LUNDS UNIVERSITET
Lunds Tekniska Högskola

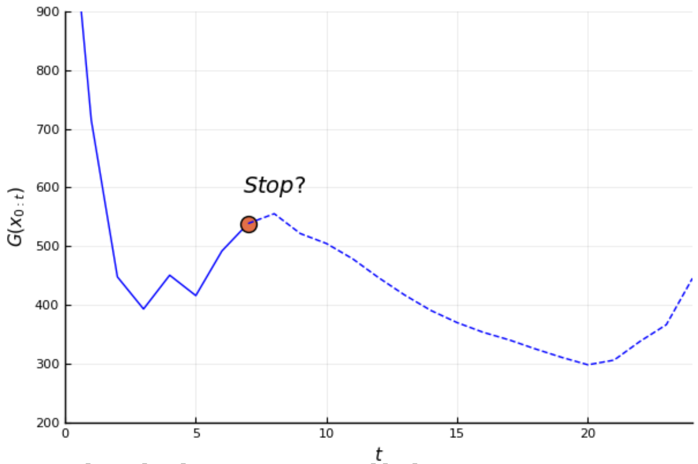
Knowing When to Stop

MARCUS THELANDER ANDRÉN

Optimal Stopping

- Stochastic control with two actions: {stop, continue}
- Find stopping policy that minimizes/maximizes cost/reward

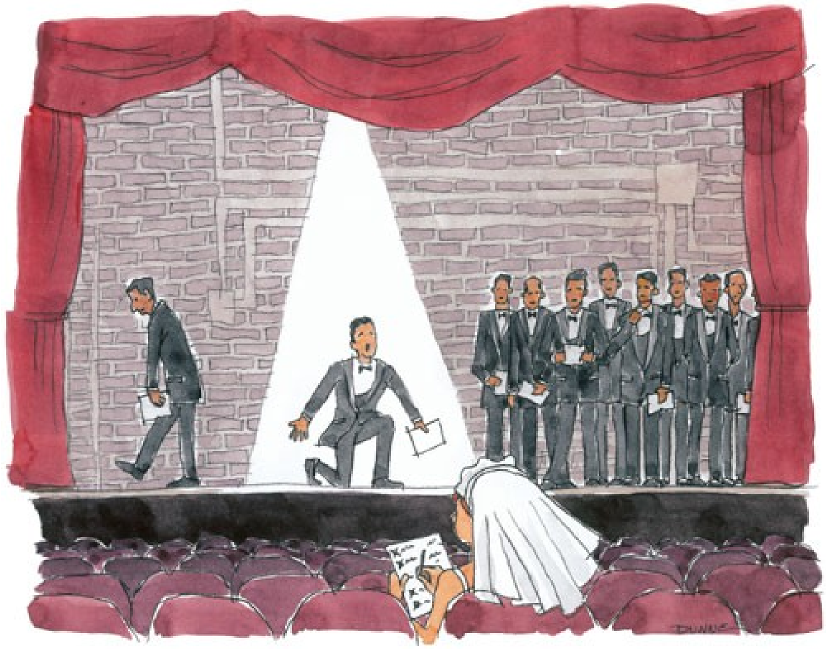
$$\min_{0 \leq \tau \leq T} \mathbb{E}[G(x_{0:\tau})]$$



- Examples:
  - When to stop interviewing and pick a candidate
  - When to stop experimenting and accept/reject a hypothesis
  - When to stop gambling and collect your winnings
  - When to stop observing the market and sell/buy stocks
  - When to stop open-loop control and apply feedback

The Fiancée Problem

- Looking for a fiancée among N bachelors
- For each bachelor, decide:
  - Reject and continue dating
  - Stop and accept bachelor
- Maximizing probability of picking the best bachelor?



