Data-driven Distributionally Robust Optimization over Time

INFORMS Annual Meeting 2023

Kartikey Sharma

Joint work with Kevin-Martin Aigner, Kristin Braun, Sebastian Pokutta, Frauke Liers, Andreas Barman, Oscar Schneider, and Sebastian Tschupik

Motivation

Motivation

- Develop optimization under uncertainty model that incorporates new information as a part of optimization process

Motivation

- Develop optimization under uncertainty model that incorporates new information as a part of optimization process

- DRO can integrate new samples while accounting for uncertainty

$$\widehat{J}_t := \min_{x \in \mathcal{X}} \max_{p \in \mathcal{P}_t} \mathbb{E}_{s \sim p} [f(x, s)]$$

Motivation

- Develop optimization under uncertainty model that incorporates new information as a part of optimization process

- DRO can integrate new samples while accounting for uncertainty

$$\widehat{J}_t := \min_{x \in \mathcal{X}} \max_{p \in \mathcal{P}_t} \mathbb{E}_{s \sim p} [f(x,s)]$$

- DRO problems are difficult to solve

ZUSE INSTITUTE BERLIN

Motivation

- Develop optimization under uncertainty model that incorporates new information as a part of optimization process

- DRO can integrate new samples while accounting for uncertainty

$$\widehat{J}_t := \min_{x \in \mathcal{X}} \max_{p \in \mathcal{P}_t} \mathbb{E}_{s \sim p} [f(x, s)]$$

- DRO problems are difficult to solve
  - Dualization increases the size of the problem

Motivation

- Develop optimization under uncertainty model that incorporates new information as a part of optimization process

- DRO can integrate new samples while accounting for uncertainty

$$\widehat{J}_t := \min_{x \in \mathcal{X}} \max_{p \in \mathcal{P}_t} \mathbb{E}_{s \sim p} [f(x, s)]$$

- DRO problems are difficult to solve
  - Dualization increases the size of the problem
  - It also removes any special structure

Existing Work

Existing Work

DRO over Time

- Uses recomputation of exact solution with updated ambiguity set (Bayraksan and Love 2015, Esfahani and Kuhn 2018, Kirschner et. al. 2021 …)

Existing Work

DRO over Time

- Uses recomputation of exact solution with updated ambiguity set (Bayraksan and Love 2015, Esfahani and Kuhn 2018, Kirschner et. al. 2021 …)

Faster Solutions Methods

- Convert DRO problem into a regularisation problem (Namkoong and Duchi 2016, Chen et al 2017, Levy et al 2020 …)

ZUSE INSTITUTE BERLIN

Contributions

- We provide an Online Algorithm for DRO that simultaneously learns the uncertainty set and converges to the optimal solution
- We prove the consistency of our algorithm and also bound its regret over time
- We illustrate the performance of our method through numerical experiments on benchmark libraries.

Model

Model

Learning

- Observe samples from a distribution p^* (**finite dimension**)
- Use samples to construct ambiguity set $\mathscr{P}_0$ containing p^* with high probability

Model

Learning

- Observe samples from a distribution p^* (**finite dimension**)
- Use samples to construct ambiguity set $\mathscr{P}_0$ containing p^* with high probability

Optimization

- Leverage ambiguity set $\mathscr{P}_t$ to make robust decisions x_t

Model

Learning

- Observe samples from a distribution p^* (**finite dimension**)
- Use samples to construct ambiguity set $\mathscr{P}_0$ containing p^* with high probability

Optimization

- Leverage ambiguity set $\mathscr{P}_t$ to make robust decisions x_t

Learning

- Use new observations to update the set $\mathscr{P}_{t+1}$ and repeat

Ambiguity Sets

- Confidence Interval Ambiguity Set

$$\mathcal{P}_t = \left\{ p \in \mathcal{P}_0 \mid |p - \hat{p}_t| \leq \frac{z_{\frac{\delta_t}{2}}}{\sqrt{t}} \right\}$$

ZUSE INSTITUTE BERLIN

Ambiguity Sets

- Confidence Interval Ambiguity Set

$$\mathcal{P}_t = \left\{ p \in \mathcal{P}_0 \mid |p - \hat{p}_t| \leq \frac{z_{\frac{\delta_t}{2}}}{\sqrt{t}} \right\}$$

- ℓ_2−norm Ambiguity Set

$$\mathcal{P}_t = \left\{ p \in \mathcal{P}_0 \mid \|p - \hat{p}_t\|_2 \leq \frac{\sqrt{2|\mathcal{S} \log(2/\delta_t)}}{\sqrt{t}} \right\}$$

Ambiguity Sets

- Confidence Interval Ambiguity Set

$$\mathcal{P}_t = \left\{ p \in \mathcal{P}_0 \mid |p - \hat{p}_t| \leq \frac{z_{\frac{\delta_t}{2}}}{\sqrt{t}} \right\}$$

- ℓ_2−norm Ambiguity Set

$$\mathcal{P}_t = \left\{ p \in \mathcal{P}_0 \mid \|p - \hat{p}_t\|_2 \leq \frac{\sqrt{2|\mathcal{S}\log(2/\delta_t)}}{\sqrt{t}} \right\}$$

- Kernel based Ambiguity Set

$$\mathcal{P}_t = \left\{ p \in \mathcal{P}_0 \mid \|p - \hat{p}_t\|_M \leq \frac{\sqrt{C}}{\sqrt{t}}(2 + \sqrt{2\log(1/\delta_t)}) \right\}$$

Ambiguity Sets

- Confidence Interval Ambiguity Set

$$\mathcal{P}_t = \left\{ p \in \mathcal{P}_0 \mid |p - \hat{p}_t| \leq \frac{z_{\frac{\delta_t}{2}}}{\sqrt{t}} \right\}$$

- ℓ_2-norm Ambiguity Set

$$\mathcal{P}_t = \left\{ p \in \mathcal{P}_0 \mid \|p - \hat{p}_t\|_2 \leq \frac{\sqrt{2|\mathcal{S} \log(2/\delta_t)}}{\sqrt{t}} \right\}$$

- Kernel based Ambiguity Set

$$\mathcal{P}_t = \left\{ p \in \mathcal{P}_0 \mid \|p - \hat{p}_t\|_M \leq \frac{\sqrt{C}}{\sqrt{t}}(2 + \sqrt{2 \log(1/\delta_t)}) \right\}$$

- The sets include the true distribution with high probability

Ambiguity Sets

- Confidence Interval Ambiguity Set

$$\mathcal{P}_t = \left\{ p \in \mathcal{P}_0 \mid |p - \hat{p}_t| \leq \frac{z_{\frac{\delta_t}{2}}}{\sqrt{t}} \right\}$$

- ℓ_2-norm Ambiguity Set

$$\mathcal{P}_t = \left\{ p \in \mathcal{P}_0 \mid \|p - \hat{p}_t\|_2 \leq \frac{\sqrt{2|\mathcal{S} \log(2/\delta_t)}}{\sqrt{t}} \right\}$$

- Kernel based Ambiguity Set

$$\mathcal{P}_t = \left\{ p \in \mathcal{P}_0 \mid \|p - \hat{p}_t\|_M \leq \frac{\sqrt{C}}{\sqrt{t}}(2 + \sqrt{2 \log(1/\delta_t)}) \right\}$$

- The sets include the true distribution with high probability
- They shrink fast enough to compensate for the increasing probability requirements

Standard Reformulation of DRO

Standard Reformulation of DRO

- DRO Problem

$$\widehat{J}_t := \min_{x \in \mathcal{X}} \max_{p \in \mathcal{P}_t} \mathbb{E}_{s \sim p} [f(x, s)]$$

Standard Reformulation of DRO

- DRO Problem

$$\widehat{J}_t := \min_{x \in \mathcal{X}} \max_{p \in \mathcal{P}_t} \mathbb{E}_{s \sim p}\left[f(x, s)\right]$$

- Dual Reformulation for Interval Ambiguity Sets

$$\begin{aligned}
\min_{x,z,\alpha,\beta} \quad & z - \langle l_t, \alpha \rangle + \langle u_t, \beta \rangle \\
\text{s.t.} \quad & z - \alpha_k + \beta_k \geq f(x, s_k) \quad \forall k = 1, \dots, |\mathcal{S}|, \\
& \alpha, \beta \geq 0, \\
& x \in \mathcal{X}, \; z \in \mathbb{R}, \; \alpha, \beta \in \mathbb{R}^{|\mathcal{S}|}.
\end{aligned}$$

Standard Reformulation of DRO

- DRO Problem

$$\widehat{J}_t := \min_{x\in\mathcal{X}} \max_{p\in\mathcal{P}_t} \mathbb{E}_{s\sim p}\left[f(x,s)\right]$$

- Dual Reformulation for Interval Ambiguity Sets

$$\begin{aligned} \min_{x,z,\alpha,\beta} \quad & z - \langle l_t, \alpha\rangle + \langle u_t, \beta\rangle \\ \text{s.t.} \quad & z - \alpha_k + \beta_k \geq f(x, s_k) \quad \forall k = 1, \dots, |\mathcal{S}|, \\ & \alpha, \beta \geq 0, \\ & x \in \mathcal{X}, \; z \in \mathbb{R}, \; \alpha, \beta \in \mathbb{R}^{|\mathcal{S}|}. \end{aligned}$$

Algorithm 1 DRO over Time

Input: functions $f(\cdot, s)$ for $s \in \mathcal{S}$, feasible set $\mathcal{X}$, initial amb

Output: sequence of DRO solutions $x_1, \dots, x_T$

for $t = 1$ **to** T **do**

  $x_t \leftarrow$ solve Problem (1) or (2) for $\mathcal{P}_{t-1}$

  $\mathcal{P}_t \leftarrow$ observe data and update set parameters such as ambiguity set.

end for

Online Optimization

Online Optimization

- Initial solutions x_0 and p_0

Online Optimization

- Initial solutions x_0 and p_0
- Step by adversary p

Online Optimization

- Initial solutions x_0 and p_0
- Step by adversary p
- Step by decision maker x

Online Optimization

- Initial solutions x_0 and p_0
- Step by adversary p $$p_t = \arg\min_{p \in \mathcal{P}_{t-1}} \left\langle -\eta \nabla_p \mathbb{E}_{s \sim p_{t-1}} \left[f(x_{t-1}, s) \right], p \right\rangle + \frac{1}{2} \|p - p_{t-1}\|^2.$$
- Step by decision maker x

Online Optimization

- Initial solutions x_0 and p_0
- Step by adversary p $$p_t = \arg\min_{p \in \mathcal{P}_{t-1}} \left\langle -\eta \nabla_p \mathbb{E}_{s \sim p_{t-1}} \left[f(x_{t-1}, s) \right], p \right\rangle + \frac{1}{2} \|p - p_{t-1}\|^2.$$
- Step by decision maker x $$x_t = \arg\min_{x \in \mathcal{X}} \mathbb{E}_{s \sim p_t} [f(x, s)]$$

Online Optimization

Algorithm 2 DRO over Time with Online Projected Gradient Descent

1: **Input:** functions $f(\cdot, s)$ for $s \in \mathcal{S}$, feasible set $\mathcal{X}$, initial ambiguity set $\mathcal{P}_0$

2: **Output:** $x_1, \ldots, x_T$

Online Optimization

Algorithm 2 DRO over Time with Online Projected Gradient Descent

Input: functions $f(\cdot, s)$ for $s \in \mathcal{S}$, feasible set $\mathcal{X}$, initial ambiguity set $\mathcal{P}_0$

Output: $x_1, \ldots, x_T$

Set $\mathcal{P}_0 = \{p \in [0,1]^{|\mathcal{S}|} \mid \sum_{k=1}^{|\mathcal{S}|} p_k = 1\}$

