Optimal Learning: Bayesian Methods for Simulation Optimization

Peter I. Frazier

Operations Research & Information Engineering, Cornell University

Thursday April 18, 2013
AFOSR Program Review
Optimization & Discrete Mathematics
Arlington, VA

AFOSR YIP FA9550-11-1-0083
(Optimization & Discrete Mathematics)
AFOSR BRI FA9550-12-1-0200, with Warren Powell
(Natural Materials, Systems & Extremophiles)

Outline

1. Part I: Worst-case Performance Guarantees
   - Ranking and Selection with Tight Bounds on Solution Quality
2. Part II: Average-case Performance
   - Materials Science: Finding peptides with specific nanoparticle binding properties

Ranking & Selection (R&S) is a basic problem in simulation optimization

- We have k alternative choices that can be simulated.
- When we simulate alternative x, we observe random payoff $Z(x)$.
- We wish to know which x has the largest $E[Z(x)]$.
- Sampling is the only way to estimate $E[Z(x)]$ because the simulator is too complex for direct theoretical analysis.
  - e.g., a large-scale discrete-event simulation of a supply chain; or a whole-theater combat simulator.
- We make no a priori assumptions about the relationships between alternatives.

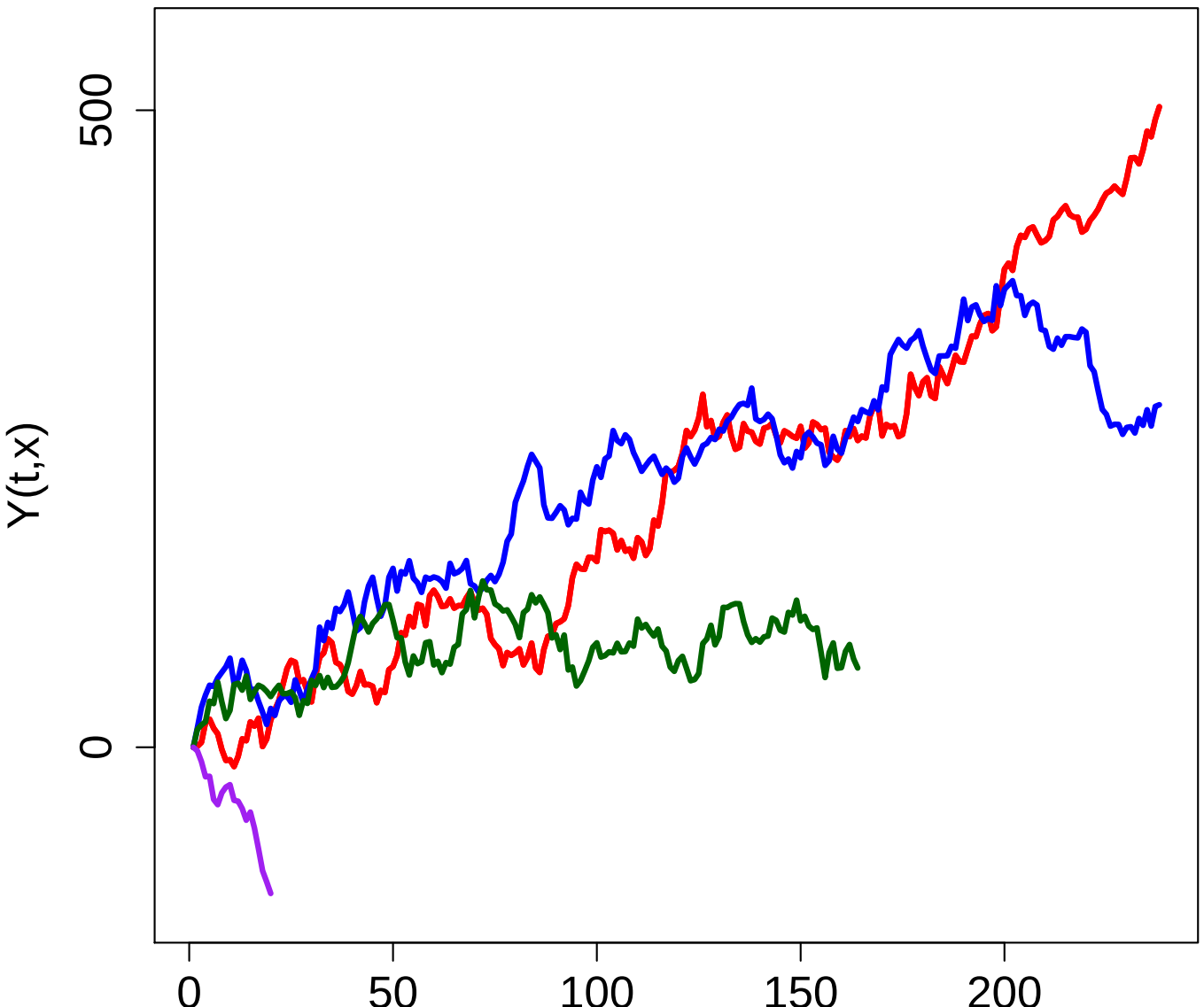
We Assume Independent Normal Samples

- The R&S literature often assumes that Z(x) is normally distributed:

$$Z(x) \sim \text{Normal}(\theta_x, \sigma_x^2) \qquad \text{(independent across } x \text{ and time)}$$

- θ_x is unknown but fixed (non-Bayesian).
- For simplicity in this talk, we assume σ_x^2 is known, but the procedure presented can also be used in the unknown variance setting.
- Approximate normality of $Z(x)$ can be checked empirically, and if it is not met, samples can be batched together.