[T. P Hill, American Scientist, 2009, 97(2)]

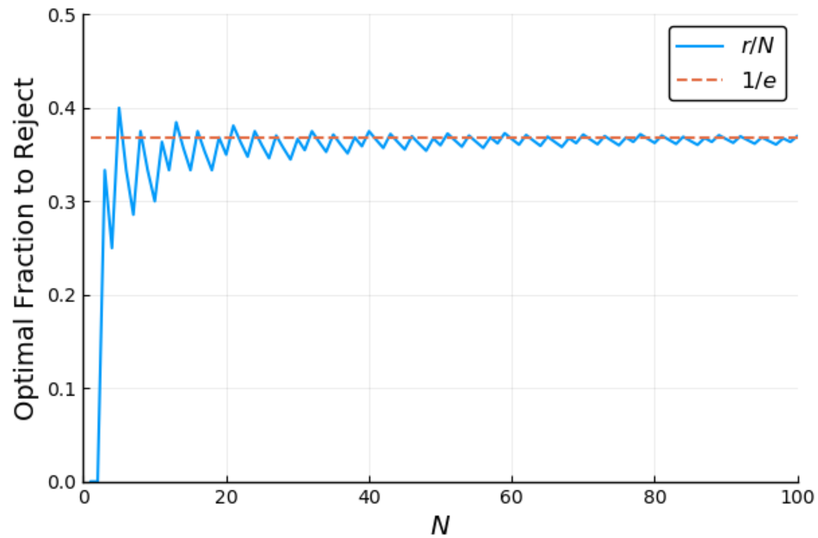
Solution

- Optimal policy:
  - Observe and reject the first r bachelors
  - Pick the first one who i better among the remaining $N - r$
- E.g reject first half, and pick first one who is better among second half – 25 % probability of picking best one

- $r = N/e$ optimal for large N

- At least $1/e \approx 37\%$

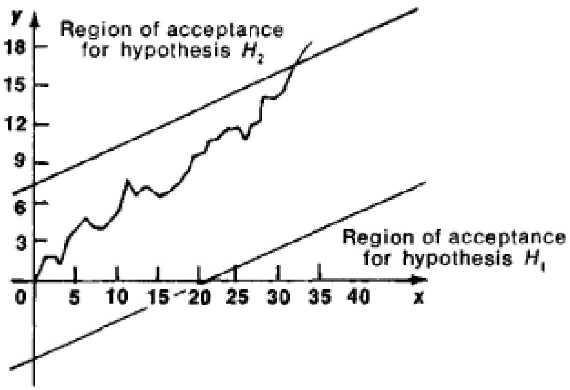
  prob. of picking the best!



[T.S Ferguson, *Who solved the secretary problem?,* 1989, Statistical Science, 4(3)]

Sequential Analysis

- Hypothesis testing with sequential experiments
- After each experiment, decide if
  - Stop, and conclude to accept or reject hypothesis
  - Continue experiments to get better estimate
- Optimal stopping rule is of threshold type