Online Optimization

Algorithm 2 DRO over Time with Online Projected Gradient Descent

Input: functions $f(\cdot, s)$ for $s \in \mathcal{S}$, feasible set $\mathcal{X}$, initial ambiguity set $\mathcal{P}_0$

Output: $x_1, \ldots, x_T$

Set $\mathcal{P}_0 = \{p \in [0,1]^{|\mathcal{S}|} \mid \sum_{k=1}^{|\mathcal{S}|} p_k = 1\}$

Set $p_0 = \left(\frac{1}{|\mathcal{S}|}, \ldots, \frac{1}{|\mathcal{S}|}\right) \in [0,1]^{|\mathcal{S}|}$

Set $x_0 = \min_{x \in \mathcal{X}} f(x, s_1)$

for $t = 1$ **to** T **do**

Online Optimization

Algorithm 2 DRO over Time with Online Projected Gradient Descent

Input: functions $f(\cdot, s)$ for $s \in \mathcal{S}$, feasible set $\mathcal{X}$, initial ambiguity set $\mathcal{P}_0$

Output: $x_1, \ldots, x_T$

Set $\mathcal{P}_0 = \{p \in [0,1]^{|\mathcal{S}|} \mid \sum_{k=1}^{|\mathcal{S}|} p_k = 1\}$

Set $p_0 = \left(\frac{1}{|\mathcal{S}|}, \ldots, \frac{1}{|\mathcal{S}|}\right) \in [0,1]^{|\mathcal{S}|}$

Set $x_0 = \min_{x \in \mathcal{X}} f(x, s_1)$

for $t = 1$ **to** T **do**

  $\tilde{p}_t \leftarrow p_{t-1} + \eta \nabla_p \mathbb{E}_{s \sim p_{t-1}} [f(x_{t-1}, s)]$

  $p_t \leftarrow \arg\min_{p \in \mathcal{P}_{t-1}} \frac{1}{2} \|p - \tilde{p}_t\|^2$

Online Optimization

Algorithm 2 DRO over Time with Online Projected Gradient Descent

Input: functions $f(\cdot, s)$ for $s \in \mathcal{S}$, feasible set $\mathcal{X}$, initial ambiguity set $\mathcal{P}_0$

Output: $x_1, \ldots, x_T$

Set $\mathcal{P}_0 = \{p \in [0,1]^{|\mathcal{S}|} \mid \sum_{k=1}^{|\mathcal{S}|} p_k = 1\}$

Set $p_0 = \left(\frac{1}{|\mathcal{S}|}, \ldots, \frac{1}{|\mathcal{S}|}\right) \in [0,1]^{|\mathcal{S}|}$

Set $x_0 = \min_{x \in \mathcal{X}} f(x, s_1)$

for $t = 1$ **to** T **do**

    $\tilde{p}_t \leftarrow p_{t-1} + \eta \nabla_p \mathbb{E}_{s \sim p_{t-1}} [f(x_{t-1}, s)]$

    $p_t \leftarrow \arg\min_{p \in \mathcal{P}_{t-1}} \frac{1}{2} \|p - \tilde{p}_t\|^2$

    $x_t \leftarrow \arg\min_{x \in \mathcal{X}} \mathbb{E}_{s \sim p_t} [f(x, s)]$

Online Optimization

Algorithm 2 DRO over Time with Online Projected Gradient Descent

Input: functions $f(\cdot, s)$ for $s \in \mathcal{S}$, feasible set $\mathcal{X}$, initial ambiguity set $\mathcal{P}_0$

Output: $x_1, \dots, x_T$

Set $\mathcal{P}_0 = \{p \in [0,1]^{|\mathcal{S}|} \mid \sum_{k=1}^{|\mathcal{S}|} p_k = 1\}$

Set $p_0 = \left(\frac{1}{|\mathcal{S}|}, \dots, \frac{1}{|\mathcal{S}|}\right) \in [0,1]^{|\mathcal{S}|}$

Set $x_0 = \min_{x \in \mathcal{X}} f(x, s_1)$

for $t = 1$ **to** T **do**

  $\tilde{p}_t \leftarrow p_{t-1} + \eta \nabla_p \mathbb{E}_{s \sim p_{t-1}} [f(x_{t-1}, s)]$

  $p_t \leftarrow \arg\min_{p \in \mathcal{P}_{t-1}} \frac{1}{2} \|p - \tilde{p}_t\|^2$

  $x_t \leftarrow \arg\min_{x \in \mathcal{X}} \mathbb{E}_{s \sim p_t} [f(x, s)]$

  $\mathcal{P}_t \leftarrow$ observe data and update set parameters such as $\hat{p}_t, l_t, u_t$ and ϵ_t as per the type of ambiguity set.

Online Optimization

Algorithm 2 DRO over Time with Online Projected Gradient Descent

Input: functions $f(\cdot, s)$ for $s \in \mathcal{S}$, feasible set $\mathcal{X}$, initial ambiguity set $\mathcal{P}_0$

Output: $x_1, \dots, x_T$

Set $\mathcal{P}_0 = \{p \in [0,1]^{|\mathcal{S}|} \mid \sum_{k=1}^{|\mathcal{S}|} p_k = 1\}$

Set $p_0 = \left(\frac{1}{|\mathcal{S}|}, \dots, \frac{1}{|\mathcal{S}|}\right) \in [0,1]^{|\mathcal{S}|}$

Set $x_0 = \min_{x \in \mathcal{X}} f(x, s_1)$

for $t = 1$ **to** T **do**

  $\tilde{p}_t \leftarrow p_{t-1} + \eta \nabla_p \mathbb{E}_{s \sim p_{t-1}} [f(x_{t-1}, s)]$

  $p_t \leftarrow \arg\min_{p \in \mathcal{P}_{t-1}} \frac{1}{2} \|p - \tilde{p}_t\|^2$

  $x_t \leftarrow \arg\min_{x \in \mathcal{X}} \mathbb{E}_{s \sim p_t} [f(x, s)]$

  $\mathcal{P}_t \leftarrow$ observe data and update set parameters such as $\hat{p}_t, l_t, u_t$ and ϵ_t as per the type of ambiguity set.

end for

Consistency

Consistency

With probability 1 we have

$$\lim_{t\to\infty} \widehat{J}_t = J^*$$

$$\lim_{t\to\infty} x_t = x^*$$

Consistency

With probability 1 we have

$$\lim_{t \to \infty} \widehat{J}_t = J^*$$

$$\lim_{t \to \infty} x_t = x^*$$

Proof technique:

- True distribution inside ambiguity set
- The ambiguity set converges to the true distribution

ZUSE INSTITUTE BERLIN

Regret Bound on Solutions

With probability at least $1 - \delta$ we have

$$\frac{1}{T}\sum_{t=1}^{T}\left(\max_{p\in\mathcal{P}_t}\mathbb{E}_{s\sim p}\left[f(x_t,s)\right]-\min_{x\in\mathcal{X}}\max_{p\in\mathcal{P}_t}\mathbb{E}_{s\sim p}\left[f(x,s)\right]\right)\leq G\sqrt{\frac{|\mathcal{S}|h(T)}{2T}}+\frac{2G}{T},$$

Regret Bound on Solutions

With probability at least $1-\delta$ we have

$$\frac{1}{T}\sum_{t=1}^{T}\left(\max_{p\in\mathcal{P}_t}\mathbb{E}_{s\sim p}\left[f(x_t,s)\right]-\min_{x\in\mathcal{X}}\max_{p\in\mathcal{P}_t}\mathbb{E}_{s\sim p}\left[f(x,s)\right]\right)\leq G\sqrt{\frac{|\mathcal{S}|h(T)}{2T}}+\frac{2G}{T},$$

Here

$$h(T)\in\mathcal{O}(|\mathcal{S}|\log^2(T))$$

Regret Bound on Solutions

With probability at least $1 - \delta$ we have

$$\frac{1}{T}\sum_{t=1}^{T}\left(\max_{p\in\mathcal{P}_t}\mathbb{E}_{s\sim p}\left[f(x_t,s)\right]-\min_{x\in\mathcal{X}}\max_{p\in\mathcal{P}_t}\mathbb{E}_{s\sim p}\left[f(x,s)\right]\right)\leq G\sqrt{\frac{|\mathcal{S}|h(T)}{2T}}+\frac{2G}{T},$$

Here

$$h(T) \in \mathcal{O}(|\mathcal{S}|\log^2(T))$$

- Linear Dependence on Scenarios

Proof

Proof

- Bound the regret term by the linear drop in the function value

$$\sum_{t=1}^{T}\left(\max_{p\in\mathcal{P}_t}\mathbb{E}_{s\sim p}\left[f(x_t,s)\right]-\min_{x\in\mathcal{X}}\mathbb{E}_{s\sim p_t}\left[f(x,s)\right]\right)=\sum_{t=1}^{T}\langle\nabla g_t(p_t),p_t-u_t\rangle$$

Proof

- Bound the regret term by the linear drop in the function value

$$\sum_{t=1}^{T}\left(\max_{p\in\mathcal{P}_t}\mathbb{E}_{s\sim p}\left[f(x_t,s)\right]-\min_{x\in\mathcal{X}}\mathbb{E}_{s\sim p_t}\left[f(x,s)\right]\right)=\sum_{t=1}^{T}\langle\nabla g_t(p_t),p_t-u_t\rangle$$

- Bound the cumulative linear drop in function value by a bound on the gradient and the cumulative length of the steps taken

$$\sum_{t=1}^{T}\langle\eta\nabla g_t(p_t),p_t-u_t\rangle\leq\sum_{t=1}^{T}\frac{\eta^2}{2}\|\nabla g_t(p_t)\|^2+\sum_{t=1}^{T}\left(\frac{1}{2}\|p_t-u_t\|^2-\frac{1}{2}\|p_{t+1}-u_t\|^2\right)$$

Proof

- Bound the regret term by the linear drop in the function value

$$\sum_{t=1}^{T}\left(\max_{p\in\mathcal{P}_t}\mathbb{E}_{s\sim p}\left[f(x_t,s)\right]-\min_{x\in\mathcal{X}}\mathbb{E}_{s\sim p_t}\left[f(x,s)\right]\right)=\sum_{t=1}^{T}\langle\nabla g_t(p_t),p_t-u_t\rangle$$

- Bound the cumulative linear drop in function value by a bound on the gradient and the cumulative length of the steps taken

$$\sum_{t=1}^{T}\langle\eta\nabla g_t(p_t),p_t-u_t\rangle\leq\sum_{t=1}^{T}\frac{\eta^2}{2}\|\nabla g_t(p_t)\|^2+\sum_{t=1}^{T}\left(\frac{1}{2}\|p_t-u_t\|^2-\frac{1}{2}\|p_{t+1}-u_t\|^2\right)$$

- Bound the cumulative length of the steps on the basis of the uncertainty set size.

$$\sum_{t=1}^{T}\left(\frac{1}{2}\|p_t-u_t\|^2-\frac{1}{2}\|p_{t+1}-u_t\|^2\right)\leq\sum_{t=1}^{T}\frac{1}{2}\|p_t-u_t\|^2$$

Regret Bound on Solutions

Regret Bound on Solutions

- Primarily focuses on the gap between adversarial solutions p_t

Regret Bound on Solutions

- Primarily focuses on the gap between adversarial solutions p_t
- Uses construction of ambiguity sets to bound the gap on the sequence of generated adversarial solutions

