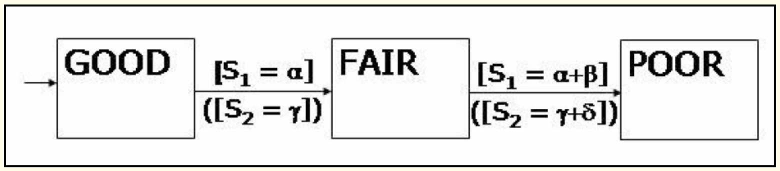
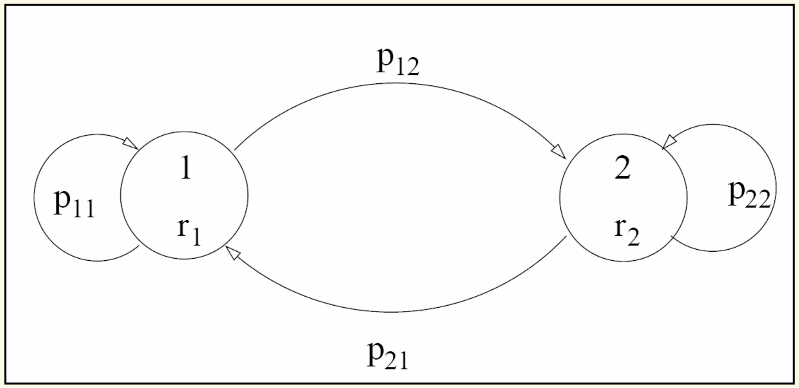
Approximation Methods and Safety Analysis for Stochastic Processes

David Thorsley
University of Washington, Seattle
October 27, 2008

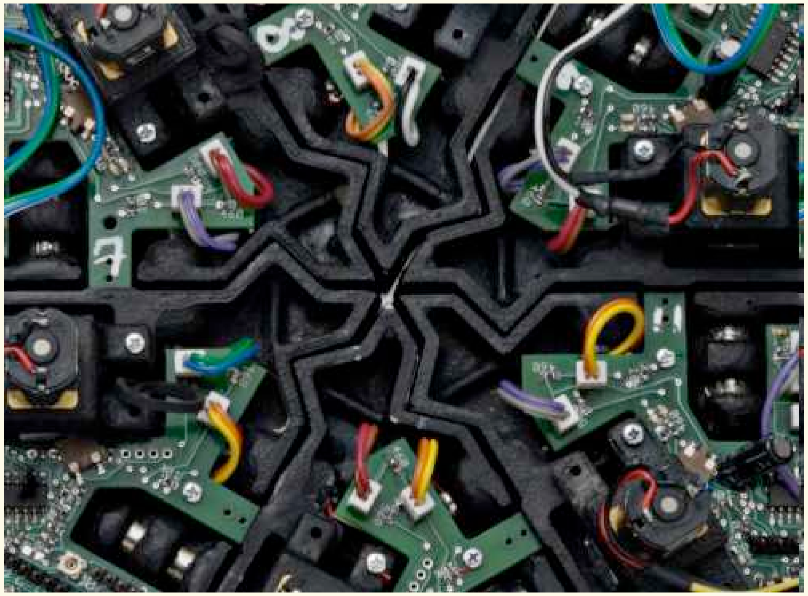
Stochastic phenomena and models for distributed systems



stochastic failure models for
parts and subsystems



stochastic games



stochastically interacting components

unpredictable behavior of enemy agents

Issues in modeling and analyzing stochastic behavior

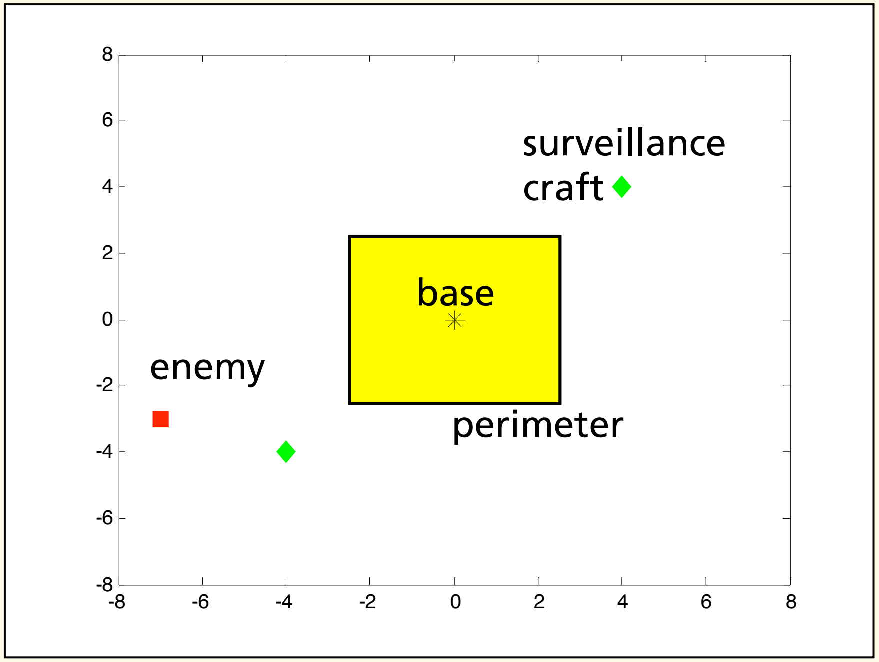
- observation of stochastic processes is limited and decisions have to be made based on limited observations
- stochastic models of phenomena (e.g. failure models, enemy models) have to be estimated from data
  - the most realistic models may be difficult to analyze

- this talk has two parts to address these two main issues:
  - calculating failure probabilities based on limited observations using a fixed class of stochastic model (continuous time Markov processes)
  - using Wasserstein pseudometrics to approximate arbitrary stochastic processes by simpler processes that are easier to analyze

Approximation Methods and Safety Analysis for Stochastic Processes

Part I: Diagnosing Special Events in Continuous Time Markov Processes

Example: Stochastic Pursuit and Evasion



The base will go on alert if the probability of the enemy having breached the perimeter crosses a given threshold.

