Non-Myopic Approaches to Adaptive Sensing: Challenges and New Results

David Castañón

Boston University

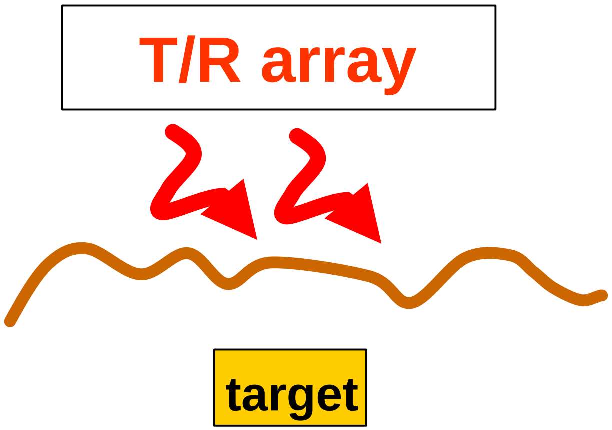
GPR Imaging under Unknown Rough Surfaces

V. Galdi, L. Felsen, D. Castañón

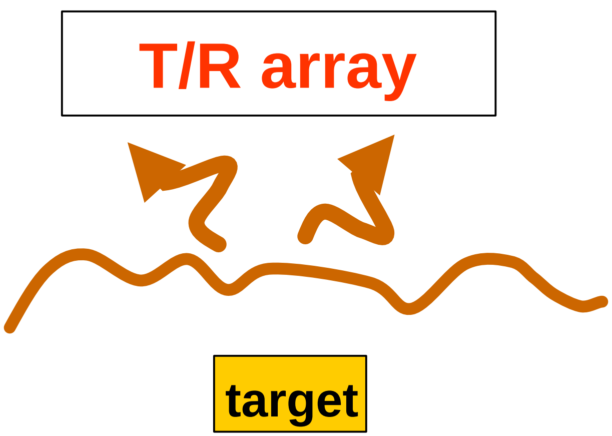
Moderately rough surface underground imaging via short-pulse quasi-ray Gaussian beams, IEEE Trans. Antennas and Propagation, Vol. 51, No. 9, pg. 2304-2318, Sept. 2003, plus previous ...

- **Interested in GPR underground imaging to estimate size and shape for subsurface object classification**
  - Air-ground rough interface is a major source of clutter
- ***Prior work focused on statistical characterization of ground clutter as additive noise***
- Proposed approach**: *Adaptive* compensation for the coarse scale roughness effect**
  - Estimate both surface and subsurface scatterers
  - Use of approximate fast Gaussian beam models for inferencing
  - Subsurface imaging which compensates for surface scattering

Outline of Adaptive Imaging Approach

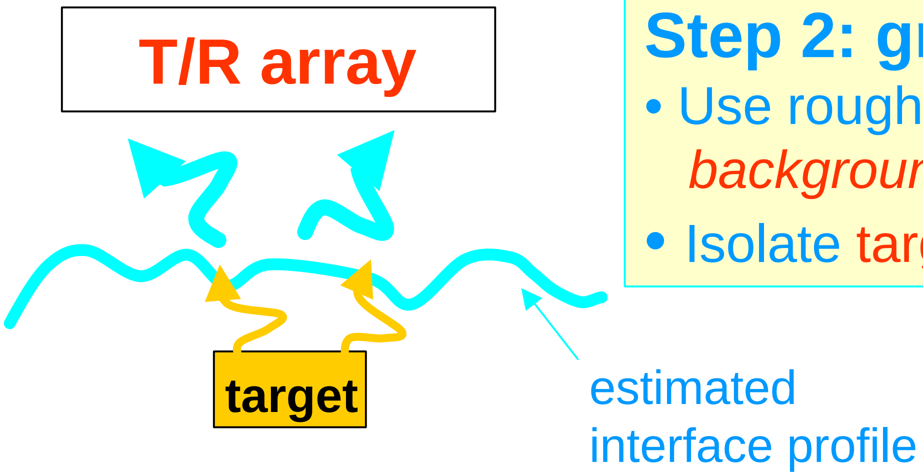


Simplified problem
- 2D geometry (1D rough profile, 2D fields)
- Tapered pulsed illumination
- Weak-contrast (e.g., plastic) target
- Wide-band illumination



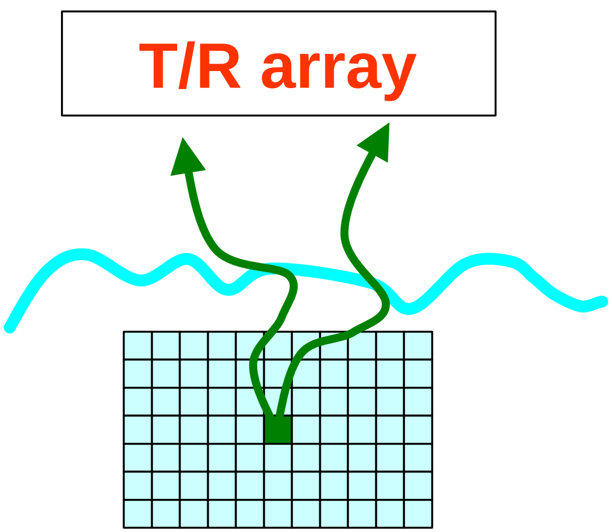
Step 1: rough profile estimation
- Compact (spline) rough profile parameterization
- Fast (beam) forward solver
- Exploit early signal returns from reflection
- Spline parameter estimation via fitting scattered field measurements and model predictions (nonlinear optimization)

Outline (cont'd)



Step 2: ground bounce compensation

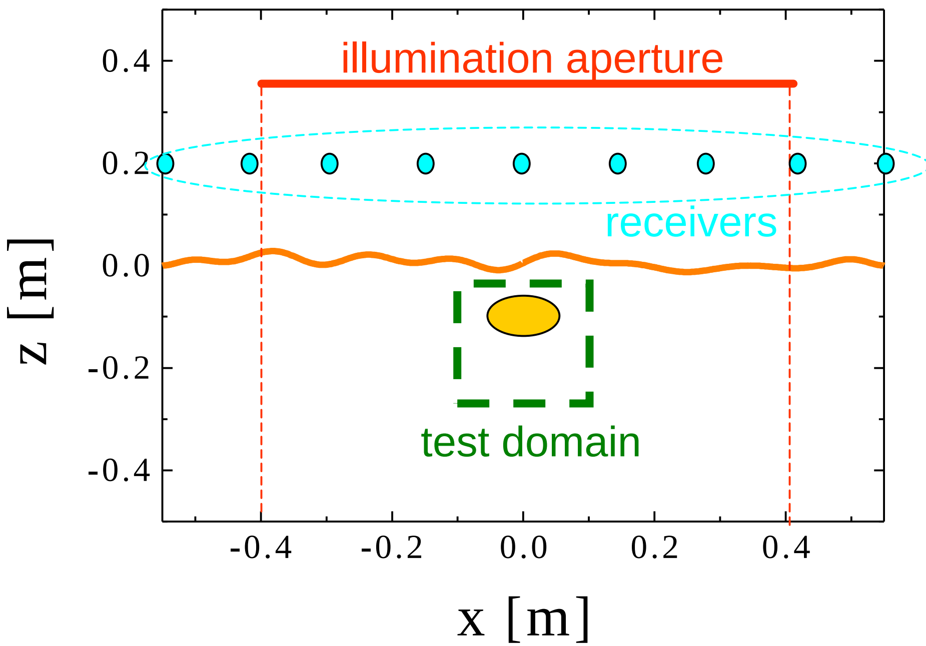
- Use rough interface estimation to compute *background field* (i.e., in the absence of target)
- Isolate target contribution



