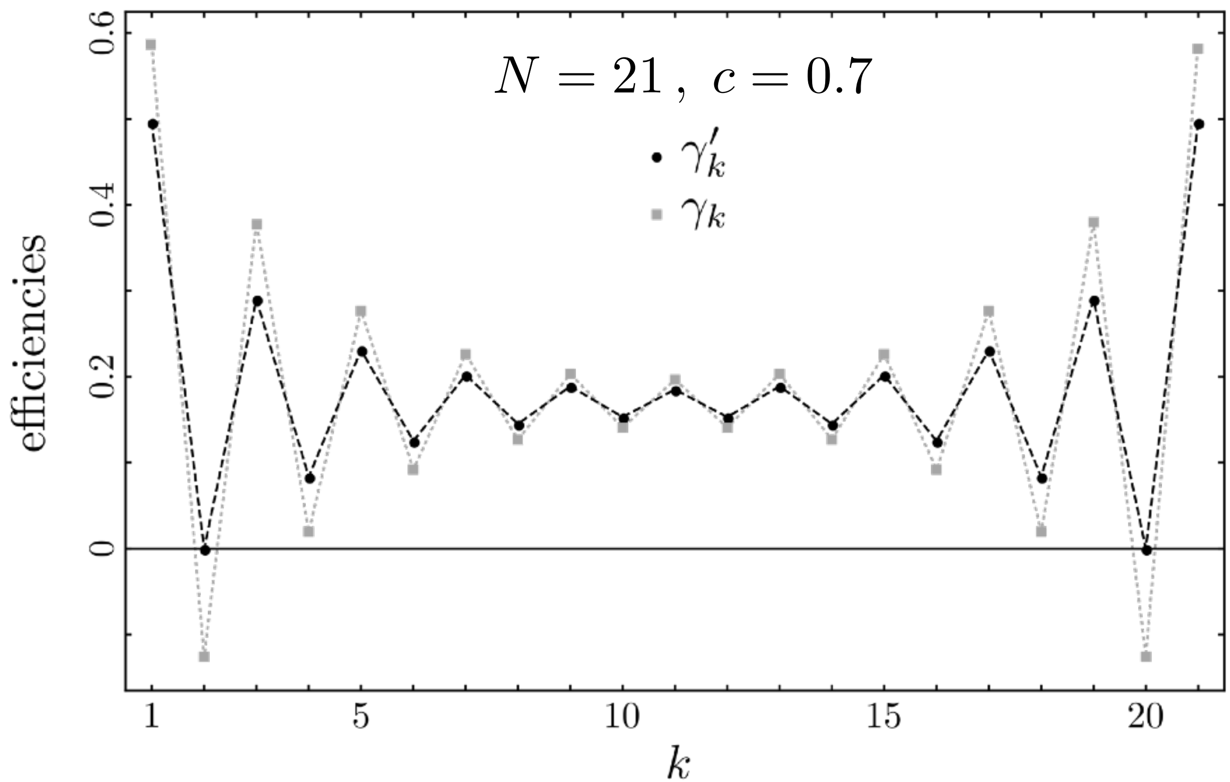
s.t.

$$G - \Gamma_D \geq 0$$

$$\Gamma \geq 0$$

$$\Gamma_D = \operatorname{diag}(\Gamma) = \begin{pmatrix} \gamma_1 & & & \\ & \gamma_2 & & \\ & & \ddots & \\ & & & \gamma_N \end{pmatrix}$$

Unambiguous QCP



The sets of efficiencies,

$$\Gamma = \{\gamma_k\} \quad \text{if} \quad c \leq c^*$$

$$\Gamma' = \{\gamma'_k\} \quad \text{if} \quad c > c^*$$

satisfy the feasibility conditions for all c:

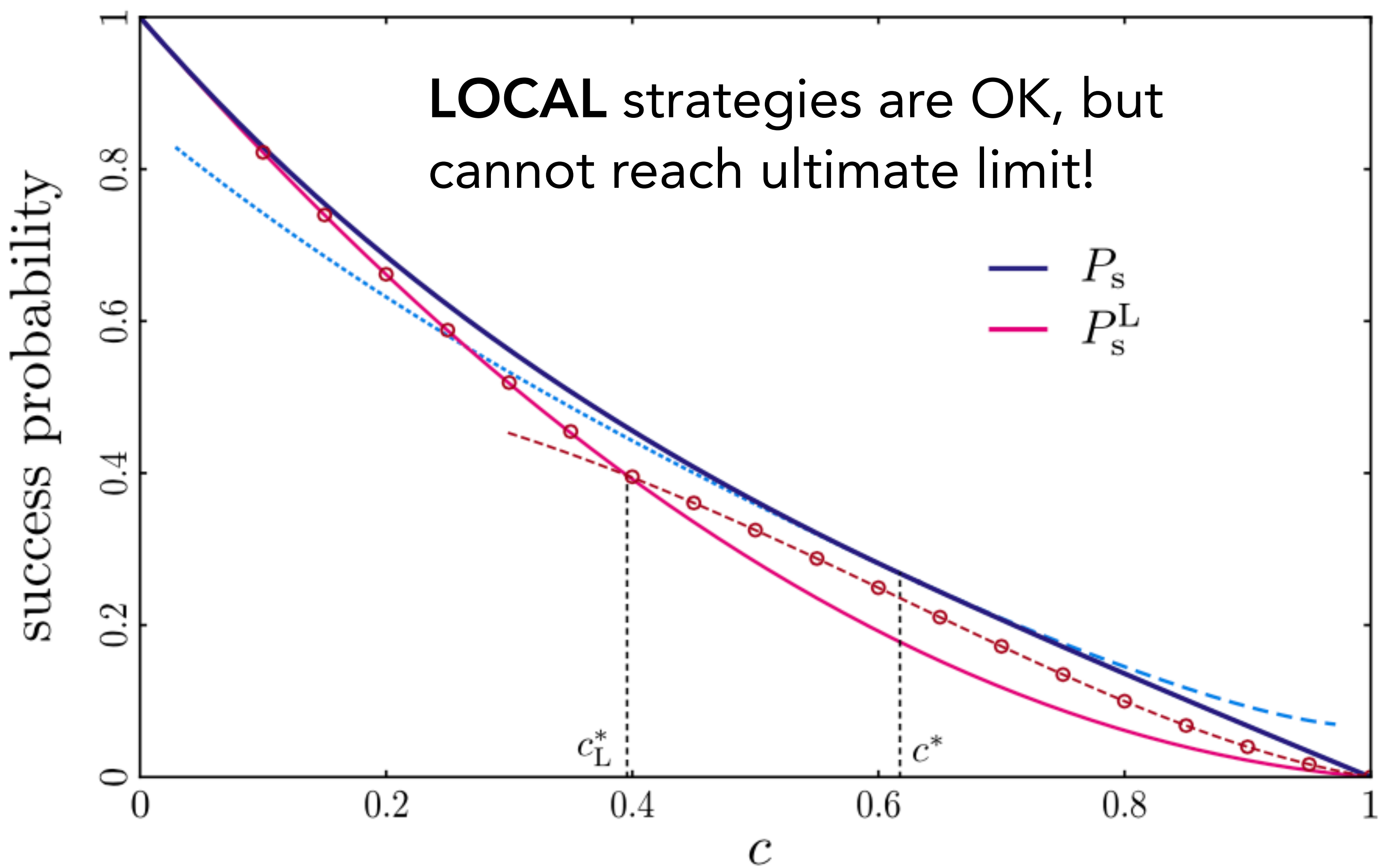
$$G - \Gamma_D \geq 0 \qquad \Gamma \geq 0$$

lengthy proof...

$$P_s^{\mathrm{U}} = \begin{cases} \frac{1-c}{1+c} + \frac{1}{N}\,\frac{2c[1-(-c)^N]}{(1+c)^2} & 0 \leq c \leq c^* \\ \frac{1-c}{1+c} + \frac{1}{N}\left(\frac{2c[1-(-c)^N]}{(1+c)^2} - \frac{2\gamma_2^2}{1+(-c)^{N-3}}\right) & c^* < c \leq 1 \end{cases}$$

Exact solution for arbitrary N!