Objective: set the base on alert before the enemy reaches the base.

Need to know: What is the probability of the enemy having crossed the threshold given the observations the surveillance craft generate along their flight paths?

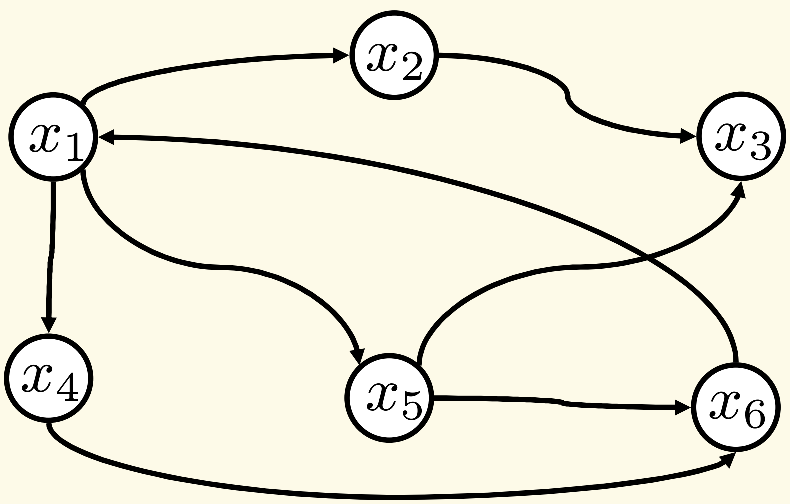
System Model

- Continuous-time Markov process S = (X,Q,π)
  - X is a finite set of states
  - Q is a transition rate matrix
  - π is an initial distribution

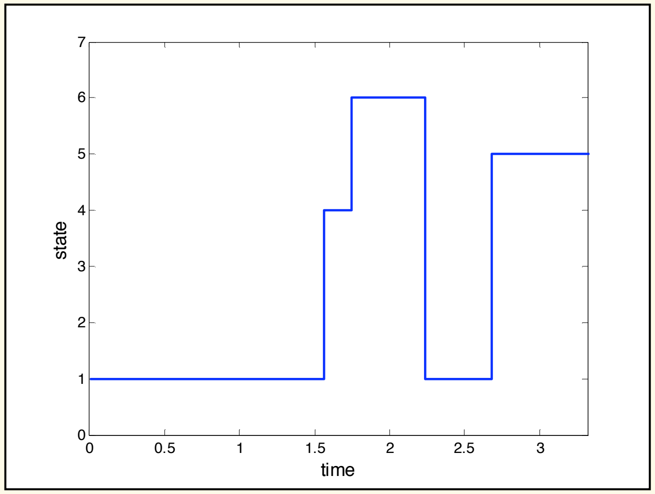
$$X = \{x_1, x_2, x_3, x_4, x_5, x_6\}$$