Step 3: Born underground imaging

- Discretize test domain into pixels
- Use Born approximation (weak contrast)
- Compute field transmitted into test domain and field reirradiated by each pixel via beam algorithm
- Compute object function (i.e., permittivity contrast) at each pixel (linear problem)

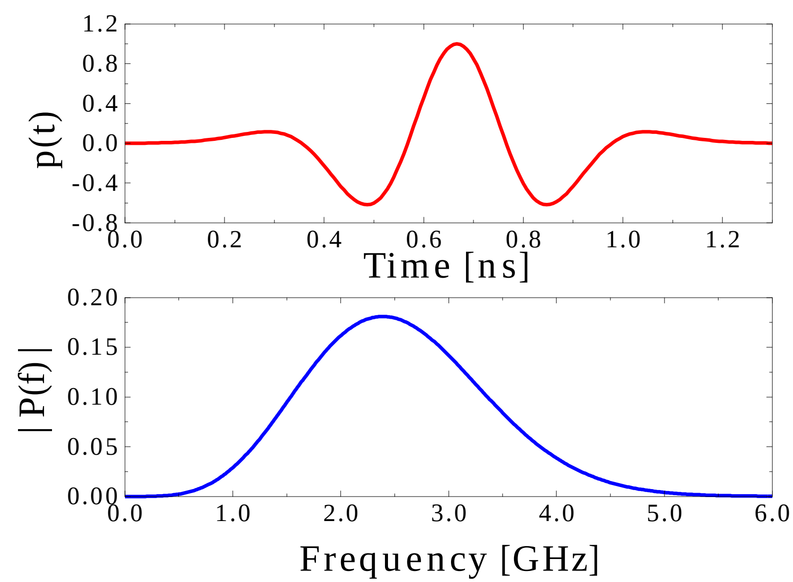
Underground imaging: setup geometry



Soil: $e_r = 3, s = 0$

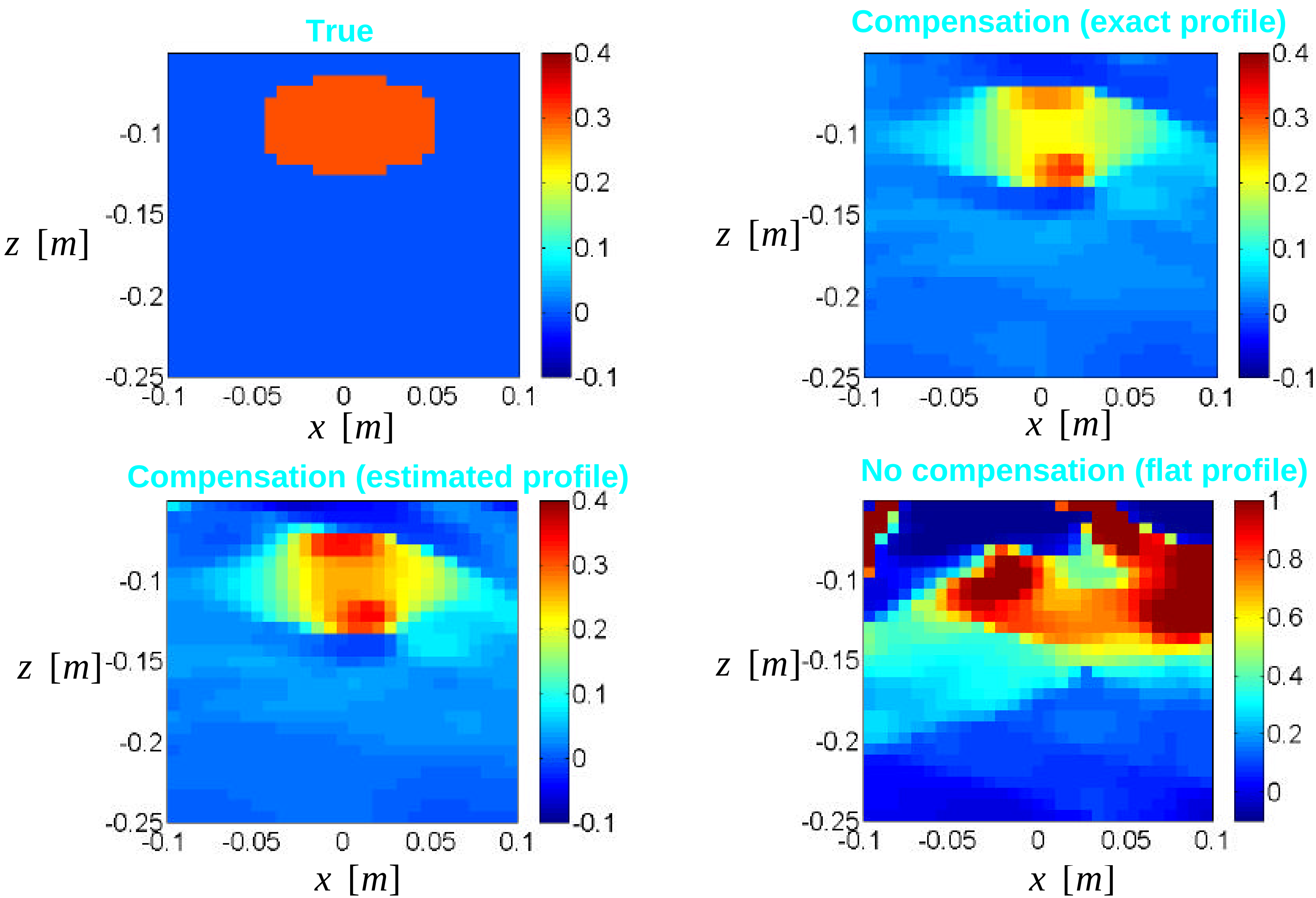
Interface: $a_{\max} = 31^0, h_{\max} = 4cm$

Target: 10cmx6cm ellipse
(center @ 10cm below ground)
$e_r^{(tar)} = 3.3, s^{(tar)} = 0$



Excitation: cosine tapering (.8m aperture)
Data: 11 receivers (20cm above ground),
300 time samples each
Test domain: 20cmx20cm square,
30x30 pixels

