A Polynomial-Time Approximation Scheme for Sequential Batch Testing of Series Systems

Danny Segev
Department of Statistics, Haifa

joint work with Yaron Shaposhnik (Rochester)

OUTLINE

MODEL DESCRIPTION

KNOWNS & UNKNOWNS

MAIN RESULT

TECHNICAL OVERVIEW

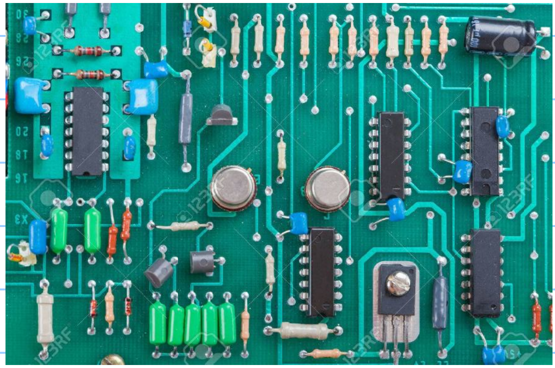
MODEL DESCRIPTION

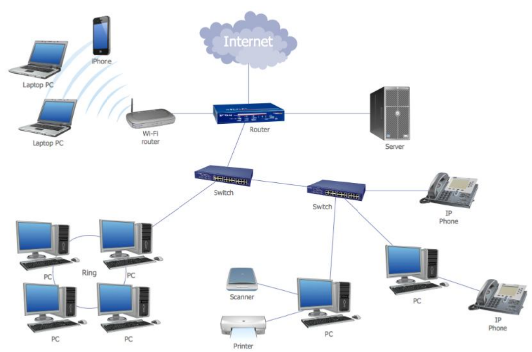
SEQUENTIAL BATCH TESTING

  . . . .  . . . . 

component | 1 | 2 | . . . . | i | . . . . | n
operational? | $X_1 \sim B(p_1)$ | $X_2 \sim B(p_2)$ | . . . | $X_i \sim B(p_i)$ | . . . . | $X_n \sim B(p_n)$

wish to decide whether $\prod_{i=1}^{n} X_i = 1$ or not



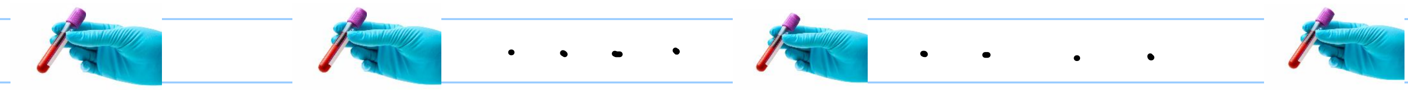


SEQUENTIAL BATCH TESTING

  . . . .  . . . . 

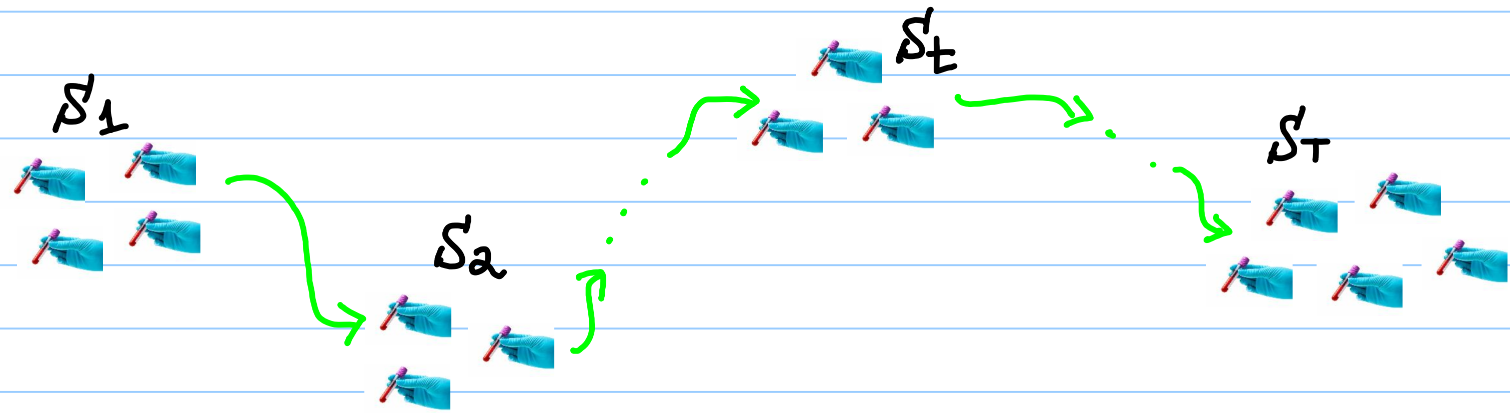
component	1	2	$\cdots$	i	$\cdots$	n
operational?	$X_1 \sim B(p_1)$	$X_2 \sim B(p_2)$	$\cdots$	$X_i \sim B(p_i)$	$\cdots$	$X_n \sim B(p_n)$
testing cost	C_1	C_2	$\cdots$	C_i	$\cdots$	C_n
batch cost	β					

SEQUENTIAL BATCH TESTING

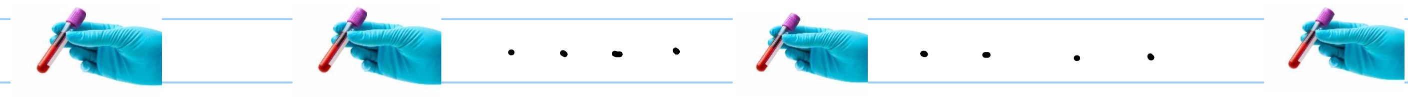


component	1	2	. . . .	i	. . . .	n
operational?	$X_1 \sim B(p_1)$	$X_2 \sim B(p_2)$	. . .	$X_i \sim B(p_i)$	. . . .	$X_n \sim B(p_n)$
testing cost	c_1	c_2	. . . .	c_i	. . . . .	c_n
batch cost	β					

sequential testing scheme



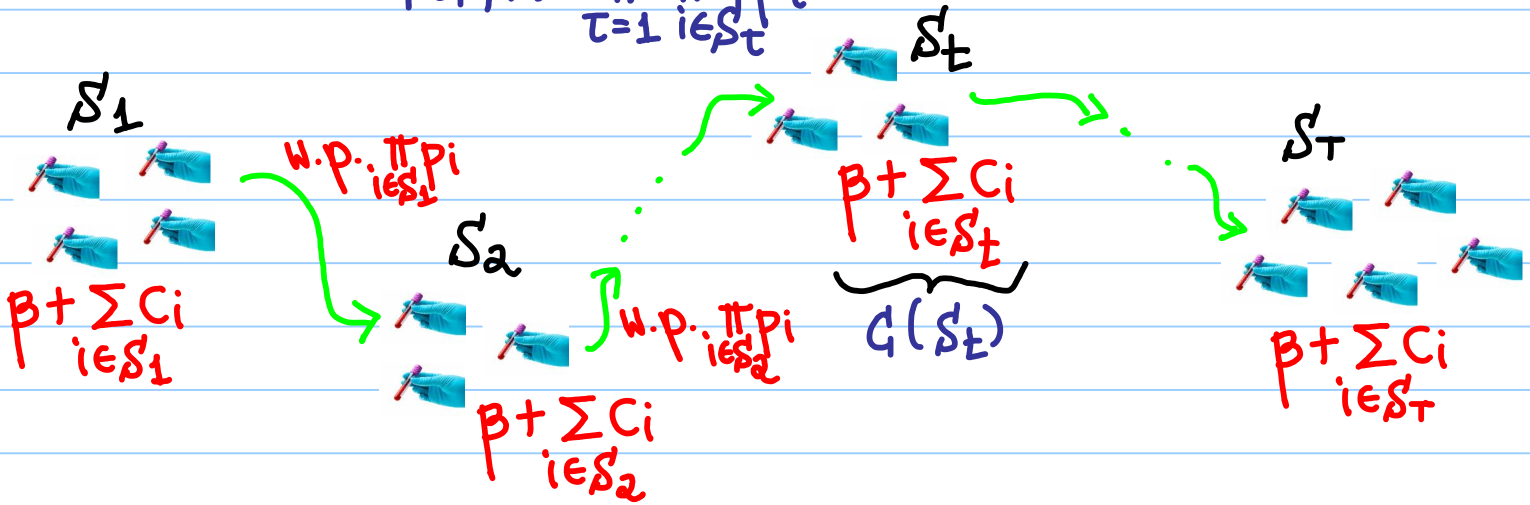
SEQUENTIAL BATCH TESTING



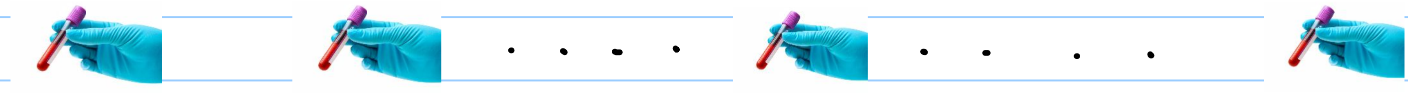
component	1	2	. . . .	i	. . . .	n
operational?	$X_1 \sim B(p_1)$	$X_2 \sim B(p_2)$	. . .	$X_i \sim B(p_i)$	. . . .	$X_n \sim B(p_n)$
testing cost	c_1	c_2	. . . .	c_i	. . . . .	c_n
batch cost	β					