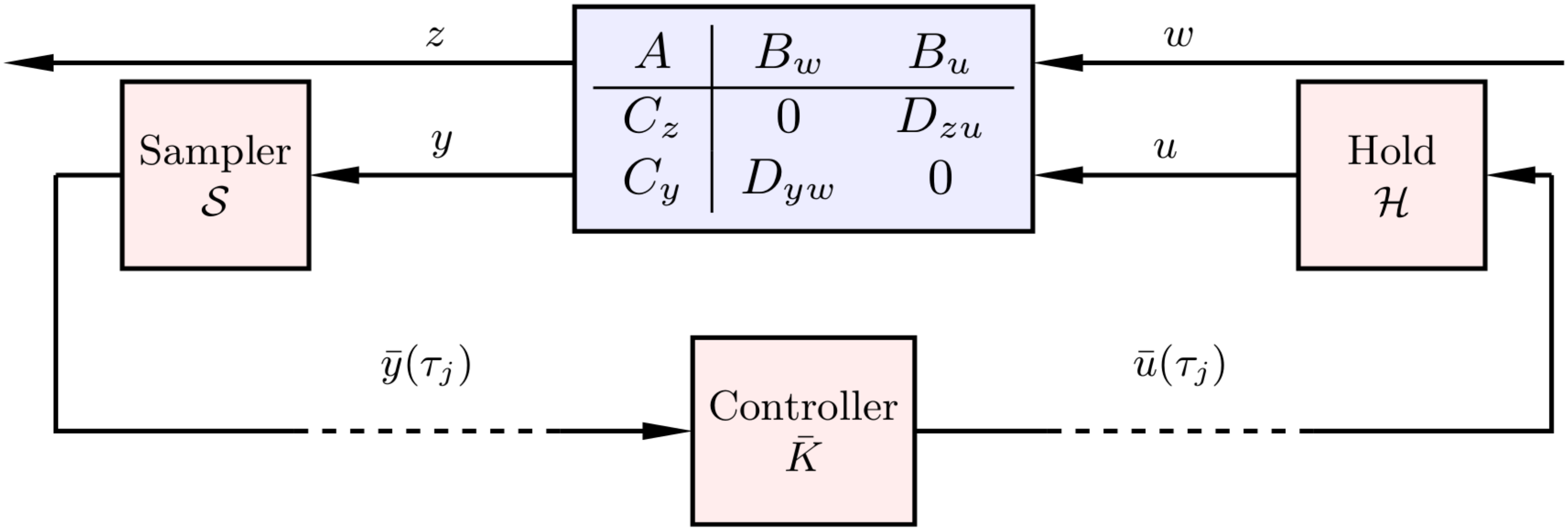
- Pioneered by A. Wald et.al during 2nd world war
- Useful when experiments are costly, e.g clinical trials

[A. Wald, Sequential Tests of Statistical Hypotheses, 1945, Ann. Math. Statist., 16(2)]

Option Pricing

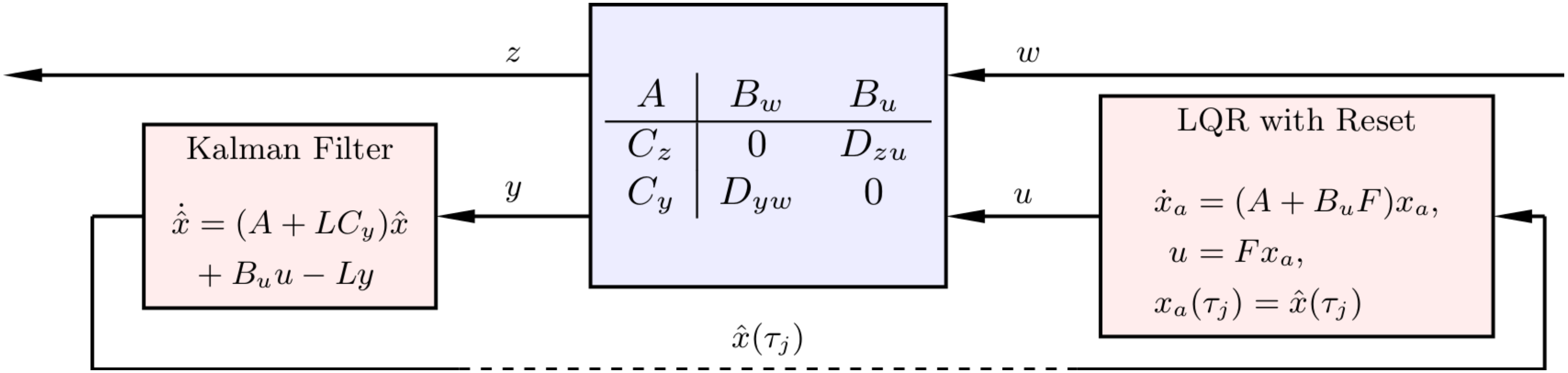
- Option = Financial contract
- Allows you to sell an asset S (e.g a stock) for a predetermined price K (strike).
- American option:
  - At $t = 0$, buy an option that allows you to sell S at any time $0 \leq t \leq T$ for price K
  - If $S(t) \geq K$, option is worthless
  - If $S(t) < K$, net-profit $K - S(t) > 0$
  - Pay-off function: $G(S(t)) = \max(K - S(t), 0)$
- "Fair" option price: $\max_{0 \leq t \leq T} \mathbb{E}[G(S(t))]$
- Black, Scholes, Merton – Nobel Prize '97