A Typical Ranking & Selection Procedure



We focus on ranking & selection procedures with statistical guarantees on solution quality

- We consider one type of ranking & selection procedures: **indifference zone** ranking and selection procedures.
- These procedures provide solutions that have an associated statistical guarantee.
- This statistical guarantee is a lower bound on probaility of correct selection, over problem instances in a certain to-be-described set.

Studying ranking & selection is useful for three reasons

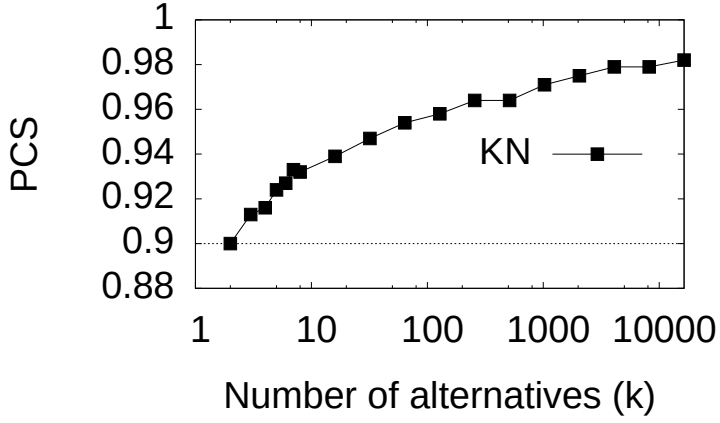
1. R&S procedures can be used directly to solve smaller problems ($k < 1000$).
   - Pro: statistical guarantee on solution quality;
   - Pro: no assumptions required on the relationships between alternatives
   - Con: lack of assumptions makes sample sizes big when k is big.
2. R&S procedures are used as subroutines by algorithms for bigger problems ($k > 1000$ or $k = \infty$):
   - Cleaning up after optimization [Boesel et al., 2003].
   - Finding a direction of local increase [Hong and Nelson, 2006, Xu et al., 2010].
3. R&S is a core problem in sim-opt problem: Understanding it supports understanding of other problems.

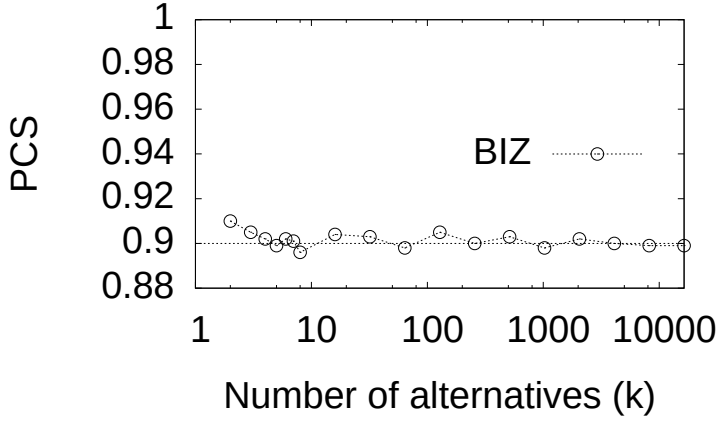
Ranking & Selection is a well-studied problem

- Indifference-zone Ranking and Selection:
  [Bechhofer, 1954, Paulson, 1964, Bechhofer et al., 1968,
  Fabian, 1974, Dudewicz and Dalal, 1975, Rinott, 1978,
  Bechhofer and Goldsman, 1987, Hartmann, 1988, Hartmann, 1991,
  Paulson, 1994, Bechhofer et al., 1995, Kim and Nelson, 2001,
  Nelson et al., 2001, Malone et al., 2005, Hong, 2006,
  Dieker and Kim, 2011, Wang and Kim, 2011, Dieker and Kim, 2012]
- This does not include all the work on other problem formulations (Bayesian, large devations, OCBA).
- There is also a recent interest body of work in theoretical computer science, under the name "pure exploration multi-armed bandit"
  [Madani et al., 2004, Bubeck et al., 2009, Audibert et al., 2010,
  Bubeck et al., 2011]

Existing procedures have loose bounds, which lead them to sample more than necessary

- **Problem 1:** Existing procedures have loose theoretical bounds on solution quality, when $k > 2$.
- **Problem 2:** This leads them to sample more than necessary, to guarantee minimum solution quality.
- **Our Contribution:** We construct a procedure (BIZ, or **B**ayes-inspired **I**ndifference **Z**one) with tight bounds on solution quality. BIZ guarantees minimum solution quality while sampling less than existing procedures.





Notation & Terminology

- Let $\mathrm{PCS}(\theta)$ be the **Probability of Correct Selection**, where $\theta = (\theta_1, \ldots, \theta_k)$ is the vector of sampling means. This is the probability that the sampling procedure selects an alternative in $\arg\max_x \theta_x$.
  (The dependence on σ_x^2 is suppressed in the notation.)
- Let $\delta > 0$ represent the smallest meaningful difference in performance.
- Let P^* be a pre-specified lower bound with which we would like to select the best.