$$\mathbf{Q} \in \mathbb{R}^{6\times 6}$$

$$\pi_0 \in [0,1]^6$$

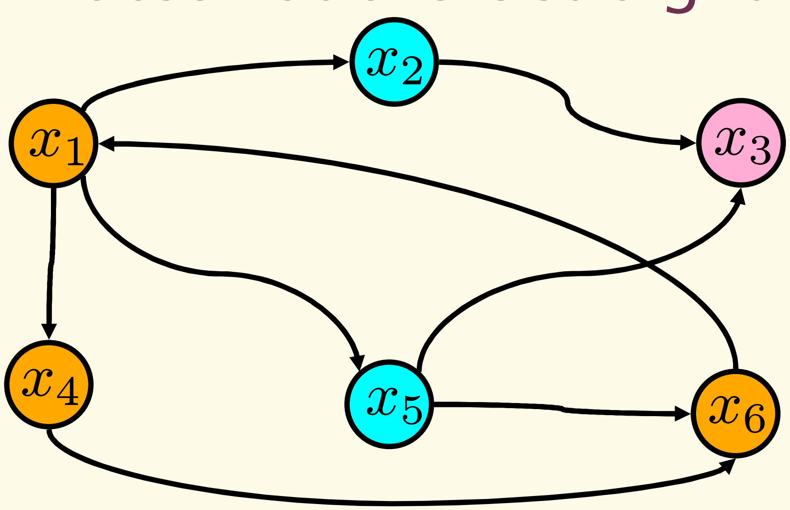


Trajectories of a CTMP are right-continuous and piecewise constant



Observation Model

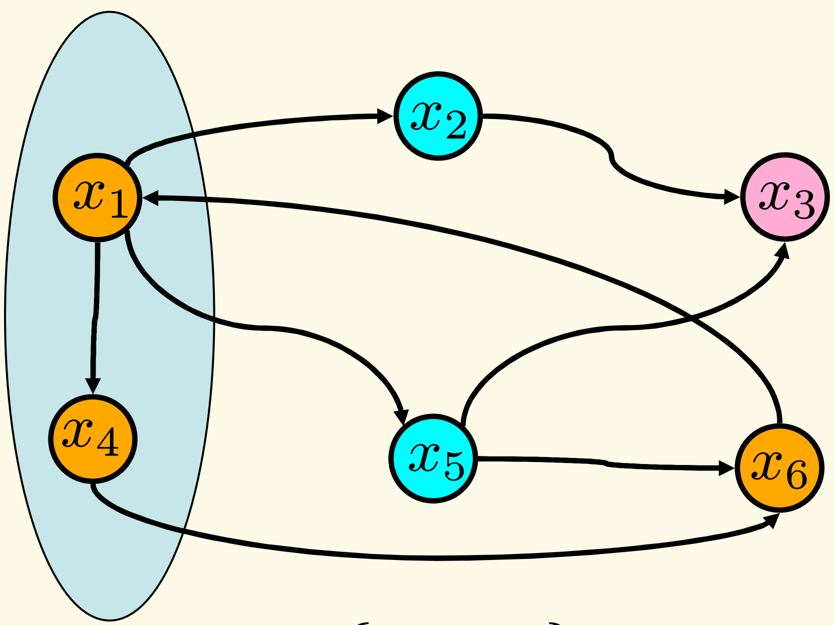
- The system state is partially observed through a state output function $h : X \to Y$
  - Many different states may produce the same observation
- The system can be observed:
  - continuously – we observe an output at all times $y^{[0,t]}$
  - intermittently – we observe a sequence $y^n = \{y_{\tau_1}, y_{\tau_2}, \ldots y_{\tau_n}\}$
- Extending the intermittent observation case to noisy observations is straightforward



$$h(x_1) = h(x_4) = h(x_6) = y_1$$
$$h(x_2) = h(x_5) = y_2$$
$$h(x_3) = y_3$$

Diagnosis Problem

- Some states are special states
  - Entering these states corresponds with rare, significant events. (e.g. enemy breaches the perimeter)
- Detecting that the system has reached a special state is non-trivial because special states can appear identical to normal states



$$X_s = \{x_1, x_4\}$$

Given a set of observations y^n (or $y^{[o,t]}$), what is the *a posteriori* probability that the system has visited a special state at some point along the interval (o,T)?

A Priori Probabilities

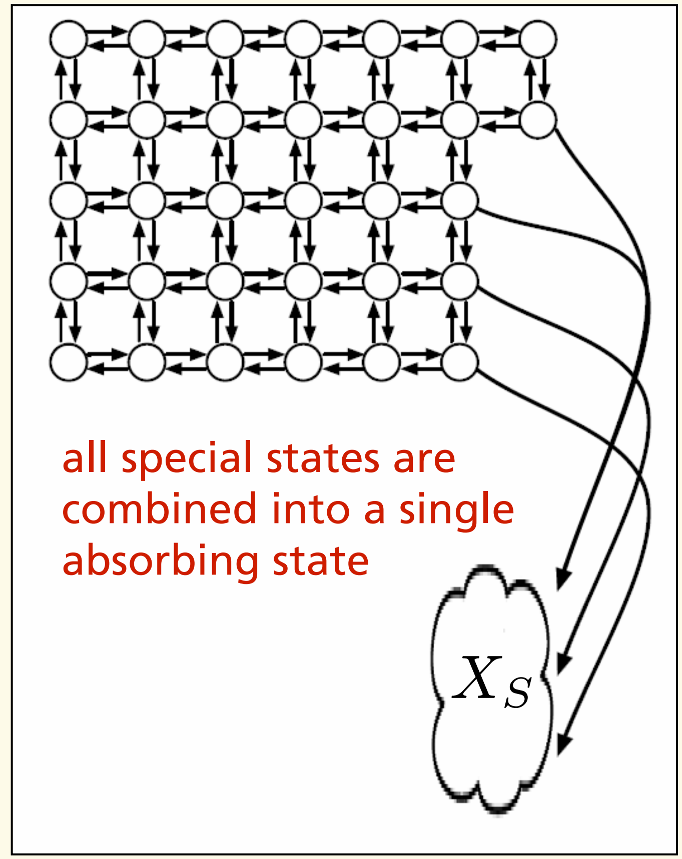
- Can be found using the master equation

$$\mathbf{p}_i(t) \triangleq \Pr(\omega(t) = x_i)$$

$$\dot{\mathbf{p}}(t) = \mathbf{Q}'\mathbf{p}(t)$$

$$\mathbf{p}(t) = e^{\mathbf{Q}'t}\pi_0$$

$$\mathbf{Q}' \triangleq \begin{bmatrix} \mathbf{Q}_{NN} & 0 \\ \mathbf{Q}_{NS} & 0 \end{bmatrix}$$



A Posteriori Solution for Intermittent Observations

- Between observations:

$$\mathbf{p}(t \mid y^n) = e^{\mathbf{Q}(t-\tau_n)}\mathbf{p}(\tau_n \mid y^n)$$

System evolves according as in *a priori* case

- When an observation occurs:

$$\mathbf{p}(\tau_n^+ \mid y^n) = \frac{1}{K}\left(\mathbf{H}_{y_n}\mathbf{p}(\tau_n^- \mid y^{n-1})\right)$$

normalization constant

Probabilities of being in states are weighted based on the observation

$$\mathbf{H}_y = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

indicator of the output
all non-diagonal elements are zero
diagonal element (i,i) is 1 if $h(x_i) = y$