Regret Bound on Solutions

- Primarily focuses on the gap between adversarial solutions p_t
- Uses construction of ambiguity sets to bound the gap on the sequence of generated adversarial solutions

$$\frac{1}{2}\sum_{t=1}^{T}\|p_t - q_t\|^2 \leq h(T)$$

$$p_t \in \mathcal{P}_{t-1}, q_t \in \mathcal{P}_t$$

Regret Bound on Solutions

- Primarily focuses on the gap between adversarial solutions p_t
- Uses construction of ambiguity sets to bound the gap on the sequence of generated adversarial solutions

$$\frac{1}{2}\sum_{t=1}^{T}\|p_t - q_t\|^2 \leq h(T)$$

$$p_t \in \mathcal{P}_{t-1}, q_t \in \mathcal{P}_t$$

Confidence Intervals $h(T) = 8|\mathcal{S}|\log(\pi T)(1+\log T)$

ℓ_2-norm Sets $h(T) = 8|\mathcal{S}|\log\frac{\pi T}{\sqrt{3\delta}}(1+\log T)$

Kernel based Sets $h(T) = \frac{32C}{\lambda^2} + \frac{32C}{\lambda^2}\log\frac{\pi T}{\sqrt{6\delta}}(1+\log T)$

Numerical Experiments

Benchmark Libraries

- MILPs and MIQPs from the MIPLIB set of benchmark Instances
- Comparisons against other methods

Distributionally Robust Network Design

- Network design with uncertain demands. Instances by Altin et. al. (2007)

Optimal Route Choice

- *ChicagoSketch* model from Transportation Networks library
- Illustration of impact on solutions

Benchmark Libraries

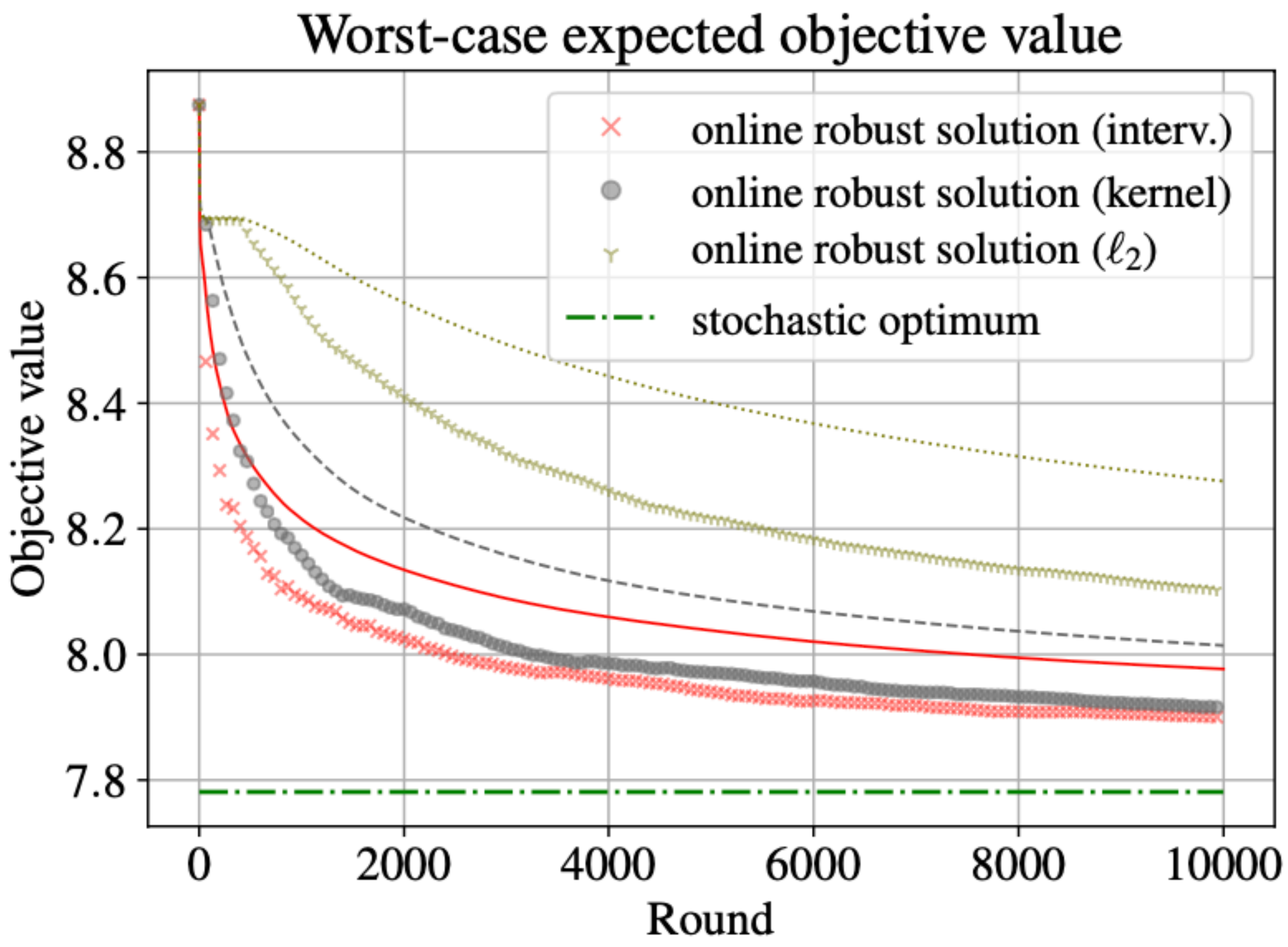
- Takes MILP and MIQP instances from MIPLIB library
- Instances are of the form

$$f(x, s) = x^\top Q x + (c + s)^\top x + d$$

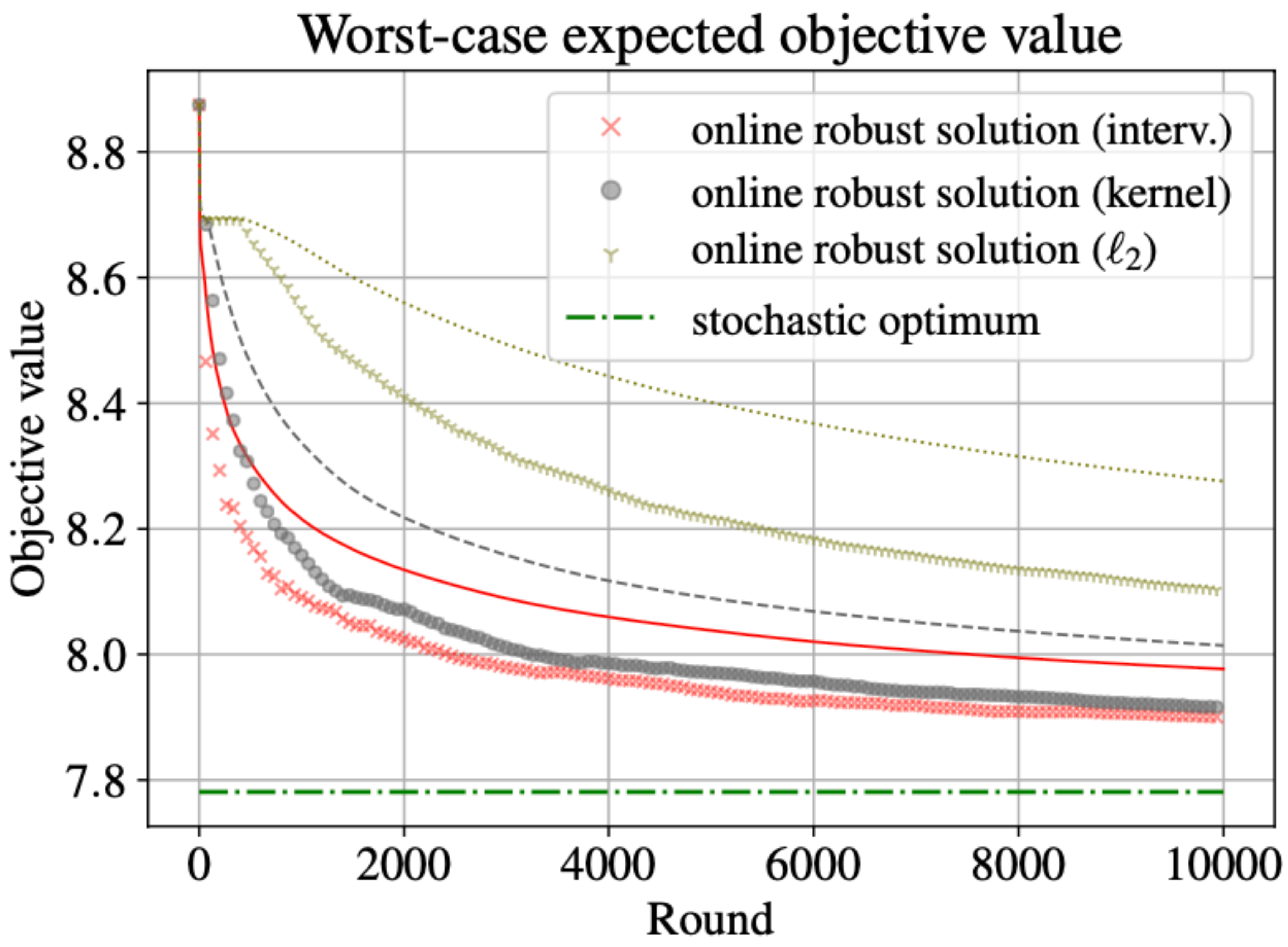
- **Objective** uncertainty in the instances through scenarios $s \in \mathcal{S}$ with $|\mathcal{S}| = (2, 10, 15)$

Benchmark Instances: Different Ambiguity Sets

Benchmark Instances: Different Ambiguity Sets



Benchmark Instances: Different Ambiguity Sets



The objective value shrinks for all set types

Fastest reduction for confidence intervals

Benchmark Instances: Running Times

Benchmark Instances: Running Times

	$\lvert\mathcal{S}\rvert$	Online Robust	Exact DRO
MIP (I)	10	52.4s	115.8s
MIP (ℓ_2)	10	49.4s	127.5s
MIP (K)	10	56.3s	129.5s
MIP (I)	50	57.7s	176.7s*
MIP (ℓ_2)	50	60.4s	206.1s*
MIP (K)	50	67.0s	244.4s*
MIQP (I)	2	170.2s	271.4s*
MIQP (ℓ_2)	2	186.3s	329.5s*
MIQP (K)	2	188.6s	359.6s*

Benchmark Instances: Running Times

	$\|\mathcal{S}\|$	Online Robust	Exact DRO
MIP (I)	10	52.4s	115.8s
MIP (ℓ_2)	10	49.4s	127.5s
MIP (K)	10	56.3s	129.5s
MIP (I)	50	57.7s	176.7s*
MIP (ℓ_2)	50	60.4s	206.1s*
MIP (K)	50	67.0s	244.4s*
MIQP (I)	2	170.2s	271.4s*
MIQP (ℓ_2)	2	186.3s	329.5s*
MIQP (K)	2	188.6s	359.6s*

More time savings for large and non linear problems

Benchmark Instances: Running Times

Benchmark Instances: Running Times

	MIP $\lvert\mathcal{S}\rvert = 10$	MIP $\lvert\mathcal{S}\rvert = 50$	MIQP $\lvert\mathcal{S}\rvert = 2$
DRO	45.6s	55.9s*	271.4s*
Wassertein	52.3s	59.1s	299.9s
DRBO	42.7s**	66.1s**	738.3s*
Online robust	26.8s	27.1s	170.2s
Running SO	26.6s	26.9s	172.6s

Benchmark Instances: Running Times

	MIP $\lvert\mathcal{S}\rvert = 10$	MIP $\lvert\mathcal{S}\rvert = 50$	MIQP $\lvert\mathcal{S}\rvert = 2$
DRO	45.6s	55.9s*	271.4s*
Wassertein	52.3s	59.1s	299.9s
DRBO	42.7s**	66.1s**	738.3s*
Online robust	26.8s	27.1s	170.2s
Running SO	26.6s	26.9s	172.6s