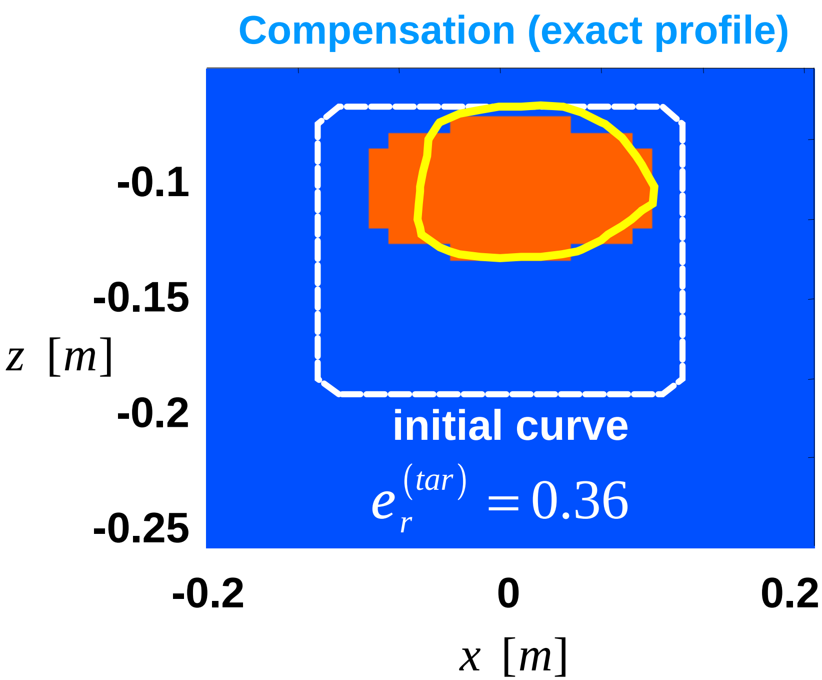
Underground imaging results

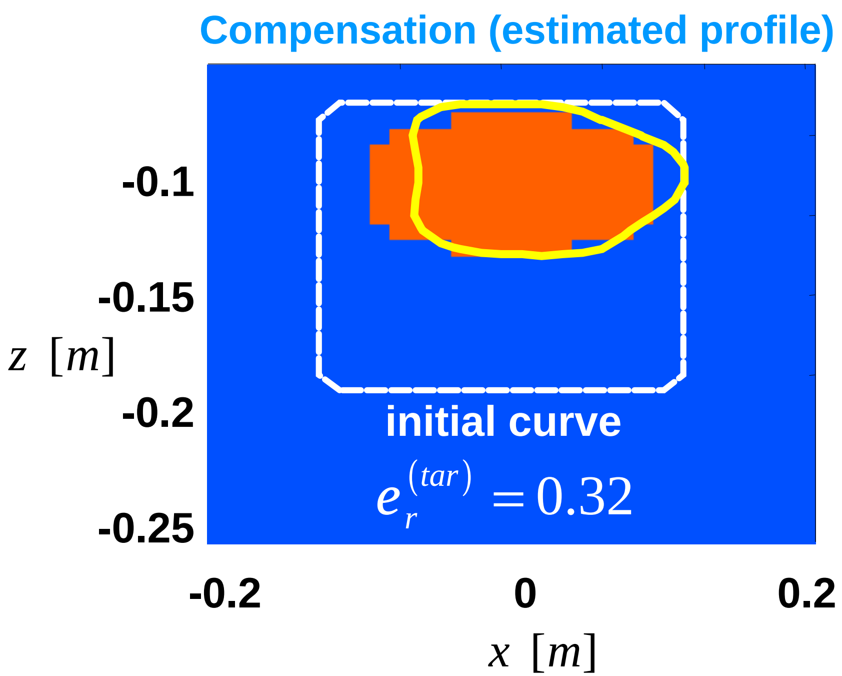


Underground imaging results
(curve evolution)

A Curve Evolution Approach to Tomography,
IEEE Transactions on Image Processing,
V. 12, No. 1, Jan. 2003, pp. 44-57.

Idea: design a gradient flow that attracts any initial closed curve toward the true boundary of the target region