$$\phi(S,t) = \prod_{\tau=1}^{t-1} \prod_{i \in S_\tau} p_i$$

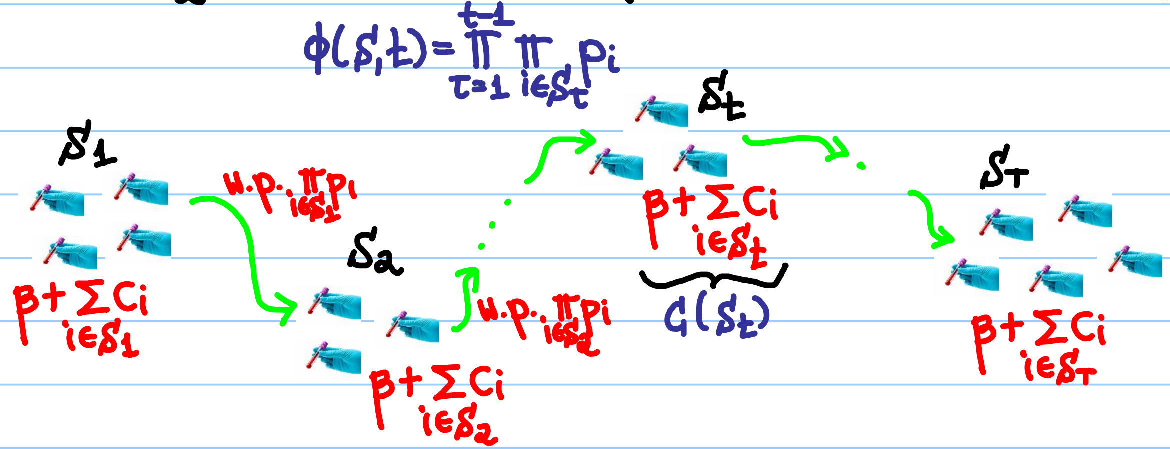
sequential testing scheme



SEQUENTIAL BATCH TESTING



component	1	2	$\cdots$	i	$\cdots$	n
operational?	$X_1 \sim B(p_1)$	$X_2 \sim B(p_2)$	$\cdots$	$X_i \sim B(p_i)$	$\cdots$	$X_n \sim B(p_n)$
testing cost	c_1	c_2	$\cdots$	c_i	$\cdots$	c_n
batch cost	β					



Objective: compute a testing scheme $S=(S_1,\ldots,S_T)$ whose expected cost $E(S)=\sum_{t=1}^{T}\phi(S,t)\cdot C(S_t)$ is minimized

KNOWNS & UNKNOWNS

$\beta = 0$

$\Downarrow$

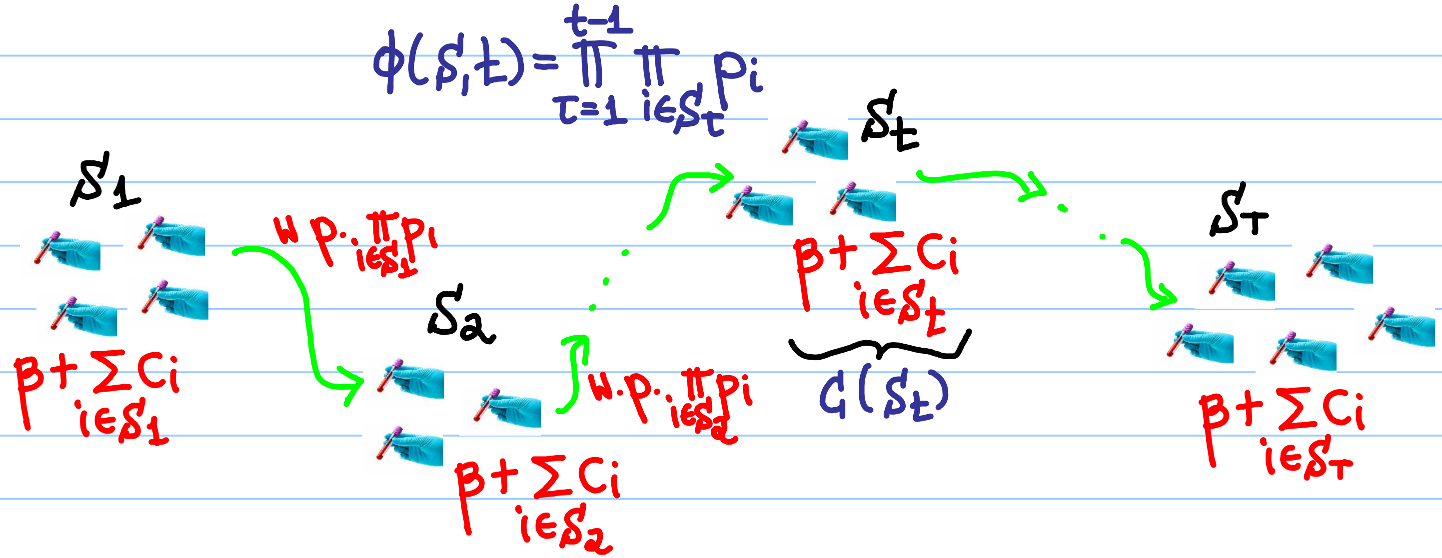
THE CLASSIC SEQUENTIAL TESTING PROBLEM

- optimal sequence: test in non-decreasing order of $\frac{c_i}{1-p_i}$
  [Mitten '60] [Butterworth '72] [Simon & Kadane '75] [Natarajan '86]

- applications: medical diagnostic procedures, quality inspection, project scheduling, telecommunications, screening employees for positions, and artificial intelligence [Greiner et al. '06] [Dufuaa & Raouf '90] [Qiu et al. '92] [De Reyck & Leus '92] [Cox et al. '89] [Garey '73] [Nilson '71] [Simon & Kadane '75] [Smith '89]

- extensions and variants: precedence constraints, k-out-of-n systems, series-parallel systems, threshold functions, certain DNF formulas, and general distributions of $X_1, \ldots, X_n$ [Garey '73] [De Reyck & Leus '92] [Berend et al. '14] [Chang et al. '90] [Ben-Dov '81] [Boros & Ünlüyurt '00] [Greiner et al '06] [Deshpande et al. '14] [Allen et al. '17] [Kaplan et al. '05]

SEQUENTIAL BATCH TESTING



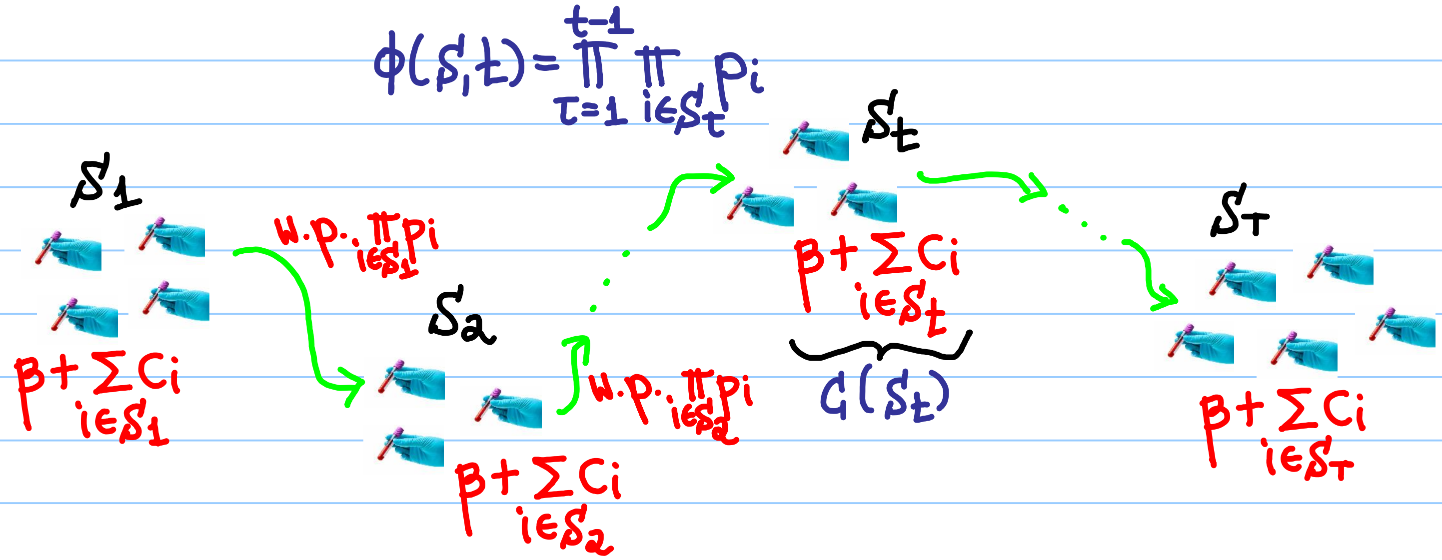
- [Daldal et al.'15]: introduced batch testing, mostly experimental results

Segev
- [Daldal et al.'16]: $(6.829+\varepsilon)$-approximation, ε-optimal IP formulation, NP-hardness for $T=2$ subsets, numerical experiments

- MOTIVATING OPEN QUESTION: improved approximation guarantees?

