Reliable Cost Predictions for Finding Optimal Solutions to LABS Problem: Evolutionary and Alternative Algorithms

F. Brglez, X.Y. Li, M. F. Stallmann, B. Militzer*

Dept. of Computer Science, NC State University
Raleigh, North Carolina, USA

*Carnegie Institution of Washington,
Washington, DC 20015, USA

Outline

- LABS Problem Background
- Our SLS Heuristic
- Two SLS Solvers Under Test
  - ES-solver
  - KL-solver
- Experimental Methodology
  - Asymptotic Termination-A Experiments
  - Asymptotic Termination-A Experiments
  - Reliability of Termination-B (KL-solver)
- Conclusions & Future Work

About the LABS Problem

LABS (low auto-correlation binary sequence) problem:

… find a binary sequence of length L with the minimum autocorrelation energy or maximum <u>merit factor</u>

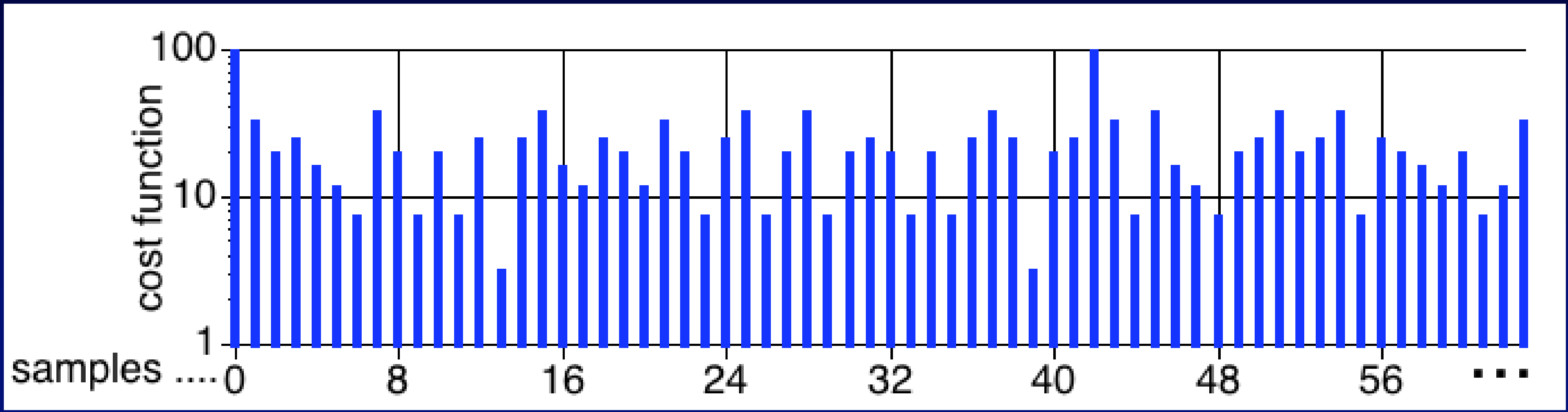
L-bit binary sequence $S = s_1 s_2 s_3 \dots s_L$

autocorrelations $C_k(S) = \text{sum}\{s_i s_{i+k}\}$ $k=1, \dots, L-1$

cost function $E(S) = \text{sum}\{C_k(S) * C_k(S)\}$

merit factor $F(S) = L*L/2E(S)$

LABS Problem Landscape for L = 7



- 4 global minima @ cost = 3
- 24 local minima @ cost = 7
- 12 local minima @ cost = 11
- ....
- 4 global maxima @ cost = 91 [= L*(L-1)*(2L-1)/6]

No more than 4 global minima observed for L = 7, 11, 13, 21, 27, 41, 43, 45, 51, 57, 63, ...

(... much harder to solve, compared to adj. problem sizes!)

LABS Problem Background

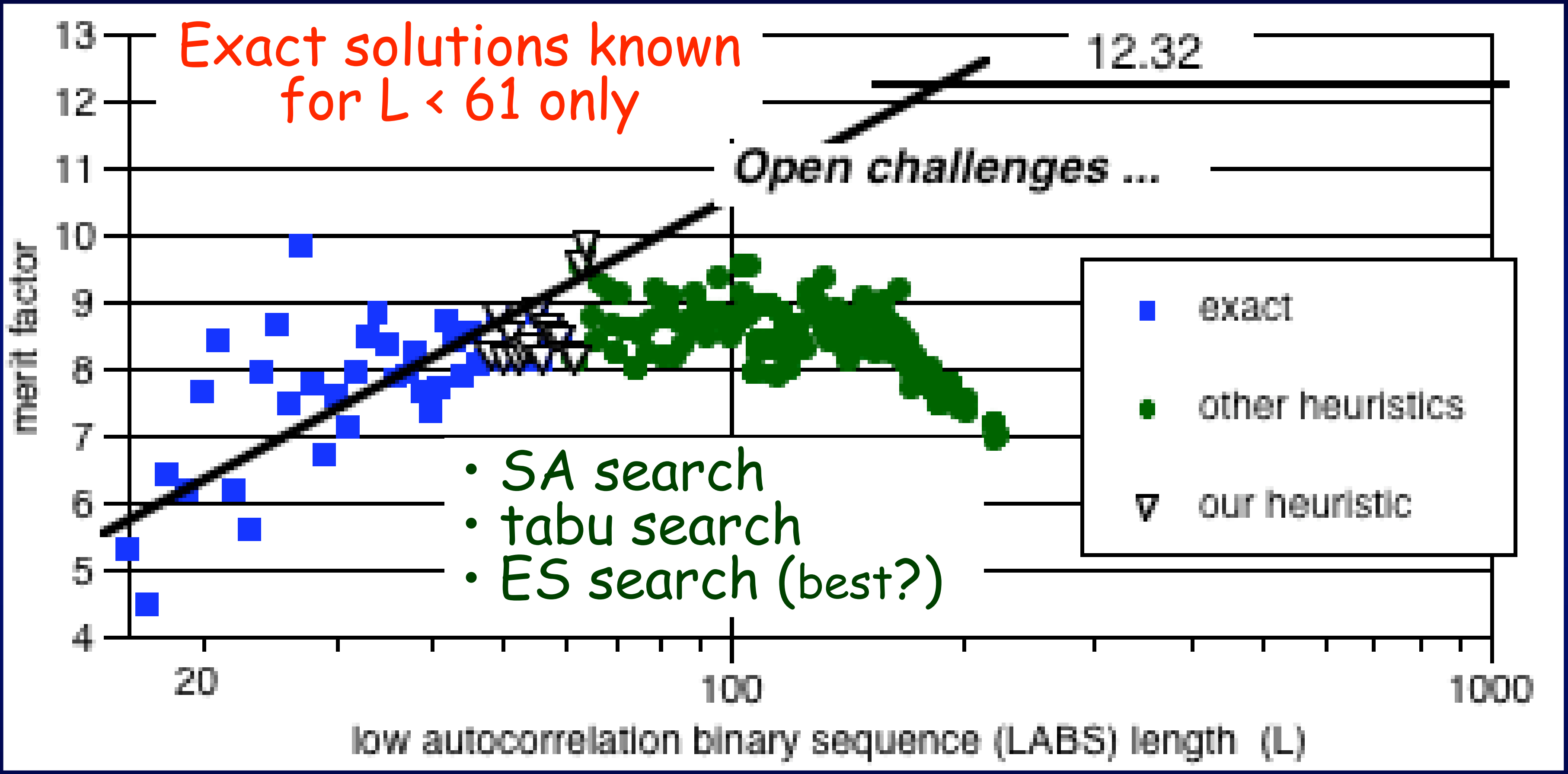
- a class of highly "frustrated problems", the global minimum is extremely hard to find