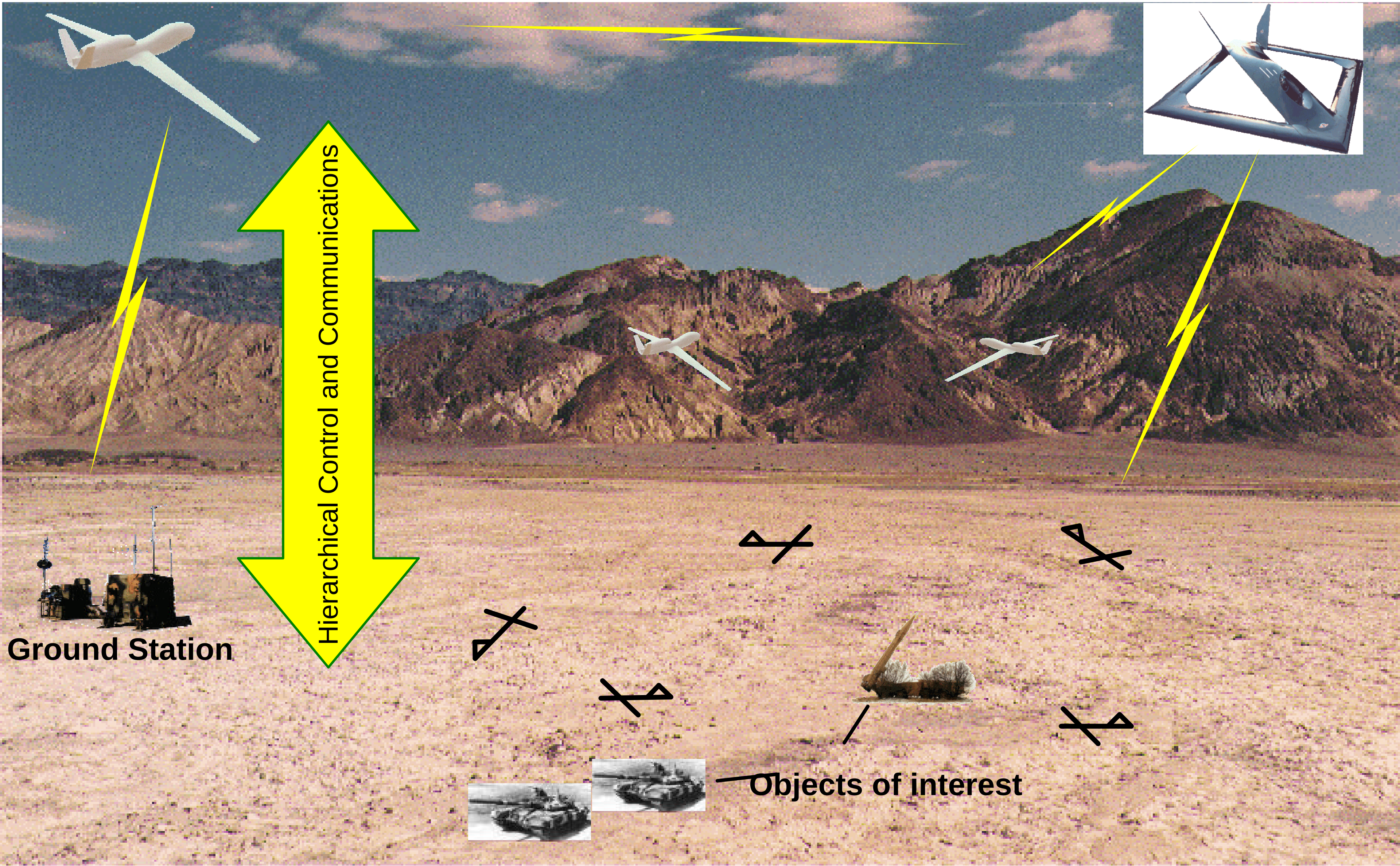


Non-Myopic Approaches to Adaptive Sensing: Challenges and New Results

David Castañón

Boston University

Surveillance and Reconnaissance

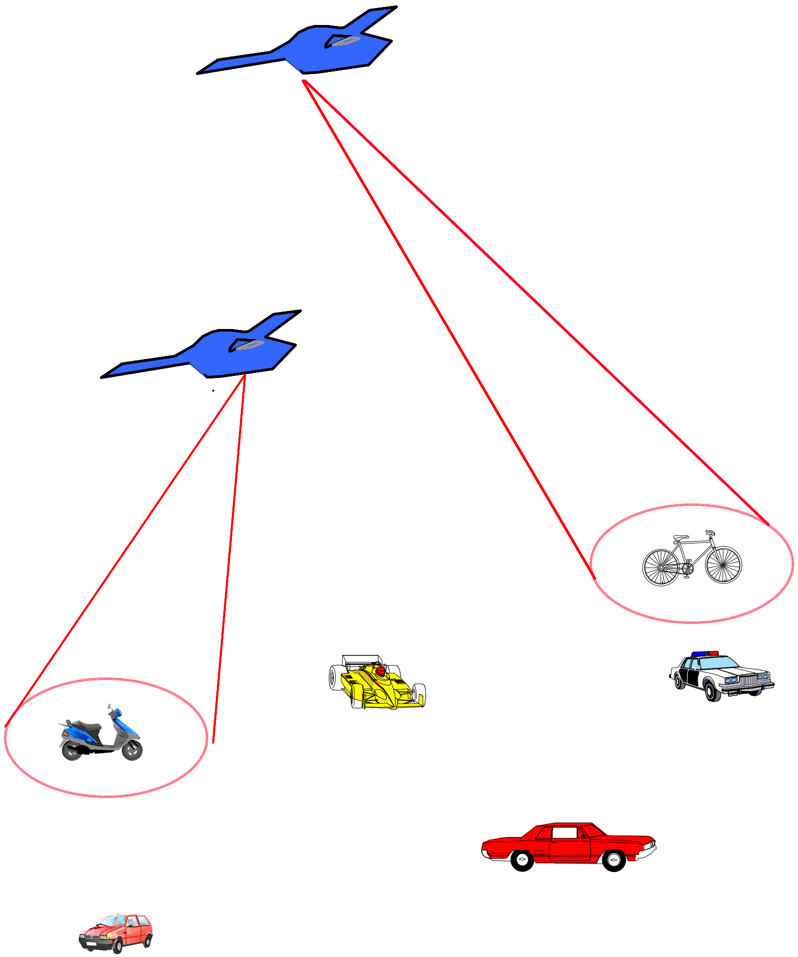


Important Considerations

- **Choices of performance objective**
  - Search, detection, track, classification have different objectives
  - Terminal cost or continuous cost
- **Time Horizon**
  - Finite horizon or infinite discounted horizon
- **Uncertainty Models**
  - Noisy measurements, uncertain object location, motion, type
  - Effectiveness of multiple diverse modes, actions
- **Constraints**
  - Potential visibility constraints depending on time, sensor trajectory
  - Limitation on radar duty, or equivalent energy transmitted
- **Moving vs. stationary objects**
  - Loss of information over time

Focus on Classification

- **Multimode imaging radars**
  - Low Resolution reduced aperture: faster but inaccurate
  - Synthetic aperture radar: slow (integration over many pulses) imaging mode, more accurate
  - Combinations (partial imaging)
- **Problem: Allocation of radar energy over time to obtain information on objects**
  - Accurate classification
  - *Resource use and quality* depend on mode, geometry, sensor, object type
- **Problem constrained by object visibility, radar duty cycle allocated to function**



Sensor Management Objective: Manage Information Acquisition

- **Reduce uncertainty**
  - Representation of uncertainty: Probability measures over object types
  - Measure: entropy (different types), information distances or pseudometrics (Kullback-Leibler, …)
  - When? At a given time, for all time with future discounting, …
- **Support better decisions**
  - Improved classification
  - Improved parameter estimation
- **These are not the same!**
  - No need to separate irrelevant alternatives…
  - Want adaptive solutions: exploit past information for selection of future actions

Classification Problem: Notation

- **N objects, 1 per site, with K possible types**
  - x_i is type of object i; Probability vector over types for object i: π_i
- **Multiple sensors J, M sensor modes per sensor**
  - Each mode m from sensor j produces finite valued observations y_{ijm} of object i
  - Likelihood $P(y_{ijm}|x_i, u_{ijm})$ known
  - Each use of mode m on object i from sensor j requires R_{ijm} sensor time
  - Maximum of one mode per sensor at a given time
- **Decisions: which modes are assigned to each object at a given time**
  - Must remain in mode and assignment for required time
  - $u_{ijm}(t) = 1$ if mode m from sensor j is applied to target i at time t
- **Finite total observation time per sensor j: C_j**
- **Information collected at area i only if observed:**
  - Observations conditionally independent over time and area given true value x_i and mode m, sensor j

Classification Problem: Uncertainty Representation and Dynamics

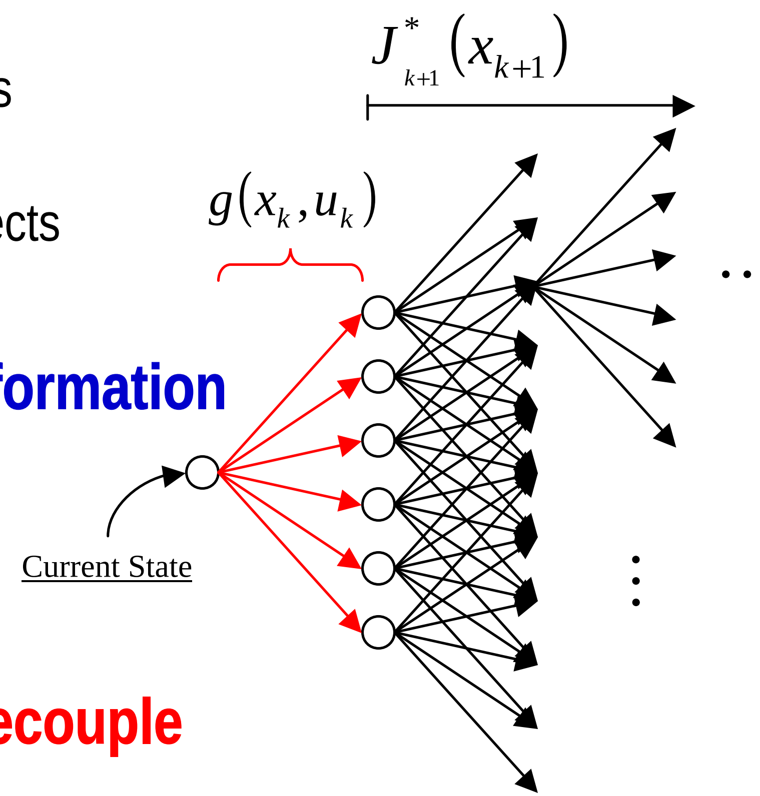
- **Uncertainty representation**
  - N objects, each of k types
  - Assume independence of object types
  - Assume conditional independence of measurement outcomes across objects
  - Sufficient statistic: information state Π (t) = {π_1(t), …, π_N(t)}, where π_i 2 S_k is conditional probability of object i's type given past information
- **Dynamics**
  - Act locally on objects: Use Bayesian inference
  - Likelihood of measurement outcomes given modes, true object types: p(y|x_i,u)

$$
\begin{aligned}
\pi_i^k(t+1) \;&\equiv\; P(x_i = k | Y_i(t+1)) \\
&= \frac{\pi_i^k(t) \prod_{j,m} P(y_{ijm}(t) | x_i = k, u_{ijm}(t))}{\sum_{k'} \pi_i^{k'}(t) \prod_{j,m} P(y_{ijm}(t) | x_i = k', u_{ijm}(t))}
\end{aligned}
$$

General Framework: Dynamic Decision Problem

- **Proposed Approach: Solve as dynamic decision problem with uncertainty and feedback**
  - Optimization over time with noisy information
- **Formulation: resource-constrained Partially Observed Markov Decision Problem (POMDP)**
  - States: true types plus available sensor resources
  - Observations: data from sensors
  - Controls: Sensor modes assigned to specific objects
  - State dynamics: mostly deterministic
- **Convert to Perfectly Observed MDP in Information State**
  - States $\Pi(t)$ 2 S_k^N
  - Bayesian dynamics