We consider the following form of statistical guarantee

- Define the **preference zone** as the set of θ where the best is better than the second best by at least δ:

$$\mathrm{PZ}(\delta) = \left\{ \theta \in \mathbb{R}^k : \theta_{[1]} - \theta_{[2]} \geq \delta \right\},$$

  where $\theta_{[1]} \geq \theta_{[2]} \geq \ldots \geq \theta_{[k]}$ are the order statistics of θ.
- The **indifference zone (IZ)** is the preference zone's complement.
- A procedure has an **IZ guarantee** with parameters δ and P^* if

$$\mathrm{PCS}(\theta) \geq P^* \quad \text{for all } \theta \in \mathrm{PZ}(\delta).$$

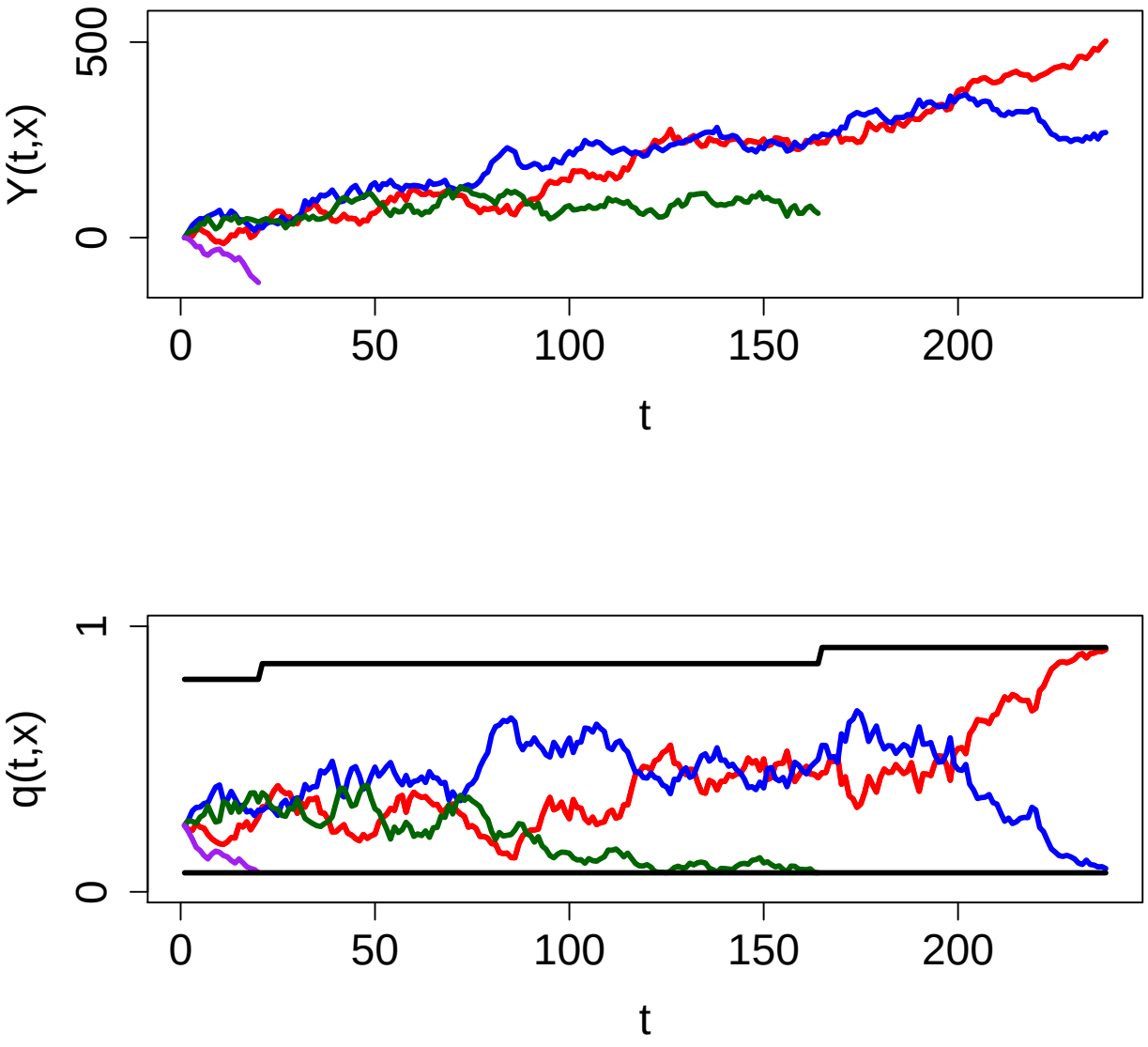
We construct our procedure with Bayesian ideas

- Let Y_{tx} be the sum of all observations from alternative x by time t.
- We use a prior distribution that is concentrated on slippage configurations.
- The posterior probability that an alternative x is the best, given that the best is in a specified set A, is

$$q_{tx}(A) = \exp\left(\frac{\delta}{\sigma^2} Y_{tx}\right) \Big/ \sum_{x' \in A} \exp\left(\frac{\delta}{\sigma^2} Y_{tx'}\right).$$

- We watch $q_{tx}(A)$, where A is the set of alternatives in contention. An alternative is eliminated if it hits a lower threshold, and selected as best if it hits an upper threshold.
- The upper threshold increases with each elimination, to compensate for the possibility that we may mistakenly eliminate the best.

The Bayes-inspired Indifference Zone (BIZ) Procedure



The BIZ Procedure

Fix parameters $c \leq 1 - (P^*)^{1/(k-1)}$, $\delta > 0$, $P^* > 1/k$.

1. Let $A \leftarrow \{1, \ldots, k\}$, $t \leftarrow 0$, $P \leftarrow P^*$.
2. While $\max_{x \in A} q_{tx}(A) < P$
   - 2a. While $\min_{x \in A} q_{tx}(A) \leq c$
     - Let $x \in \arg\min_x q_{tx}(A)$.
     - Let $P \leftarrow P/(1 - q_{tx}(A))$.
     - Remove x from A.
   - 2b. Sample from each $x \in A$ to obtain $Y_{t+1,x}$. Then increment t.
3. Select $\hat{x} \in \arg\max_{x \in A} Y_{tx}$ as our estimate of the best.