Online robust methods are significantly faster

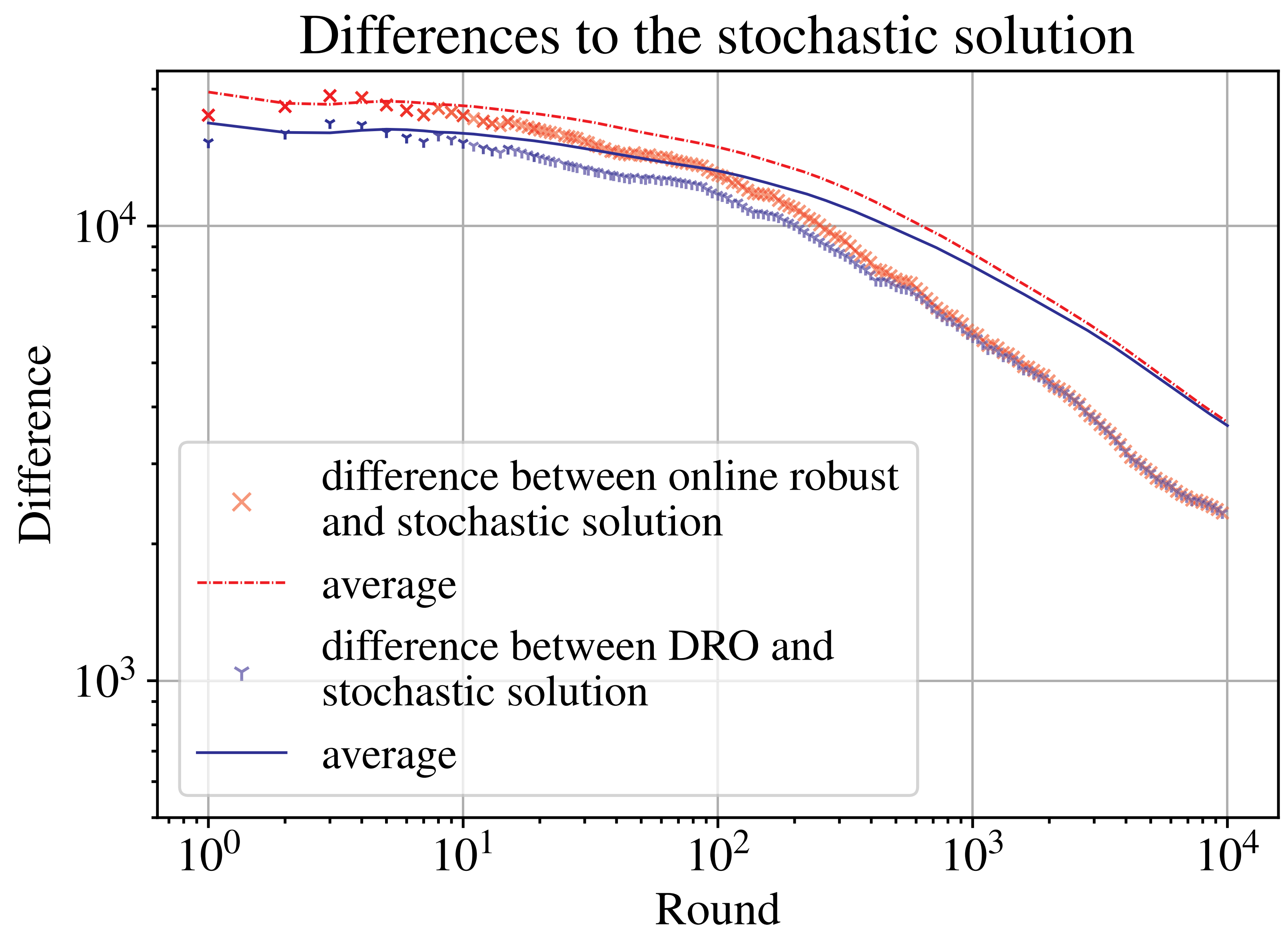
ZUSE INSTITUTE BERLIN

Distributionally Robust Network Design

- Compute minimum cost network topology and edge capacity to satisfy demand
- Demand is uncertain. Interval ambiguity sets.
- **Instances**
  - res8: V = 50, E = 77
  - w1_100: V = 100, E = 207
  - w1_200: V = 200, E = 775

Running Times

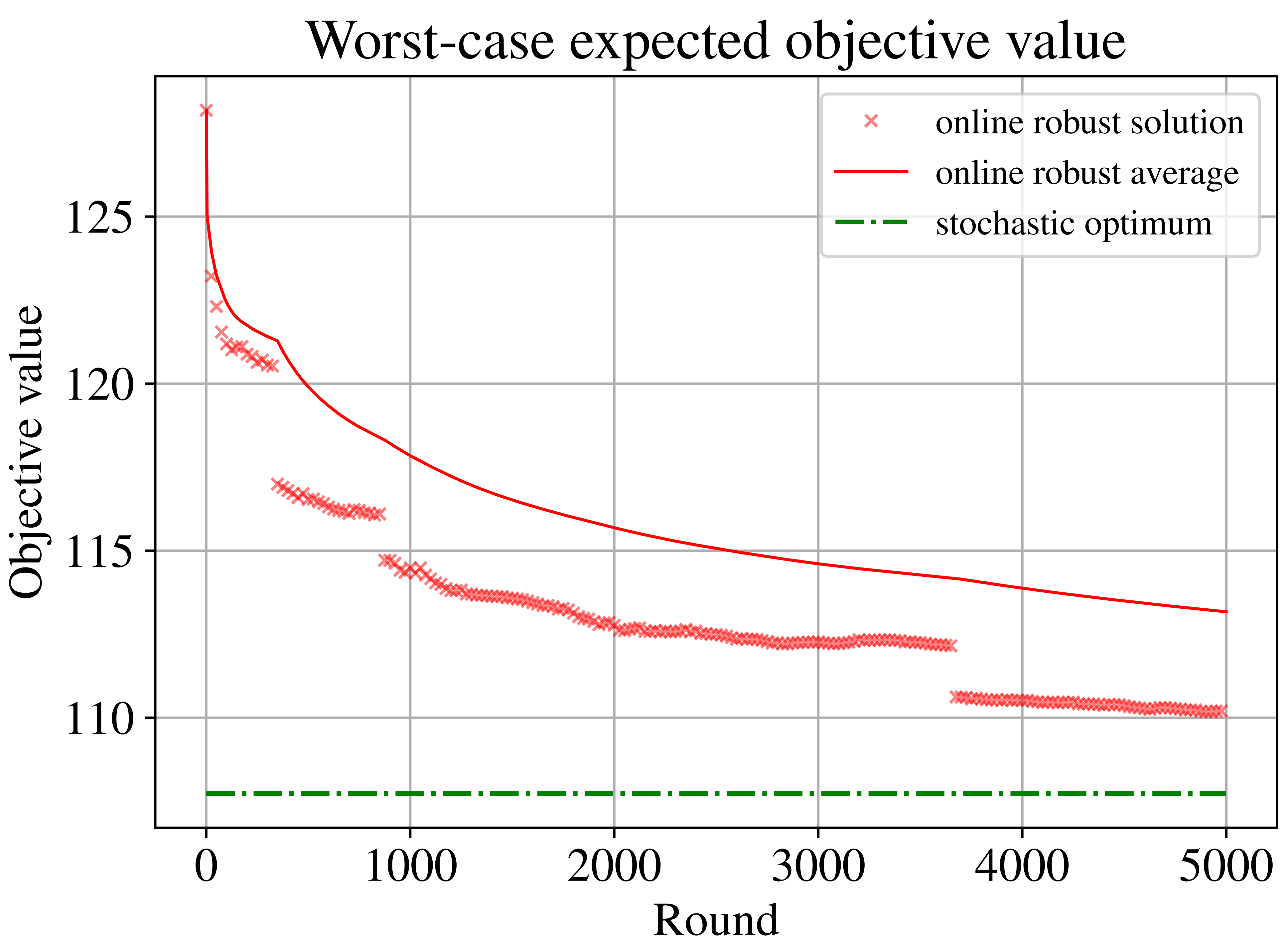
	$\lvert\mathcal{S}\rvert$	Online robust	Exact DRO
res8	10	0.2s	0.5s
res8	50	0.6s	11.6s
w1_100	10	0.3s	32.0s
w1_100	50	1.5s	95.6s
w1_200	10	1.2s	38.7s
w1_200	50	4.7s	1282.2s