MAIN RESULT

MAIN RESULT



theorem: the sequential batch testing problem can be approximated within factor $1+\varepsilon$ in time $O\left((|I|/\varepsilon)^{O(1/\varepsilon^3)}\right)$

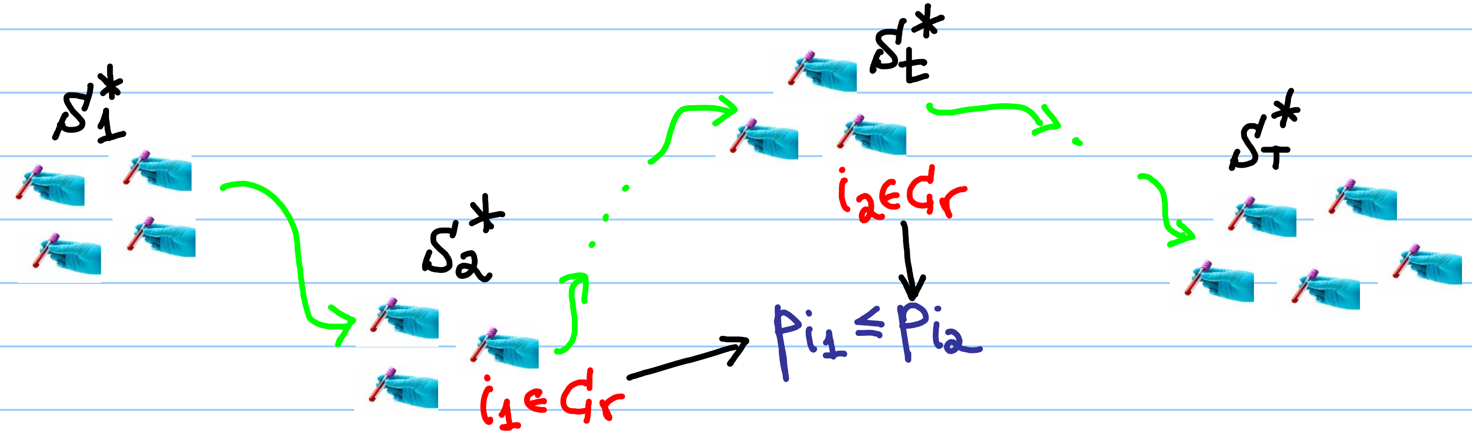
TECHNICAL OVERVIEW

PRELIMINARIES

- WLOG, $\min_i c_i = 1$
- round up each testing cost c_i to nearest power of $1+\varepsilon$
- cost class $G_r = \{ i \in [n] : c_i = (1+\varepsilon)^r \}$ $\quad 0 \leq r \leq R = O\left(\frac{1}{\varepsilon} \log \frac{c_{max}}{c_{min}}\right)$

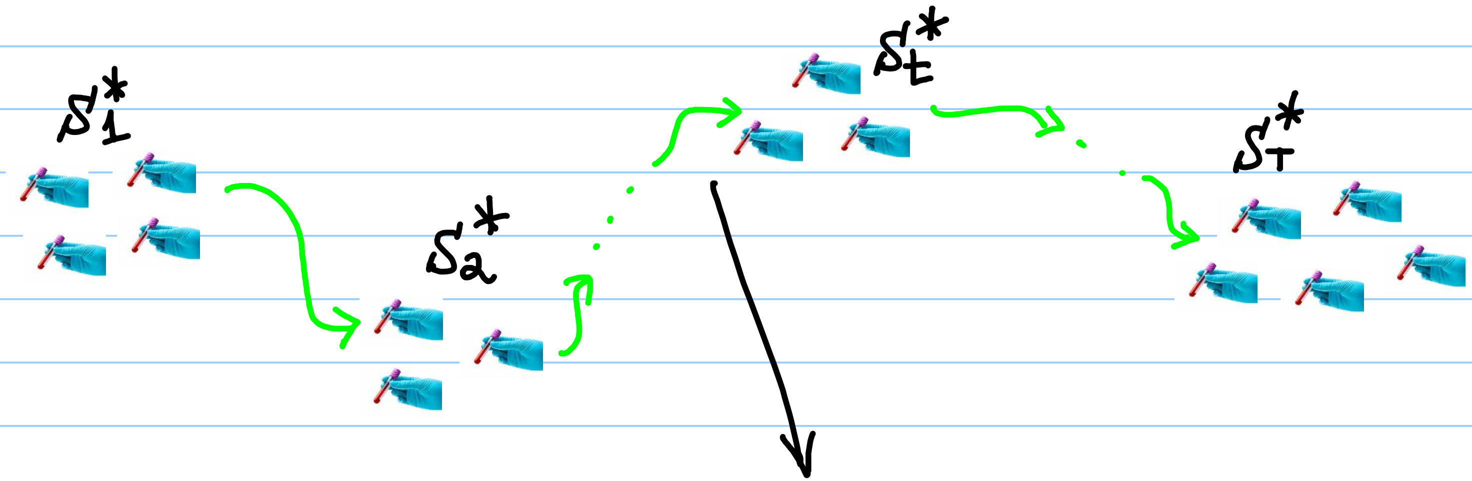
PRELIMINARIES

- WLOG, $\min_i c_i = 1$
- round up each testing cost c_i to nearest power of $1+\varepsilon$
- cost class $G_r = \{ i \in [n] : c_i = (1+\varepsilon)^r \}$ $\qquad 0 \leq r \leq R = O\left(\frac{1}{\varepsilon} \log \frac{c_{max}}{c_{min}}\right)$
- greedy-within-class property: there exists an optimal testing scheme S^* where, for every r, the G_r-variables appear by non-decreasing probabilities



SYSTEM STATES

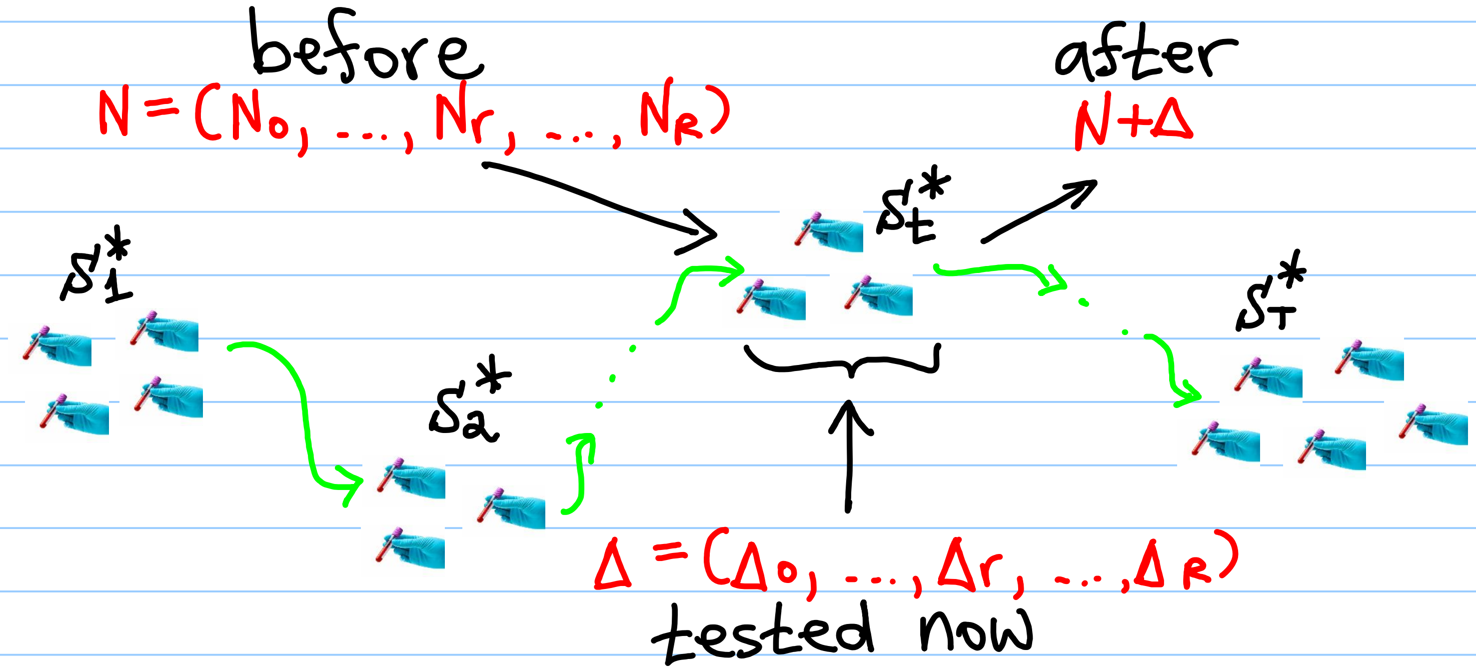
- greedy-within-class property: there exists an optimal testing scheme S^* where, for every r, the G_r-variables appear by non-decreasing probabilities



