# Robust Partitioning for Stochastic Multi-Vehicle Routing

Erick Delage, John G. Carlsson, Yinyu Ye
Stanford University

October 2008

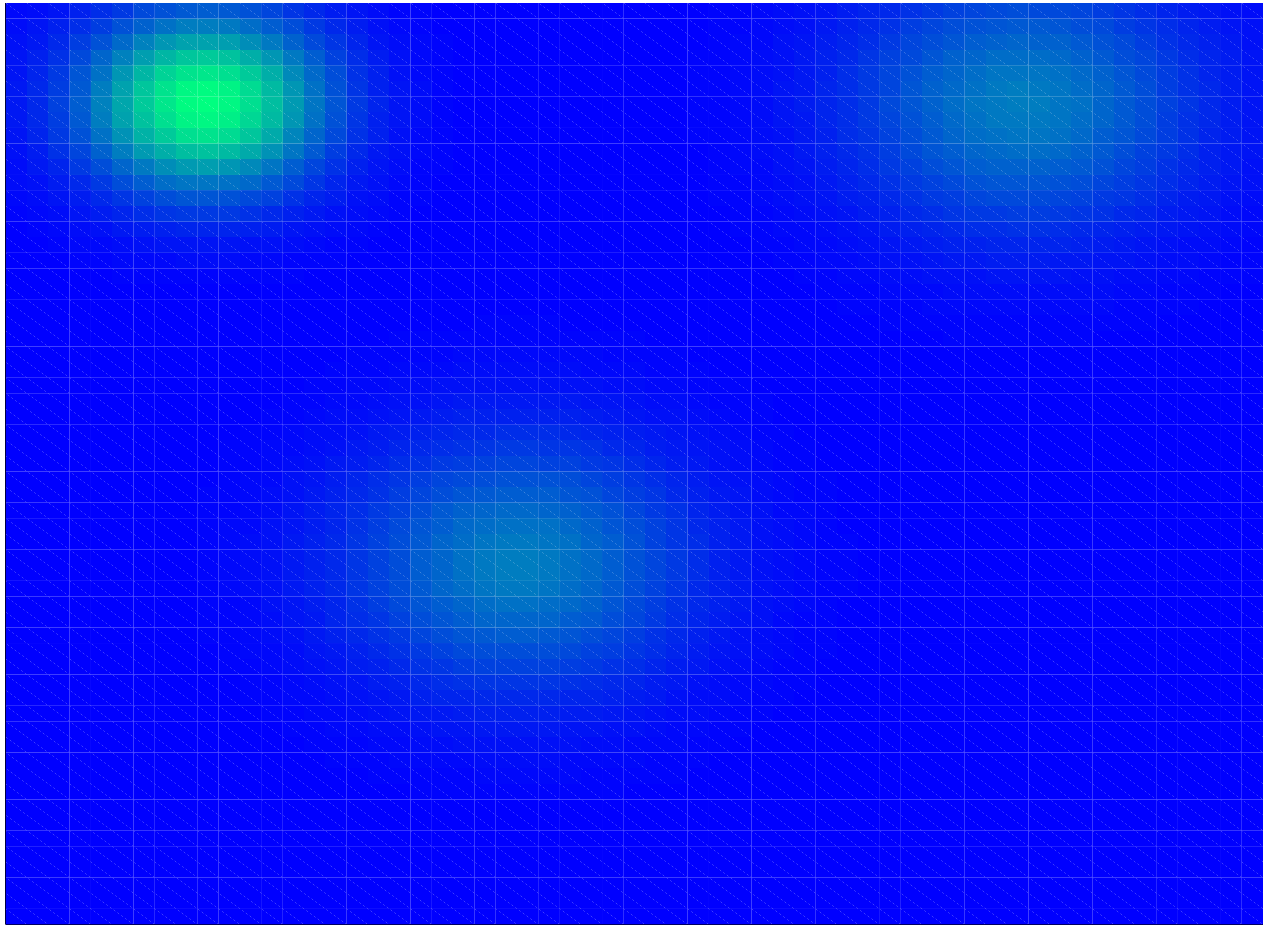
# The Region Partitioning Problem

Consider the following Two-Stage Stochastic Vehicle Routing problem:

- On a daily basis, a fleet of $k$ vehicles need to service $N$ demand points on a given territory
- The geographic location of each demand point is generated randomly from $f(x)$
- In a first stage, the territory needs to be divided into $k$ non-overlapping service regions which will each be assigned to a vehicle
- In a second stage, the vehicles will learn the location of $N$ demand points and each one will visit the demand points in his own region following the shortest (TSP) tour

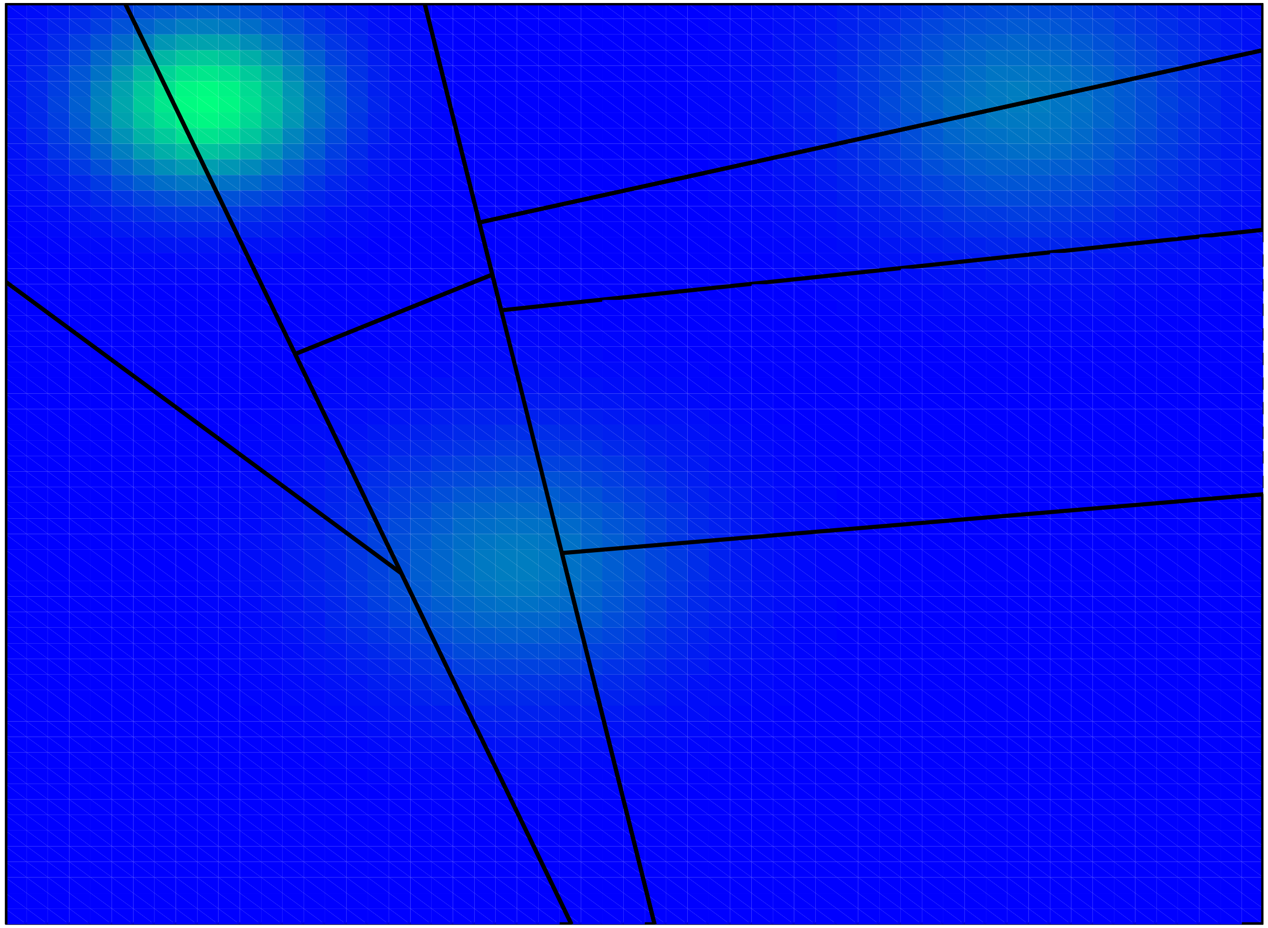
# Example : Stage 1

On a given territory, assume that demand is generated according to a distribution $f(x)$