- solutions are relevant to
  physicists (Ising spin model)
  coding specialists (Barker codes)
  security specialists ??

- exact (branch-and-bound) solutions known for $L <= 60$ only

- took over a year to report an exact minimum for L=60 (Mertens 2003)

LABS Solutions and Challenges



Wrong trend direction for L > 100:

- ineffective algorithms?
- premature termination?
- both?

Our SLS Heuristic

(1) Termination criterion A (for L < 48)

(a) terminate the SLS (stochastic local search) solver when known optimum is found for the first time.

(b) repeat (a) to find a reliable estimates of solver performance statistics (mean, std, med, distr … of runtime, samples, generations, restarts, etc)

(2) Termination criterion B (for L >= 48)

(a) choose the SLS solver with best performance under termination criterion A

(b) use solver performance statistics to predict the <u>required runtime</u> and/or <u>required samples</u> so we may find the new minimum with probability P_{succ}.

Two SLS Solvers Under Test

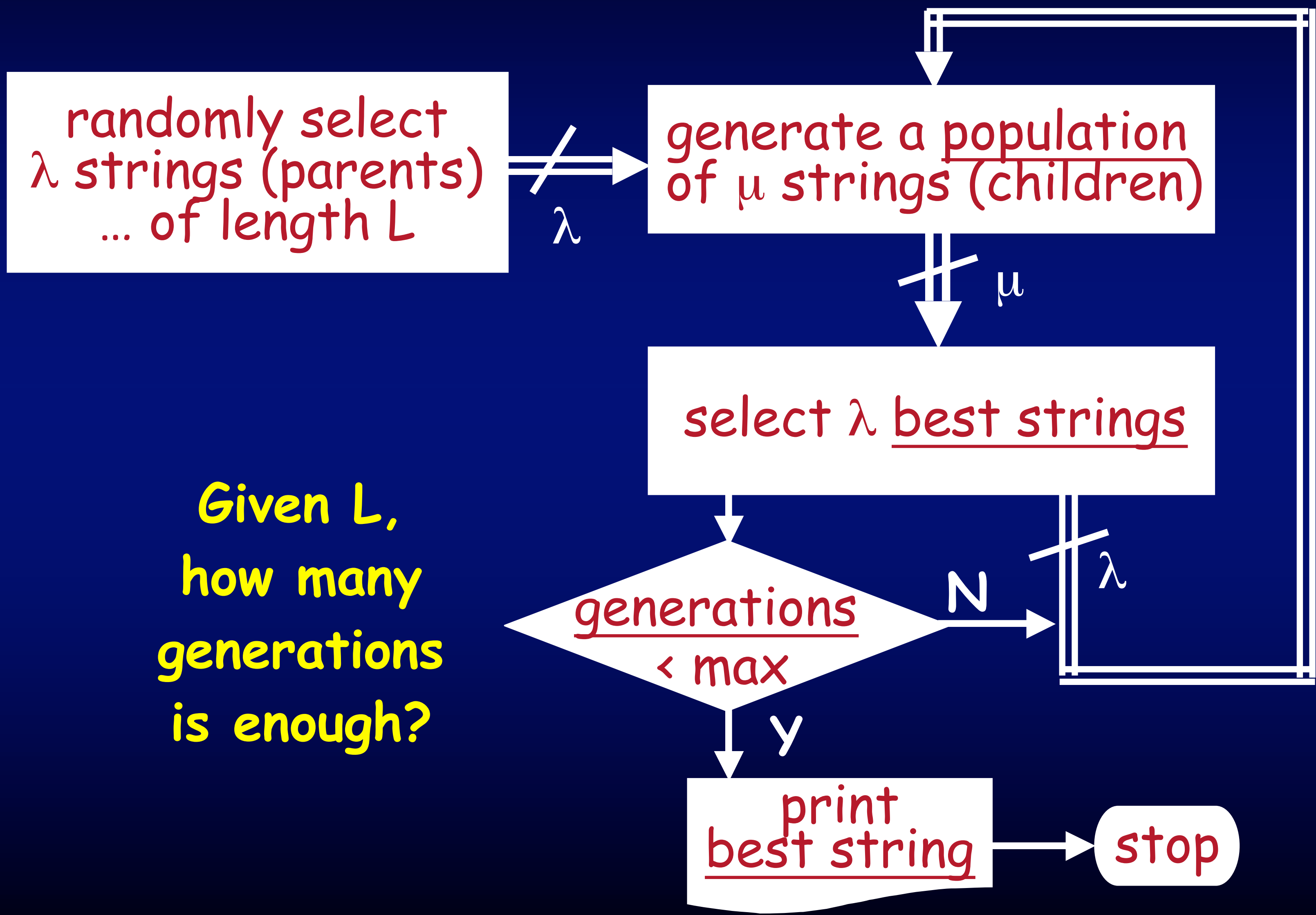
ES-solver (evolutionary strategy),
optimized for LABS problem in:

- Militzer, Zamparelli and Beule, Evolutionary Search for Low Autocorrelated Binary Sequences, IEEE Transactions on Evolutionary Computation, 1998

KL-solver (Kernighan-Lin),
adapted in this paper for LABS problem from:

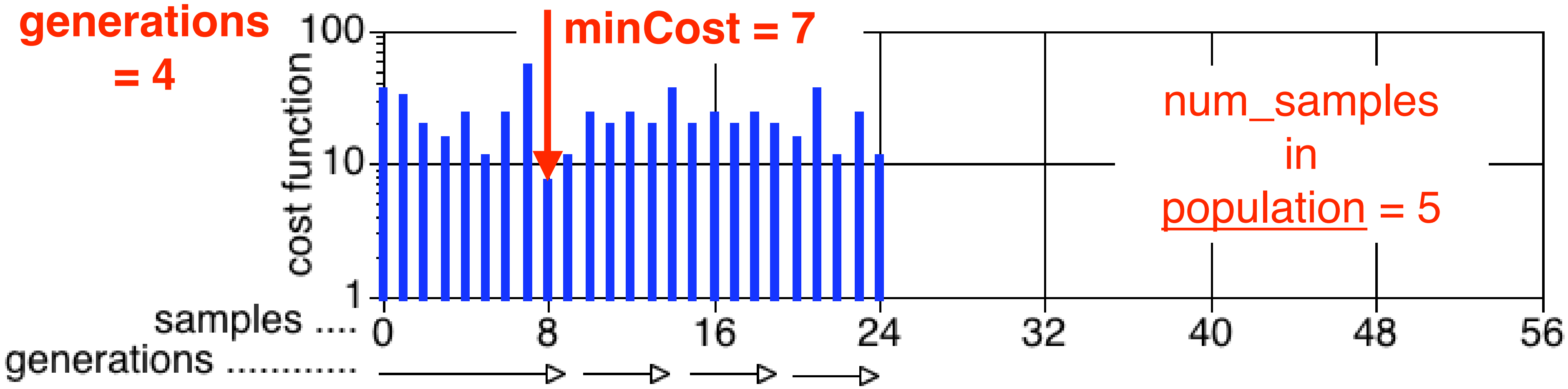
- Kernighan and Lin, An Efficient Heuristic Procedure for Partitioning Graphs, Bell System Tech. Journal, 1970
- Chen and Stallmann, Local Search Variants for Hypercube Embedding, Proc. Fifth Distributed Memory Computing Conference, 1990