$$N = (N_0, \ldots, N_r, \ldots, N_R)$$

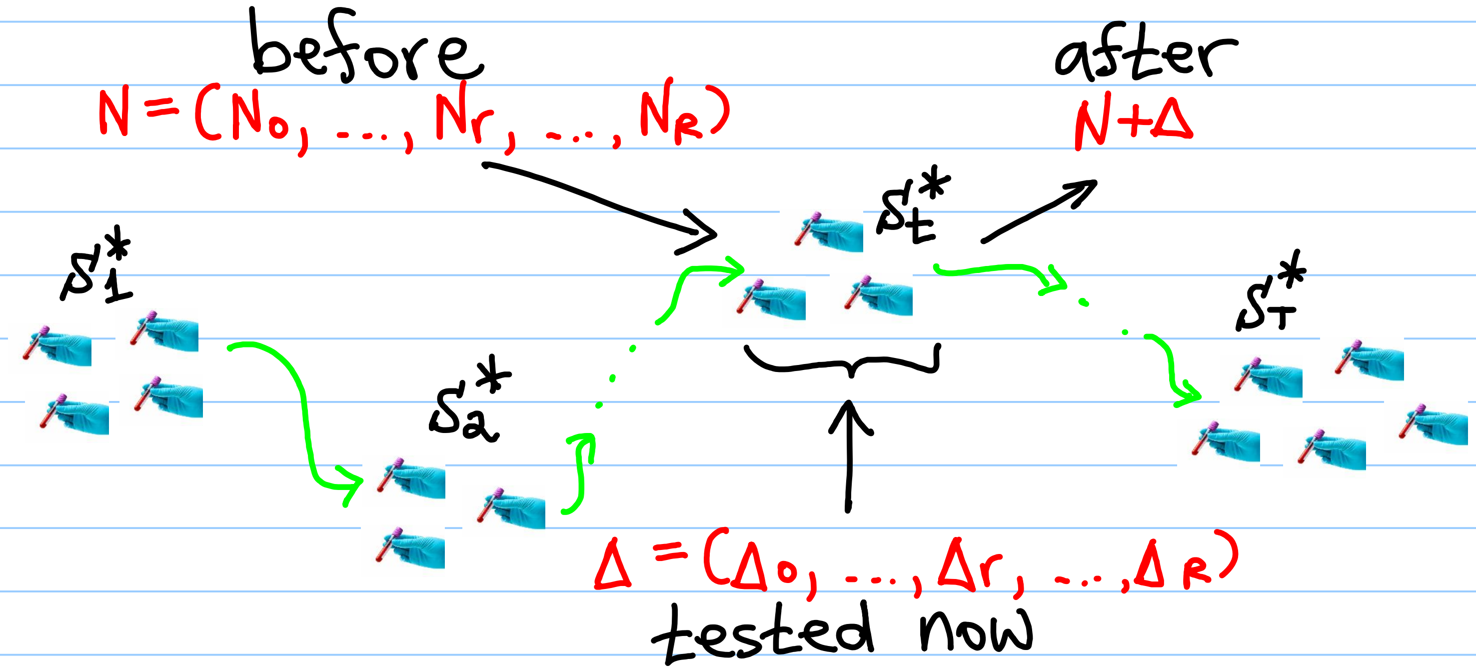
- N_r = number of G_r-variables that have already been tested
- F = collection of all feasible states

THE NAIVE DP



before
$N = (N_0, \ldots, N_r, \ldots, N_R)$
after
$N+\Delta$
S_t^*
S_1^*
S_2^*
S_T^*
$\Delta = (\Delta_0, \ldots, \Delta_r, \ldots, \Delta_R)$
tested now

THE NAIVE DP



$$OPT_F(N) = \min_{\Delta : N+\Delta \in F} \left\{ \underbrace{\beta + \sum_r \Delta_r \cdot (1+\varepsilon)^r}_{\text{immediate cost}} + \overbrace{\left(\prod_r \prod_{i \in G_r[N_r+1,\, N_r+\Delta_r]} p_i\right)}^{\text{probability to move forward}} \cdot \underbrace{OPT_F(N+\Delta)}_{\text{future cost}} \right\}$$

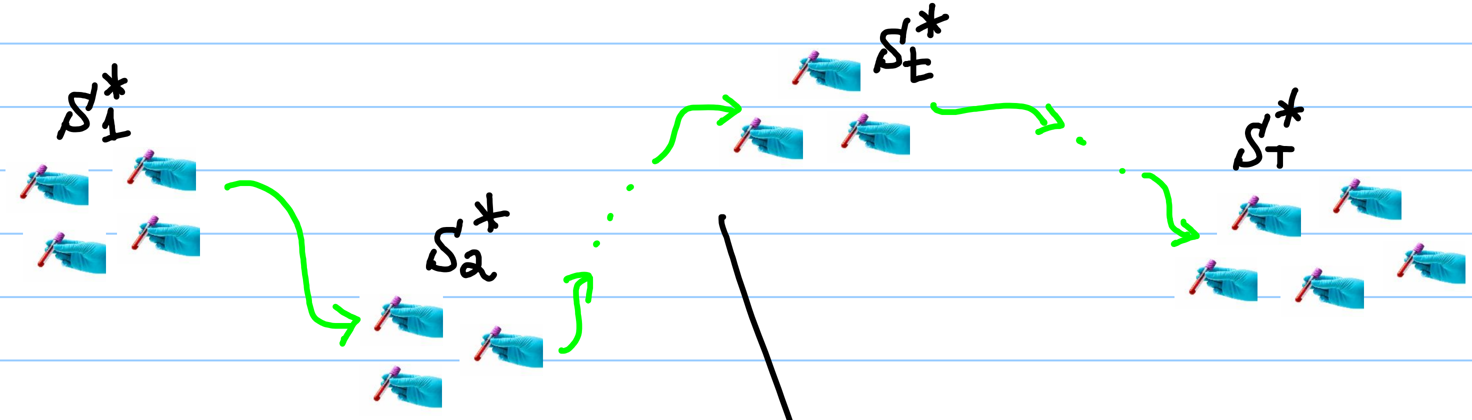
STATE-SPACE COLLAPSE?

$$OPT_F(N) = \min_{\Delta : N+\Delta \in F} \Big\{ \underbrace{\beta + \sum_r \Delta_r \cdot (1+\varepsilon)^r}_{\text{immediate cost}} + \overbrace{\Big(\prod_r \prod_{i \in G_r[N_r+1,\, N_r+\Delta_r]} p_i \Big)}^{\text{probability to move forward}} \cdot \underbrace{OPT_F(N+\Delta)}_{\text{future cost}} \Big\}$$

- but $|F| = O(n^{O(R)})$! 
- QUESTION: can we find a subspace $\tilde{F} \subseteq F$ such that

$$\underbrace{|\tilde{F}| << |F|}_{\text{efficiency}} \quad \text{and} \quad \underbrace{OPT_{\tilde{F}}(0) \approx OPT_F(0)}_{\text{near-optimality}} \;?$$

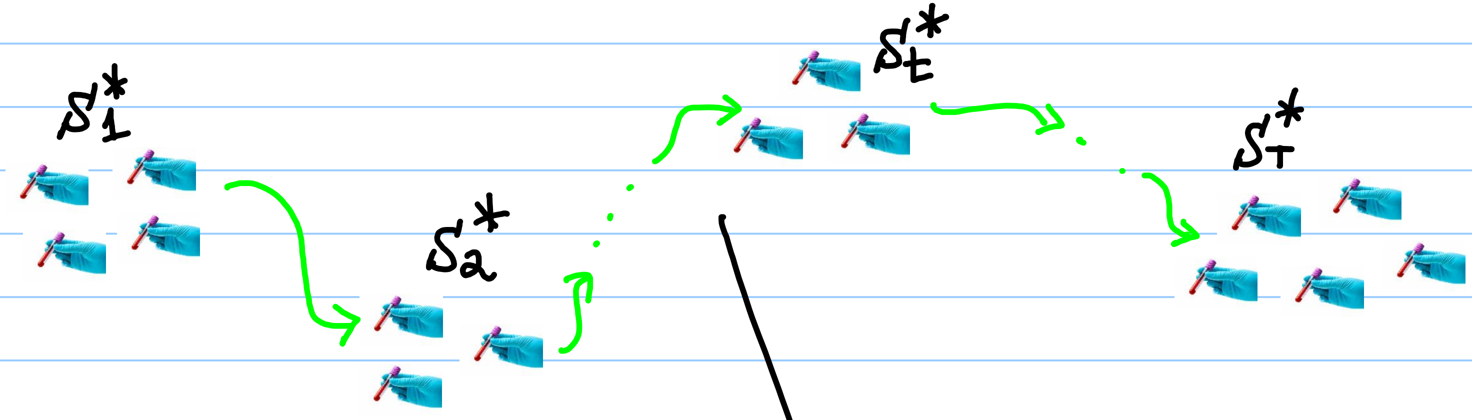
GOOD STATES $\Rightarrow$ QPTAS



$K = \lceil \log_{1+\varepsilon}(\frac{n}{\varepsilon}) \rceil$

- good states: $N = (\underbrace{|G_1|, \dots, |G_{\ell-K}|}_{\text{exhausted}}, \overbrace{\underbrace{N_{\ell-K+1}, \dots, N_\ell}_{\text{any value}}}^{\leftarrow K \text{ classes} \rightarrow}, \underbrace{0, \dots, 0}_{\text{untouched}})$

