Robust Discrete, Dynamic Optimization and Network Flows

Dimitris Bertsimas and Melvyn Sim
Operations Research, MIT

October, 2002

1 Structure

SLIDE 1

- Motivation
- Data Uncertainty
- Robust Mixed Integer Programming
- Robust 0-1 Programming
- Robust Approximation Algorithms
- Robust Network Flows
- Robust Inventory Control
- Experimental Results
- Summary and Conclusions

2 Motivation

SLIDE 2

- The classical paradigm in optimization is to develop a model that assumes that the input data is precisely known and equal to some nominal values. This approach, however, does not take into account the influence of data uncertainties on the quality and feasibility of the model.
- Can we design solution approaches that are immune to data uncertainty, that is they are robust?

SLIDE 3

- Ben-Tal and Nemirovski (2000):

> In real-world applications of Linear Programming (Net Lib library), one cannot ignore the possibility that a small uncertainty in the data can make the usual optimal solution completely meaningless from a practical viewpoint.

2.1 Literature

SLIDE 4

- Ellipsoidal uncertainty; Robust convex optimization Ben-Tal and Nemirovski (1997), El-Ghaoui et. al (1996)
- Flexible adjustment of conservativism
- Nonlinear convex models
- Not extendable to discrete optimization

3 Goal

SLIDE 5

Develop an approach to address data uncertainty for discrete optimization problems that:

- It allows to control the degree of conservatism of the solution;
- It is computationally tractable both practically and theoretically.

4 Data Uncertainty

SLIDE 6

$$\begin{array}{rll} \text{minimize} & \boldsymbol{c}'\boldsymbol{x} & \\ \text{subject to} & \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b} & \\ & \boldsymbol{l} \leq \boldsymbol{x} \leq \boldsymbol{u} & \\ & x_i \in \mathcal{Z}, & i = 1, \ldots, k, \end{array}$$

- WLOG data uncertainty affects only $\boldsymbol{A}$ and $\boldsymbol{c}$, but not the vector $\boldsymbol{b}$.
- **(Uncertainty for matrix $\boldsymbol{A}$)**: $a_{ij},\ j \in J_i$ is independent, symmetric and bounded random variable (but with unknown distribution) $\tilde{a}_{ij}, j \in J_i$ that takes values in $[a_{ij} - \hat{a}_{ij}, a_{ij} + \hat{a}_{ij}]$.
- **(Uncertainty for cost vector $\boldsymbol{c}$)**: $c_j,\ j \in J_0$ takes values in $[c_j, c_j + d_j]$.

5 Robust MIP

SLIDE 7

- Consider an integer $\Gamma_i \in [0, |J_i|]$, $i = 0, 1, \ldots, m$.
- Γ_i adjusts the robustness of the proposed method against the level of conservativeness of the solution.
- Speaking intuitively, it is unlikely that all of the $a_{ij},\ j \in J_i$ will change. We want to be protected against all cases that up to Γ_i of the a_{ij}'s are allowed to change.

SLIDE 8

- Nature will be restricted in its behavior, in that only a subset of the coefficients will change in order to adversely affect the solution.
- We will guarantee that if nature behaves like this then the robust solution will be feasible deterministically. Even if more than Γ_i change, then the robust solution will be feasible with very high probability.

5.1 Problem

SLIDE 9

$$
\begin{array}{ll}
\text{minimize} & \boldsymbol{c}'\boldsymbol{x} + \max\limits_{\{S_0 |\ S_0 \subseteq J_0, |S_0| = \Gamma_0\}} \left\{ \sum\limits_{j \in S_0} d_j |x_j| \right\} \\
\text{subject to} & \sum\limits_j a_{ij} x_j + \max\limits_{\{S_i |\ S_i \subseteq J_i, |S_i| = \Gamma_i\}} \left\{ \sum\limits_{j \in S_i} \hat{a}_{ij} |x_j| \right\} \leq b_i, \quad \forall i \\
& \boldsymbol{l} \leq \boldsymbol{x} \leq \boldsymbol{u} \\
& x_i \in \mathcal{Z}, \qquad \forall i = 1, \ldots k.
\end{array}
$$