LQG-Optimal Sampled-Data Control



$$\min_{\mathcal{S},\mathcal{H},\bar{K},\{\tau_j\}} \mathbb{E}[z^{\mathsf{T}} z] + \rho f$$

Optimal Controller Structure



Optimal for all sampling sequences!

$$\tilde{x} = \hat{x} - x_a$$

$$\mathbb{E}[z^\mathsf{T} z] = \gamma_0 + \mathbb{E}[\tilde{x}^\mathsf{T} Q \tilde{x}]$$

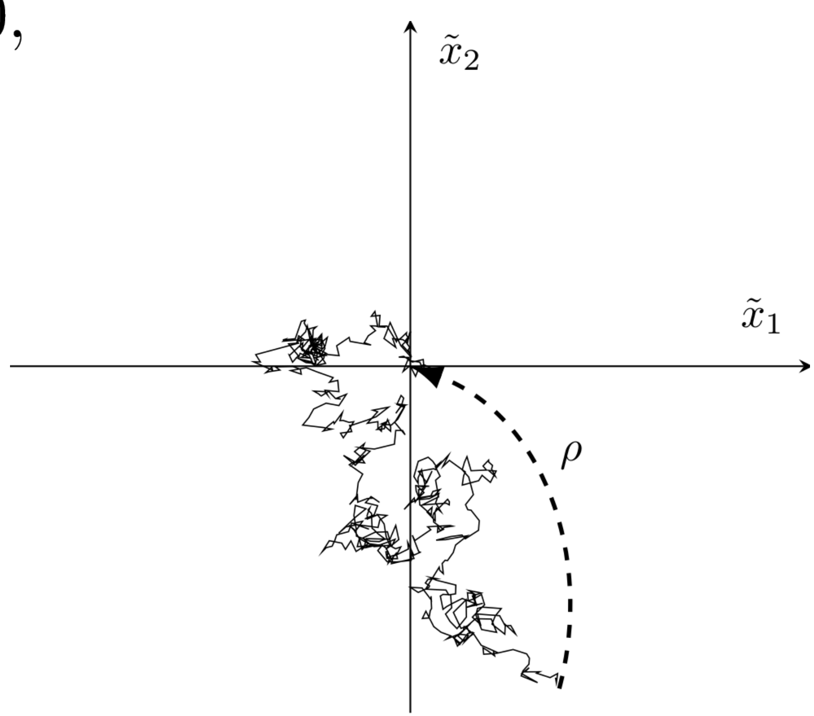
[A. Goldenschluger & L. Mirkin, *On Minimum-Variance Event-Triggered Control,* 2017, Control System Letters, 1(1)]