… Intractable! Need approaches that decouple across objects

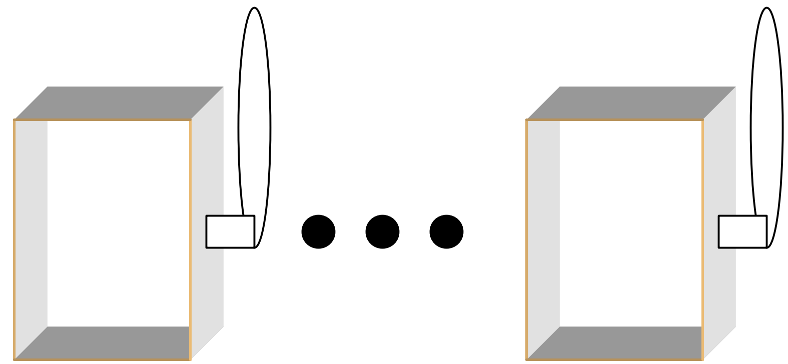


Computational Issues

- **State Space**
  - Typical state per site: probability distribution over site content
  - True state is dimension is product over sites, plus sensor states
  - Problem in storing as well as computing optimal strategy
- **Decision Spaces: Combinatorial (deterministic instances NP-hard)**
  - Deterministic equivalents: multidimensional knapsack, vehicle routing problems, scheduling problems.
- **Expectations**
  - Computing expected performance of a policy from a state averages over sequences of dependent events
  - Exponentially many sequences possible

Special Dynamic Models

- **Multi-armed bandit problems**
  - Infinite horizon
  - Discounted future cost objective
  - Only one mode per sensor
  - Same resource use for all modes, sensors
  - Decision: u(t) = machine to play at time t
  - State: x_i(t) = state of machine i at time t
  - Unplayed machines do not change state, played machine transitions randomly

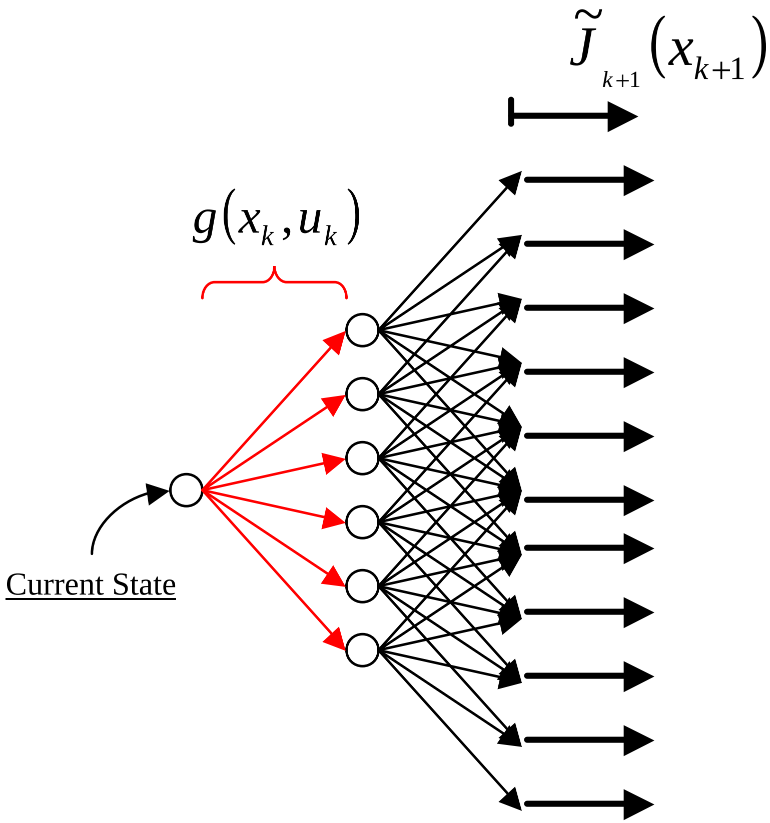


$$J = \sum_{t=1}^{\infty} \alpha^t \sum_{i=1}^{N} c_i(x_i(t))\delta(u(t) - i)$$

- **Theory (Gittins): There exist index functions $I_i(x_i)$ computable only from a and $c_i(x_i)$ such that the optimal sensing strategy is to play machine with largest index given current states**
  - Decoupling among objects

Practical algorithms: Approximate Dynamic Programming

- **Stochastic Dynamic Programming**
  - *Optimal* approach to solving dynamic, stochastic decision problems
  - *Solves* decision problem for all future states
- **Approximate Dynamic Programming**
  - Focuses on solving decision problem from current state and likely future states
  - Increased focus on real-time computation using receding horizon control
  - Key issue: approximation of future cost



Approximation Approaches for Future Costs

- **Simple functions**
  - E.g. 0 yields greedy approaches
- **Functional approximations trained off-line based on on-line features (e.g. Neural Nets)**
  - Main difficulty: Learning functional approximations which are accurate for classes of problems vs. problem instances
  - Reinforcement learning (TD-lambda, Q-learning) instead of direct training because correct actions are not known
- **On-line evaluation of suboptimal policies (Rollout algorithm)**
  - Can use either open-loop policies where future controls do not depend on future states (model-predictive control) or closed-loop policies
  - Key: Determining a suboptimal policy, and evaluating expected future performance over horizon.
    - Evaluate using Monte Carlo simulation or analytical approximations
- **Model Predictive Control**
  - Use approximate optimization model for current state to simultaneously solve for current best policy plus approximate cost-to-go
  - Needs good approximation

Missing Element

- **Have many alternative approximation approaches**
- **There is a need for theory that characterizes achievable performance to determine when approximations are accurate enough**
  - Performance bounds
  - Sensitive to decision problem parameters
  - Similar to Cramer-Rao bounds for parameter estimation algorithms
- **Recent result: algorithm for computing bounds on optimal performance for classification problems**

Lower Bound Formulation

- **Key assumptions: POMDP structure**
  - Finite number of possible classes per object
  - Finite number of objects
  - Finite number of possible measurement values per sensing action on an object
  - Bounded resources per sensor
  - Conditionally independent measurements across objects, sensors, time, given true object classes
  - Simplify notation: 1 sensor, multiple modes
- **Goal: accurate classification at end of horizon**
  - Cost: Minimize expected weighted classification error at final action time T
    - Classification decision for object i: $v_i(T)$
  - Can include stage costs, other forms

$$J = \sum_{i=1}^{N} E\{\min_{v_i} c(x_i(T), v_i(T))\}$$

POMDP Constraints

- **Constraints:** *for all observation sample paths and all sensors*
  - **Cannot exceed total sensor capacity**

$$\sum_{t=0}^{T-1} \sum_{i=1}^{N} \sum_{m=1}^{M} R_{im} u_{im}(t) \leq C \quad \text{(sensor resource constraint)}$$

  - **Major limitation: one constraint per sample path!**
- **Lower Bound: Relax constraints to average value to get new POMDP problem**

$$\sum_{i=1}^{N} \sum_{m=1}^{M} E\{R_{im} u_{im}(t)\} \leq C$$

  - No constraints total actions per interval (optimistic)
  - Expands admissible strategies, so obtains *lower bound*
- **Also expand strategy space of relaxed problem to allow for mixed strategies**
  - Simplifies the integer programming nature of the relaxed problem
  - Convexifies problem and *maintains lower bound*