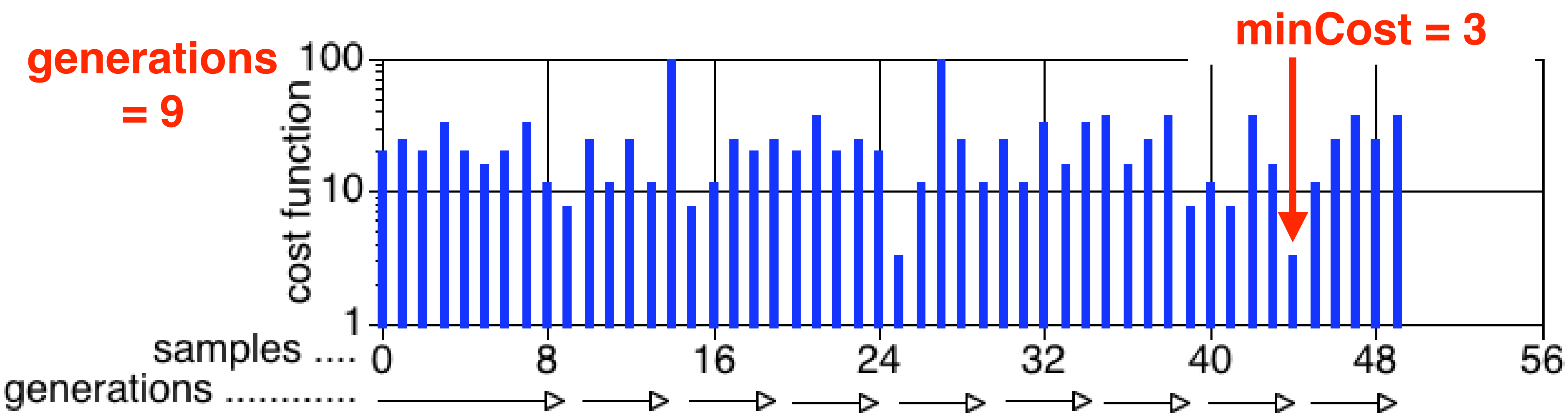
ES-solver: High Level Outline



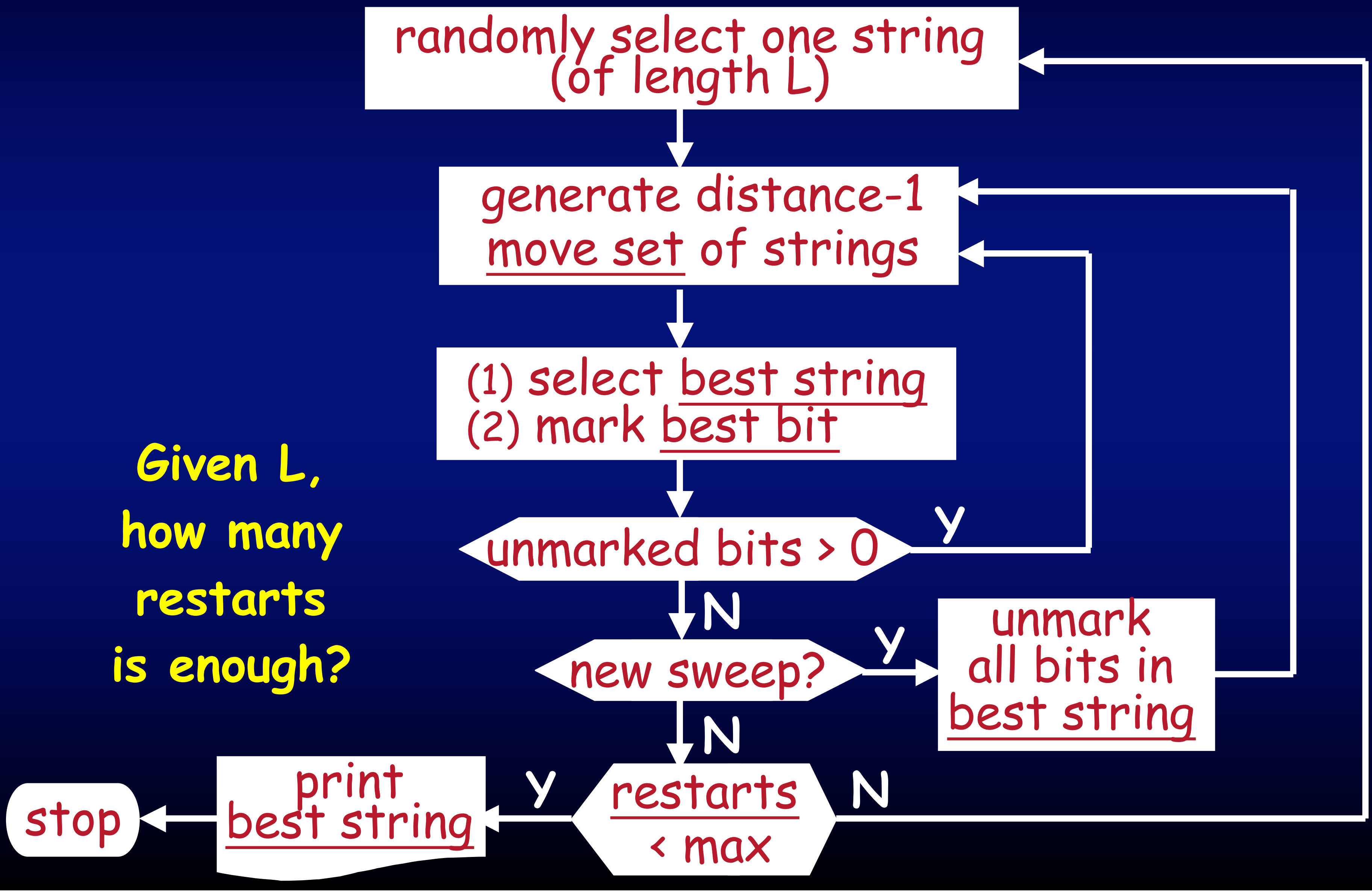
Given L, how many generations is enough?