GOOD STATES ⇒ QPTAS



$K = \lceil \log_{1+\varepsilon}(\frac{n}{\varepsilon}) \rceil$

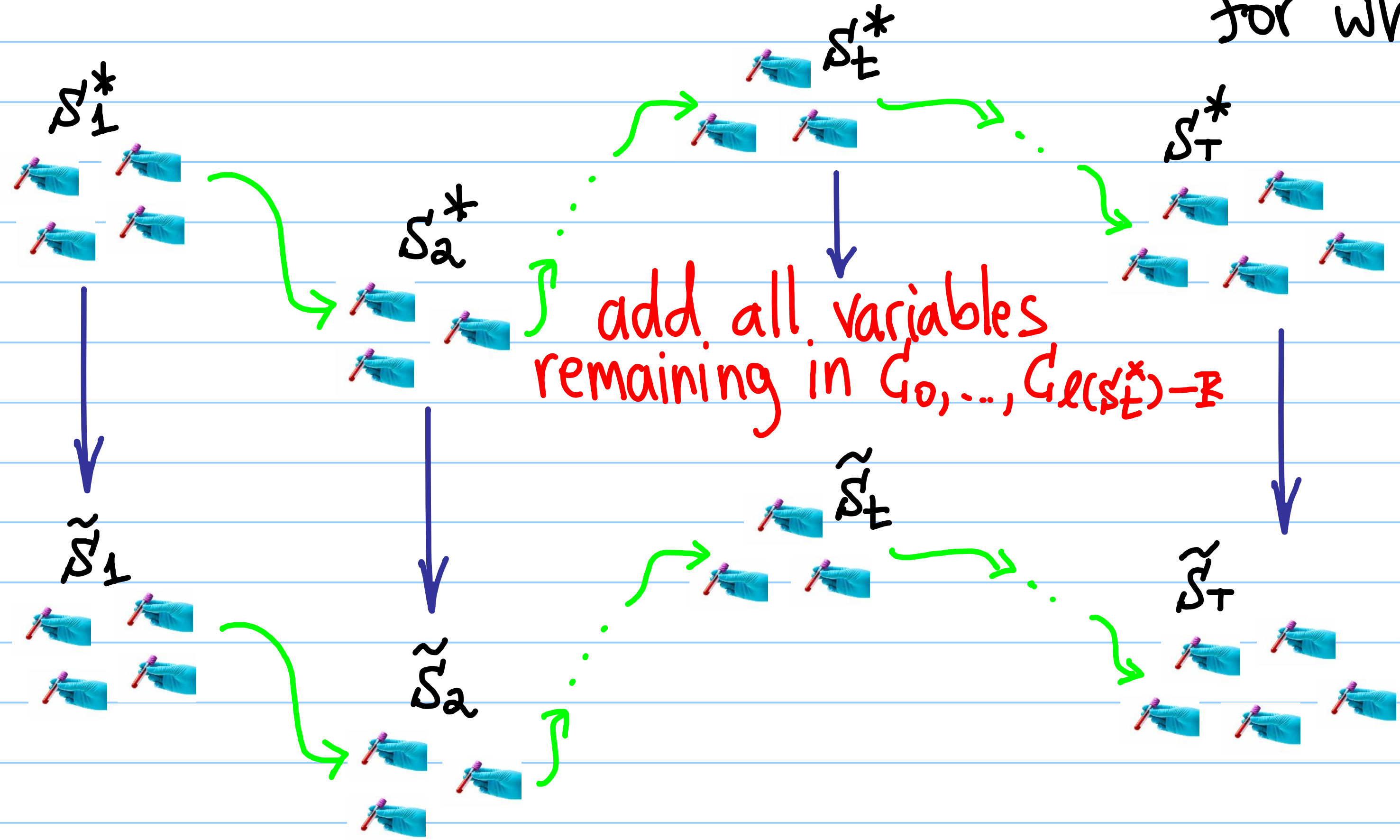
← K classes → 1,1

- good states: $N = (\underbrace{|G_1|, \dots, |G_{\ell-K}|}_{\text{exhausted}}, \underbrace{N_{\ell-K+1}, \dots, N_\ell}_{\text{any value}}, \underbrace{0, \dots, 0}_{\text{untouched}})$