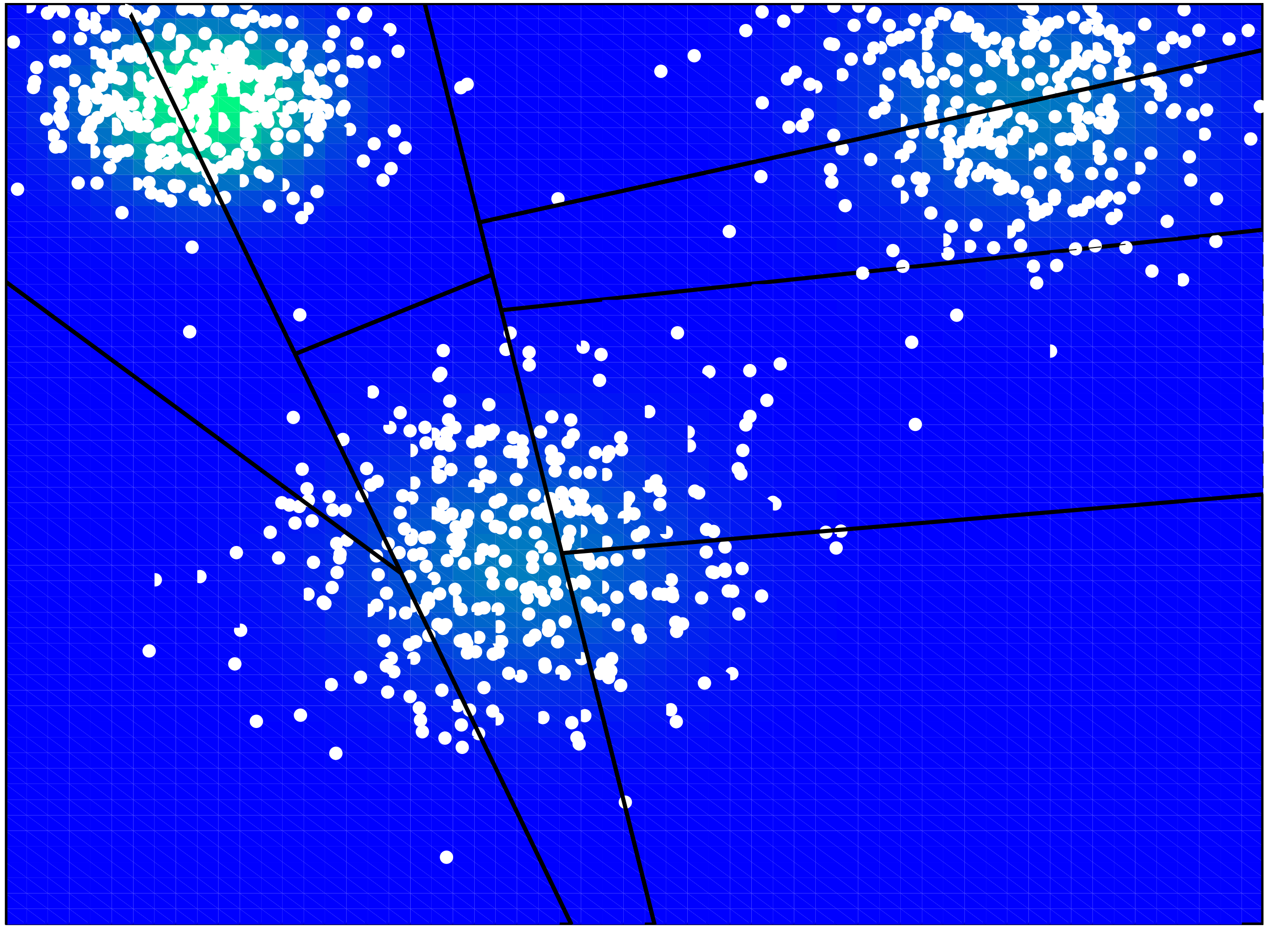
# Example : Stage 1

Partition the territory into $k$ sub-regions which are to be assign to $k$ vehicles