Trace with ES-solver for L=7



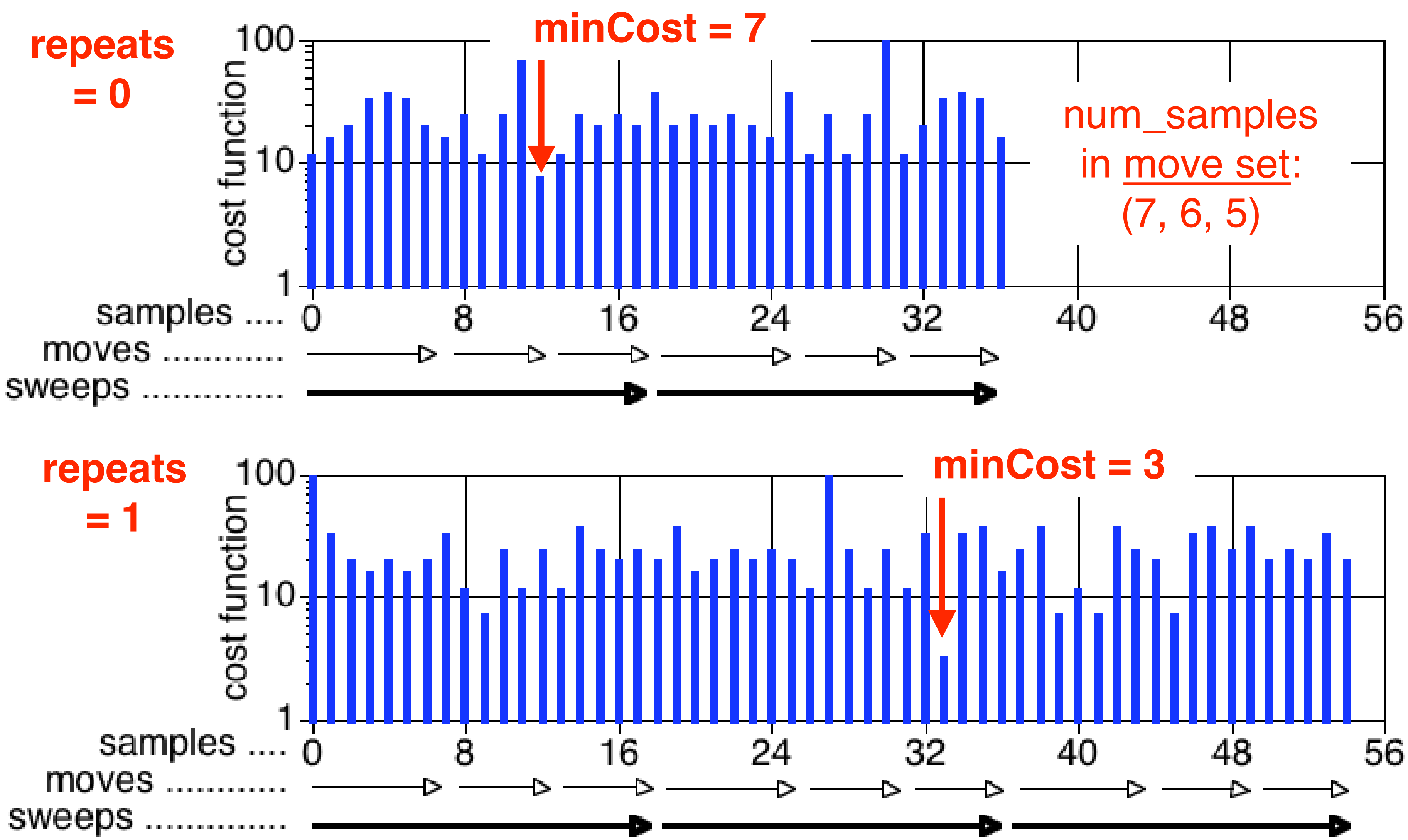


KL-solver: High Level Outline



**Given L,
how many
restarts
is enough?**

Trace with KL-solver for L=7



… To Restart Or Not To Restart (KL)?

vs … One More Population to Generate (ES)?

- these two parameters are not directly comparable
- preliminary runtime comparisons favor KL by a wide margin, but provide insufficient insights
- sample counts are THE universal common denominator:

samples(KL) ~ restarts*sweeps*moves

… sweeps ~ 2--4, moves ~ O(L*L/2)

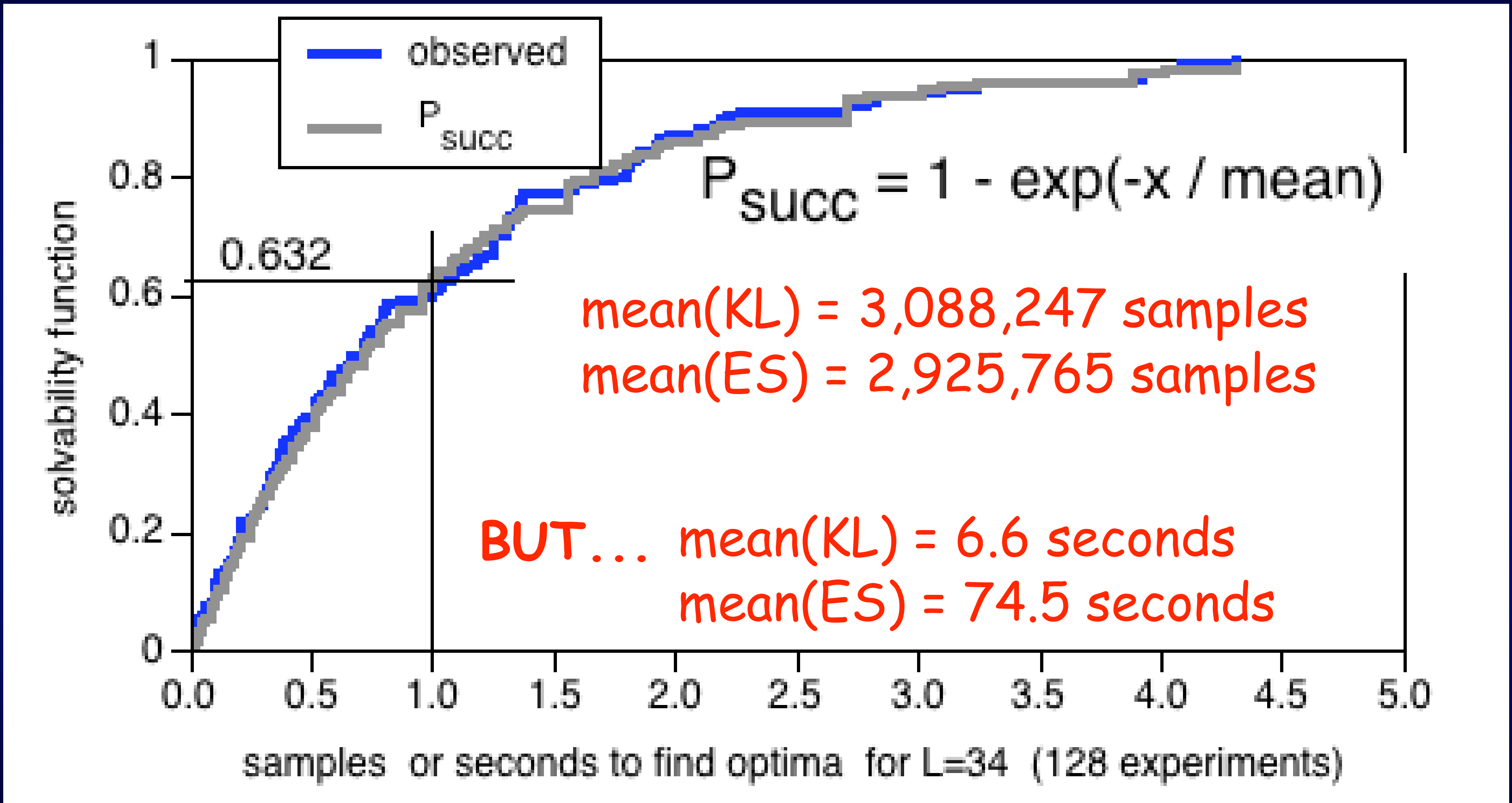
samples(ES) ~ parents + generations*children

… $\lambda \sim 10$, $\mu \sim 3*L$

Experimental Methodology (Term A)