- $|F_{good}| = O(n^{O(K)})$, so DP runs in quasi-polynomial time 

GOOD STATES $\Rightarrow$ QPTAS

$\ell(S_t^*)$ = maximal r for which $S_t^* \cap G_r \neq \emptyset$

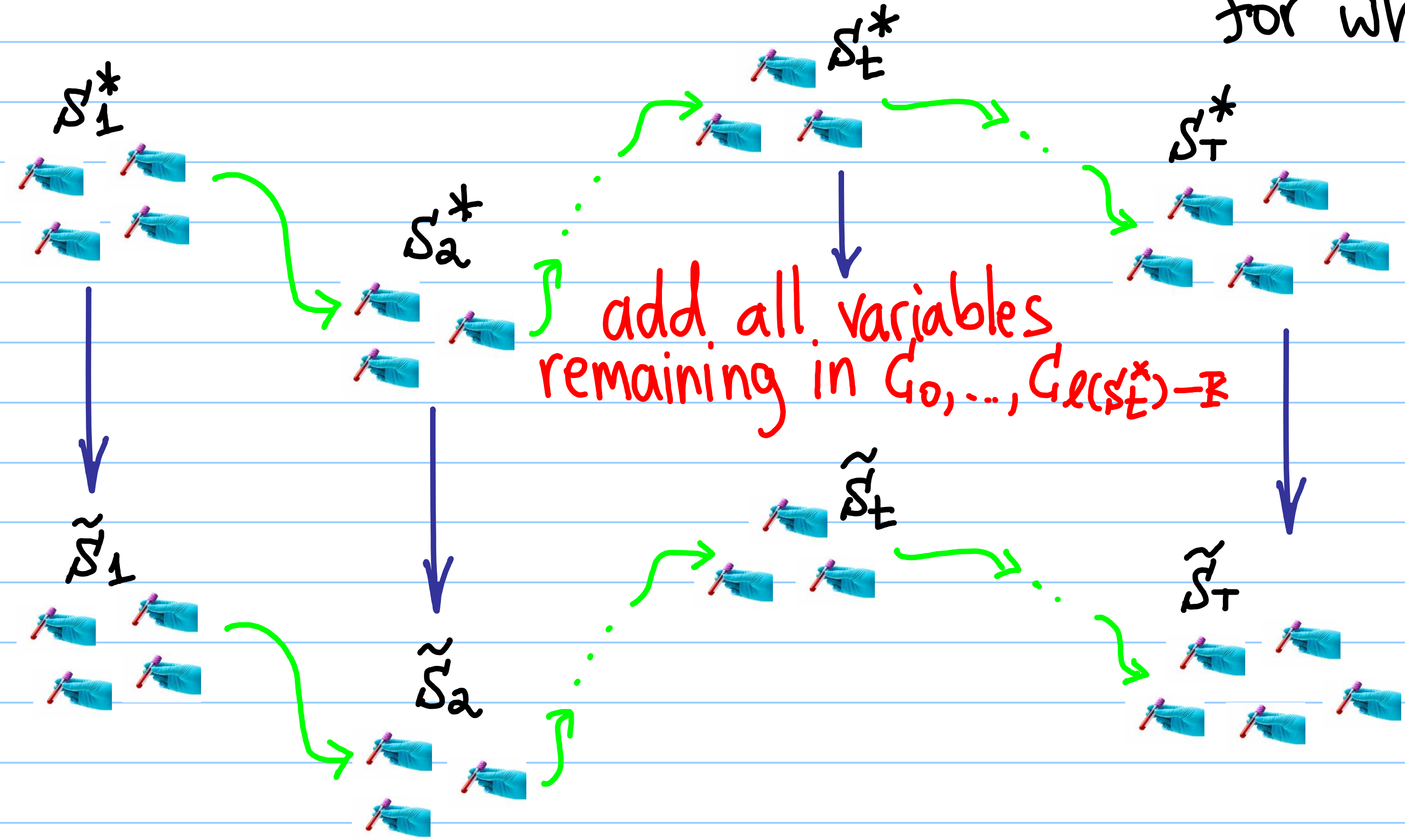


add all variables remaining in $G_0, \ldots, G_{\ell(S_t^*)-K}$

$K = \lceil \log_{1+\varepsilon}(\frac{n}{\varepsilon}) \rceil$

GOOD STATES $\Rightarrow$ QPTAS

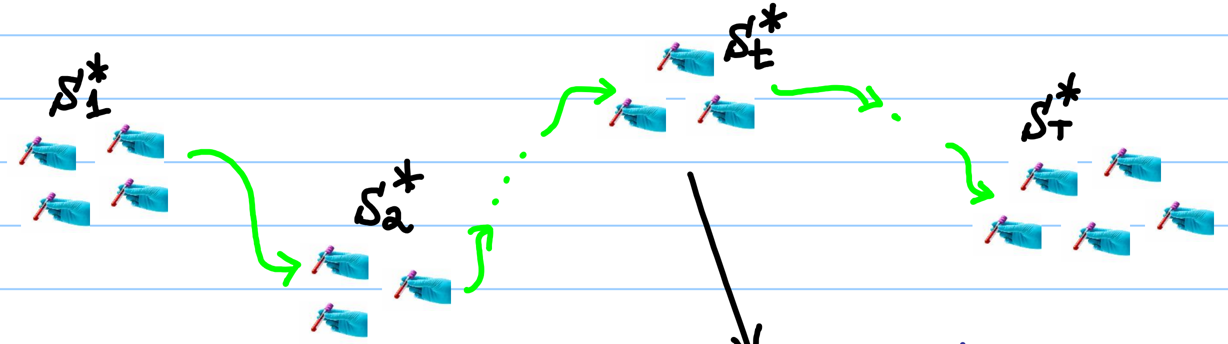
$\ell(S_t^*)$ = maximal r for which $S_t^* \cap G_r \neq \emptyset$



$K = \lceil \log_{1+\varepsilon}(\frac{n}{\varepsilon}) \rceil$

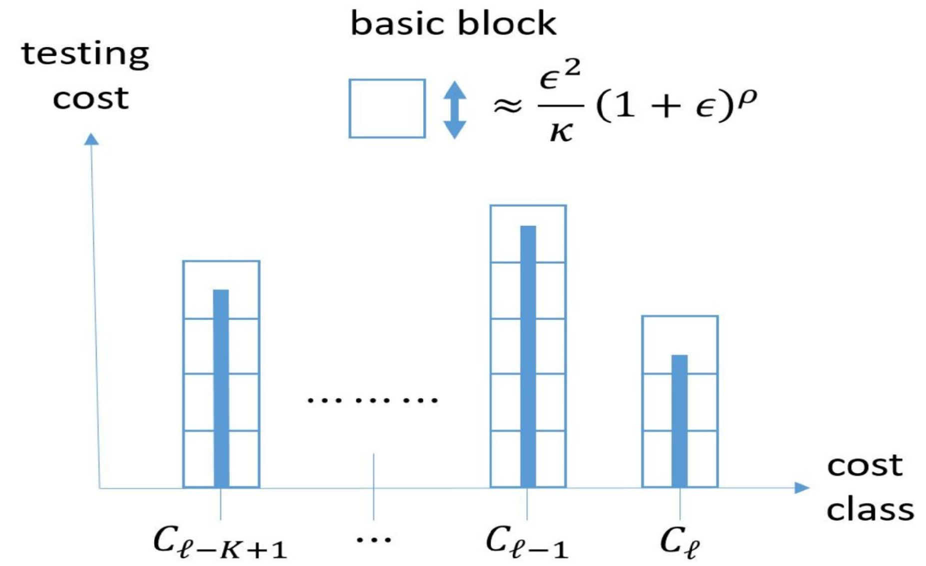
- immediate cost: $\beta + \sum_r \Delta_r \cdot (1+\varepsilon)^r$ blows up by at most $1+\varepsilon$
- probability to move forward: $\left(\prod_r \prod_{i \in G_r[N_r+1, N_r+\Delta_r]} p_i\right)$ can only decrease

GREAT STATES ⇒ QPTAS

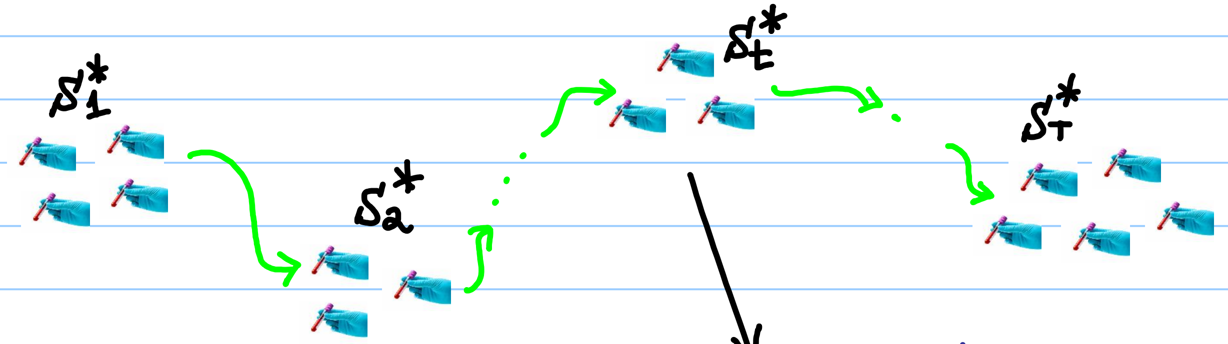


$$K = \lceil \log_{1+\varepsilon}\left(\tfrac{n}{\varepsilon}\right) \rceil$$

- great states: $N = (\underbrace{|G_1|, \dots, |G_{\ell-K}|}_{\text{exhausted}}, \overbrace{\underbrace{N_{\ell-K+1}, \dots, N_\ell}}^{\leftarrow K \text{ classes} \rightarrow}, \underbrace{0, \dots, 0}_{\text{untouched}})$



GREAT STATES ⇒ QPTAS

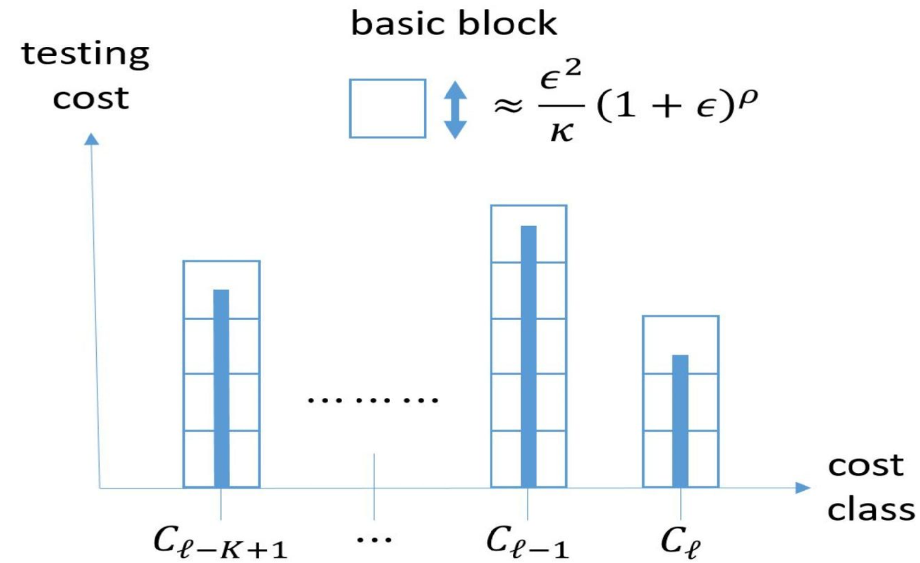


$K = \lceil \log_{1+\varepsilon}(\frac{n}{\varepsilon}) \rceil$

- great states: $N = (\underbrace{|G_1|, \dots, |G_{\ell-K}|}_{\text{exhausted}}, \overbrace{N_{\ell-K+1}, \dots, N_\ell}^{\leftarrow K \text{ classes} \rightarrow}, \underbrace{0, \dots, 0}_{\text{untouched}})$