Unambiguous QCP



n=15

$$c_L^* = \sqrt{2} - 1$$

$$c^* = \frac{\sqrt{5} - 1}{2}$$

Anomaly detection | Malfunctioning sources

Translational symmetry (and more!) $\left\{ |\psi_k\rangle \equiv |0,\ldots,0,\underset{k}{\phi},0,\ldots,0\rangle \right\}_{k=1}^{N}$

Equal overlaps $c = \langle \phi | 0 \rangle = \cos\frac{\phi}{2}$

$$G = \begin{pmatrix} 1 & c^2 & \ldots & c^2 \\ c^2 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & c^2 \\ c^2 & \ldots & c^2 & 1 \end{pmatrix} = \frac{(1-c^2)}{N}\mathbb{1} + c^2\,|\mathbf{1}\rangle\langle\mathbf{1}|$$

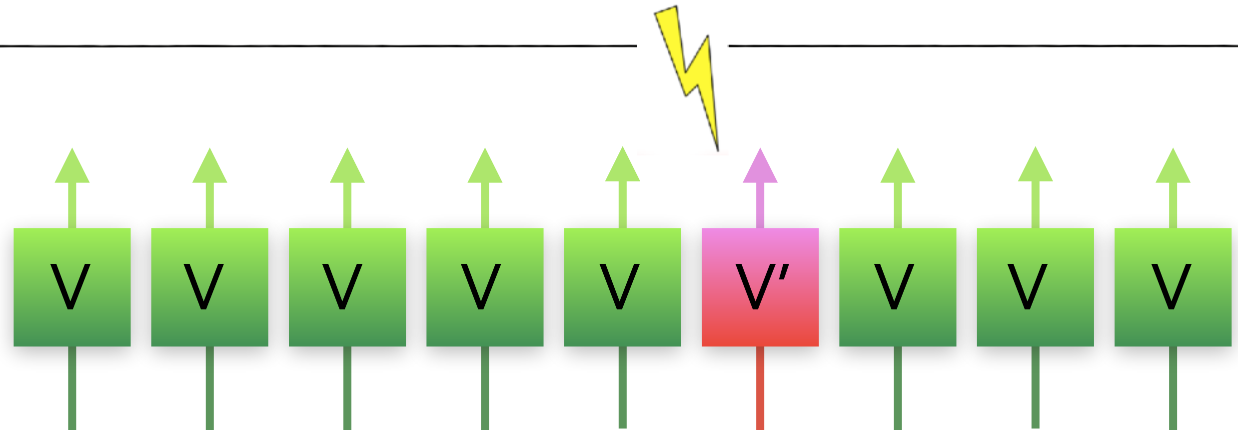
SRM is optimal if Gram matrix is circulant!