Recall:

$$q_{tx}(A) = \exp\left(\frac{\delta}{\sigma^2} Y_{tx}\right) \Bigg/ \sum_{x' \in A} \exp\left(\frac{\delta}{\sigma^2} Y_{tx'}\right).$$

BIZ has Tight PCS Bounds

Theorem

Assume that $\sigma_x^2 = \sigma^2$ is known. Fix any $\delta > 0$, $P^* \in (1/k, 1)$, $c \leq 1 - (P^*)^{1/(k-1)}$, $\mathbb{T} \in \{\mathbb{R}_+, \mathbb{Z}_+\}$, and let sampling occur under BIZ. Then,

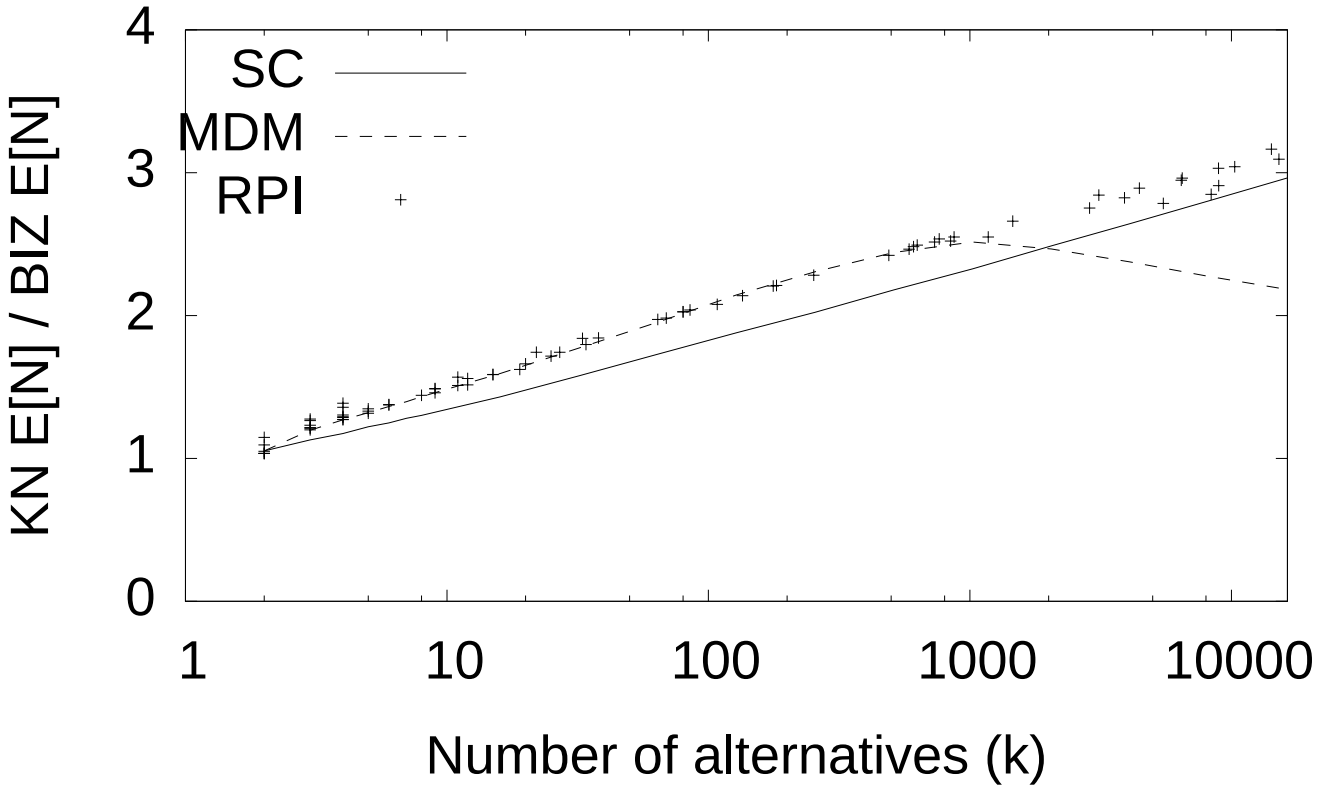
$$\mathrm{PCS}(\theta) \geq P^* \quad \forall \theta \in \mathrm{PZ}(\delta)$$

Moreover, for a generalization of BIZ to continuous time,

$$\inf_{\theta \in \mathrm{PZ}(\delta)} \mathrm{PCS}(\theta) = P^*$$

This is the first sequential elimination indifference-zone procedure with tight bounds for $k > 2$.