Optimal Sampling = Optimal Stopping

$$d\tilde{x} = A\tilde{x}dt + dv, \quad \tilde{x}(\tau_j) = 0,$$

$$\mathbb{E}[dv dv^\intercal] = Rdt$$

$$\min_{\{\tau_j\}} \mathbb{E}[\tilde{x}^\intercal Q \tilde{x}] + \rho f$$



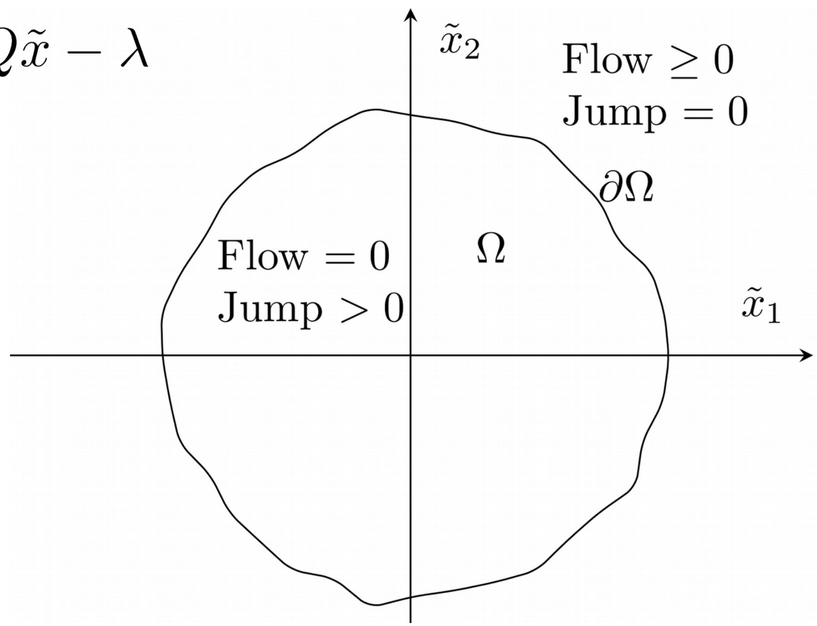
A Free Boundary Problem

Value function satisfies Hamilton-Jacobi-Bellman eq:

$$\min\{\text{Flow}, \text{Jump}\} = 0$$

$$\text{Flow} = \frac{1}{2}\text{tr}(R\nabla^2 V) + \tilde{x}^\intercal A^\intercal \nabla V + \tilde{x}^\intercal Q \tilde{x} - \lambda$$

$$\text{Jump} = V(0) - V(\tilde{x}) + \rho$$



- Optimal cost: λ
- Optimal policy: $\partial\Omega$
- There exists a unique solution

Numerical Solution

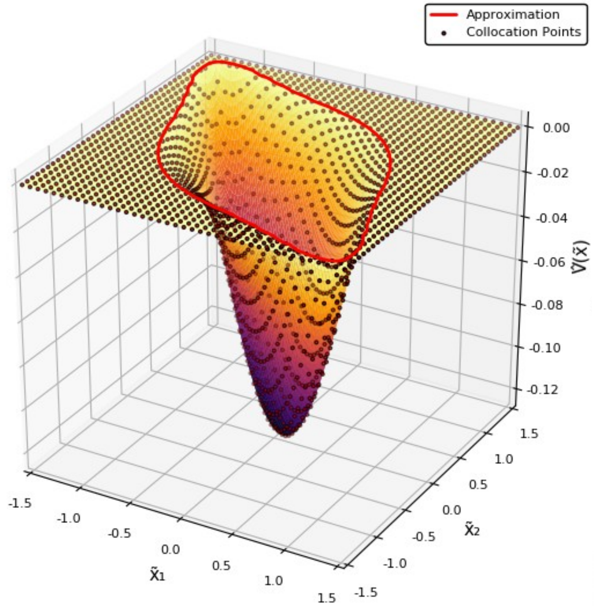
Complementarity form:

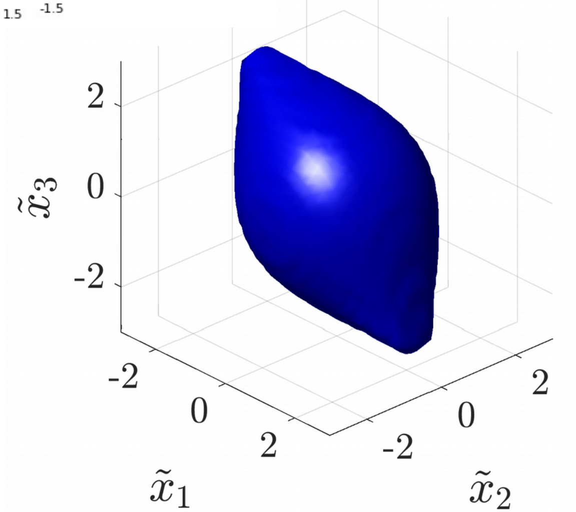
$$\text{Flow} \cdot \text{Jump} = 0,$$
$$\text{s.t Flow} \geq 0, \ \text{Jump} \geq 0$$

Quadratic program:

$$\min_{\hat{v}} \hat{v}^\intercal (\Lambda \hat{v} + b),$$
$$\text{s.t} \ -\hat{v} \geq 0, \ \Lambda \hat{v} + b \geq 0.$$