Anomaly detection | Malfunctioning sources

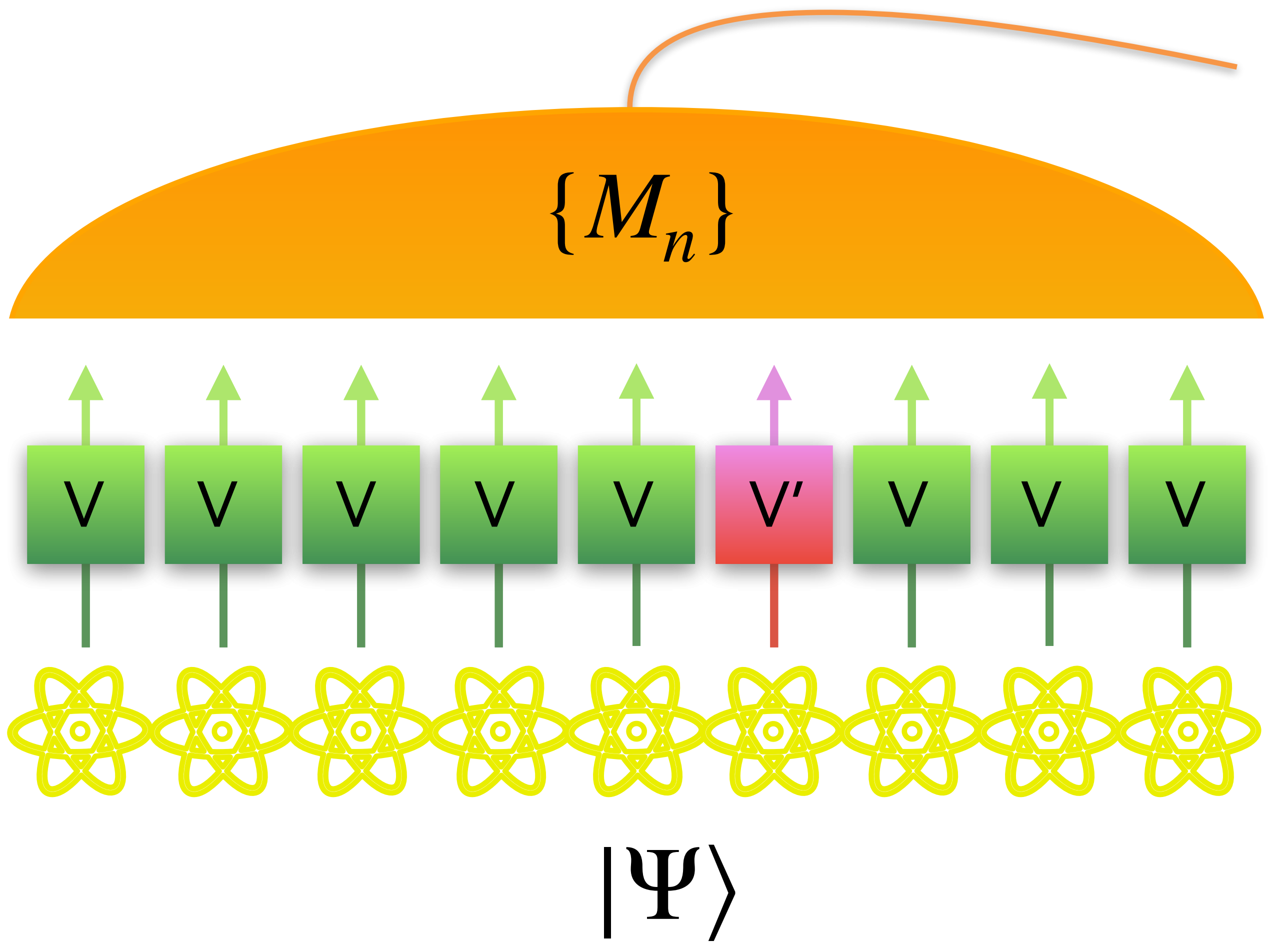
$$P_s = \left(\frac{1}{N} \sum_l \sqrt{\lambda_l} \right)^2 = \left(\frac{\sqrt{1 + (N-1)\cos^2\frac{\phi}{2}} + (N-1)\sin\frac{\phi}{2}}{N} \right)^2$$

$$\approx \sin^2\frac{\phi}{2} + \frac{\sin\phi}{\sqrt{N}} + \frac{\cos\phi}{N}$$

Anomaly detection | Malfunctioning gates



Anomaly detection | Malfunctioning gates



Equivalent to discrimination of N unitaries: $\{U_k = e^{i\phi\sigma_y^{(k)}}\}_{k=1}^{N}$

Anomaly detection | Malfunctioning gates

- Optimal P_S is no longer an SDP
- Observe that our problem still retains its translational symmetry
- Allows us to reduce our search for the optimal $\boldsymbol{\psi}$ to the permutationally invariant subspace of N qubits

$$|\Psi\rangle = \sum_{m=0}^{N} \sqrt{c_m} |N, m\rangle$$

$$|N, m\rangle = \binom{N}{m}^{-1/2} \sum_{g \in S_N} \pi_g \left(|1\rangle^{\otimes m} |0\rangle^{\otimes (N-m)} \right)$$