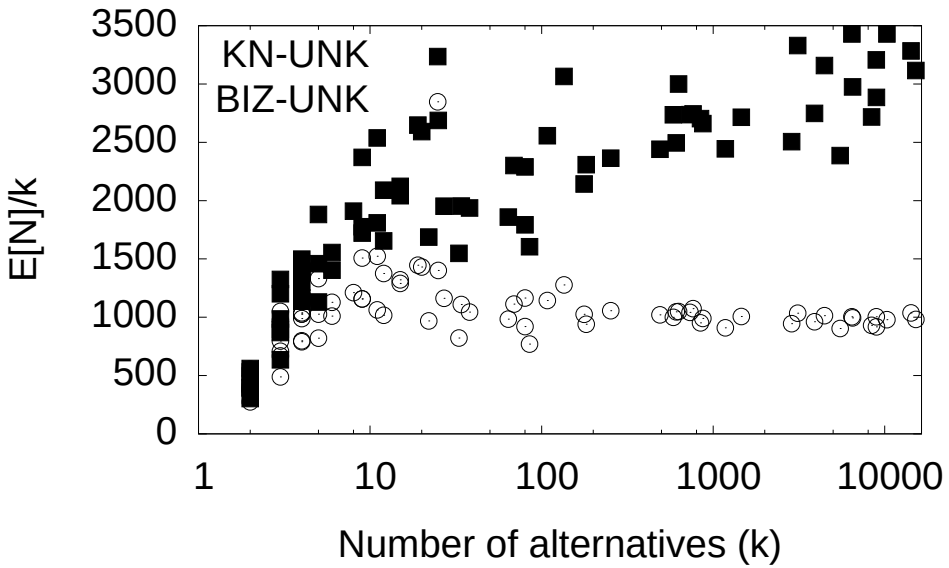
Tight bounds allow taking fewer samples

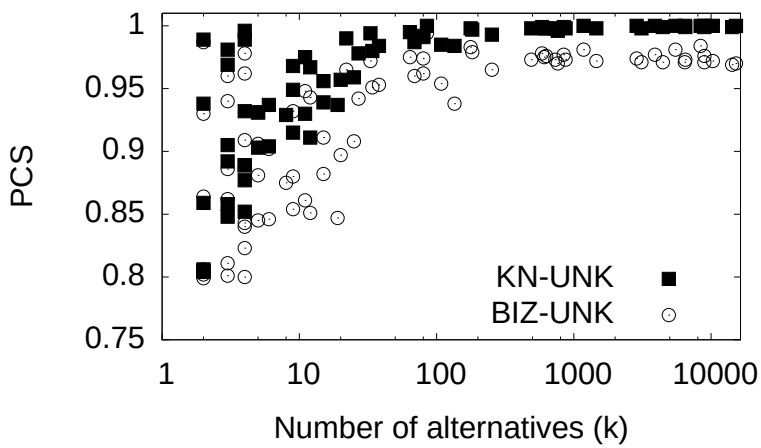


- On many of the largest problems considered in our numerical experiments, BIZ requires between 1/3 and 1/2 as many samples as KN [Kim and Nelson, 2001], the $\mathscr{P}$ procedure of [Wang and Kim, 2011], and Paulson's procedure [Paulson, 1964].

BIZ can be generalized to Heterogeneous and Unknown Sampling Variance

- In practice, sampling variances are unknown and heterogeneous.
- The BIZ procedure has been generalized to this more general setting.
- Theoretical results can also be generalized, but require restrictive assumptions unlikely to be met in practice, e.g., a continuous-time observation process.
- Empirical results, however, in the heterogeneous unknown variance case are similar to what is predicted by the continuous-time theory.





Part I Conclusion

- BIZ is a fully sequential IZ procedure with elimination that delivers **exactly** P^* (in continuous time, and under the worst θ in the preference zone).
- To my knowledge, this is the **first fully sequential elimination IZ procedure with this property** for $k > 2$.
- Theoretical results require unrealistic assumptions on the sampling variances, but empirical results suggest that behavior is robust to violations of these assumptions in the problem regimes tested.

Outline

1. Part I: Worst-case Performance Guarantees
   - Ranking and Selection with Tight Bounds on Solution Quality

2. Part II: Average-case Performance
   - Materials Science: Finding peptides with specific nanoparticle binding properties

Bayesian methods are useful for designing methods with good average-case performance

- The previous part of the talk used Bayesian methods as a theoretical tool for designing a procedure with a **worst-case** performance guarantee.
- Bayesian methods are also extremely useful for designing methods with **good average-case** performance.

Current Applications of Optimal Learning

- Design of cardiovascular bypass grafts (Alison Marsden, UCSD)
- Large-scale simulation optimization with common random numbers (Steve Chick, INSEAD)
- Materials Science: finding minimal peptide-based substrates (Nathan Giannesci, UCSD)
- **Materials Science: finding peptides with specific nanoparticle binding properties** (Paras Prasad, Buffalo)
- Materials Science: optimization of continuous parameters (several groups, at Princeton, Cornell and AFRL)