A Posteriori Solution for Continuous Observations

- Between jumps (if output y_i is observed):

$$\mathbf{p}(t \mid y^{[0,t]}) = \frac{1}{K} e^{\mathbf{Q}_{II} t} \mathbf{p}(\tau_n \mid y^{[0,\tau_n]})$$

System must remain in the set of states with output y_i

- When a jump from y_i to y_j occurs:

$$\mathbf{p}(\tau_n^+ \mid y^{[0,\tau_n]}) = \frac{1}{K} \mathbf{Q}_{JI} \mathbf{p}(\tau_n^- \mid y^{[0,\tau_n)})$$

Probabilities are weighted based on how likely it is to jump from y_i to y_j

Normalize, Wait, and Jump

- The results of the two cases can be summed up using three functions:
- normalize: set the probability vector to length 1
- wait: describes the evolution between observations/jumps
- jump: describes the evolution at observable jumps

Intermittent Observations	Continuous Observations
$normalize(\mathbf{p}) = \frac{\mathbf{p}}{\mathbf{1}^T\mathbf{p}}$	$normalize(\mathbf{p}) = \frac{\mathbf{p}}{\mathbf{1}^T\mathbf{p}}$
$wait(\mathbf{p}) = e^{\mathbf{Q}t}\mathbf{p}$	$wait(\mathbf{p}) = e^{\mathbf{Q}_{II}t}\mathbf{p}$
$jump(\mathbf{p}) = \mathbf{H}_y\mathbf{p}$	$jump(\mathbf{p}) = \mathbf{Q}_{JI}\mathbf{p}$

Theorem: In both cases, the *a posteriori* probability distribution is:

$$\mathbf{p}(t \mid y) = normalize(wait_{t_n}(jump(wait_{t_{n-1}}(jump(\ldots\pi_0))))$$

Hybrid Systems Representation

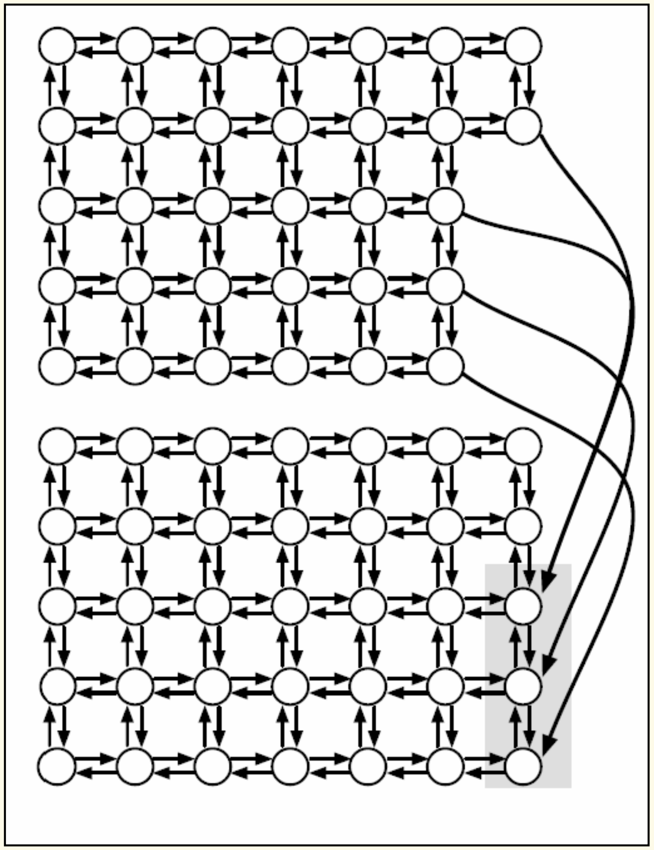
- because normalization can be done once at the end or whenever during the observation trajectory, there are HS representations of the *a posteriori* behavior

Intermittent	Continuous
$\dot{\mathbf{p}} = \mathbf{Q}\mathbf{p}$ $\mathbf{p} \mapsto \dfrac{\mathbf{H}_y \mathbf{p}}{\mathbf{1}^T \mathbf{H}_y \mathbf{p}}$	$\dot{\mathbf{p}} = \mathbf{Q}_{II}\mathbf{p} - \left(\mathbf{1}^T \mathbf{Q}_{II}\mathbf{p}\right)\mathbf{p}$ $\mathbf{p} \mapsto \dfrac{\mathbf{Q}_{JI}\mathbf{p}}{\mathbf{1}^T \mathbf{Q}_{JI}\mathbf{p}}$