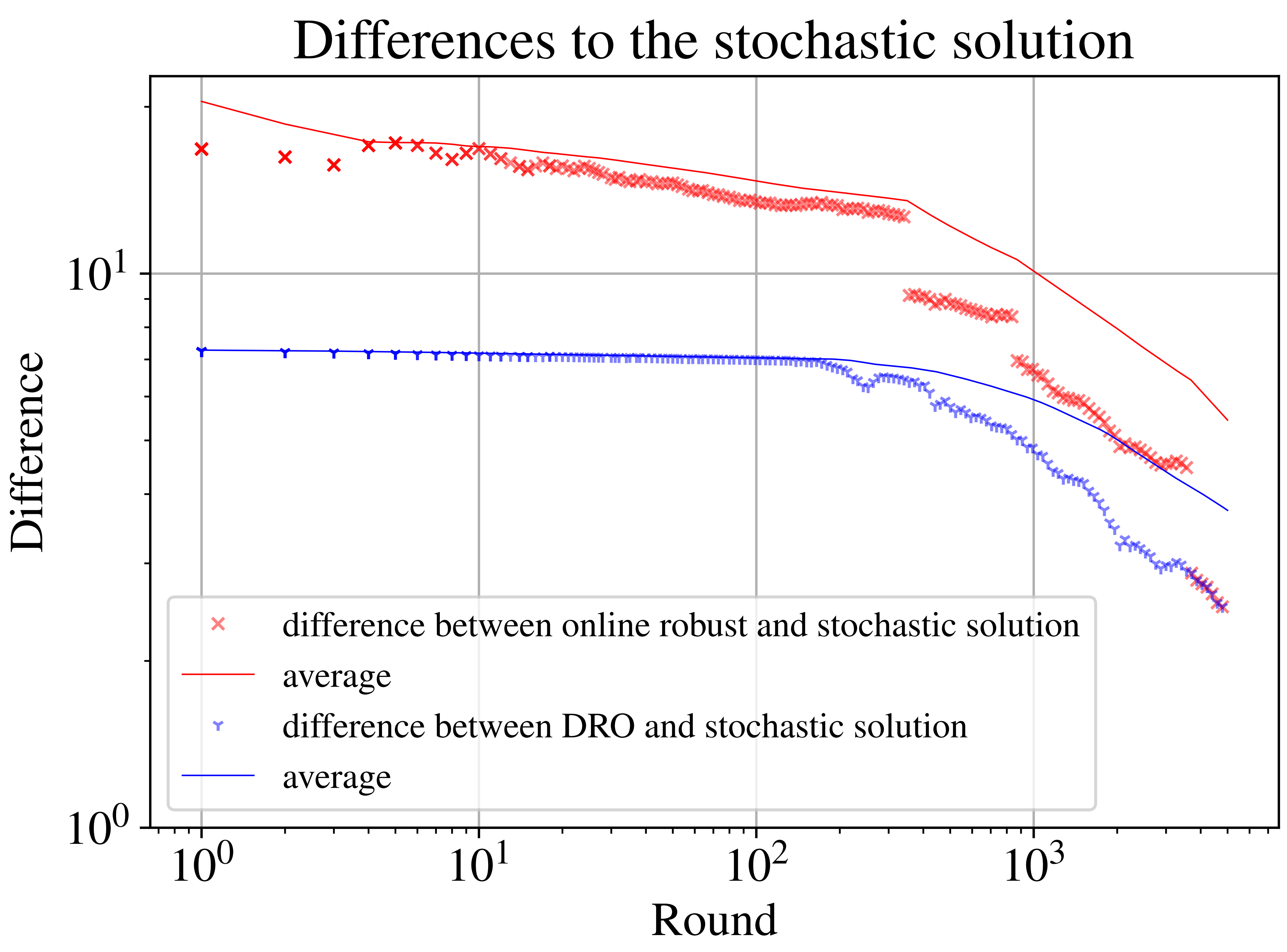


Optimal Route Choice

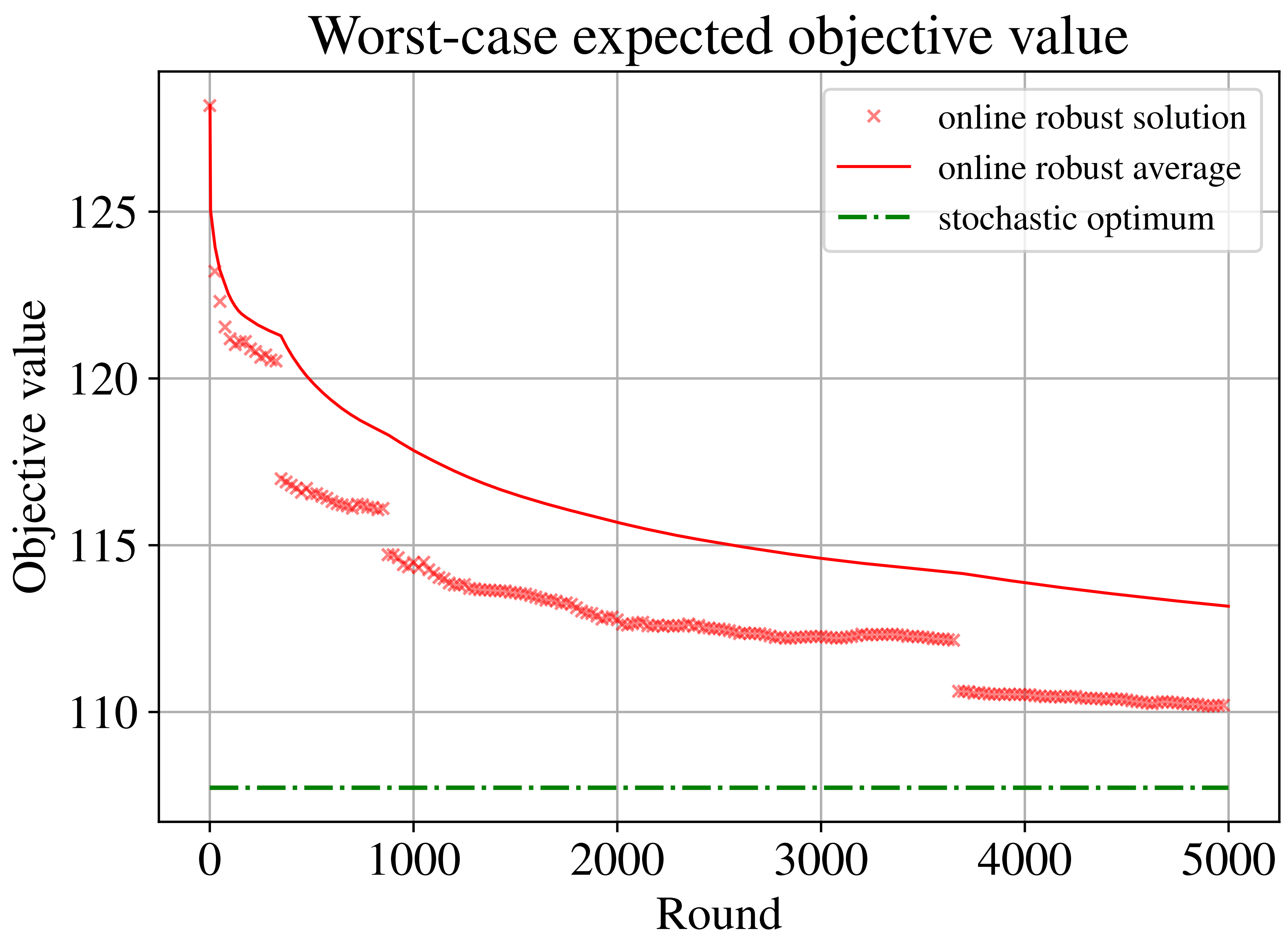
- Choose the shortest paths in a street network with uncertain arc times
- Model: *ChicagoSketch* with 933 nodes and 2950 arcs
- Randomly generated *true* probability distribution for arcs lengths
- Solving directly eliminates structure