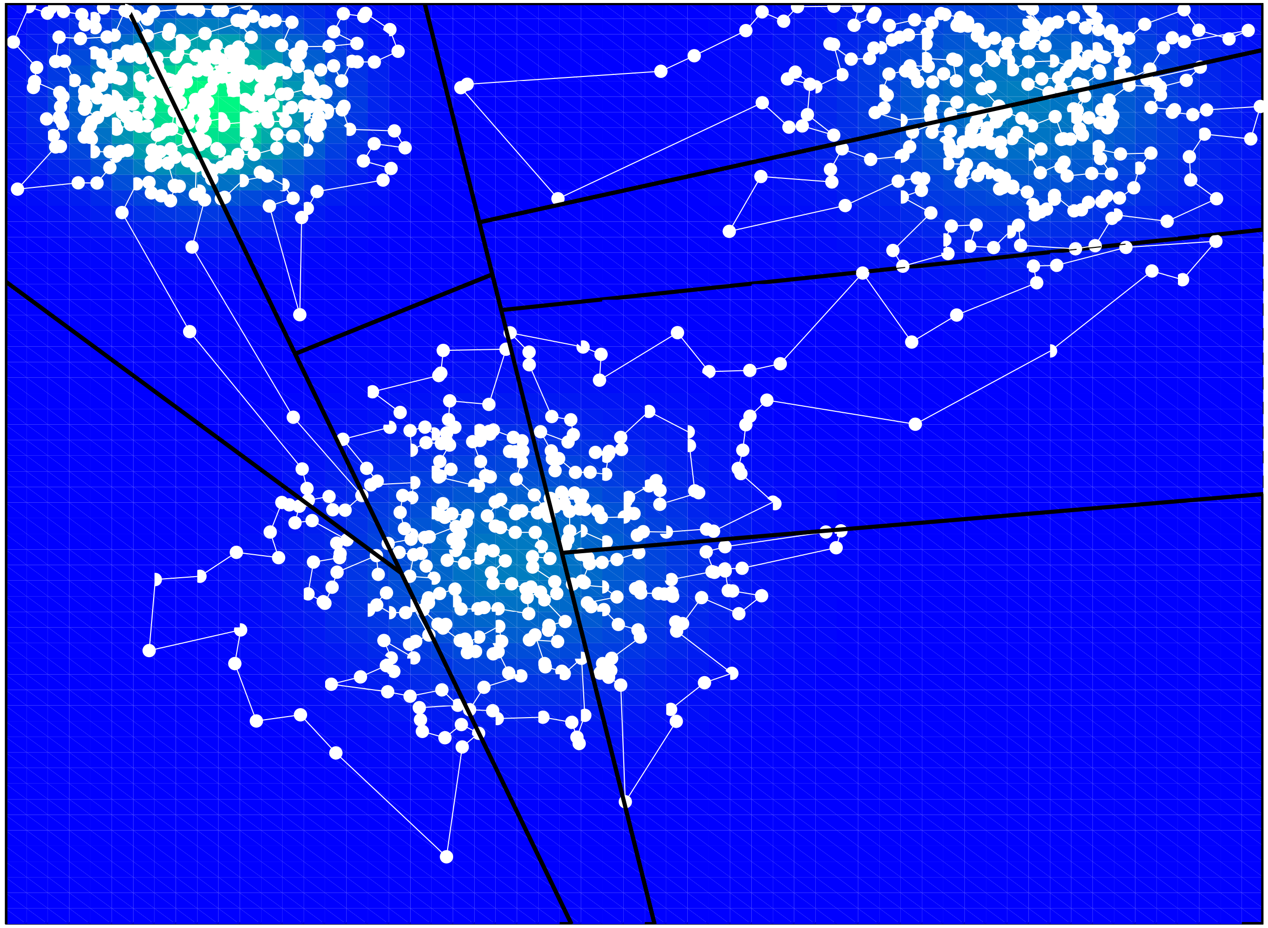
# Example : Stage 2

Every day, $N$ demand locations are generated and ask for the visit of a vehicle



# Example : Stage 2

The $k$ vehicles solve a TSP problem for demand points in their assigned region and satisfy the demand