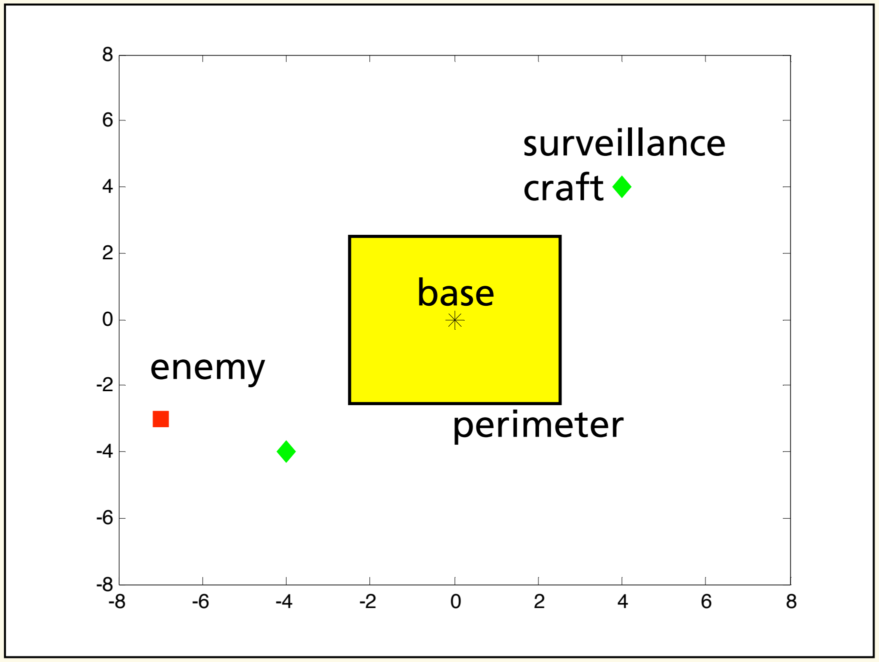
Solving the diagnosis problem

- We cannot treat the special states as a sink when computing the *a posteriori probabilities*
- the probability of being in the lower half of the state space is the probability that a special state was reached

$$\bar{Q} \triangleq \begin{bmatrix} \mathbf{Q}_{NN} & 0 & 0 \\ \mathbf{Q}_{NS} & \mathbf{Q}_{SS} & \mathbf{Q}_{NS} \\ 0 & \mathbf{Q}_{SN} & \mathbf{Q}_{NN} \end{bmatrix}$$



Stochastic Pursuit and Evasion



The base will go on alert if the probability of the enemy having breached the perimeter crosses a given threshold.

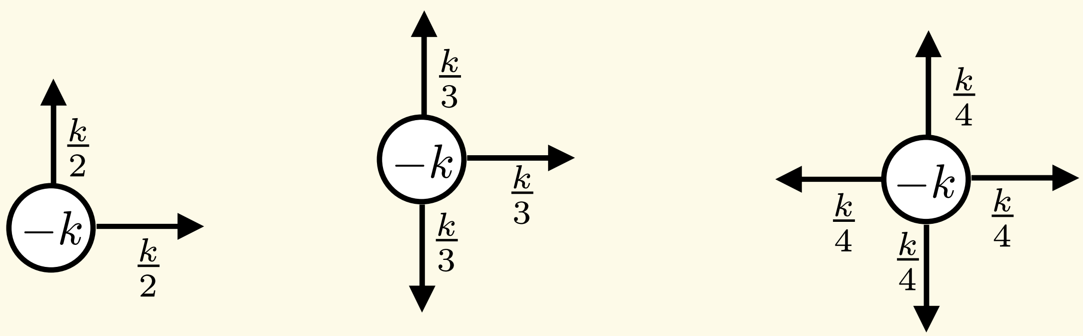
Objective: set the base on alert before the enemy reaches the base.

Need to know: What is the probability of the enemy having crossed the threshold given the observations the surveillance craft generate along their flight paths?

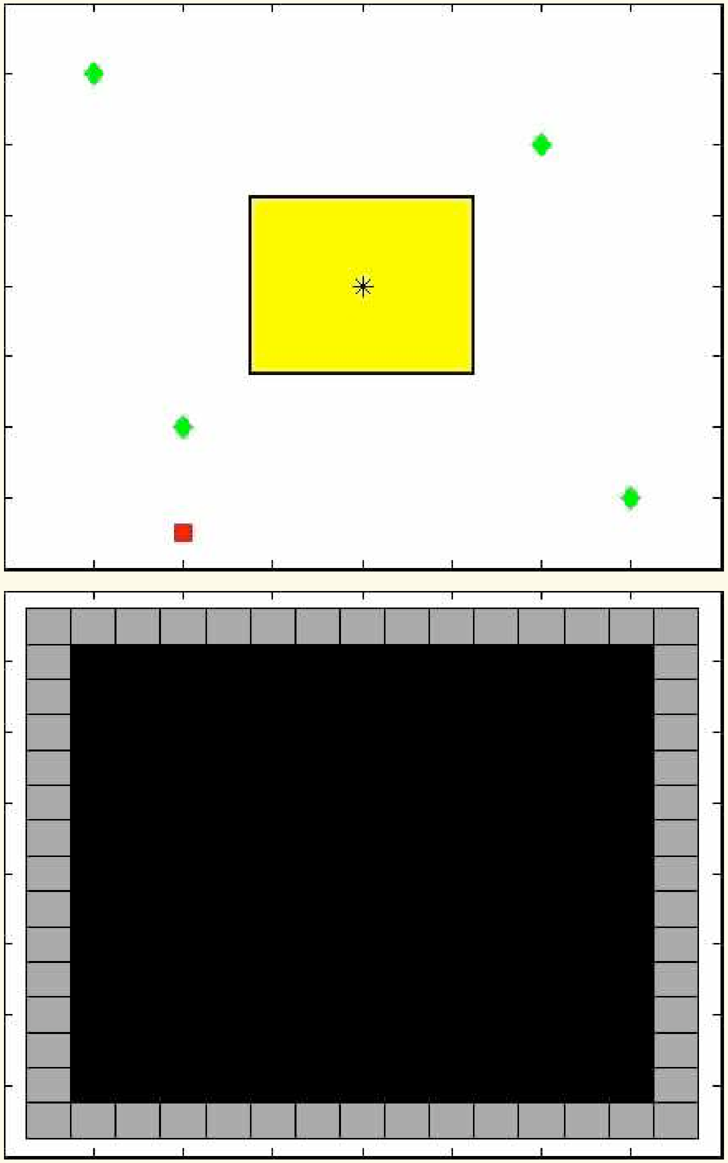
Can solve: Using the proposed diagnosis method. (continuous observations)