The Gram matrix is again circulant:

$$P_S(\phi) = \frac{1}{N^2}\left(\sqrt{1 + (N-1)\left(\sum_{m=0}^{N} c_m b_m(\phi)\right)} + (N-1)\sqrt{1 - \sum_{m=0}^{N} c_m b_m(\phi)}\right)^2$$

$$b_m(\phi) = \langle N, m | e^{\mathrm{i}\phi(\sigma_y^{(1)} - \sigma_y^{(2)})} | N, m\rangle = 1 - \frac{2m(N-m)(1-\cos\phi)}{N(N-1)}$$

- P_S is maximal if the overlaps $b_m(\phi)$ are minimal
- $0 \leq b_{N/2}(\phi) \leq b_m(\phi)$ for all angles in the range

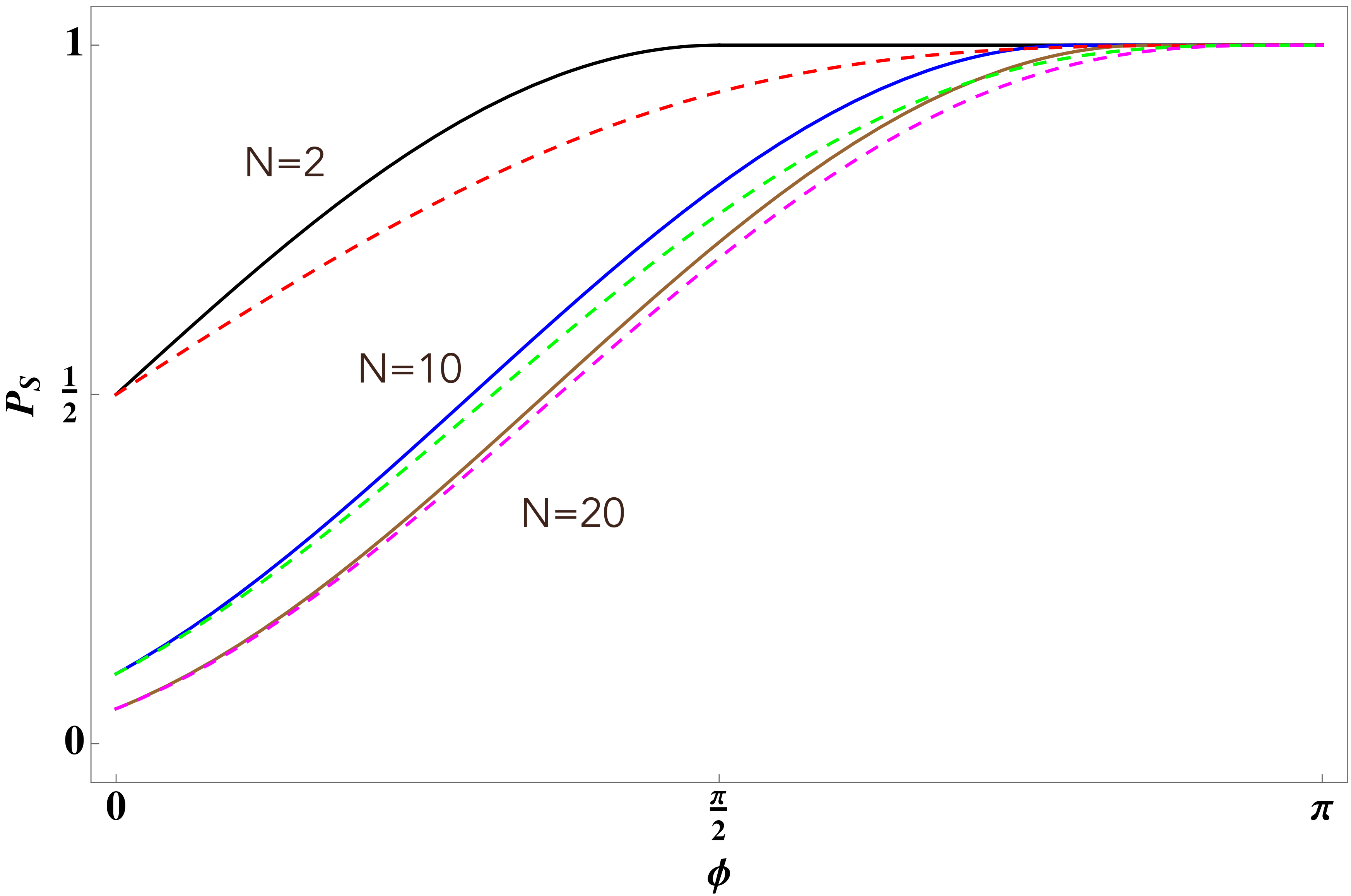
$$0 \leq \phi \leq \cos^{-1}\left(-1 + \frac{1}{\lceil \frac{N}{2} \rceil}\right) =: \phi^*$$