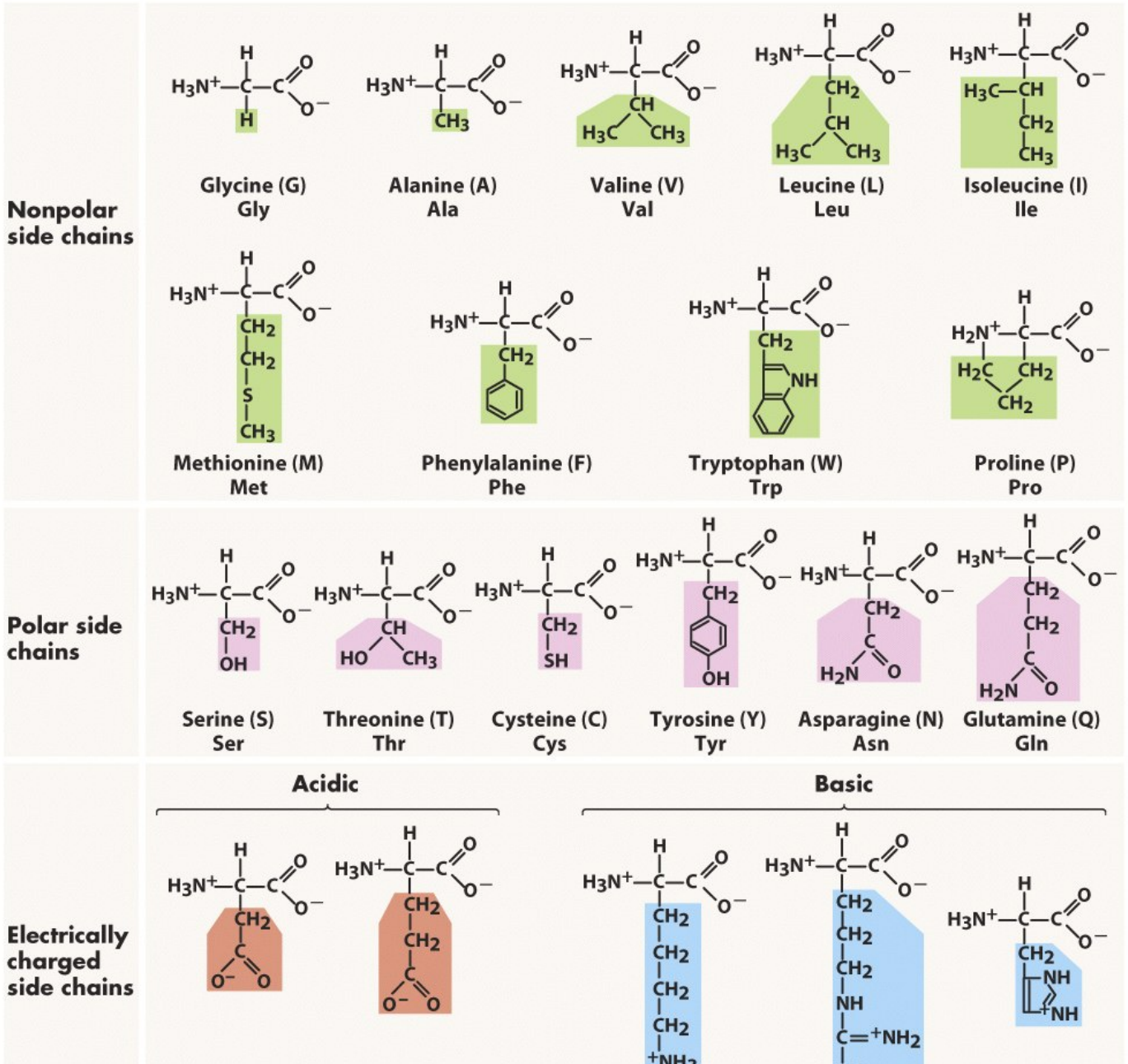
The materials science collaborations are funded by AFOSR BRI FA9550-12-1-0200, from Natural Materials, Systems & Extremophiles, with Warren Powell as PI.

Finding peptides with specific nanoparticle binding properties

Our **collaborators**: Paras Prasad (PI, Buffalo); Tiff Walsh (Deakin); Marc Knecht (Miami); Aidong Zhang (Buffalo); Mark Swihart (Buffalo).

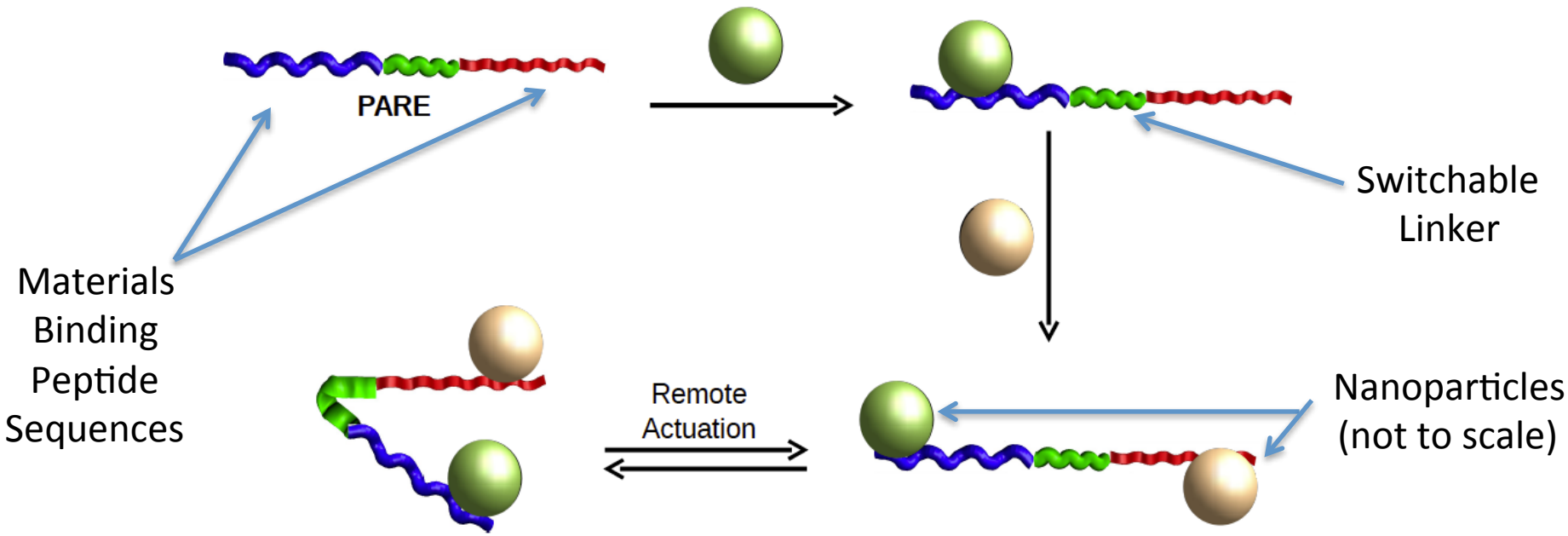
Project Goal: create novel nano-materials with useful photonic, electronic, and plasmonic properties, by creating peptide links between various nano-size metals and oxides.

Background: What's a Peptide? What's an Amino Acid?