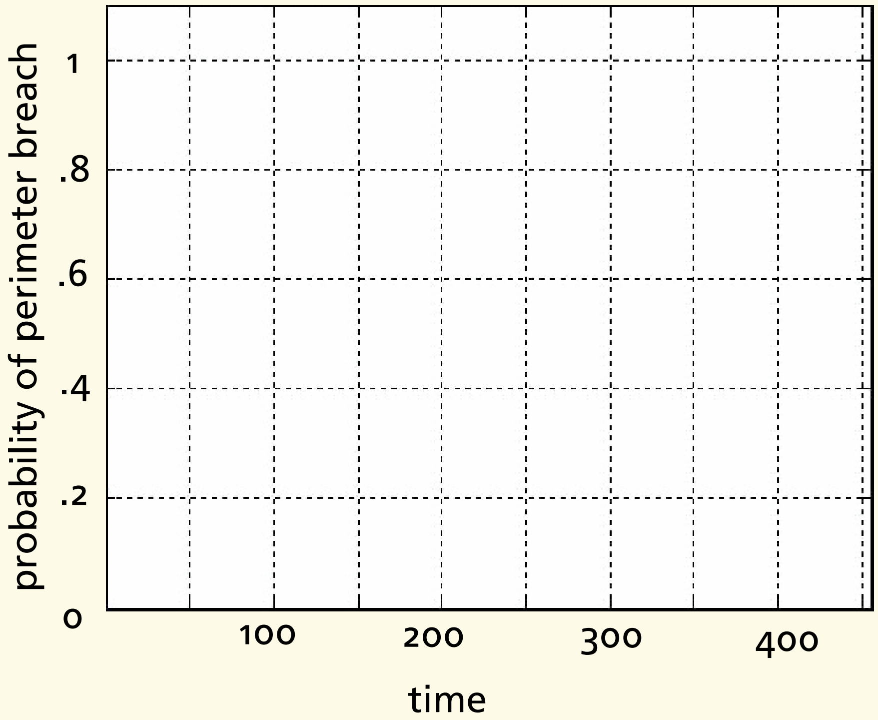
Enemy Model

- The enemy searches each grid square at a given rate k
- After completing the search, the enemy is equally likely to move to any adjacent square



Diagnostic dynamics along a sample trajectory





Modeling Assumptions

Modeling assumptions	Necessary?	
Enemy moves described by a Markov process	✓	
Enemy is equally likely to move in any valid direction		✗
Surveillance craft follow fixed paths		✗
Playing field is rectangular		✗
Enemy starts at the edge of the playing field		✗
Base is at the center of the playing field		✗
Perimeter is rectangular and centered at the base		✗
Surveillance craft do not track the sighted enemy		✗

Approximation Methods and Safety Analysis for Stochastic Processes

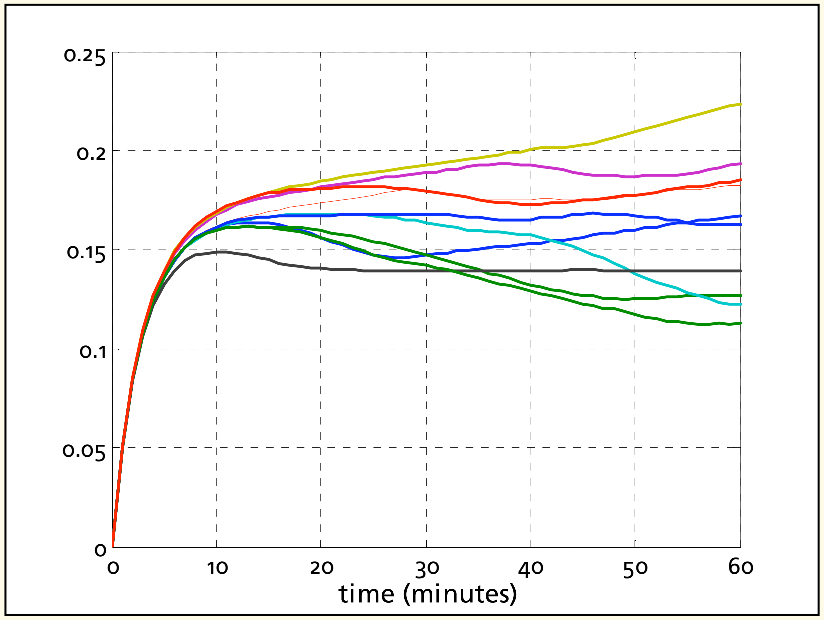
Part II: Approximating Stochastic Processes Using Wasserstein Pseudometrics