# Partitioning under Perfect Information

When the demand distribution $f(x)$ is known, this partitioning problem can be formulated as:

$$\underset{\text{Partition:}\{\mathcal{C}_1,\ldots,\mathcal{C}_k\}}{\text{minimize}} \quad \left(\max_i \; ETSP(\mathcal{C}_i, f, N)\right) \quad (PSMVR)$$

- $ETSP(\mathcal{C}_i, f, N)$ is the Expected TSP tour length associated with region $\mathcal{C}_i$ given that $N$ points are drawn from $f(x)$
- For case where $N$ is large enough, Carlsson et al. (2007) propose an efficient solution

# Partitioning under Perfect Information

Given that $N$ is large enough, Beardwood et al. (1959) show:

$$ETSP(\mathcal{C}_i, f, N) = \sqrt{N} \int_{\mathcal{C}_i} \sqrt{f(x)}dx + o(\sqrt{N}) \ .$$

- For $f(x)$ uniform, solution is a partition with all $C_i$ having equal area : $ETSP(\mathcal{C}_i, f, N) \approx \sqrt{N} \cdot \text{Area}(\mathcal{C}_i)$
- If $f(x)$ is non-uniform, then one can sample $k \times M$ points from $\sqrt{f(x)}$ and choose $C_i$ such that each contain $M$ points
- For any $k$, Carlsson et al. (2007) show that both solutions can be found efficiently

What if demand distribution $f(x)$ is not exactly known?

# Partitioning under Limited Information

- In many applications, one cannot commit to a single demand distribution : e.g., historical data, non-stationary distribution
- It is usually easier to assume that $f(x)$ is a member of a set of distribution $\mathcal{D}$
- Then, one should consider a robust form:

$$\underset{\text{Partition:}\{\mathcal{C}_1,\ldots,\mathcal{C}_k\}}{\text{minimize}} \quad \left( \max_{f \in \mathcal{D}} \left( \max_i \; ETSP(\mathcal{C}_i, f, N) \right) \right)$$

In this talk:

- How should we describe $\mathcal{D}$?
- How can we compute worst case performance?
- How can we find good a two-partition?

# Confidence region for $f_\xi$

We propose using:

$$\mathcal{D}(\gamma) = \left\{ f(x) \,\middle|\, \begin{array}{l} \mathbb{P}(x \in \mathcal{C}) = 1 \\ (\mathbb{E}[x] - \hat{\mu})^{\mathsf{T}} \hat{\Sigma}^{-1} (\mathbb{E}[x] - \hat{\mu}) \leq \gamma \end{array} \right\}$$

In fact, as shown in Delage et al. (2008) if:

1. $\mathcal{C}$ lies in a ball of radius $R$
2. $\hat{\mu}$ and $\hat{\Sigma}$ are empirical estimates based on a set $\{x_i\}_{i=1}^m$ drawn i.i.d. from $f(x)$

Then, for some $\bar{\gamma} = O(\frac{R^2}{m} \log(1/\delta))$,

$$\mathbb{P}(f(x) \in \mathcal{D}(\bar{\gamma})) \geq 1 - \delta \quad .$$

# Worst Case Performance

For a given partition, one needs to evaluate the worst case ETSP load for each region $i$:

$$\max_{f\in\mathcal{D}} \; ETSP(\mathcal{C}_i, f, N) \approx \sqrt{N}\int_{\mathcal{C}_i}\sqrt{f(x)}dx$$

- Using duality theory, we can show that it is equivalent to:

$$\begin{array}{ll} \underset{r,\mathbf{q}}{\text{minimize}} & \sqrt{N}\left(1/4\displaystyle\int_{\mathcal{C}_i}\frac{1}{r+x^T\mathbf{q}}dx + r + \hat{\mu}^T\mathbf{q} + \sqrt{\gamma}\|\hat{\Sigma}^{1/2}\mathbf{q}\|_2\right) \\ \text{subject to} & r + x^T\mathbf{q} \geq 0 \;\; \forall x \in \mathcal{V} \end{array} ,$$