Challenge

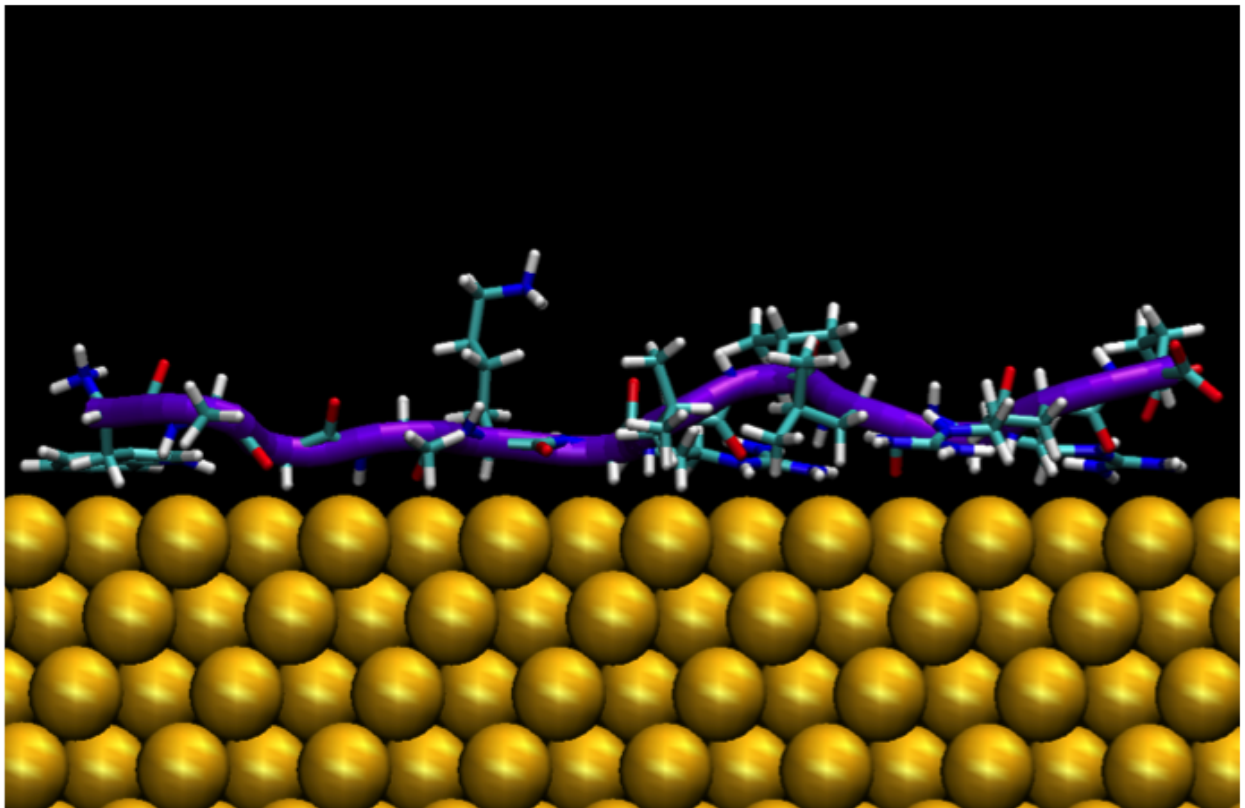
- Given two materials, e.g., gold and iron oxide, find a peptide that binds strongly to the first material (e.g., gold), and weakly to the second (e.g., iron oxide).
- The role of mathematics is to suggest which experiments to perform.



There are two kinds of experiments we can run

- **Physical (lab) experiments**: Given a peptide, measure the Gibbs free energy of binding (a measure of binding strength).
- **Computer experiments**: Given a peptide, use a physics-based simulation to calculate to calculate, for each amino acid, the percentage of time it is in contact with the target material.

Each experiment (lab or computer) takes ≈ 1 week.



Our preliminary dataset required $\approx$ 2.5 months of lab effort, and $\approx$ 2.5 months of computer time

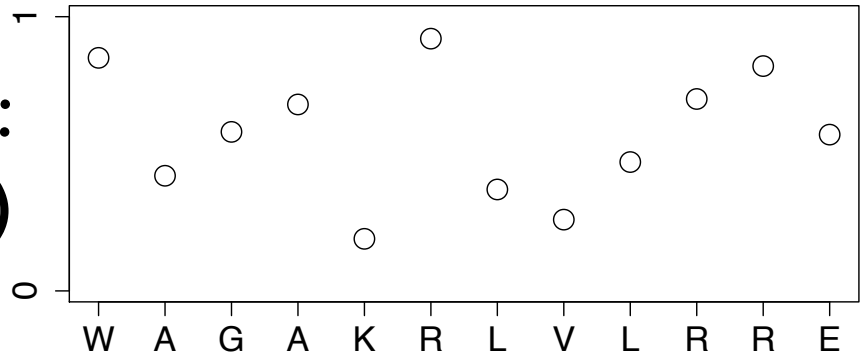
Peptide Name	Sequence	Experimental ΔG_{abs} (kJ/mol)
A3	AYSSGAPPMPPF	-31.8 ±0.3
AgBP1	TGIFKSARAMRN	-31.6 ±0.2
AgBP2	EQLGVRKELRGV	-35.3 ±1.2
AuBP1	WAGAKRLVLRRE	-37.6 ±0.9
AuBP2	WALRRSIRRQSY	-36.4±0.3
B1	LKAHLPPSRLPS	-36.6±1.2
GBP1	MHGKTQATSGTIQS	-37.6 ±1.0
Midas2	TGTSVLIATPYV	-35.7 ±1.2
Pd4	TSNAVHPTLRHL	-30.3 ±0.2
QBP1	PPPWLPYMPPWS	-35.0 ±1.1