Definitions

- A ***state feedback policy*** **is a map from the current information state of all objects at a given time into a set of sensor decisions at that time**
  - Stochastic dynamic programming: there is an optimal solution in the set of information state feedback policies

- A ***local state feedback policy*** **is a state feedback policy where the feedback actions chosen for object i at time t depend only on the local information state of object i and not on the information state of other objects**

- **Local state feedback → will lead to local subproblems**
  - Identify what to do for each object

Duality

- **Lagrangian, for l >= 0:**

$$J(\lambda,\gamma) = E_\gamma\{\sum_{i=1}^{N}[c(v_i,x_i) + \lambda \sum_{t=0}^{T-1}\sum_{m=1}^{M} R_{im}u_{im}(t))\} - \lambda C$$

- **Lower bounds given by weak duality**

$$\min_\gamma J(\lambda,\gamma) \le \max_{\lambda\ge 0}\min_\gamma J(\lambda,\gamma) \le \min_\gamma J(\gamma)$$

- **Lagrangian problem is almost separable over objects**
  - Coupled by feedback strategies!
  - Initial lower bounds: weak duality: For all λ,

Convexity

- **Convexity of objective in strategy mixture**
  - Expected cost of strategy γ

$$J^{\gamma} = E^{\gamma}[\sum_{i=1}^{N} c(x_i, v_i(T))] = \sum_{i=1}^{N} J_i^{\gamma}$$

  - Expected cost of mixture strategy with mixing probability p(γ)

$$J^{p} = \sum_{\gamma} p(\gamma) E^{\gamma}[\sum_{i=1}^{N} c(x_i, v_i(T))] = \sum_{\gamma} \sum_{i=1}^{N} p(\gamma) J_i^{\gamma}$$

→ Achievable performance is *convex hull* of performance of pure strategies

  - Expected resource use for sensor j of same strategy

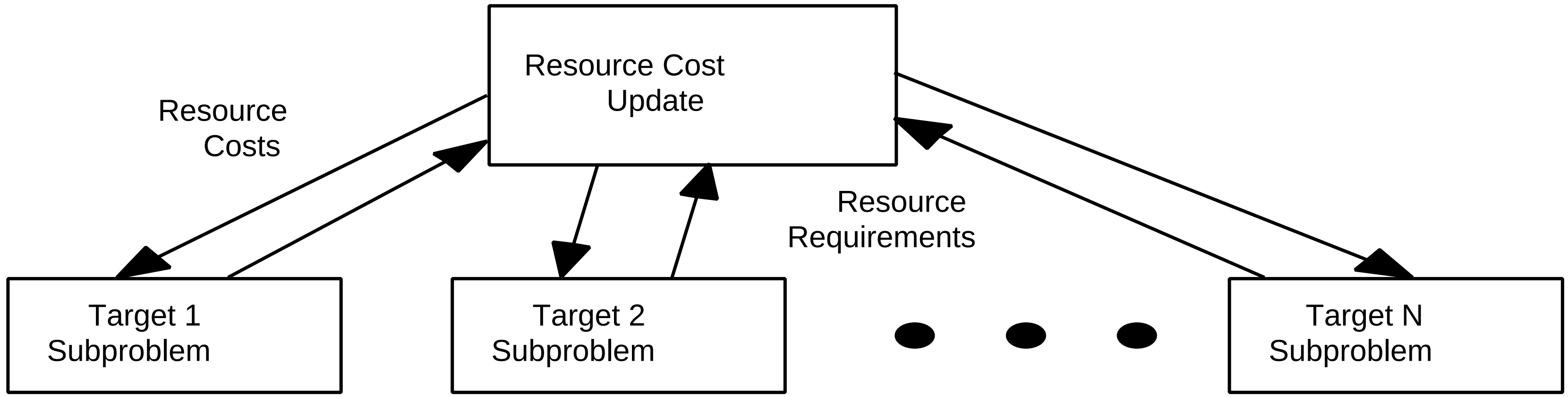
$$R_j^{p} = \sum_{\gamma} p(\gamma) \sum_{i=1}^{N} \sum_{m=1}^{M} E^{\gamma}\{R_{ijm} u_{ijm}(t)\} = \sum_{\gamma} p(\gamma) \sum_{i=1}^{N} R_{ij}^{\gamma}$$

Principal Results

- **Given any state feedback policy, there is a mixed strategy using only local state feedback policies that achieves the same performance and the same expected resource constraints**
  - Prove by construction
  - Exploit additive decomposition across objects of performance and constraints
  - Use mixed strategies to average effect of information on other objects

- **Therefore, the optimal value of the relaxed problem is achievable with a mixture of local feedback strategies…**
  - Implication: Relaxed problem is nearly decoupled across sites, with only one average constraint per sensor
  - Hierarchical decomposition algorithms
  - Can get fast exact algorithms for approximate POMDP problem

Hierarchical Approach



$$\min_{p} L(p,\lambda) = \sum_i \min_{p_i} p_i(\gamma_i)(J_i^{\gamma_i} - \sum_j \lambda_j R_{ij}^{\gamma_i}) + \sum_j C_j \lambda_j$$

Note: minimum is achieved in pure strategies for each price vector λ

- **Resource costs: dual variables for consuming sensor time for different sensors**
  - Replace sample path constraints with average constraint
  - Subproblems solved optimally using POMDP single site algorithms
  - Resource costs obtained from dual optimization
  - NK-dimensional POMDP reduced to N single object K-dimensional POMDPs + dual

Algorithm: Local Problem

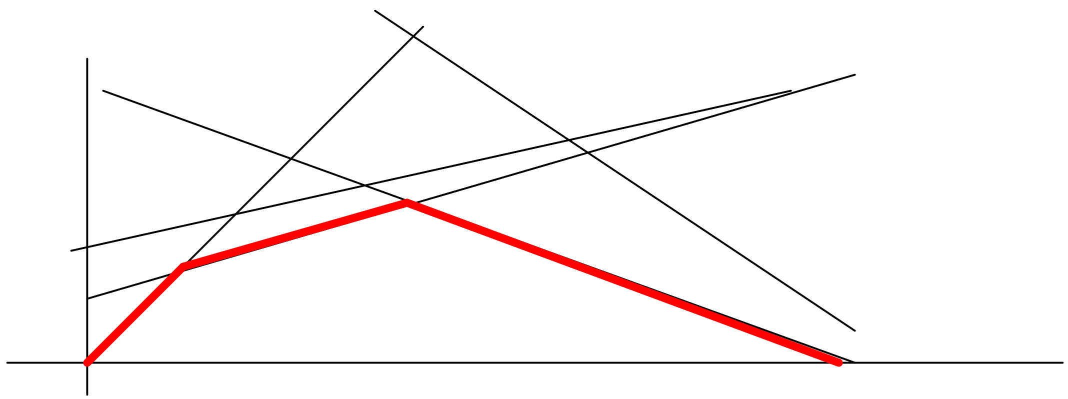
- **Need to find sensor activities per site given sensor resource "prices"**
  - Solution of single site Partially-Observed Markov Decision Problem with sensor costs
  - Small dimension to information state: number of object types
- **Major assumptions:**
  - Finite number of observation values for any mode
  - Finite horizon: number of recourse actions per object
- **Basic idea: for finite horizon problems, one can write a representation for cost to go as a function of information state using stochastic dynamic programming**

$$J^*(\pi_i(t)) = \min_{u_{ijm}(t)} E[J^*(\pi_i(t+1)) | u_{ijm}(t)]$$

POMDP Background

- **Given information state $P(t)$, there is a unique representation for the cost to go from stage *t***
  - Smallwood & Sondik (71)

$$J(\pi_i(t)) = \min_k < \gamma_k, \pi_i(t) >$$



- **There is a simple backward recursion for computing the support hyperplanes g_k**
  - Based on Dynamic Programming
  - Exploits finite values of possible observations, finite actions