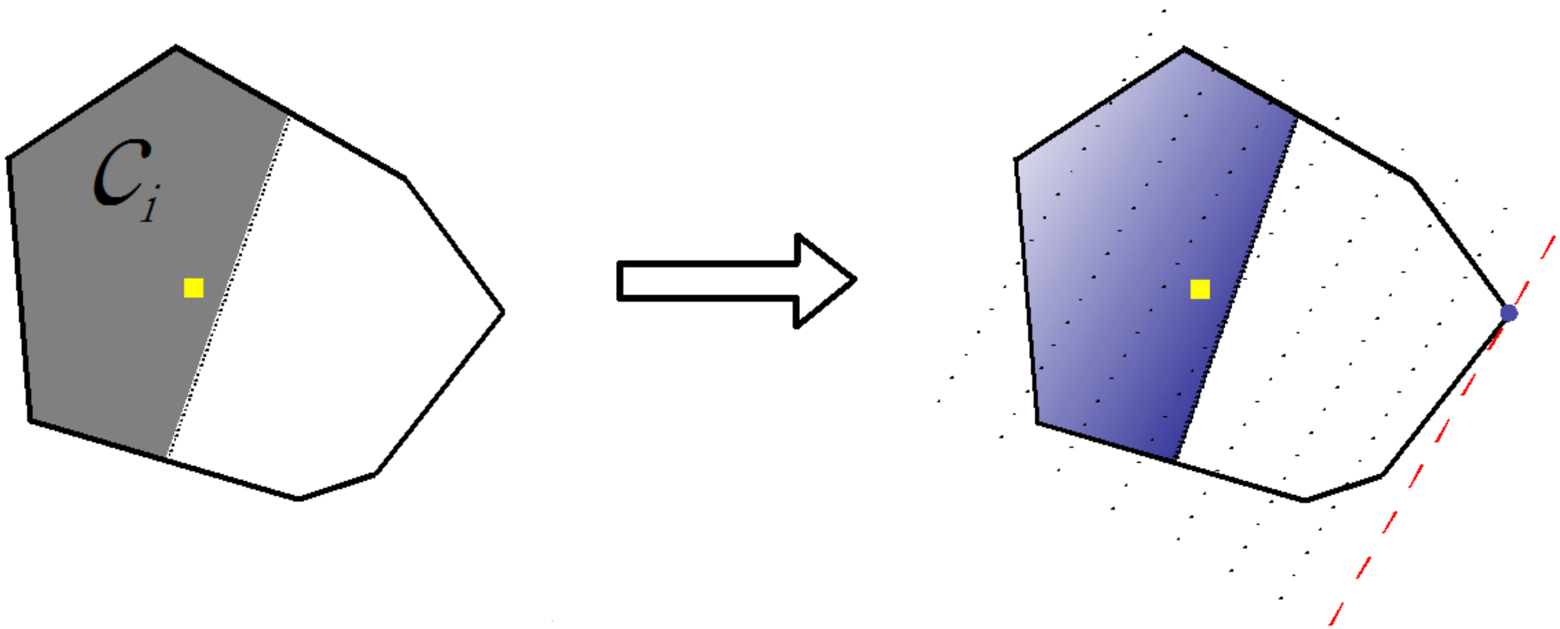
  where $\mathcal{V}$ is the set of vertices of $\mathcal{C}$.

- This problem can be solved using interior point methods after approximation of the integral by a finite sum.

# Worst Case Distribution

A worst case distribution has the form:

$$f(x) = \frac{1}{(r + x^T\mathbf{q})^2}\mathbb{1}\{x \in \mathcal{C}_i\} + \sum_{i=1}^{\|\mathcal{V}\|} p_i \delta_{x_i}$$



# Optimal Robust 2-Partition

Consider partitioning a region with a line:

$$x \cos(\theta) + y \sin(\theta) = b$$

Hence, we wish to solve:

$$\underset{\theta,b}{\text{minimize}} \quad \left( \max_{f \in \mathcal{D}} \max_{i \in \{1,2\}} ETSP(\mathcal{C}_i(\theta, b), f, N) \right)$$