5.2 Theorem 1

SLIDE 10

The robust problem can be reformulated has an equivalent MIP:

$$
\begin{array}{lll}
\text{minimize} & \boldsymbol{c}'\boldsymbol{x} + z_0 \Gamma_0 + \sum_{j \in J_0} p_{0j} & \\
\text{subject to} & \sum\limits_j a_{ij} x_j + z_i \Gamma_i + \sum\limits_{j \in J_i} p_{ij} \leq b_i & \forall i \\
& z_0 + p_{0j} \geq d_j y_j & \forall j \in J_0 \\
& z_i + p_{ij} \geq \hat{a}_{ij} y_j & \forall i \neq 0, j \in J_i \\
& p_{ij}, y_j, z_i \geq 0 & \forall i, j \in J_i \\
& -y_j \leq x_j \leq y_j & \forall j \\
& l_j \leq x_j \leq u_j & \forall j \\
& x_i \in \mathcal{Z} & i = 1, \ldots, k.
\end{array}
$$

5.3 Proof

SLIDE 11

Given a vector $\boldsymbol{x}^*$, we define:

$$
\beta_i(\boldsymbol{x}^*) = \max_{\{S_i |\ S_i \subseteq J_i, |S_i| = \Gamma_i\}} \left\{ \sum_{j \in S_i} \hat{a}_{ij} |x_j^*| \right\}.
$$

This equals to:

$$
\begin{array}{rll}
\beta_i(\boldsymbol{x}^*) = \max & \sum\limits_{j \in J_i} \hat{a}_{ij} |x_j^*| z_{ij} & \\
\text{s.t.} & \sum\limits_{j \in J_i} z_{ij} \leq \Gamma_i & \\
& 0 \leq z_{ij} \leq 1 & \forall i, j \in J_i.
\end{array}
$$

SLIDE 12

Dual:

$$
\begin{array}{rll}
\beta_i(\boldsymbol{x}^*) = \min & \sum\limits_{j \in J_i} p_{ij} + \Gamma_i z_i & \\
\text{s.t.} & z_i + p_{ij} \geq \hat{a}_{ij} |x_j^*| & \forall j \in J_i \\
& p_{ij} \geq 0 & \forall j \in J_i \\
& z_i \geq 0 & \forall i.
\end{array}
$$

$\lvert J_i \rvert$	Γ_i
5	5
10	8.3565
100	24.263
200	33.899

Table 1: Choice of Γ_i as a function of $|J_i|$ so that the probability of constraint violation is less than 1%.

5.4 Size

SLIDE 13

- Original Problem has n variables and m constraints
- Robust counterpart has $2n + m + l$ variables, where $l = \sum_{i=0}^{m} |J_i|$ is the number of uncertain coefficients, and $2n + m + l$ constraints.

5.5 Probabilistic Guarantee

5.5.1 Theorem 2

SLIDE 14

Let $\boldsymbol{x}^*$ be an optimal solution of robust MIP.
(a) If $\boldsymbol{A}$ is subject to the model of data uncertainty U:

$$\Pr\left(\sum_j \tilde{a}_{ij} x_j^* > b_i\right) \le \frac{1}{2^n}\left\{(1-\mu)\sum_{l=\lfloor\nu\rfloor}^{n}\binom{n}{l} + \mu \sum_{l=\lfloor\nu\rfloor+1}^{n}\binom{n}{l}\right\},$$

$n = |J_i|$, $\nu = \frac{\Gamma_i + n}{2}$ and $\mu = \nu - \lfloor\nu\rfloor$; bound is tight.
(b) As $n \to \infty$

$$\frac{1}{2^n}\left\{(1-\mu)\sum_{l=\lfloor\nu\rfloor}^{n}\binom{n}{l} + \mu \sum_{l=\lfloor\nu\rfloor+1}^{n}\binom{n}{l}\right\} \sim 1 - \Phi\left(\frac{\Gamma_i - 1}{\sqrt{n}}\right),$$

$$\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} \exp\left(-\frac{y^2}{2}\right) dy.$$

SLIDE 15
SLIDE 16

6 Robust 0-1 Optimization

SLIDE 17

- Nominal combinatorial optimization:

$$\begin{aligned} &\text{minimize} && \boldsymbol{c}'\boldsymbol{x} \\ &\text{subject to} && \boldsymbol{x} \in X \subset \{0,1\}^n. \end{aligned}$$