- Thus if $\phi \leq \phi^*$ the optimal quantum strategy uses the state $|\Psi\rangle = |N, \lceil \frac{N}{2} \rceil\rangle$ and the SQRM

$$P_S(\phi) = \sin^2 \frac{\phi}{2} + \sin\phi \sqrt{\frac{N-1}{N^2}} + \frac{\cos\phi}{N}$$

- If, however, $\phi > \phi^*$ $b_{N/2}(\phi) < 0$ and superpositions between symmetric states can help. In fact **$P_S(\phi)=1$** using the state

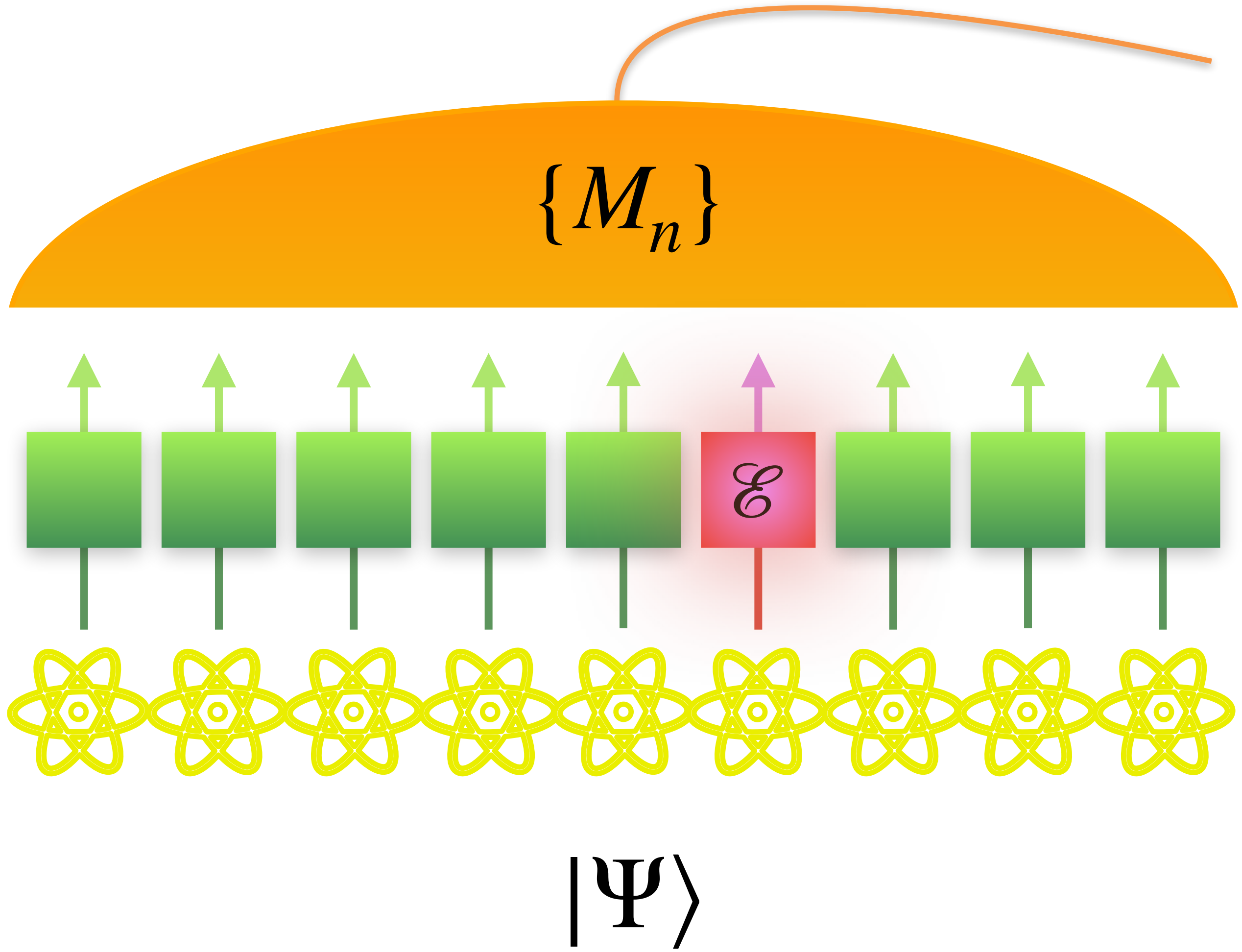
$$|\Psi\rangle = \sqrt{c_0}|N,0\rangle + \sqrt{c_{\lceil \frac{N}{2} \rceil}}|N, \lceil \frac{N}{2} \rceil\rangle \qquad c_{\lceil \frac{N}{2} \rceil} = \frac{N-1}{N\sin^2\frac{\phi}{2}}$$