POMDP Algorithms

- **Important aspect: Control exponential growth of support hyperplanes**
  - Each backward induction can increase the number of hyperplanes by a factor of number of measurements times number of controls
- **Practical algorithms**
  - Prune redundant hyperplanes at each stage using heuristics, linear programming
- **Useful algorithms**
  - Witness algorithm (Cassandra, '98)
  - Incremental Pruning (Cassandra, '99)
  - Software available

Dual Solution: Finding Optimal Prices

- **Approach: Optimize dual function**
  - Given current set of prices, get optimal local strategies γ_i, find a subgradient

$$\frac{\partial}{\partial \lambda_j} L(\gamma, \lambda) = C_j - \sum_i R_{ij}^{\gamma_i}$$

- **Use nondifferentiable optimization techniques to optimize dual**
  - Slow for many constraints

Dual Solution: Column Generation

- **Alternative approach: Use linear programming theory for dual solution of prices (Yost-Washburn)**
  - Faster convergence for multisensor problems
  - Avoids nondifferentiable optimization algorithms
- **Idea: Each set of prices l^k generates a candidate strategy g^k_i for each object i**
  - Results in expected total resource use R^k_j for sensor j and expected performance J^k across all objects
  - Given a finite set of ***previous*** strategies indexed by k, find best linear combination of strategies that satisfy expected resource constraints

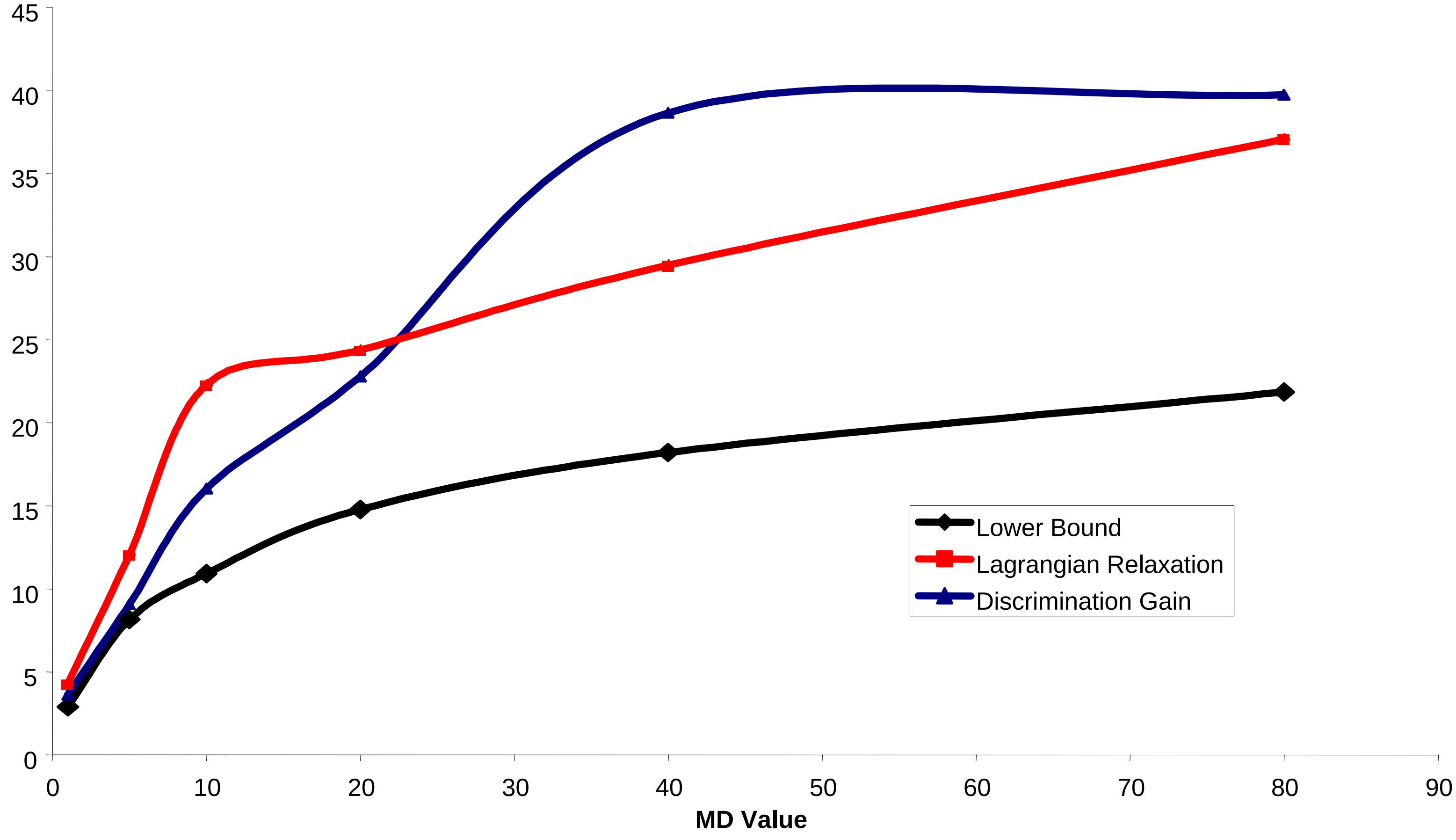
$$\min \sum_k \alpha_k J_k$$
$$\sum_k \alpha_k = 1, \quad \sum_k \alpha_k R^j_k \leq C^j$$

  - Note: this is an upper bound: considers mixtures over smaller set of strategies
  - Optimal linear programming prices used to generate new candidate strategy
    - Continues until no new strategies generated