- Robust Counterpart:

$$\begin{aligned} Z^* = \quad &\text{minimize} && \boldsymbol{c}'\boldsymbol{x} + \max_{\{S|\ S\subseteq J, |S|=\Gamma\}} \sum_{j\in S} d_j x_j \\ &\text{subject to} && \boldsymbol{x} \in X, \end{aligned}$$

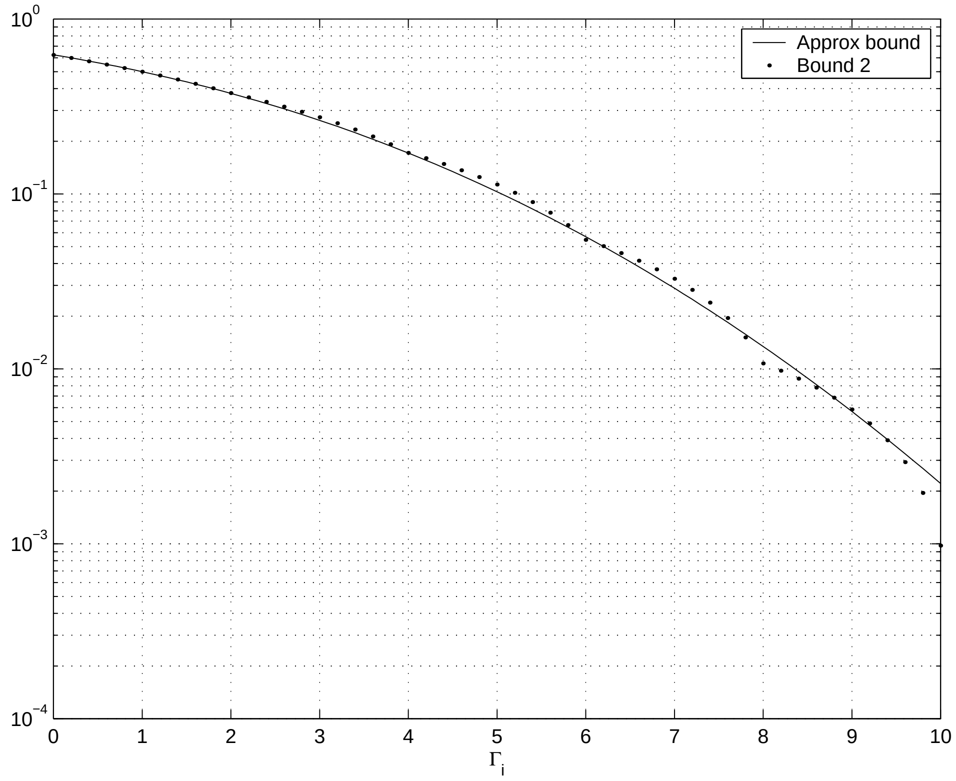

Figure 1: Quality of approximation

- WLOG $d_1 \geq d_2 \geq \ldots \geq d_n$.

6.1 Remarks

SLIDE 18

- Examples: the shortest path, the minimum spanning tree, the minimum assignment, the traveling salesman, the vehicle routing and matroid intersection problems.

- Other approaches to robustness are hard. Scenario based uncertainty:

$$\begin{array}{ll} \text{minimize} & \max(\boldsymbol{c_1'x}, \boldsymbol{c_2'x}) \\ \text{subject to} & \boldsymbol{x} \in X. \end{array}$$

  is NP-hard for the shortest path problem.

6.2 Approach

SLIDE 19

$$\begin{aligned} \text{Primal} : Z^* = \min_{\boldsymbol{x} \in X} \boldsymbol{c}'\boldsymbol{x} + \max \quad & \sum_j d_j x_j u_j \\ \text{s.t.} \quad & 0 \le u_j \le 1, \qquad \forall\, j \\ & \sum_j u_j \le \Gamma \end{aligned}$$

$$\begin{aligned} \text{Dual} : Z^* = \min_{\boldsymbol{x} \in X} \boldsymbol{c}'\boldsymbol{x} + \min \quad & \theta\Gamma + \sum_j y_j \\ \text{s.t.} \quad & y_j + \theta \ge d_j x_j, \qquad \forall\, j \\ & y_j, \theta \ge 0 \end{aligned}$$