Optimal Route Choice

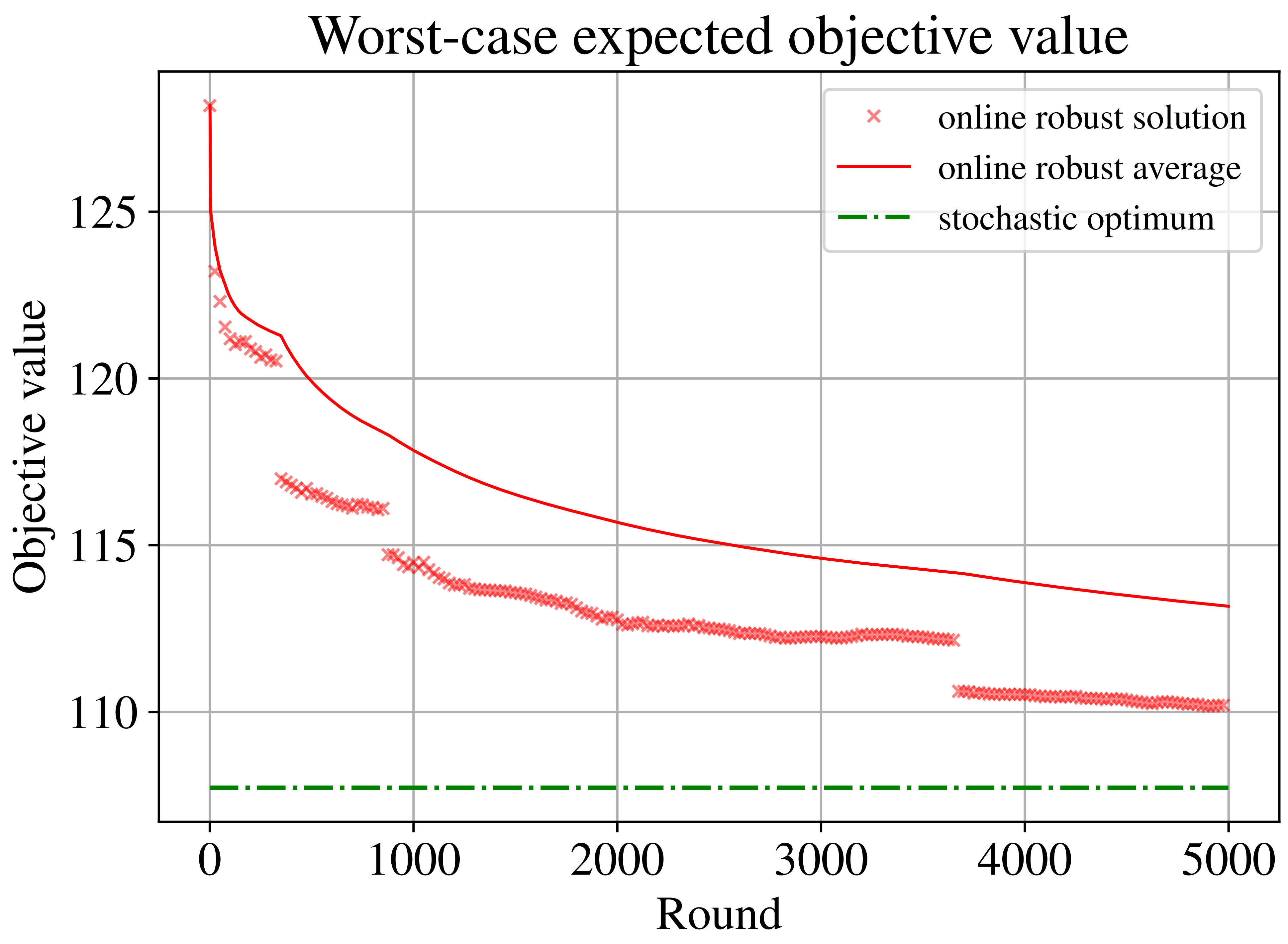


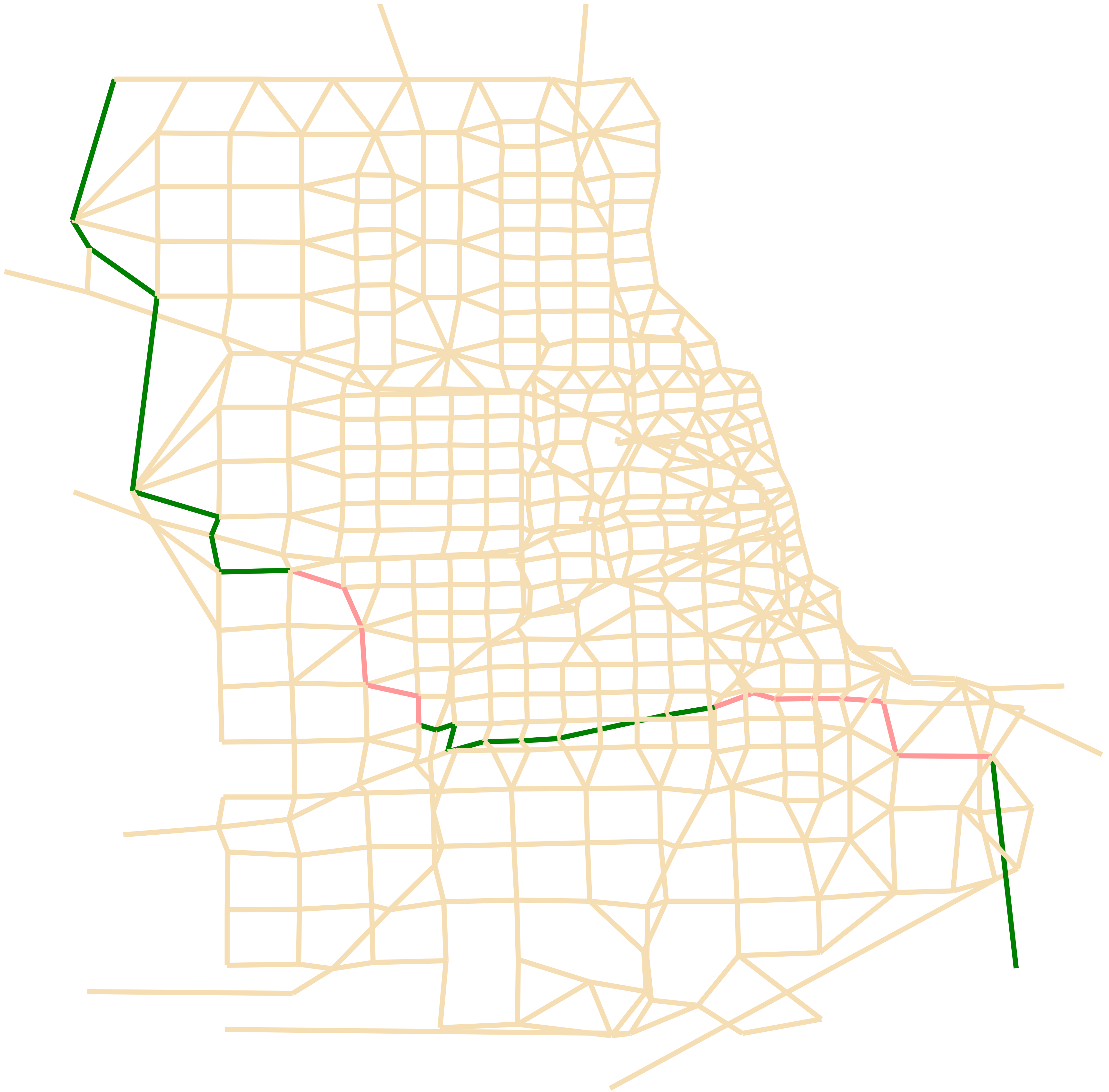


Optimal Route Choice

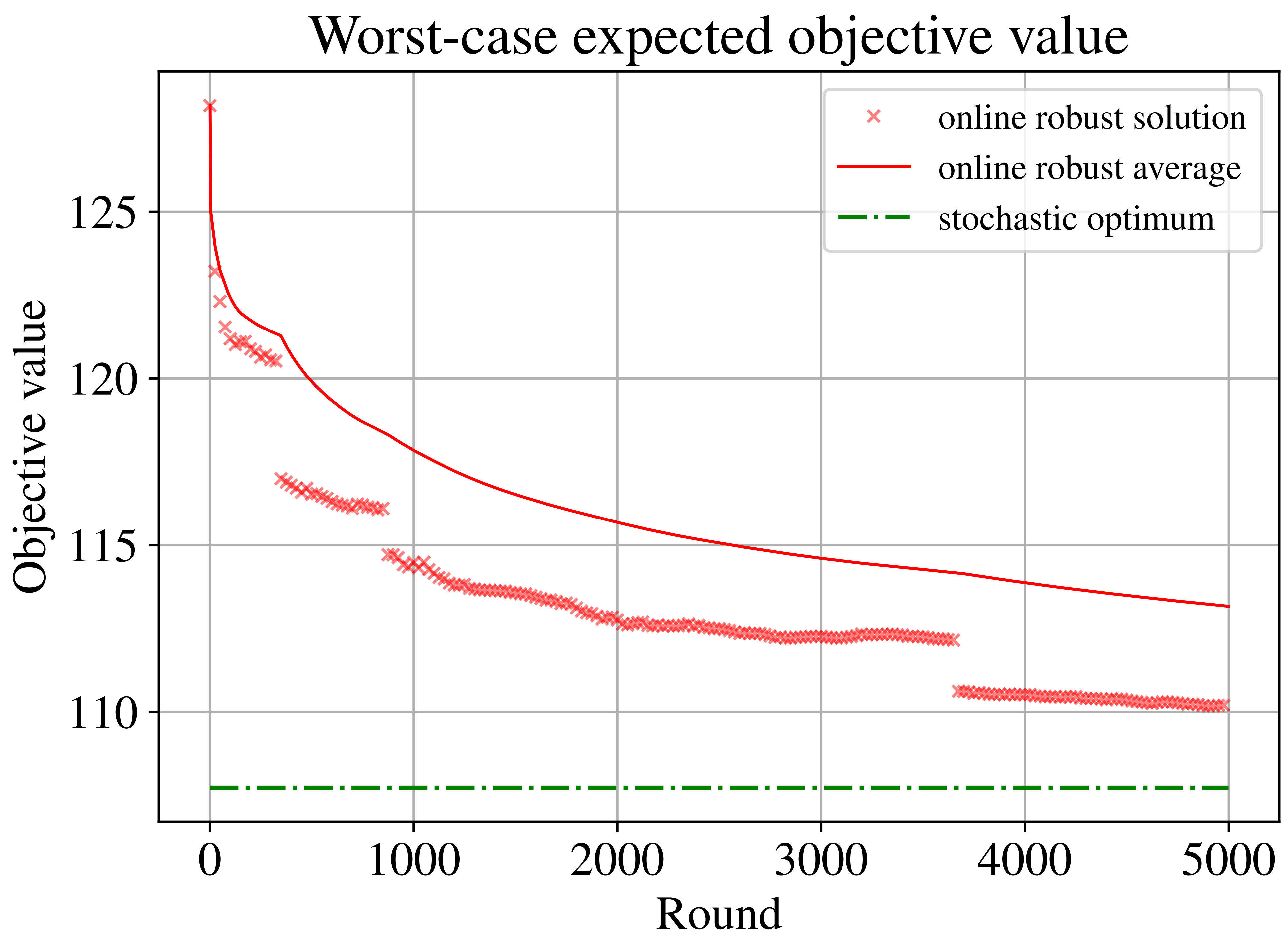


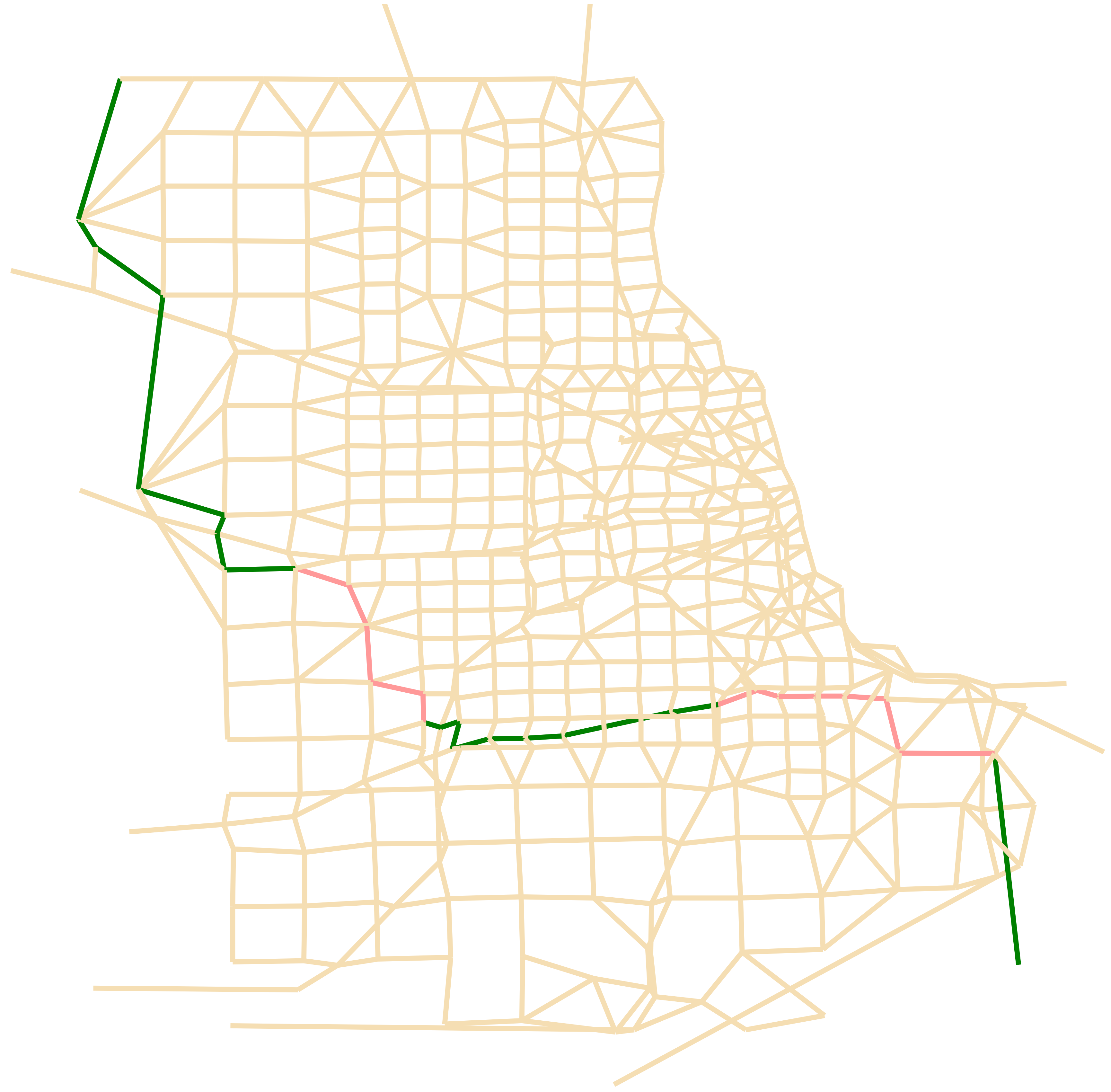