Works for low-dim (~3D)

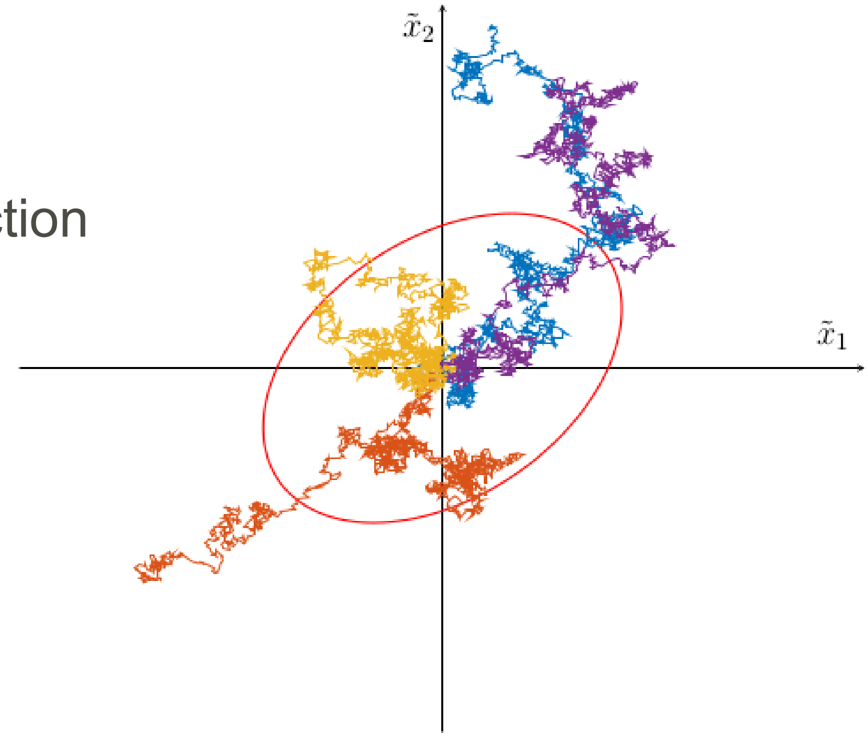




WIP: Monte-Carlo Approach

Motivation:

- Targeting higher-order systems
- Parametrize policy, not value function
- ML/RL methods applicable
- Generalize beyond LQG



[illegible]
descent
and evaluate cost and sampling period

Stopping Policy Representation

Stopping Problem in Discrete-Time

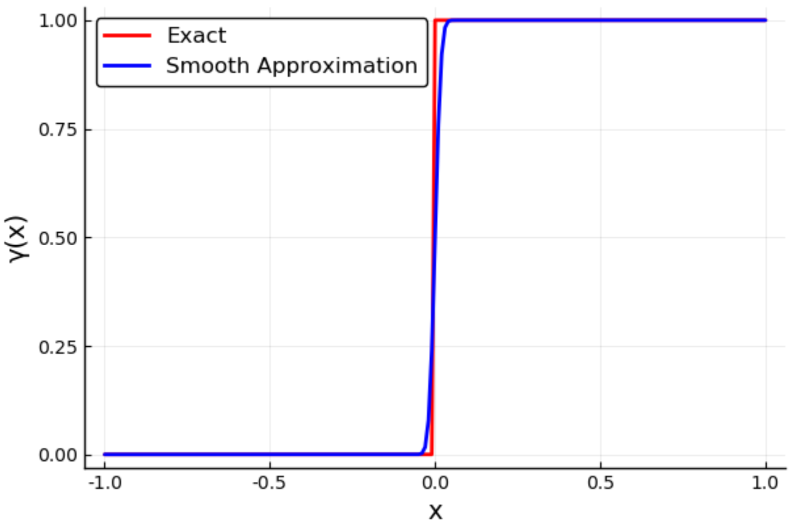
$$\min_{\tau} \frac{\mathbb{E}[\sum_{k=0}^{\tau} x_k^{\intercal} Q x_k] + \rho}{\mathbb{E}[\tau]}$$