6.3 Algorithm A

SLIDE 20

- Solution: $y_j = \max(d_j x_j - \theta, 0)$
- $$Z^* = \min_{\boldsymbol{x} \in X, \theta \ge 0} \theta\Gamma + \sum_j \left(c_j x_j + \max(d_j x_j - \theta, 0)\right)$$
- Since $X \subset \{0,1\}^n$,

  $$\max(d_j x_j - \theta, 0) = \max(d_j - \theta, 0)\ x_j$$
- $$Z^* = \min_{\boldsymbol{x} \in X, \theta \ge 0} \theta\Gamma + \sum_j \left(c_j + \max(d_j - \theta, 0)\right) x_j$$

SLIDE 21

- $d_1 \ge d_2 \ge \ldots \ge d_n \ge d_{n+1} = 0$.
- For $d_l \ge \theta \ge d_{l+1}$,

  $$\min_{\boldsymbol{x} \in X, d_l \ge \theta \ge d_{l+1}} \theta\Gamma + \sum_{j=1}^n c_j x_j + \sum_{j=1}^l (d_j - \theta) x_j =$$

  $$d_l\Gamma + \min_{\boldsymbol{x} \in X} \sum_{j=1}^n c_j x_j + \sum_{j=1}^l (d_j - d_l) x_j = Z_l$$
- $$Z^* = \min_{l=1,\ldots,n+1} d_l\Gamma + \min_{\boldsymbol{x} \in X} \sum_{j=1}^n c_j x_j + \sum_{j=1}^l (d_j - d_l) x_j.$$

6.4 Theorem 3

SLIDE 22

- Algorithm A correctly solves the robust 0-1 optimization problem.
- It requires at most $|J| + 1$ solutions of nominal problems. Thus, If the nominal problem is polynomially time solvable, then the robust 0-1 counterpart is also polynomially solvable.
- Robust minimum spanning tree, minimum assignment, minimum matching, shortest path and matroid intersection, are polynomially solvable.

7 Robust Approximation Algorithms

SLIDE 23

- If the nominal problem is α-approximable, is the robust counterpart also α-approximable?
- **Input:** Vectors $\boldsymbol{c}, \boldsymbol{d} \in \Re_+^n$, an integer Γ, and a polynomial time algorithm H that returns a solution Z_H: $Z_H \le \alpha Z$, where $Z = \min \boldsymbol{c}'\boldsymbol{x}$ subject to $\boldsymbol{x} \in X \subseteq \{0,1\}^n$ for all $\boldsymbol{c} \geq \boldsymbol{0}$.
- **Output:** A solution $\boldsymbol{x^B} \in X$ such that $Z_B = \boldsymbol{c}'\boldsymbol{x^B} + \max_{\{S|\ S \subseteq J, |S| = \Gamma\}} \sum_{j \in S} d_j x_j^B$ satisfies $Z^* \le Z_B \le \alpha Z^*$.

7.1 Algorithm B

SLIDE 24

- Use an α-approximate solution to

$$\min_{\boldsymbol{x} \in X} \sum_{j=1}^{n} c_j x_j + \sum_{j=1}^{l} (d_j - d_l) x_j .$$

- Theorem 4: Overall algorithm is α-approximate.

8 Robust Network Flows

SLIDE 25

- Nominal

$$\begin{aligned} \min \quad & \sum_{(i,j) \in \mathcal{A}} c_{ij} x_{ij} & \\ \text{s.t.} \quad & \sum_{\{j:(i,j) \in \mathcal{A}\}} x_{ij} - \sum_{\{j:(j,i) \in \mathcal{A}\}} x_{ji} = b_i & \forall i \in \mathcal{N} \\ & 0 \le x_{ij} \le u_{ij} & \forall (i,j) \in \boldsymbol{\mathcal{A}}. \end{aligned}$$

- X set of feasible solutions flows.
- Robust

$$\begin{aligned} Z^* = \min \quad & \boldsymbol{c}'\boldsymbol{x} + \max_{\{S|\ S \subseteq \mathcal{A}, |S| \le \Gamma\}} \sum_{(i,j) \in S} d_{ij} x_{ij} \\ \text{subject to} \quad & \boldsymbol{x} \in X. \end{aligned}$$