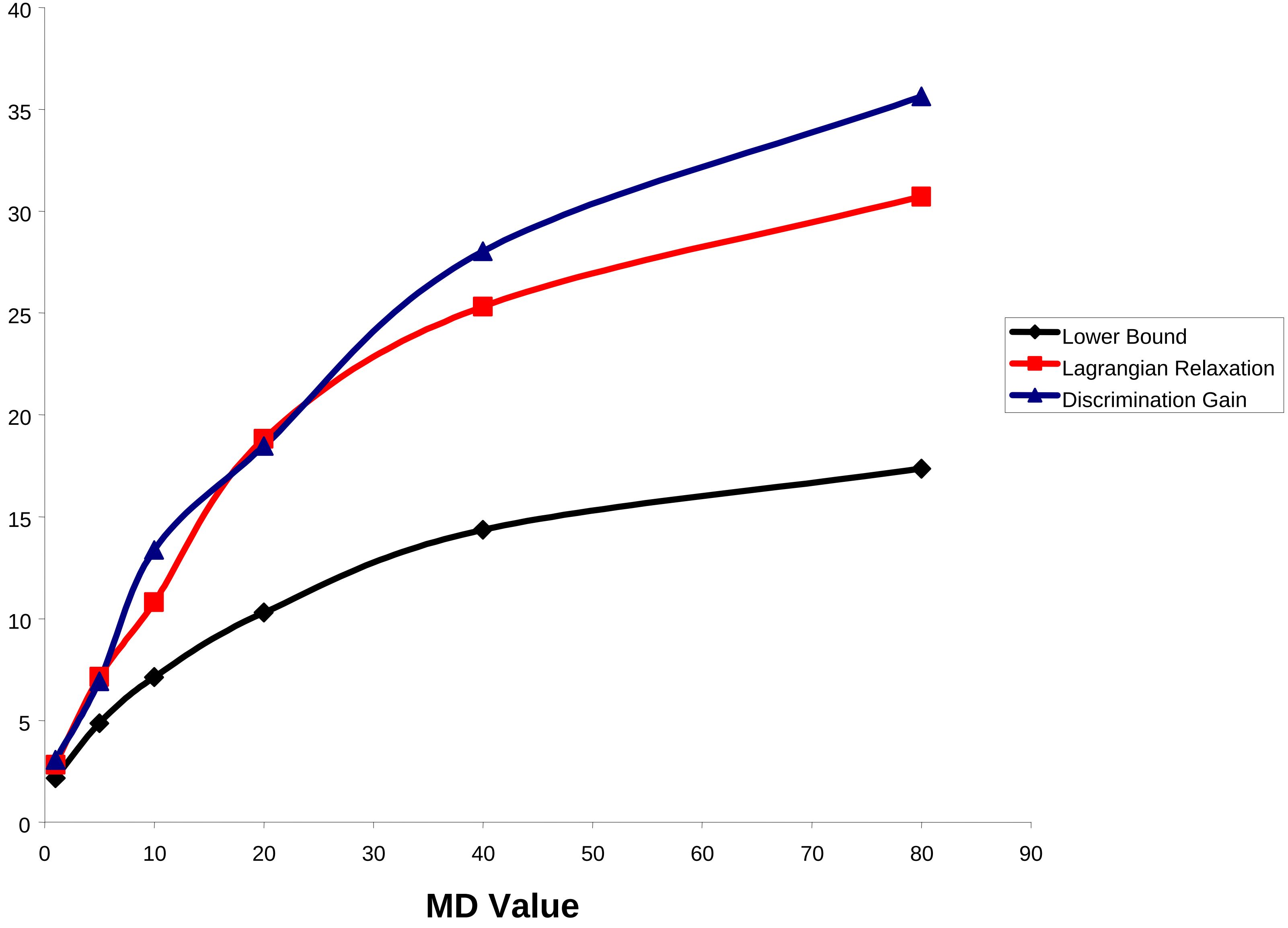
Numerical Examples

- **Classification problem**
  - Objects: 100 sites with 3 types of objects: cars, military vehicles, trucks
  - Sensors
    - Two radar modes: low-baseline SAR and high-resolution SAR, No visibility restrictions
    - High res. SAR mode requires 5 seconds, Low Res 1 seconds
    - Binary-valued measurements: military or not military
    - Low-Res separates cars from others, trucks; High-Res separates others, cars and trucks
  - Constraints: 300 – 700 seconds of radar time
  - Objective: MD for error of declaring military vehicle as car or truck, 1 for declaring car or truck as military vehicle, all after terminal time
  - Prior distribution: 10 % military vehicles, 20 % trucks, 70 % cars
- **Algorithms:**
  - Dynamic algorithm based on 3 sensing actions per object lookahead horizon
  - Greedy, selecting sensor action to maximize Kullback-Leibler discrimination gain per sensor time
  - Lower bound

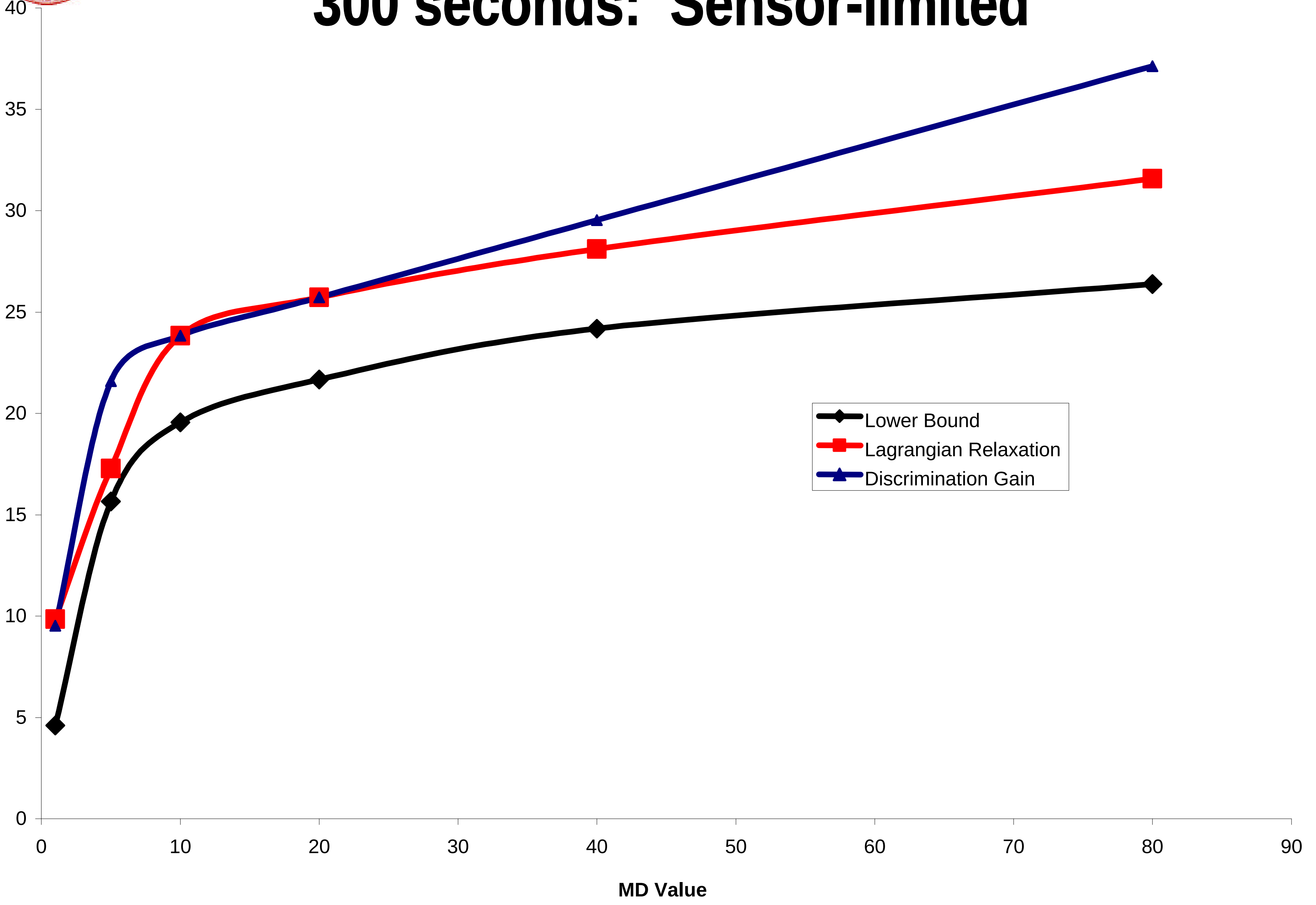
500 seconds: Sensor-balanced



700 seconds: Sensor-rich



300 seconds: Sensor-limited



Conclusion

- **Investigated *new formulation* for closed-loop sensor management for classification with multiple sensors, multiple modes**
  - Finite horizon POMDP
- **Developed *exact solution* of approximate stochastic dynamic program**
  - Approximate SDP is lower bound to original Bayes' cost
  - Hierarchical iteration between single object POMDP subproblems and resource coordination using dual variables
  - Combines ideas from linear programming, nondifferentiable optimization and stochastic control
- **Lower bound provides reference for evaluation of approximations**
  - Can also be used as part of approximations
- **Future work: Extend approach to include functions such as search and target tracking**