Optimal Route Choice



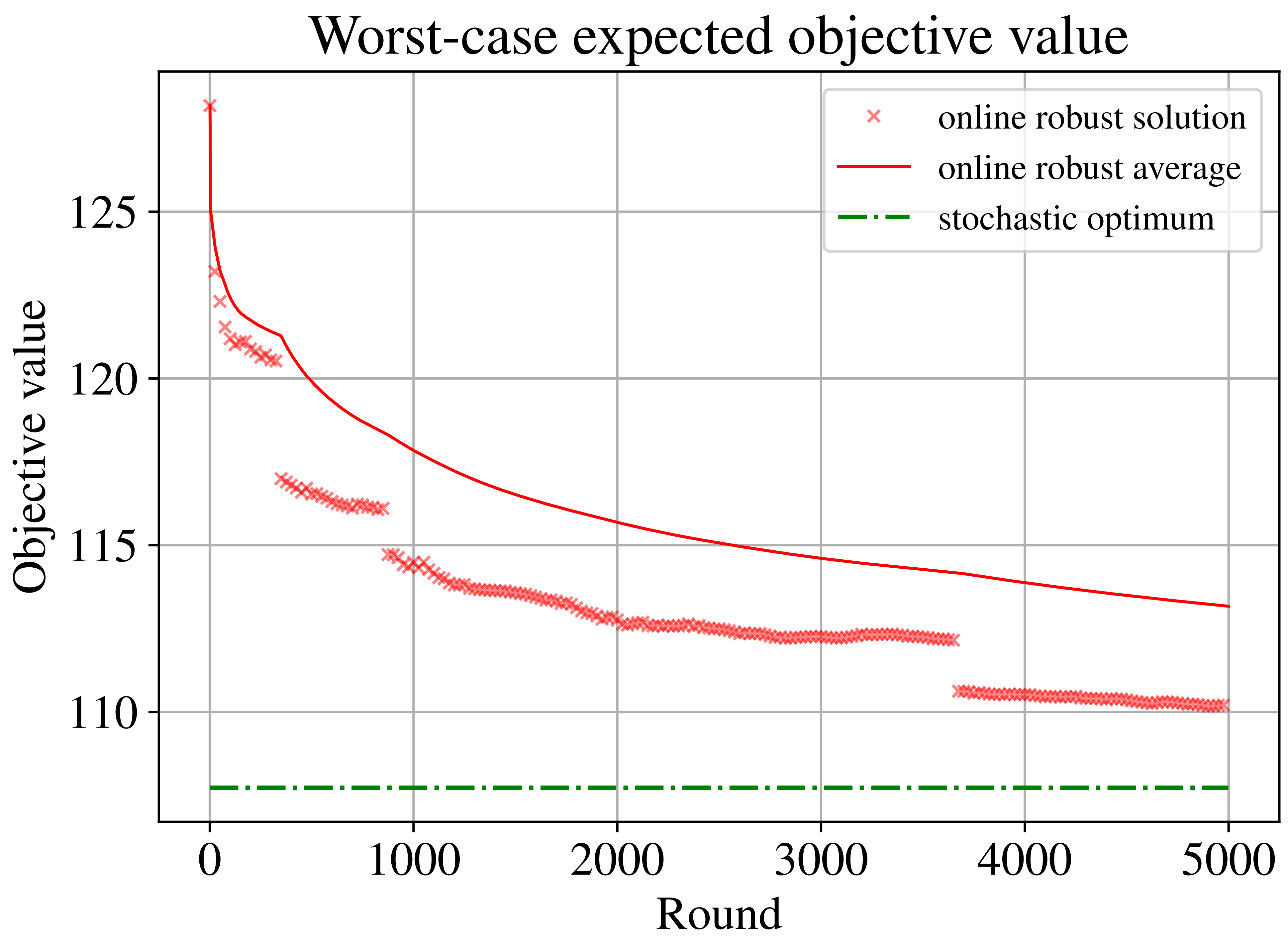


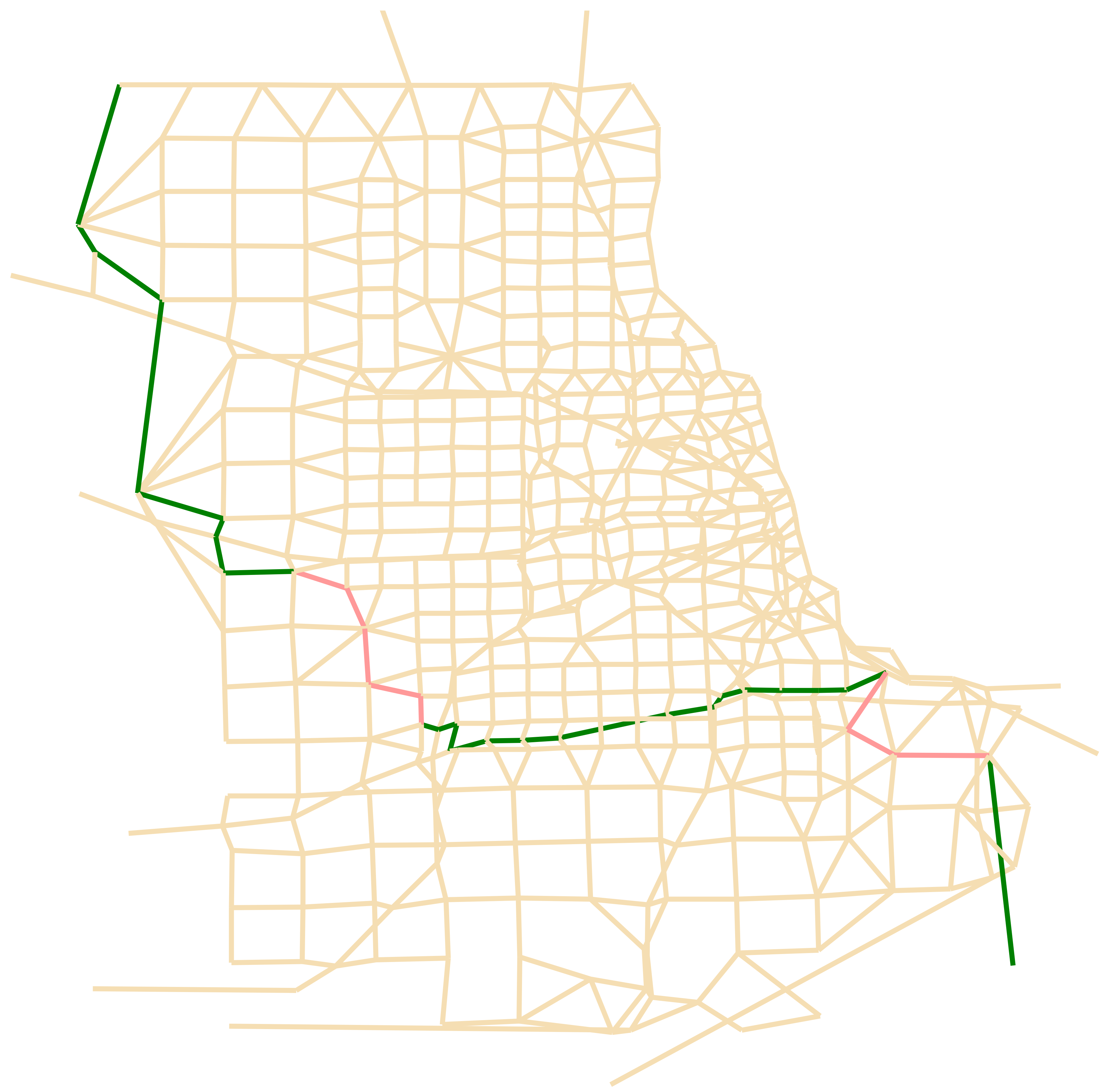
Optimal Route Choice



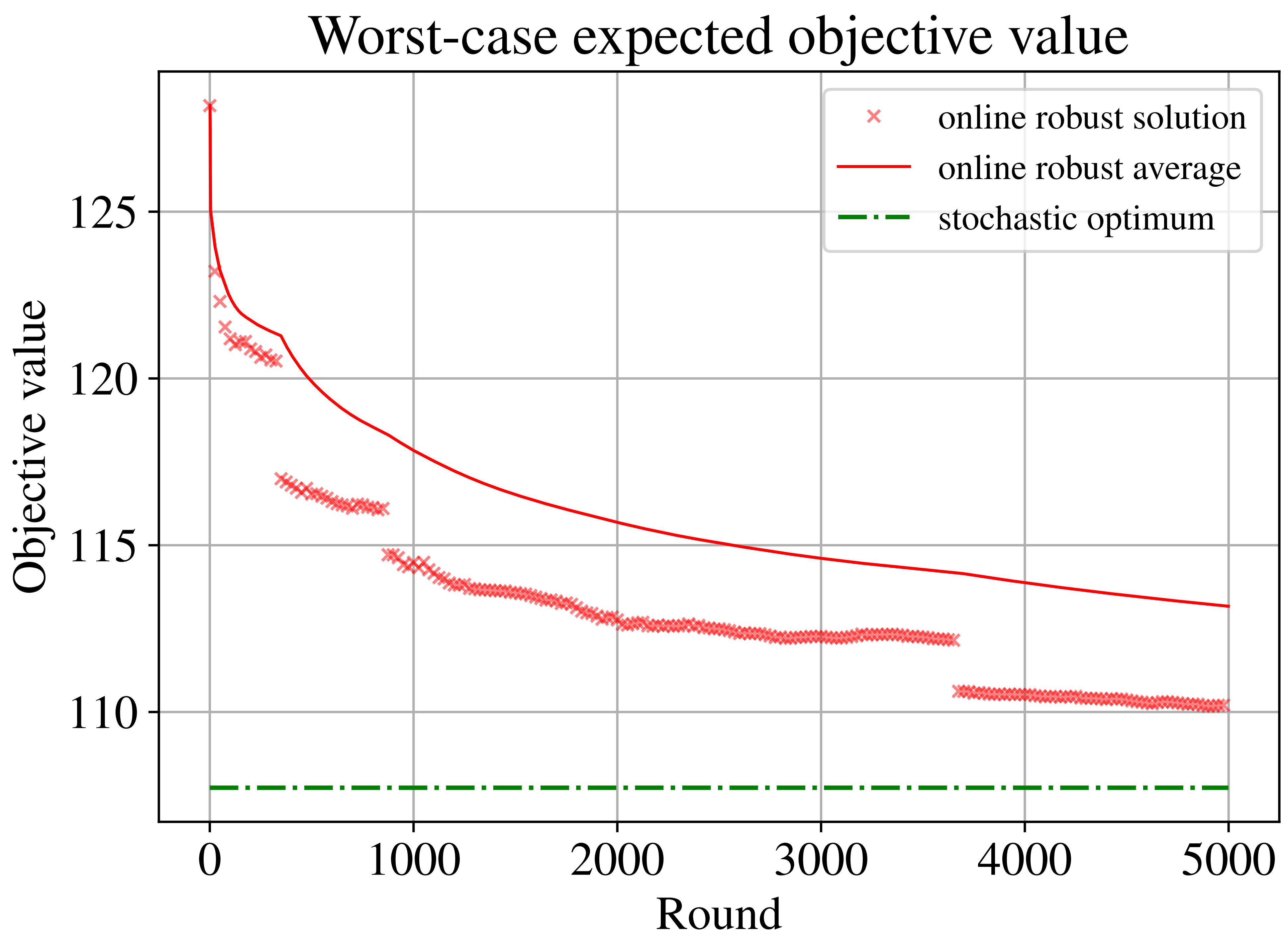