8.1 Reformulation

SLIDE 26

- $$Z^* = \min_{\theta \geq 0} Z(\theta),$$

$$\begin{aligned} Z(\theta) = \Gamma\theta + \min \quad & \boldsymbol{c}'\boldsymbol{x} + \sum_{(i,j)\in\mathcal{A}} p_{ij} \\ \text{subject to} \quad & p_{ij} \geq d_{ij}x_{ij} - \theta \quad \forall (i,j) \in \mathcal{A} \\ & p_{ij} \geq 0 \quad \forall (i,j) \in \mathcal{A} \\ & \boldsymbol{x} \in X. \end{aligned}$$

- Equivalently

$$\begin{aligned} Z(\theta) = \Gamma\theta + \min \quad & \boldsymbol{c}'\boldsymbol{x} + \sum_{(i,j)\in\mathcal{A}} d_{ij} \max\left(x_{ij} - \frac{\theta}{d_{ij}}, 0\right) \\ \text{subject to} \quad & \boldsymbol{x} \in X. \end{aligned}$$

8.2 Network Reformulation

SLIDE 27

Theorem: For fixed θ we can solve the robust problem as a network flow problem

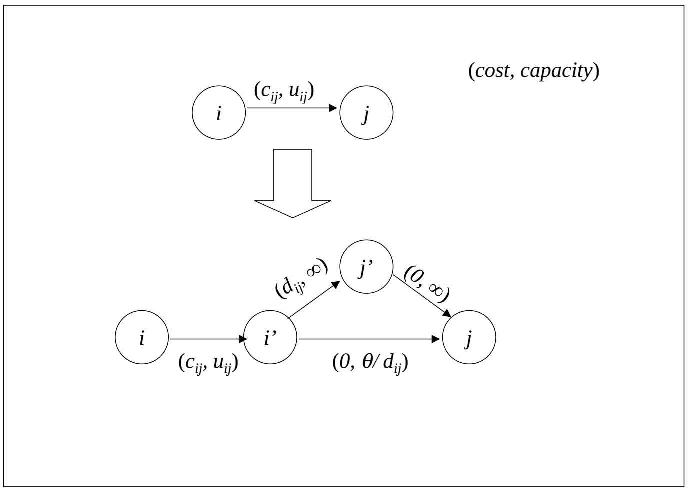

Figure 2: Construction

8.3 Complexity

SLIDE 28

- $Z(\theta)$ is a convex function and for all $\theta_1, \theta_2 \geq 0$, we have

$$|Z(\theta_1) - Z(\theta_2)| \leq |\mathcal{A}||\theta_1 - \theta_2|.$$

- For any fixed $\Gamma \leq |\mathcal{A}|$ and every $\epsilon > 0$, we can find a solution $\hat{\boldsymbol{x}} \in X$ with robust objective value
$$\hat{Z} = \boldsymbol{c}'\hat{\boldsymbol{x}} + \max_{\{S \mid S \subseteq \mathcal{A}, |S| \leq \Gamma\}} \sum_{(i,j) \in S} d_{ij} \hat{x}_{ij}$$
such that
$$Z^* \leq \hat{Z} \leq (1+\epsilon) Z^*$$
by solving $2\lceil \log_2(|\mathcal{A}|\overline{\theta}/\epsilon)\rceil + 3$ network flow problems, where $\overline{\theta} = \max\{u_{ij}d_{ij} \ : (i,j) \in \mathcal{A}\}$.

9 Robust Inventory Control

SLIDE 29

- Joint work with Aurelie Thiele, MIT
- Single station
- State x_k: stock available at the beginning of the kth period
- Control u_k: stock ordered at the beginning of the kth period
- Randomness w_k: demand during the kth period
- Dynamics: $x_{k+1} = x_k + u_k - w_k$
- Cost: $cu_k + \max(hx_{k+1}, -px_{k+1})$

9.1 Modeling Randomness

SLIDE 30

- $z_k = (w_k - \bar{w}_k)/\hat{w}_k \in [-1, 1]$.
- Uncertainty budget $\sum_{i=0}^{k} |z_i| \leq \Gamma_k$.
- Theorem: Robust problem optimal ordering policy is also (S, S), or base-stock, i.e., there exists a threshold sequence (S_k) such that, at each time period k, it is optimal to order $S_k - x_k$ if $x_k < S_k$ and 0, otherwise. S_k given in closed form.