We build a Bayesian model that takes advantage of **all** of our data: domain knowledge, computer experiments, lab experiments

Peptide: x=WAGAKRLVLRRE

$y \sim F(x,\alpha)$

% of time in contact:
y=(y(1),...,y(k))



$z \sim G(y,\beta)$

Gibbs free energy:
(smaller = more binding)

z=-37.6 ±0.9

Part II: Conclusion

- This statistical model takes advantage of all of the data available: domain knowledge, computer experiments, and lab experiments.
- This allows us to make reasonable predictions, despite having very little experimental data (only 12 datapoints).
- A KG method is then defined by this statistical model, and will suggest which experiment (computer or lab, and which peptide) we should perform next.
- Next steps:
  - We are validating the statistical model.
  - We are developing methods for computing the KG method efficiently.

Thank You!

References I

Audibert, J.-Y., Bubeck, S., et al. (2010).

Best arm identification in multi-armed bandits.

COLT 2010-Proceedings.

Bechhofer, R. E. (1954).

A Single-Sample Multiple Decision Procedure for Ranking Means of Normal Populations with known Variances.

The Annals of Mathematical Statistics, 25(1):16–39.

Bechhofer, R. E. and Goldsman, D. M. (1987).

Truncation of the Bechhofer-Kiefer-Sobel sequential procedure for selecting the normal population which has the largest mean.

Communications in Statistics-Simulation and Computation, 16(4):1067–1092.

Bechhofer, R. E., Kiefer, J., and Sobel, M. (1968).

Sequential Identification and Ranking Procedures.

University of Chicago Press, Chicago.

Bechhofer, R. E., Santner, T. J., and Goldsman, D. M. (1995).

Design and Analysis of Experiments for Statistical Selection, Screening and Multiple Comparisons.

J.Wiley & Sons, New York.

Boesel, J., Nelson, B. L., and Kim, S.-h. (2003).

USING RANKING AND SELECTION TO CLEAN UP AFTER SIMULATION OPTIMIZATION.

Operations Research, 51(5):814–825.

Bubeck, S., Munos, R., and Stoltz, G. (2009).

Pure exploration in multi-armed bandits problems.

In *Proceedings of the 20th international conference on Algorithmic learning theory*, pages 23–37. Springer-Verlag.

References II

Bubeck, S., Munos, R., and Stoltz, G. (2011).

Pure exploration in finitely-armed and continuous-armed bandits.
Theoretical Computer Science, 412(19):1832–1852.

Dieker, A. B. and Kim, S.-h. (2011).

Selecting the best by comparing simulated systems in a group of three.
In *Proceedings of the 2011 Winter Simulation Conference*, pages 3992–4002.

Dieker, A. B. and Kim, S.-h. (2012).

SELECTING THE BEST BY COMPARING SIMULATED SYSTEMS IN A GROUP OF THREE.
In *Proceedings of the 2012 Winter Simulation Conference*.

Dudewicz, E. J. and Dalal, S. R. (1975).

Allocation of observations in ranking and selection with unequal variances.
Sankhy{\=a}: The Indian Journal of Statistics, Series B, 37(1):28–78.

Fabian, V. (1974).

Note on Anderson's sequential procedures with triangular boundary.
The Annals of Statistics, pages 170–176.

Hartmann (1991).

An improvement on {P}aulson's procedure for selecting the population with the largest mean from k normal populations with a common unknown variance.
Sequential Analysis, 10(1-2):1–16.

Hartmann, M. (1988).

An improvement on {P}aulson's sequential ranking procedure.
Sequential Analysis, 7(4):363–372.

References III

Hong, J. (2006).

Fully sequential indifference-zone selection procedures with variance-dependent sampling.
Naval Research Logistics (NRL), 53(5):464–476.

Hong, L. J. and Nelson, B. L. (2006).

Discrete optimization via simulation using COMPASS.
OPERATIONS RESEARCH, 54(1):115.

Kim, S. H. and Nelson, B. L. (2001).

A fully sequential procedure for indifference-zone selection in simulation.
ACM Trans. Model. Comput. Simul., 11(3):251–273.

Madani, O., Lizotte, D. J., and Greiner, R. (2004).

The budgeted multi-armed bandit problem.
In *Learning Theory*, pages 643–645. Springer.

Malone, G. J., Kim, S.-h., Goldsman, D., and Batur, D. (2005).

Performance of Variance Updating Ranking and Selection Procedures.
In *Proceedings of the 2005 Winter Simulation Conference*. IEEE.

Nelson, B. L., Swann, J., Goldsman, D., and Song, W. (2001).

Simple Procedures for Selecting the Best Simulated System When the Number of Alternatives is Large.
Operations Research, 49(6):950–963.

Paulson, E. (1964).

A Sequential Procedure for Selecting the Population with the Largest Mean from k Normal Populations.
The Annals of Mathematical Statistics, 35(1):174–180.

References IV

Paulson, E. (1994).

Sequential procedures for selecting the best one of K Koopman-Darmois populations.
Sequential Analysis, 13(3).

Rinott, Y. (1978).

On two-stage selection procedures and related probability-inequalities.
Communications in Statistics-Theory and Methods, 7(8):799–811.

Wang, H. and Kim, S. H. (2011).

On the Conservativeness of Fully Sequential Indifference-Zone Procedures.

Xu, J., Nelson, B. L., and Hong, J. (2010).

Industrial strength COMPASS: A comprehensive algorithm and software for optimization via simulation.
ACM Transactions on Modeling and Computer Simulation (TOMACS), 20(1):3.