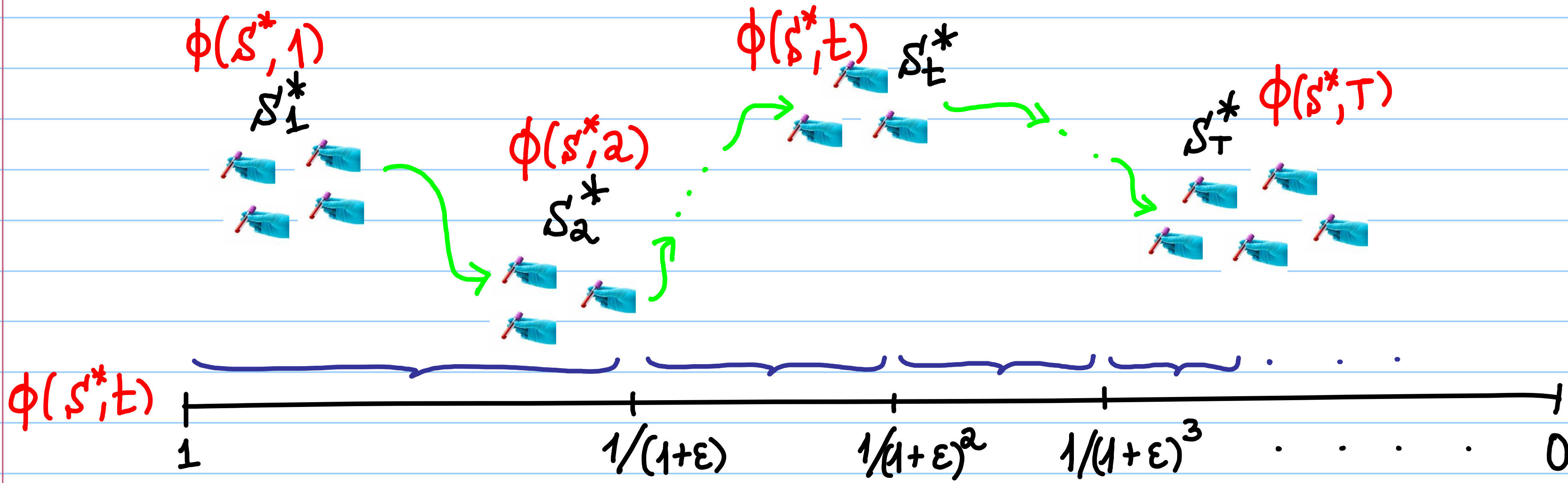
- $|F_{great}| = O(e^{O(K/\varepsilon^2)}) = O((n/\varepsilon)^{O(1/\varepsilon^3)})$,

so DP runs in polynomial time



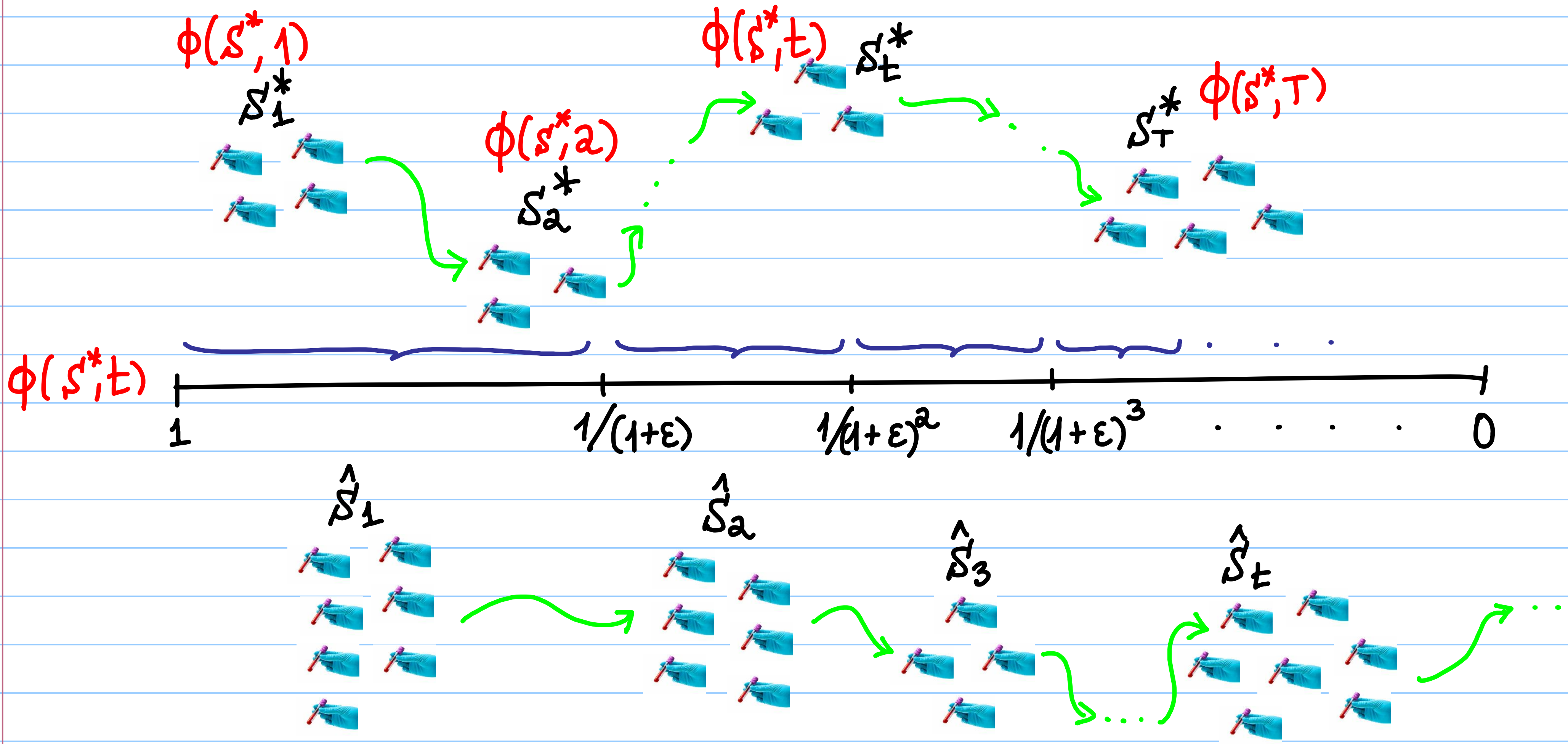
GREAT STATES $\Rightarrow$ QPTAS

- clustering of s^* by $\phi(s^*,\cdot)$-values:



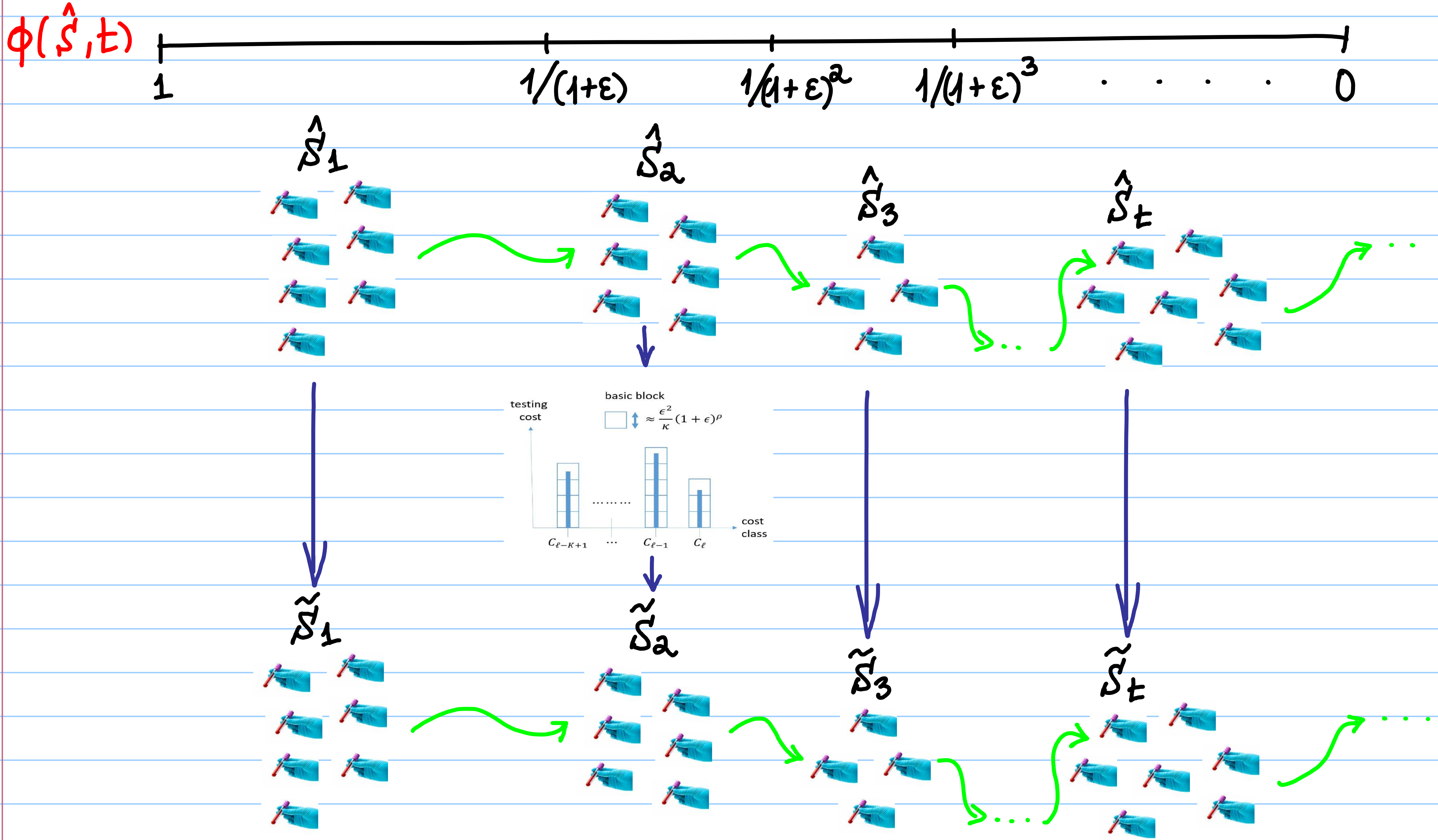
GREAT STATES $\Rightarrow$ QPTAS

- clustering of S^* by $\phi(S^*,\cdot)$-values:

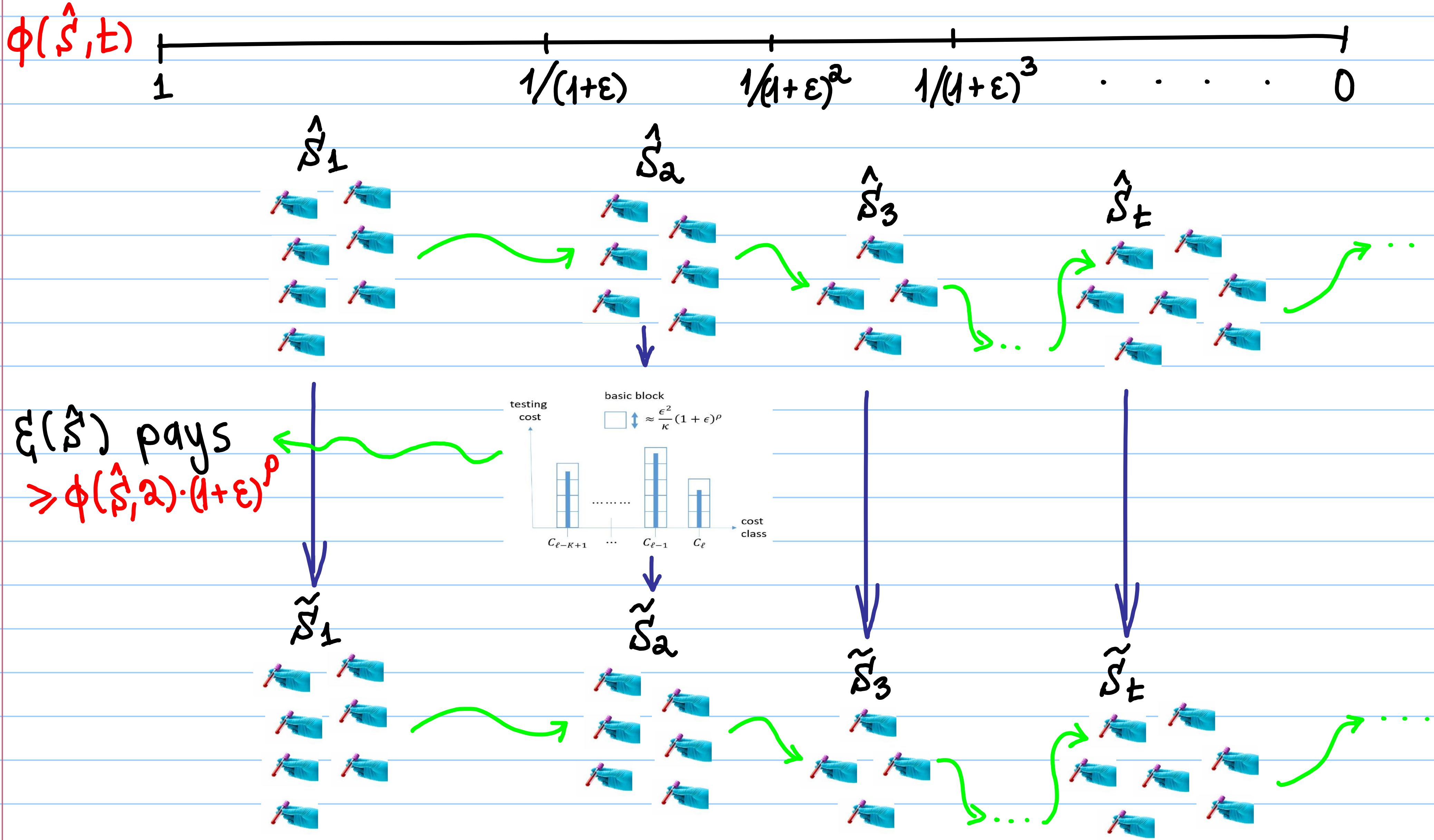


- at most $1+\varepsilon$ blow-up in cost: $\mathcal{E}(\hat{S}) \le (1+\varepsilon)\cdot\mathcal{E}(S^*)$

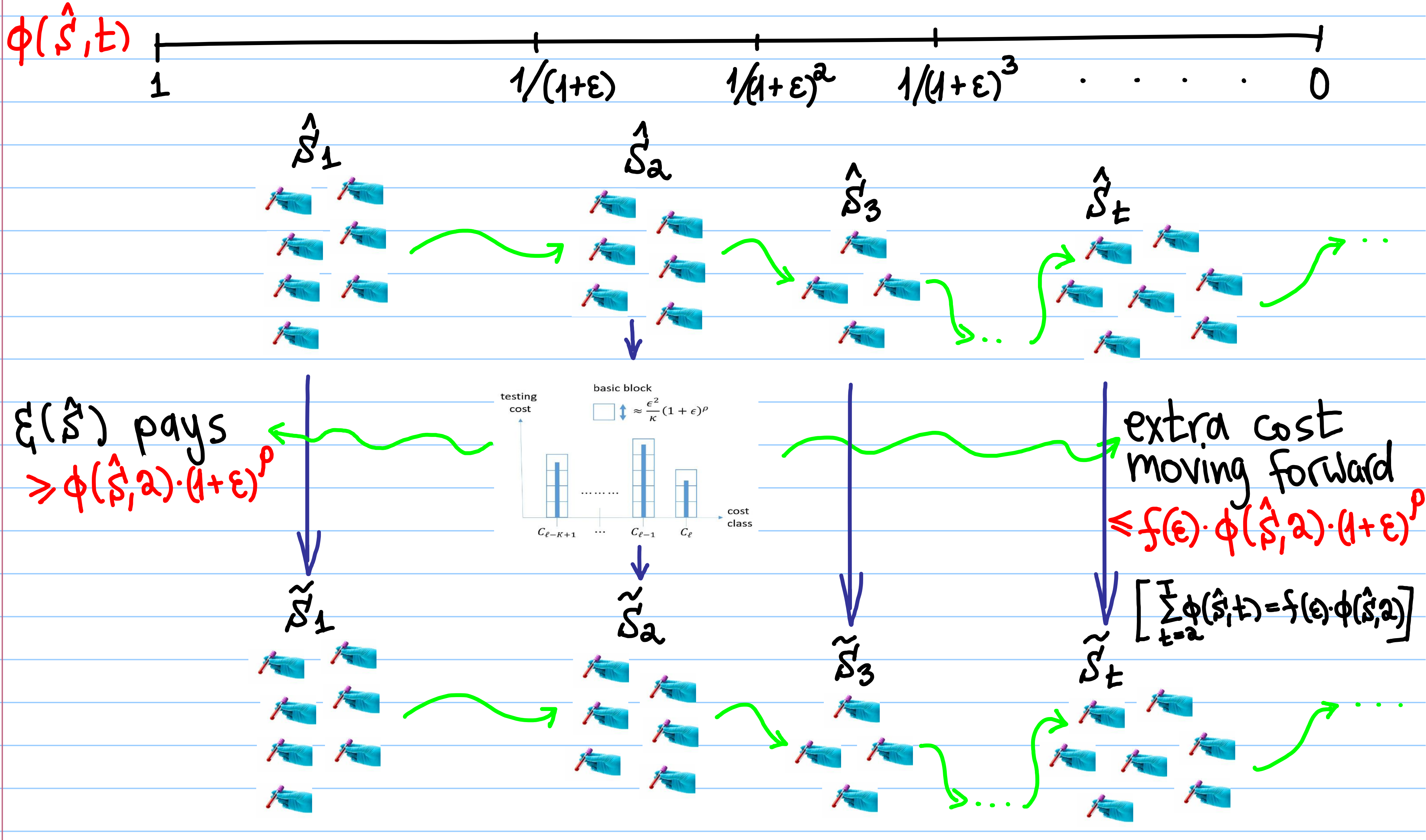
GREAT STATES ⇒ QPTAS



GREAT STATES ⇒ QPTAS



GREAT STATES $\Rightarrow$ QPTAS



CONCLUDING REMARKS

CONCLUDING REMARKS

- numerical experiments
- hardness with arbitrary number of subsets?
- applications of technical ideas in other settings?

CONCLUDING REMARKS

- numerical experiments
- hardness with arbitrary number of subsets?
- applications of technical ideas in other settings?

THANK YOU!