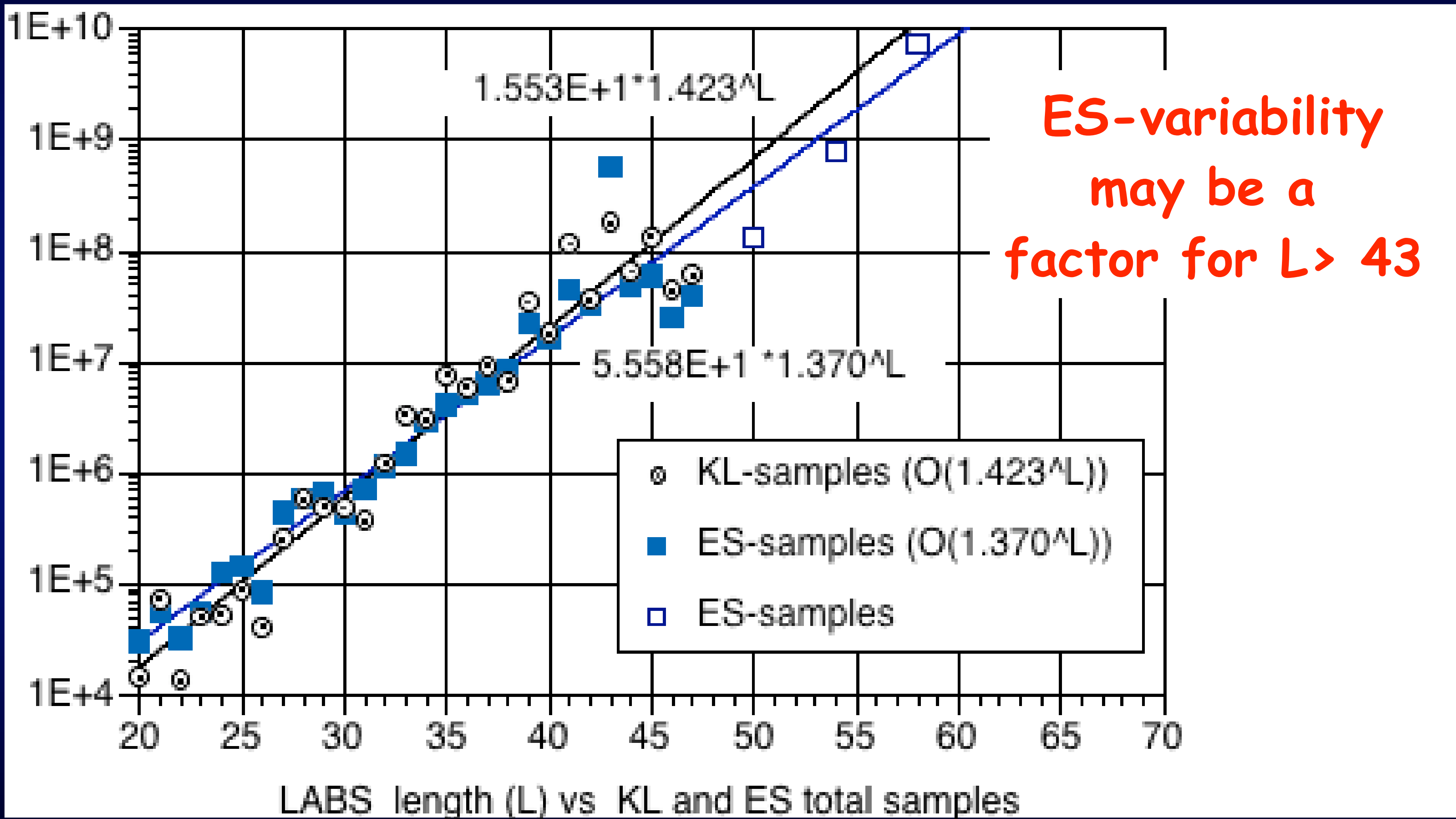
```
solverList <- {ES, KL}
randomSeedList <- {33, 1538, 1968, …} (a total of 128)
While {10 < L < 48} {
    For each seed {
        For each solver & first optimum {
            record runtime, samples
                   (generations, restarts)
        }
    }
}
Generate statistics and distributions
Analyze
```

Termination-A Experiments for L=34



And the distribution is EXPONENTIAL for BOTH solvers
… implying significant variability (mean ~ stdev)

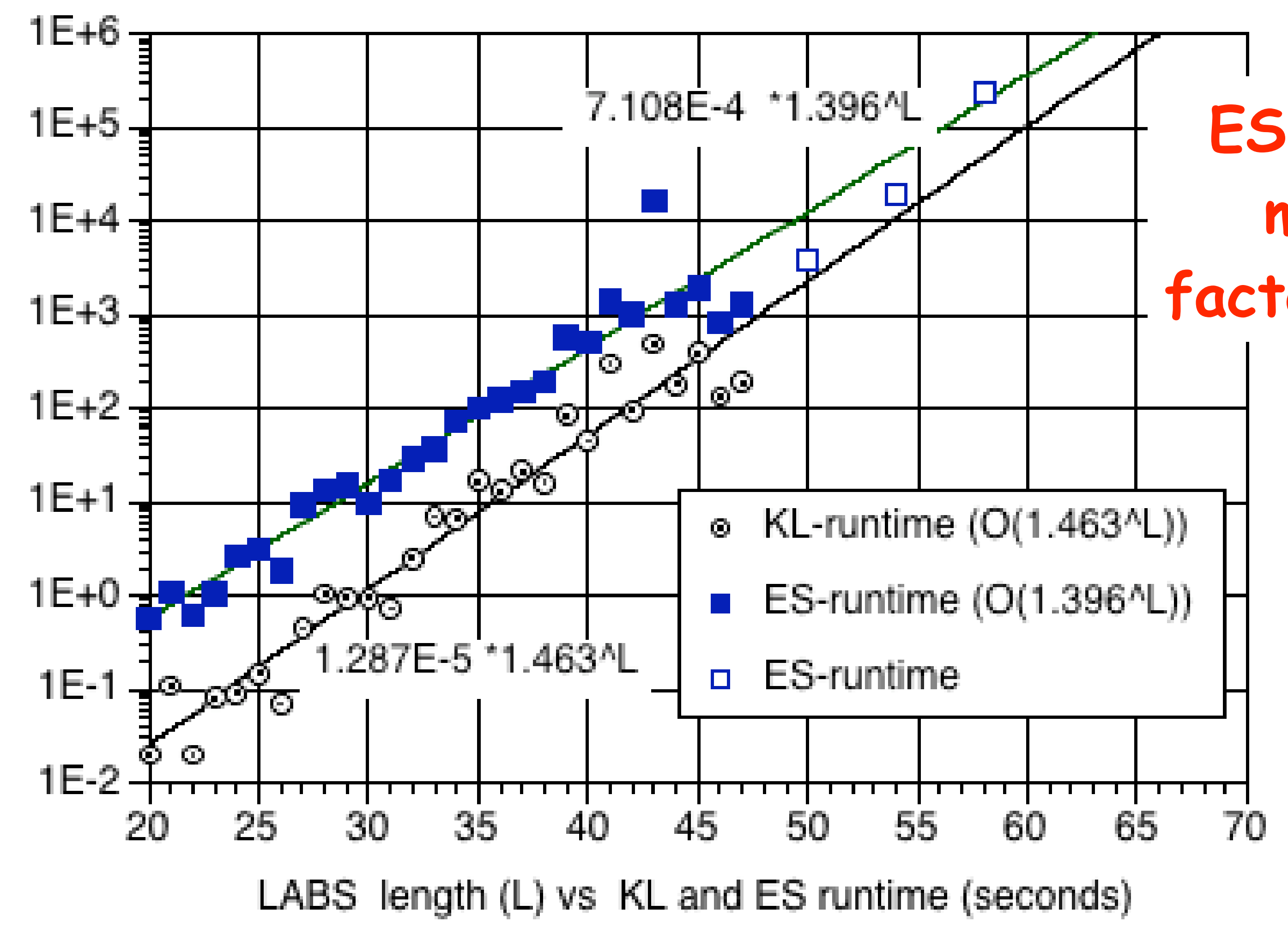
KL and ES Asymptotics (samples)