In the limit of a large number of gates

$$P_S^{(gates)} = P_S^{(sources)} = \sin^2\frac{\phi}{2} + \frac{\sin\phi}{\sqrt{N}} + \frac{\cos\phi}{N}$$

Anomaly detection | Malfunctioning gates (noisy)



Anomaly detection | Malfunctioning gates (noisy)

- General upper and lower bounds.

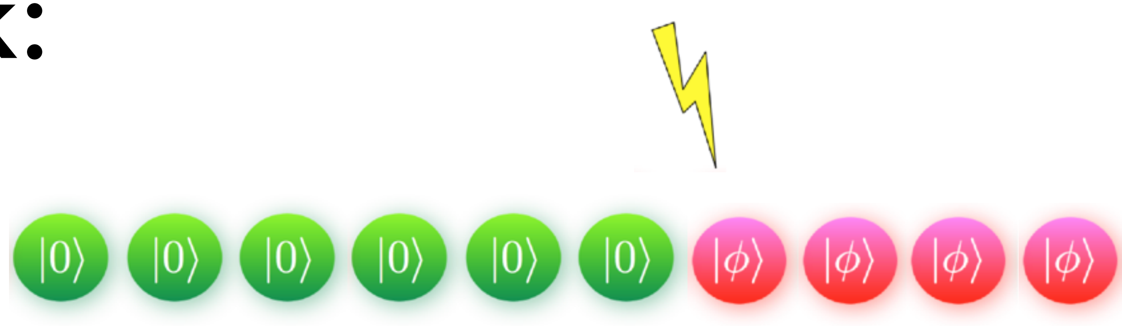
$$\mathcal{E}(\rho) = \sum_{l=1}^{m} K_l \, \rho \, K_l^{\dagger} \qquad P_S(\mathcal{E}) \leq \sum_{l=1}^{m} P_S(K_l)$$

- Upper bound attainable for rank 1&2 **Pauli channels** with product probe states.

- Closed expressions for **amplitud damping** channel

$$\gamma + \frac{1-\gamma}{N} \leq P_S \leq \gamma + \frac{\sqrt{1-\gamma}(\sqrt{1-\gamma}+1)}{2N} - \frac{\gamma}{4N^2} + \mathcal{O}(N^{-3})$$

outlook:



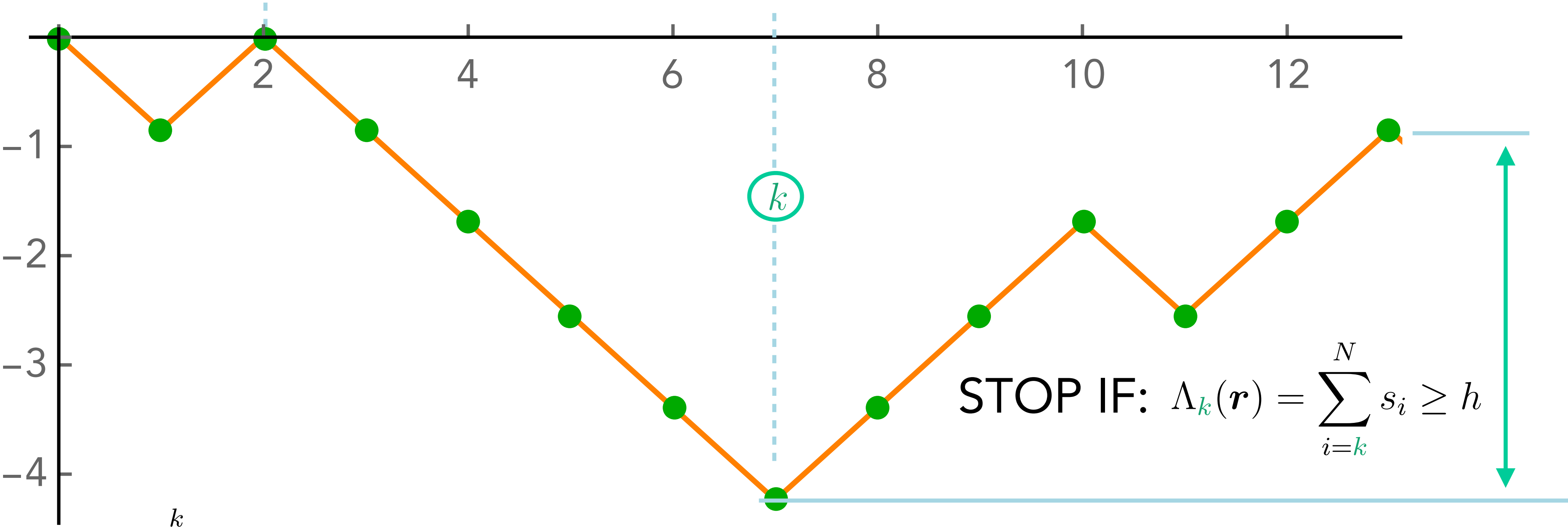

```
It is just an illusion here on Earth that
one moment follows another one, like beads
on a string, and that once a moment is
gone, it is gone forever.

Kurt Vonnegut, Slaughterhouse-Five
```

However, …we do live in this illusion, and often decisions have to be made before “moment is gone”.

Quickest change-point detection (CUSUM)

$r_i =$ 0 1 0 0 0 0 0 1 1 1 0 1 1



STOP IF: $\Lambda_k(\boldsymbol{r}) = \sum_{i=k}^{N} s_i \geq h$

$$S_k(\boldsymbol{r}) = \sum_{i=1}^{k} s_i$$

$$\langle s \rangle_c = \sum_r q(r) s(r) = \sum_r q(r) \log \frac{q(r)}{p(r)} = D(q||p)$$

QUSUM

$$\tau_d = \frac{h}{D(\rho||\sigma)} = \frac{\log T_{\mathrm{FA}}}{D(\rho||\sigma)}$$

$$\tau_d = \frac{h}{D(p||q)} = \frac{\log T_{\mathrm{FA}}}{D(p||q)}$$

work in progress!

Summary | Köszönöm!

Grup d'Informació Quàntica >

- Quantum Change-point and Anomaly detection: novel quantum task
- QCP collective outperform local measurements, not true for QAD
- Analytic asymptotic results.
- Exact unambiguous CP identification for arbitrary N : one of the few multi-hypothesis discrimination instances with exact solution
- Many possible extensions:

⚒ Unknown states/channels

⚒ Continuous variables & mixed states

⚒ Channel CP

⚒ Multiple CPs

⚒ **QUSUM**: Quantum CUSUM (**sequential problems**):

false alarm time vs minimum detection time

arXiv:1808.02729

PRL 117, 150502 (2016)

PRL 119, 140506 (2017)