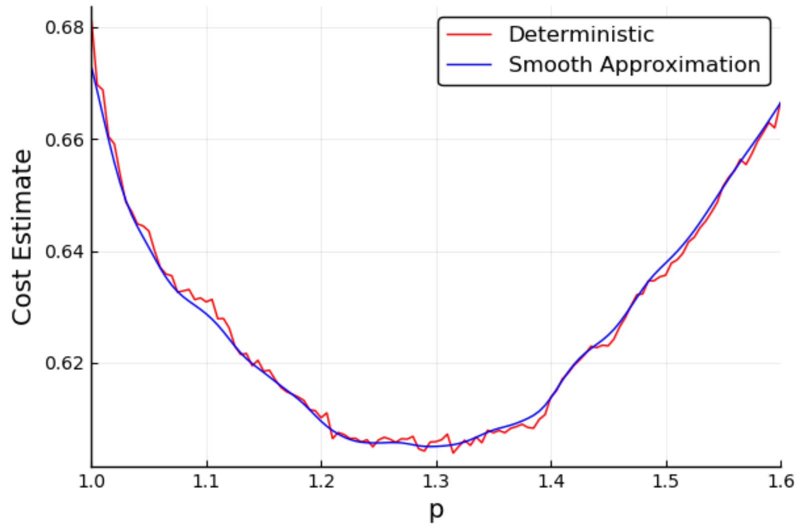
Assume threshold policy:

$$\tau = \min\{k : \gamma_p(x_k) = 1\}$$

$$\mathbb{E}[\sum_{k=0}^{\tau} x_k^{\intercal} Q x_k] = \sum_{k=0}^{\infty} \mathbb{E}[\Big(\sum_{j=0}^{k} x_j^{\intercal} Q x_j\Big) \gamma_p(x_k) \prod_{i=0}^{k-1} (1 - \gamma_p(x_i))]$$