ES-variability may be a factor for L> 43

KL and ES asymptotes crossover around L=34

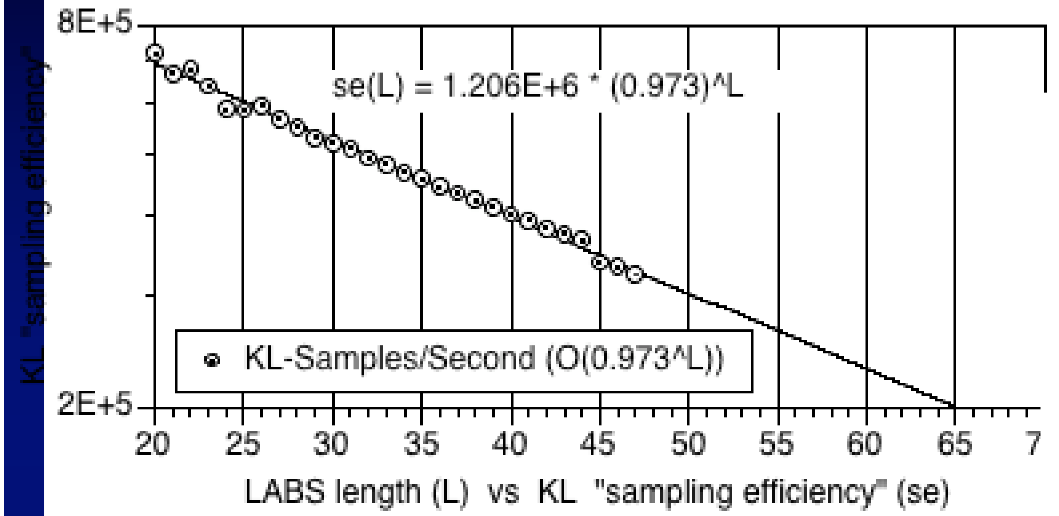
KL and ES Asymptotics (seconds)



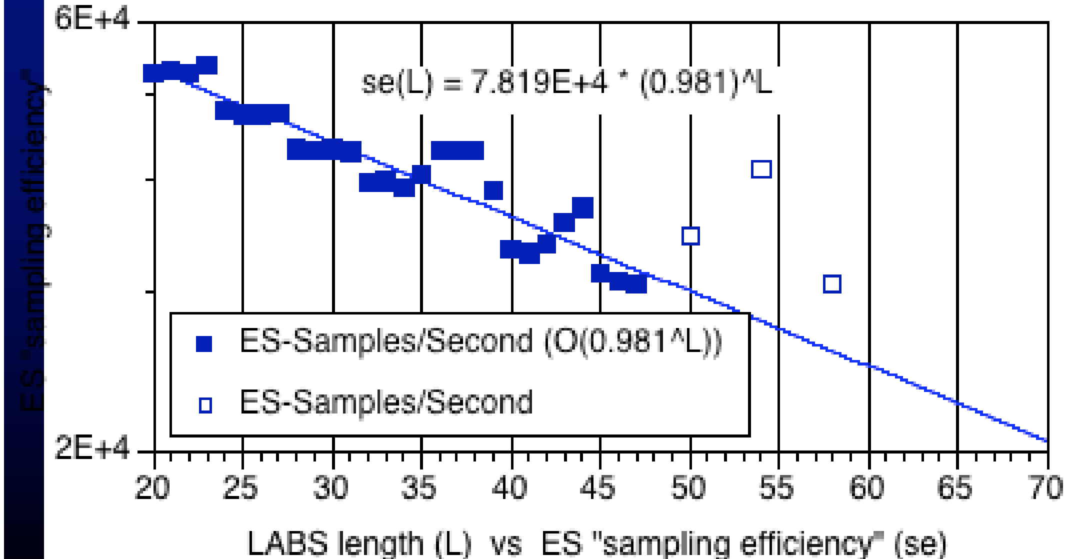
ES-variability may be a factor for L> 43

KL and ES asymptotes crossover around L= 88

KL and ES Asymptotics (sampling efficiency)