Which is equivalent to:

$$\underset{\theta,b}{\text{minimize}} \quad \left( \max_{i \in \{1,2\}} \max_{f \in \mathcal{D}} ETSP(\mathcal{C}_i(\theta, b), f, N) \right)$$

# Optimal Robust 2-Partition

Partitioning with a line $x\cos(\theta) + y\sin(\theta) = b$, we wish to solve:

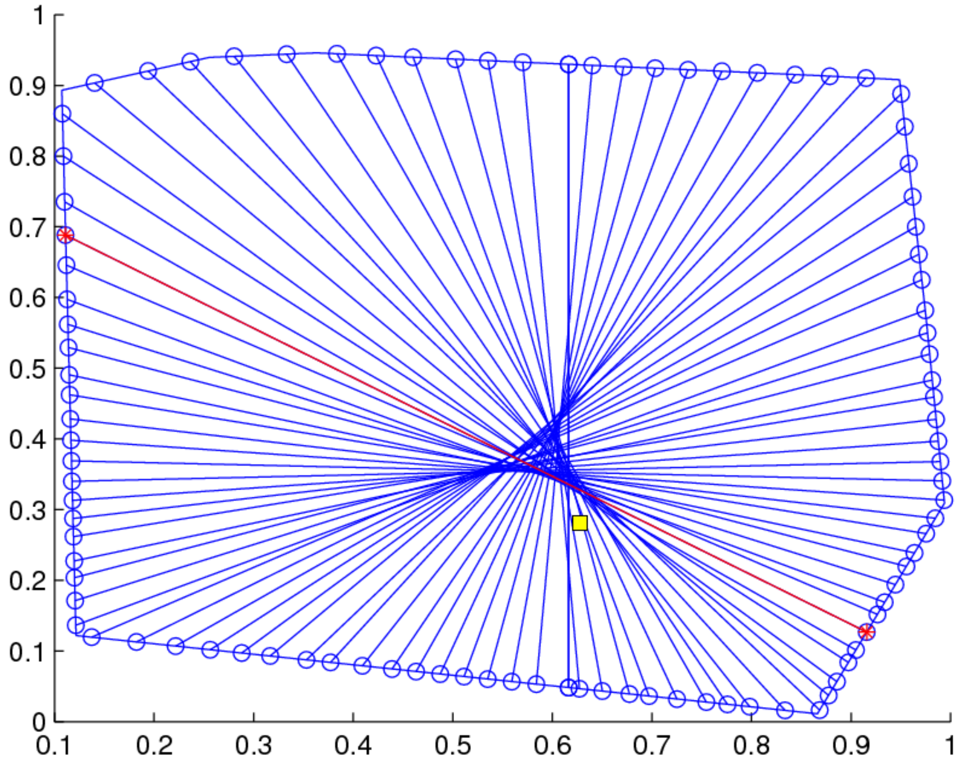
$$\underset{\theta,b}{\text{minimize}} \qquad h(\theta, b)$$

where $h(\theta, b) = \max_{i\in\{1,2\}} \ \max_{f\in\mathcal{D}} \ ETSP(\mathcal{C}_i(\theta, b), f, N)$

- We know how to evaluate $h(\theta, b)$
- For a fixed $\theta$, optimal $b$ can be found using bi-section method since $h(\theta, b)$ is quasi-convex in $b$
- We are left with a non-linear optimization problem over $\theta$: $\min_{\theta\in[0,\pi]} \ h(\theta, b^*(\theta))$

# Numerical Example

Consider the following set with mean given by the red box. We approximated the integral by discretizing over a $100 \times 100$ grid. The search over $\theta$ considered 100 equally spaced angles. A solution is obtained in less than 5 minutes.



# Summary & Open Questions

- We studied the problem of limited distribution information for partitioning in a stochastic multi-vehicle routing problem
- Under the assumption of large $N$, we can measure worst case performance and even characterize a worst case distribution
- A near-optimal two-partition can be efficiently generated

- How can we generate solutions for general $k$-partitioning problems? Does recursive two-partitioning generate good $k$-partitions?
- Can we account for more information about the distribution in formulating the distributional set: e.g. variance information?

# Questions & Comments ...

... Thank you!