Stochastic processes

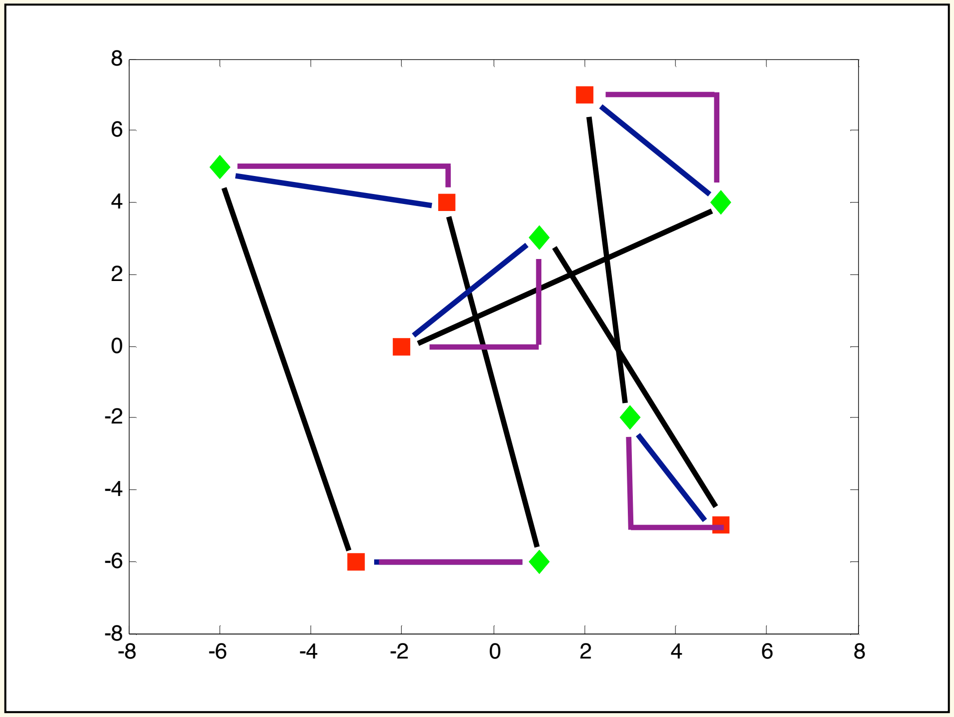
- Textbook definition:

  A stochastic process on a probability space $(\Omega, \mathcal{F}, \mathcal{P})$ is a set of random variables $\{Z_t\}_{t \in \mathcal{T}}$ indexed on a set $\mathcal{T}$.

- Ω (sample space) is the space of time-varying **trajectories**
- A method of specifying a stochastic process is just a method for defining a probability distribution $\mathcal{P}$ on Ω
- Comparing stochastic processes means comparing probability distributions



Distances between probability measures



$$dist_d(\mathcal{P}_1, \mathcal{P}_2) =$$

green squares: surveillance craft
red circles: enemies
Each set defines a random variable with a discrete probability distribution.

Matching up a enemy with an aircraft defines a joint probability distribution

A possible distance between the enemies and the aircrafts as the average distance traveled with respect to the joint distribution

There is an "optimal" joint distribution that minimizes this average distance

The distance between the probability distributions depends on how we define the distance on the plane

A general definition: Wasserstein pseudometrics

- The ***Wasserstein pseudometric*** with respect to d between two probability distributions $\mathcal{P}_1$ and $\mathcal{P}_2$ on Ω is

$$W_d(\mathcal{P}_1, \mathcal{P}_2) = \inf_{\mathcal{Q} \in J(\mathcal{P}_1, \mathcal{P}_2)} \int_{\Omega \times \Omega} d(\omega, \eta) d\mathcal{Q}(\omega, \eta)$$

- This concept has many names:
  - Monge-Kantorovich metric, Kantorovich-Rubinstein metric, Gini distance, Mallows distance, Hutchinson metric, earth mover's distance, Fortet-Mourier distance etc.

- When the sample space is $\Omega = \mathbb{R}$, the Wasserstein pseudometric is given by

$$W_d(\mathcal{P}_1, \mathcal{P}_2) = \int_{-\infty}^{\infty} |F_1(t) - F_2(t)| \, dt$$

- We can map an arbitrary sample space to the real line by defining the distance d as

$$d(\omega, \eta) = |Z(\omega) - Z(\eta)|$$

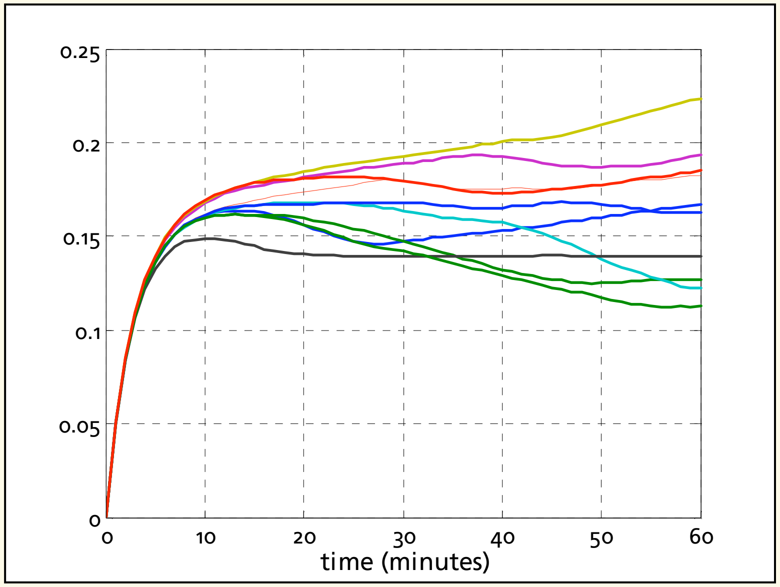
Z is an arbitrary scalar random variable