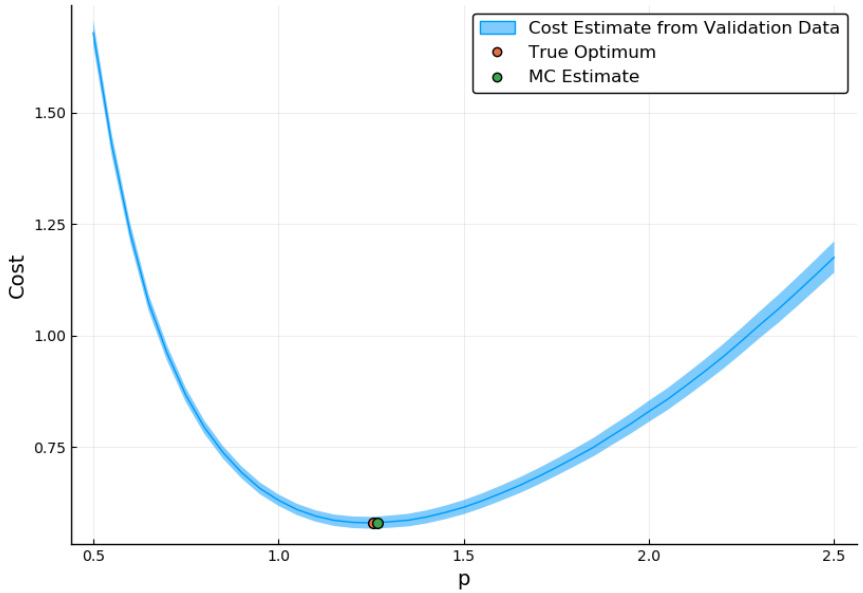
Smoothing

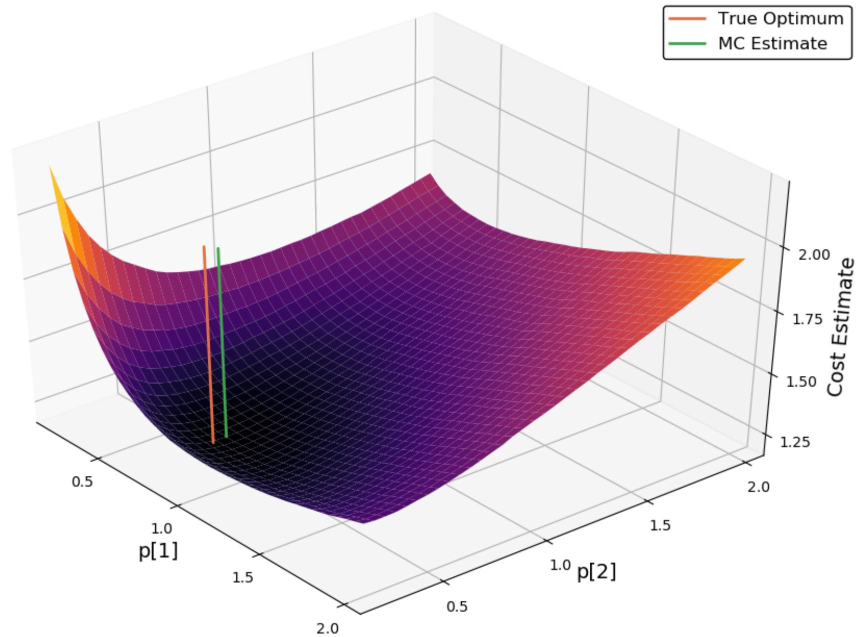




Smooth approximation = Stochastic stopping policy

Some Initial Testing

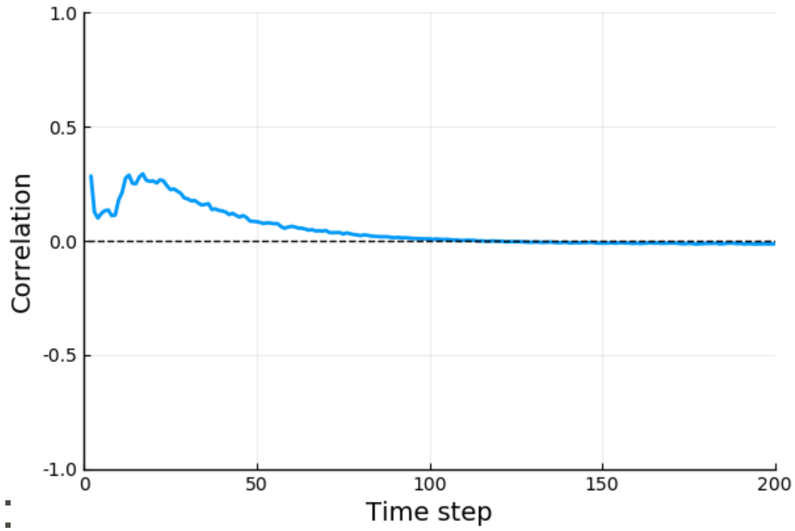




- Low-order examples with known parametrization
- 10,000 trajectories => ~1% error in cost
- Currently validating higher-order systems

Issues

- Slow to compute cost & gradient for large number of trajectories
- Local minimas due to randomness i MC estimate
  - Stochastic Gradient Descent?



- Variance reduction of MC estimate
  - Control variates method?
  - Requires variable with **known** expectation and **high correlation**.

Random variable: Control variate:

$$\left(\sum_{j=0}^{k} x_j^\intercal Q x_j\right)\gamma_p(x_k)\prod_{i=0}^{k-1}(1-\gamma_p(x_i)) \qquad \left(\sum_{j=0}^{k} x_j^\intercal Q x_j\right)$$

Conclusion

- Optimal stopping – stochastic control with two actions {stop,continue}
- Applications:
  - Sequential analysis
  - Fair option pricing
  - Event-based control
  - Finding your future spouse...
- Event-based Sampling = Optimal Stopping
  - PDE Solver methods for low-dimensional problems
  - Current work: Monte-Carlo approach for higher dimensions.