9.2 Fixed costs

SLIDE 31

- If there are fixed ordering costs, optimal policy for robust problem is (s, S), i.e., there exists a threshold sequence (s_k, S_k) such that, at each time period k, it is optimal to order $S_k - x_k$ if $x_k < s_k$ and 0 otherwise, with $s_k \leq S_k$.
- Constrast to the stochastic case.

9.3 Multiple stations

SLIDE 32

- Robust counterpart recovers the Clark and Scarf solution for echelons
- Robust solutions extends to the case of fixed costs in all stations
- Extensions to general supply chains; Problem reduces to a deterministic LP or MIP, computationally tractable.
- Major messages: numerical tractability not affected by dimensionality in sharp contrast to the stocahstic case
- Insightful policies that are the same qualitatively with the stochastic case; closed form computable.

10 Experimental Results

10.1 Knapsack Problems

SLIDE 33

- $$\begin{array}{ll} \text{maximize} & \sum_{i \in N} c_i x_i \\ \text{subject to} & \sum_{i \in N} w_i x_i \leq b \\ & \boldsymbol{x} \in \{0,1\}^n. \end{array}$$

- $\tilde{w}_i$ are independently distributed and follow symmetric distributions in $[w_i - \delta_i, w_i + \delta_i]$;
- $\boldsymbol{c}$ is not subject to data uncertainty.

10.1.1 Data

SLIDE 34

- $|N| = 200$, $b = 4000$,
- w_i randomly chosen from $\{20, 21, \ldots, 29\}$.
- c_i randomly chosen from $\{16, 17, \ldots, 77\}$.
- $\delta_i = 0.1 w_i$.

10.1.2 Results

SLIDE 35

10.2 Robust Sorting

SLIDE 36

$$\begin{array}{ll} \text{minimize} & \sum_{i \in N} c_i x_i \\ \text{subject to} & \sum_{i \in N} x_i = k \\ & \boldsymbol{x} \in \{0,1\}^n. \end{array}$$

Γ	Violation Probability	Optimal Value	Reduction
0	1	5592	0%
2.8	4.49×10^{-1}	5585	0.13%
36.8	5.71×10^{-3}	5506	1.54%
82.0	5.04×10^{-9}	5408	3.29%
200	0	5283	5.50%

Γ	$\bar{Z}(\Gamma)$	% change in $\bar{Z}(\Gamma)$	$\sigma(\Gamma)$	% change in $\sigma(\Gamma)$
0	8822	0 %	501.0	0.0 %
10	8827	0.056 %	493.1	-1.6 %
20	8923	1.145 %	471.9	-5.8 %
30	9059	2.686 %	454.3	-9.3 %
40	9627	9.125 %	396.3	-20.9 %
50	10049	13.91 %	371.6	-25.8 %
60	10146	15.00 %	365.7	-27.0 %
70	10355	17.38 %	352.9	-29.6 %
80	10619	20.37 %	342.5	-31.6 %
100	10619	20.37 %	340.1	-32.1 %

$$
\begin{aligned}
Z^*(\Gamma) = \quad & \text{minimize} \quad \boldsymbol{c}'\boldsymbol{x} + \max_{\{S \mid S \subseteq J, |S| = \Gamma\}} \sum_{j \in S} d_j x_j \\
& \text{subject to} \quad \sum_{i \in N} x_i = k \\
& \qquad\qquad\quad \boldsymbol{x} \in \{0,1\}^n.
\end{aligned}
$$

10.2.1 Data

SLIDE 37

- $|N| = 200$;
- $k = 100$;
- $c_j \sim U[50, 200]$; $d_j \sim U[20, 200]$;
- For testing robustness, generate instances such that each cost component independently deviates with probability $\rho = 0.2$ from the nominal value c_j to $c_j + d_j$.

10.2.2 Results

SLIDE 38

10.3 Shortest Path

SLIDE 39

SLIDE 40

11 Conclusions

SLIDE 41

- Robust counterpart of a MIP remains a MIP, of comparable size.