Sampling efficiency ratio
at L=34:
467,916 : 39,312
11.9 : 1



ES-variability
may be a
factor for L> 43

Reality Check

ES/KL runtime crossover at L = 88 of academic interest only

- on a 266 MHz platform (current runs):
  projected runtime L = 74 for KL 0.9 years

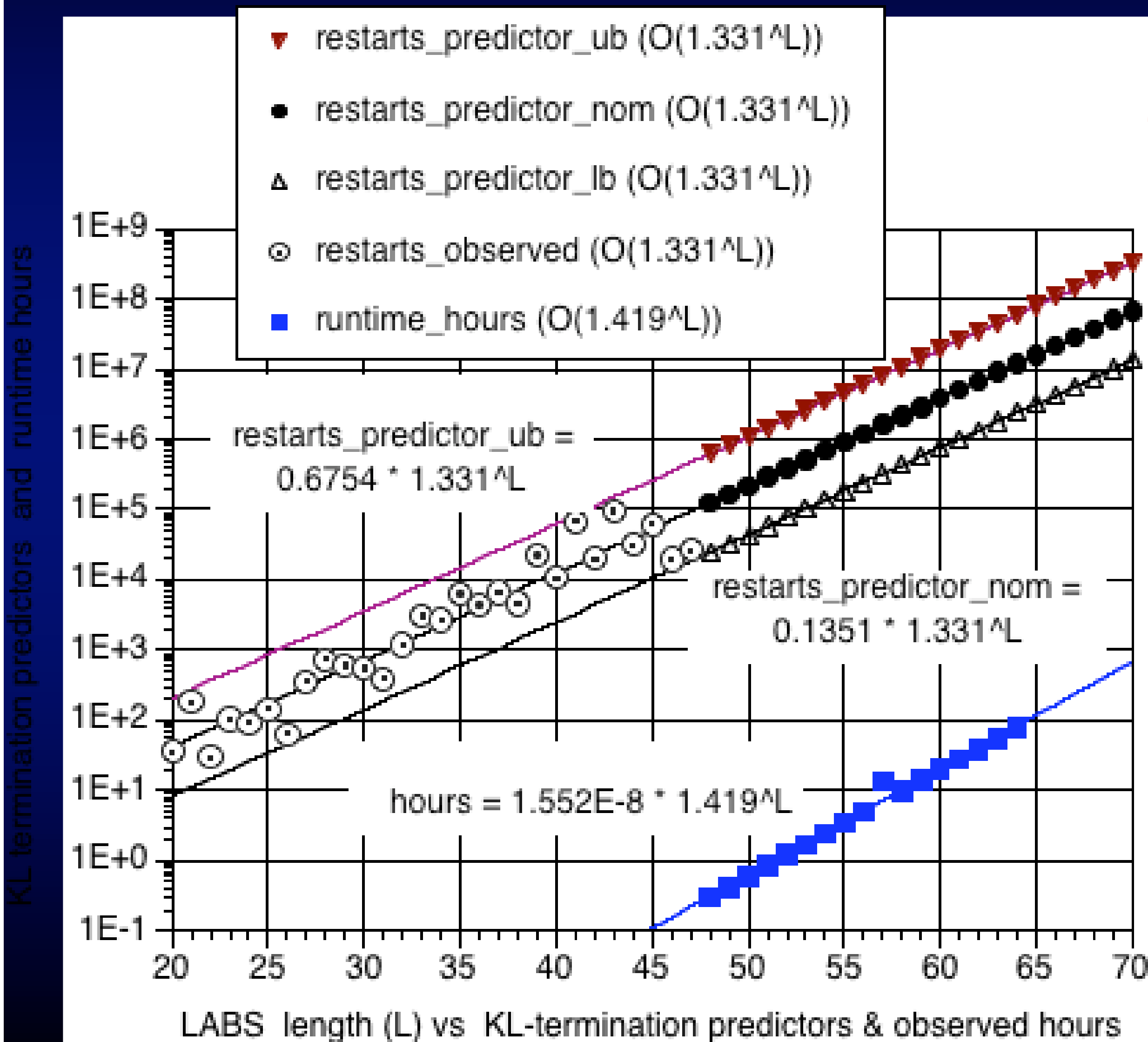
- on a 10-fold faster platform:
  projected runtime L = 84 for KL 1.1 years

We need a new LABS solver with BOTH:

(1) runtime much better than O(1.42^L)

(2) constant factor much better than 1.55e-8

Predictions for Termination-B: KL-solver



Restart predictors for the KL-solver (based on L < 47)

Experiments for L >= 48

(a) cost=$O(1.42^L)$

(b) all optima and best known solutions reached for L <=70

Reliability of Termination-B (KL-solver)

L	E_{opt}	nE	hitR	hours
48	140	16	0.69	0.3
49	136	16	0.75	0.4
50	153	16	0.94	0.6
51	153	16	0.31	0.8
52	166	16	0.75	1.2
53	170	16	0.75	1.7
54	175	16	0.63	2.4
55	171	16	0.88	3.5
56	192	16	1.00	5.0
57	188	16	0.69	13.2
58	197	16	0.81	9.8
59	205	16	1.00	13.9
60	218	16	1.00	20.0
…	…	…	….	….

16 experiments, L=48--60

Expected hit ratio of finding optimum = 0.632

For L=48, 0.3*16=4.8 h
(with BnB, ~10 days/soln)

For L=60, 20*16=320 h
(with BnB, ~1 year/soln)

Still, much better solver than KL is needed for L>77

Conclusions

- methodology to reliably evaluate performance of SLS solvers on LABS problem
  - KL needs to reduce number of samples to find optima
  - ES needs to improve the `sampling efficiency' while finding optima
- termination criterion B reliably predicts computational cost of finding optima with current generation of LABS solvers
- significant solver improvements are needed for to find optimal solutions at L>100

For longer paper, see the LABS problem home page:
http://www.cbl.ncsu.edu/OpenExperiments/LABS/

Future Work

- faster and more efficient LABS solvers
  (KL-like, ES-like, ….)

- termination criterion *C*
  (when we don't know the optimum solution)

THANK YOU FOR ATTENTION!