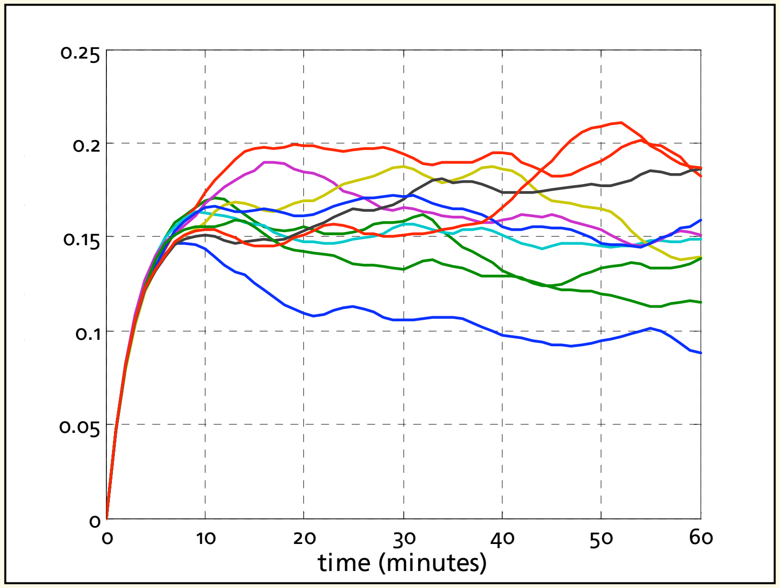
- Calculation of Wasserstein pseudometrics is difficult in general, but is simplified if the probability distributed are generated from sample data

Distances between stochastic processes

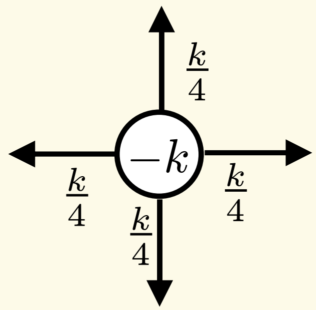
- The Wasserstein pseudometric definition is general and **does not depend** on how the probability distribution (or stochastic process) is specified

- ***Programmability of Similarity:*** By defining the underlying distance d on the space of trajectories, we can define Wasserstein pseudometrics comparing stochastic processes based on ***whatever criteria*** we deem relevant and are ***capable of measuring***

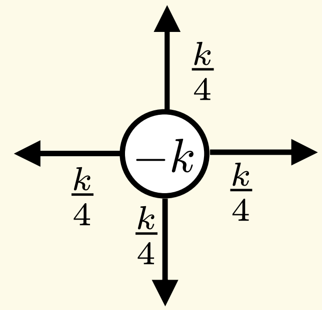




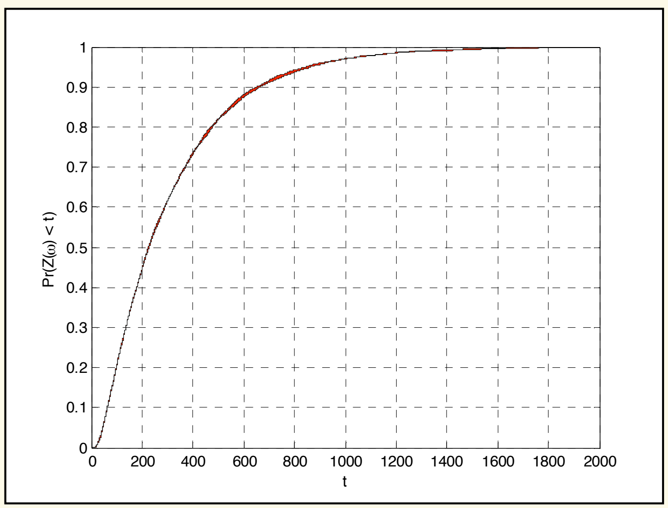
Wasserstein pseudometric examples

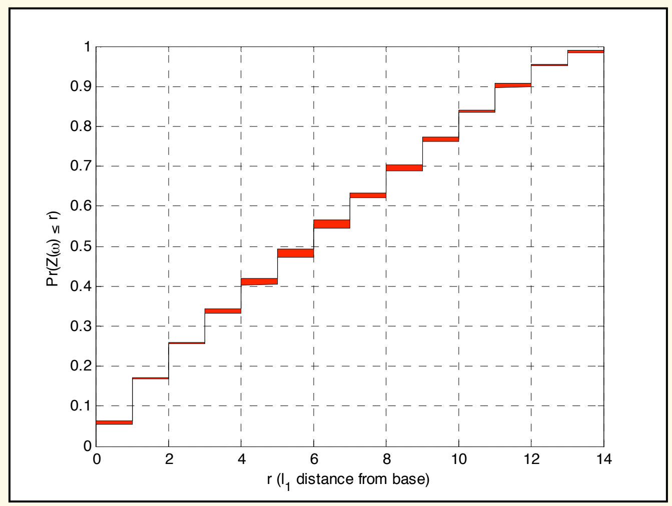


exponential dwell times



uniformly distributed
dwell times





$Z(\omega) = \min(t: \omega(t) \in X_S)$

$W_d(P_1,P_2)$ = 5.53 time units

$Z(\omega) = \| \omega(20) \|_1$

$W_d(P_1,P_2)$ = .0245 grid units

The overall picture