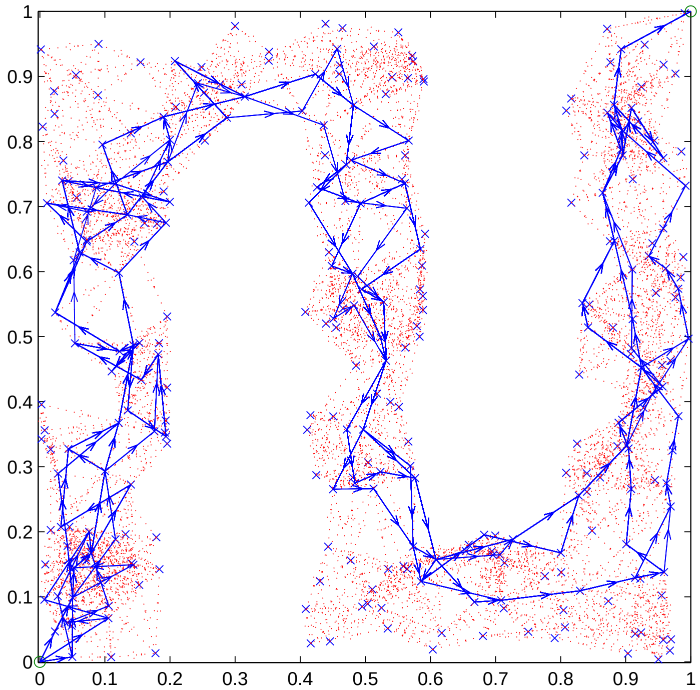

Figure 3: Randomly generated digraph.

- Approach permits flexibility of adjusting the level of conservatism in terms of probabilistic bound of constraint violation
- For polynomial solvable 0-1 optimization problems with cost uncertainty, the robust counterpart is polynomial solvable.
- For NP-hard 0-1 discrete problems that have α-approximation algorithm, the robust counterpart is also α-approximable.

SLIDE 42

- Robust network flows are solvable as a series of nominal network flow problems.
- Robust inventory control is numerically tractable even for large dimensions, and offers same policies as the stochastic case.
- Robust optimization is tractable for stochastic optimization problems without the curse of dimensionality

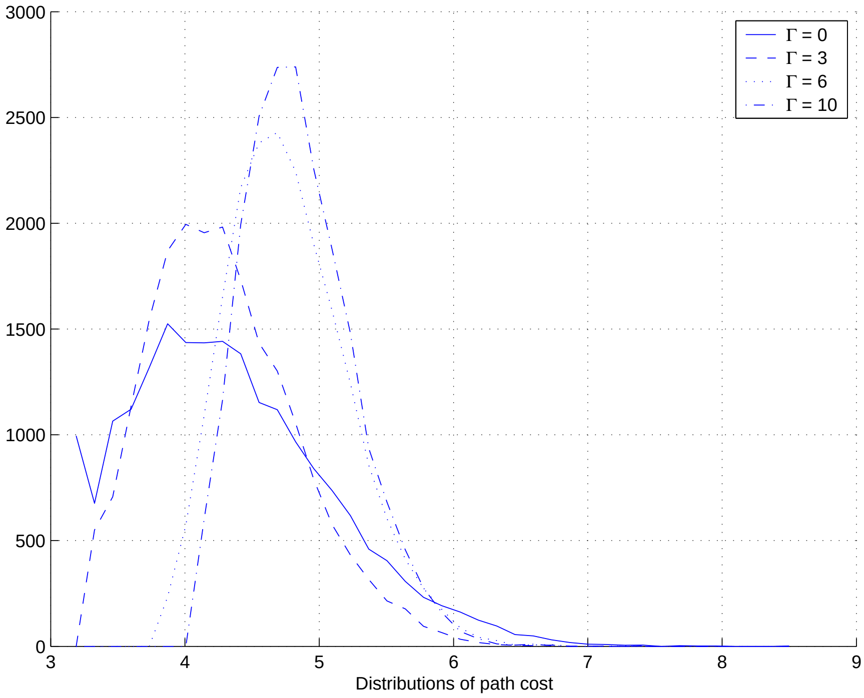

Figure 4: Influence of Γ on the distribution of path cost for $\rho = 0.1$.