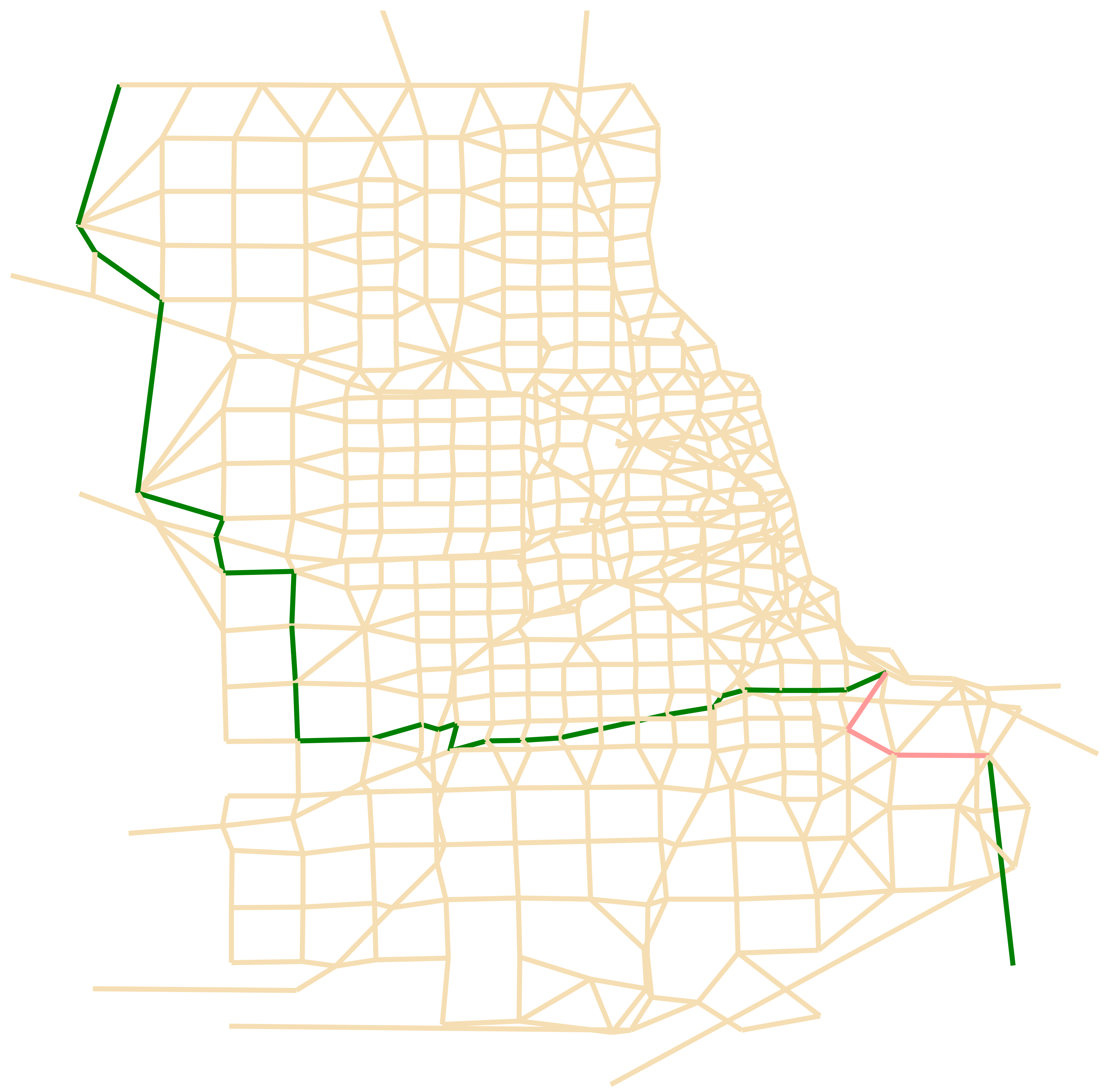
Optimal Route Choice



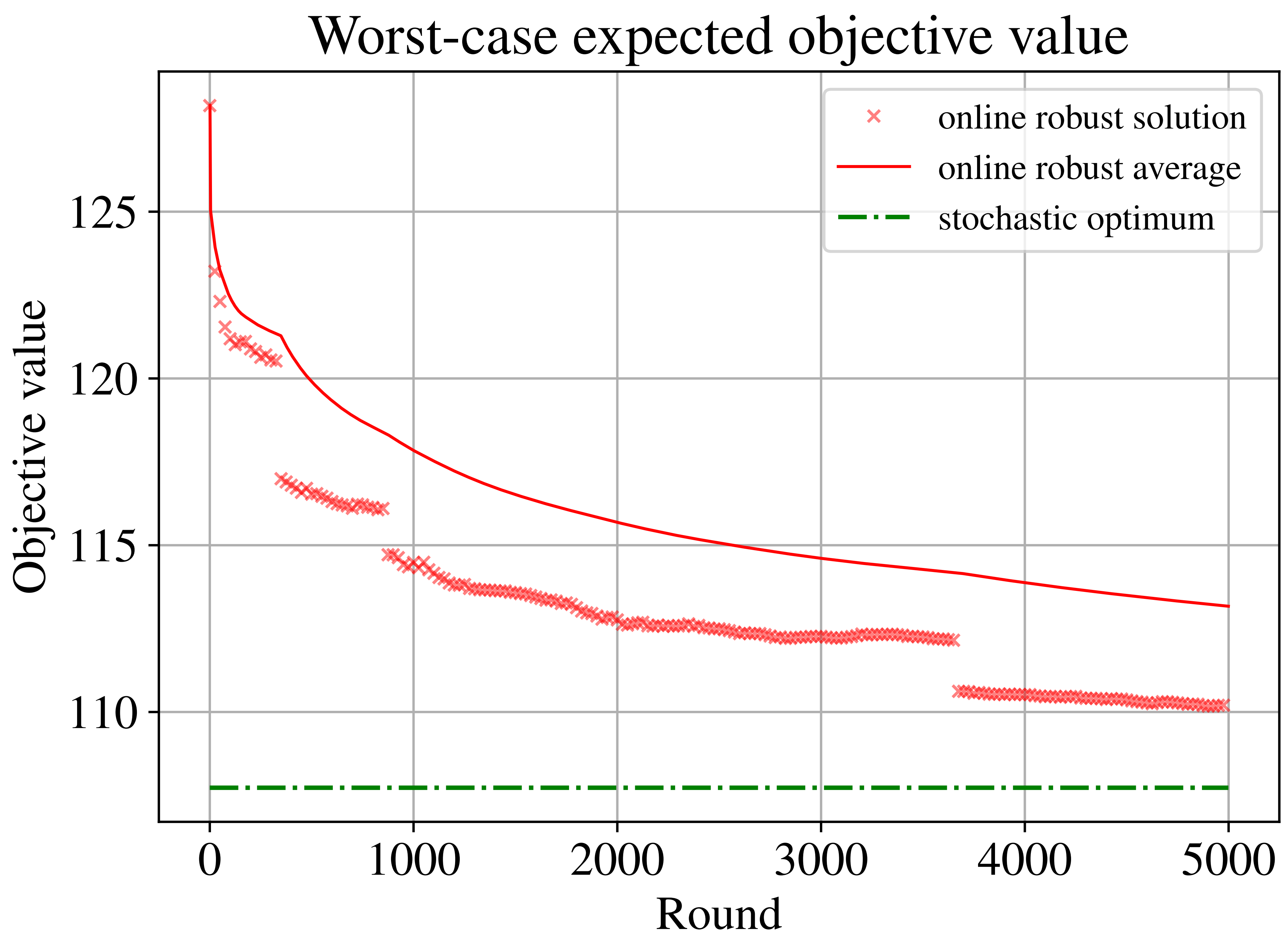


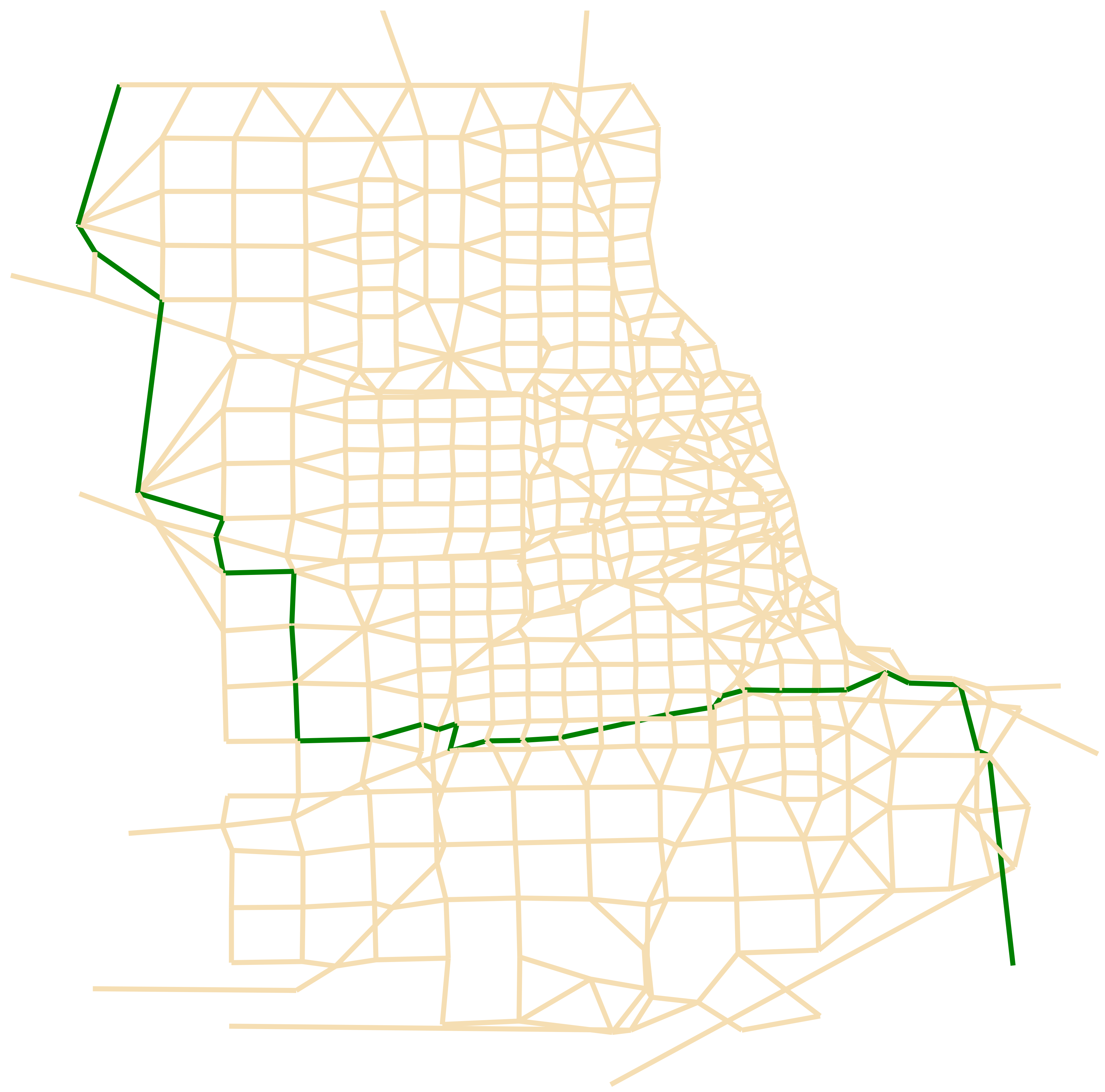
Optimal Route Choice



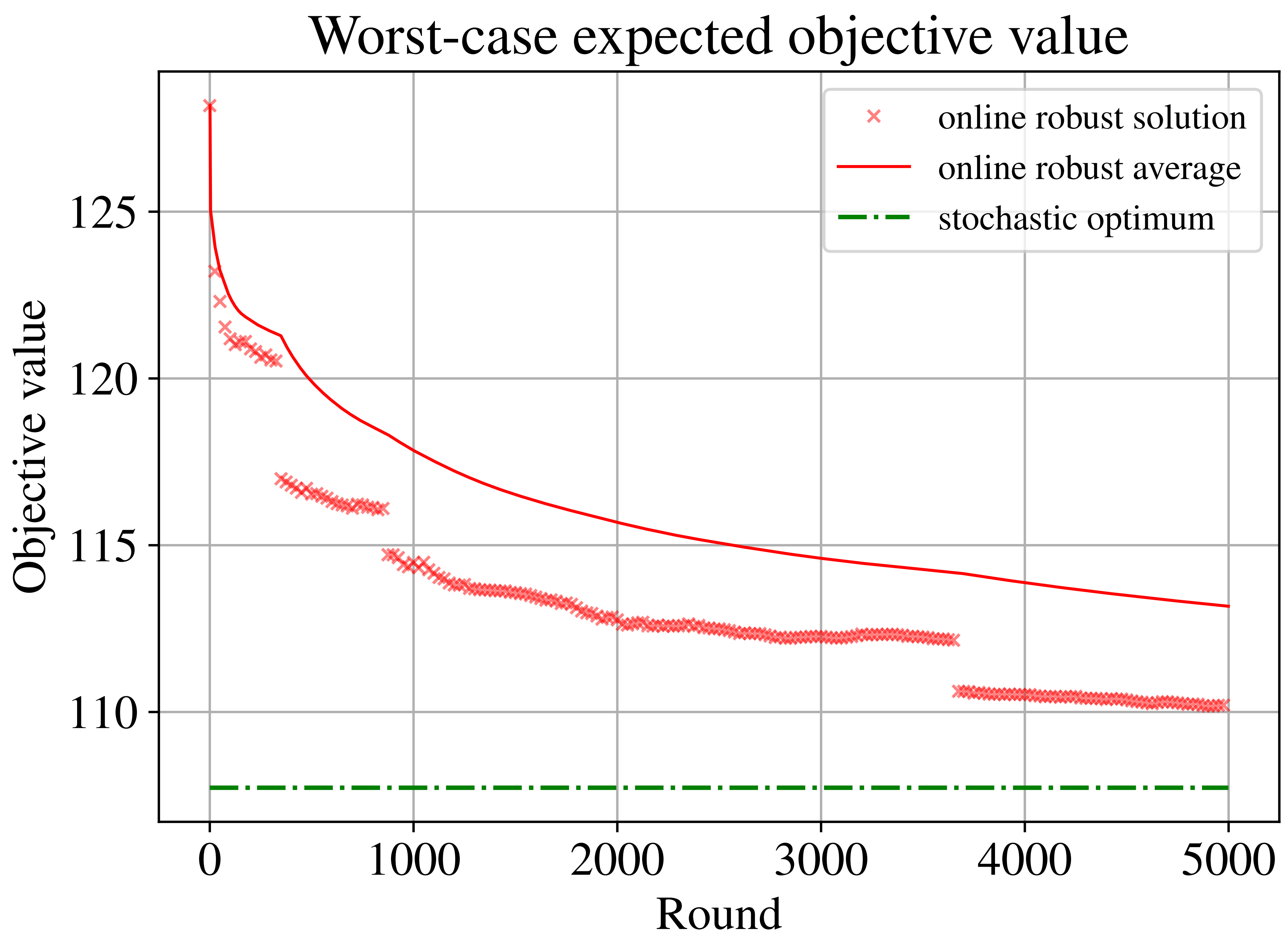


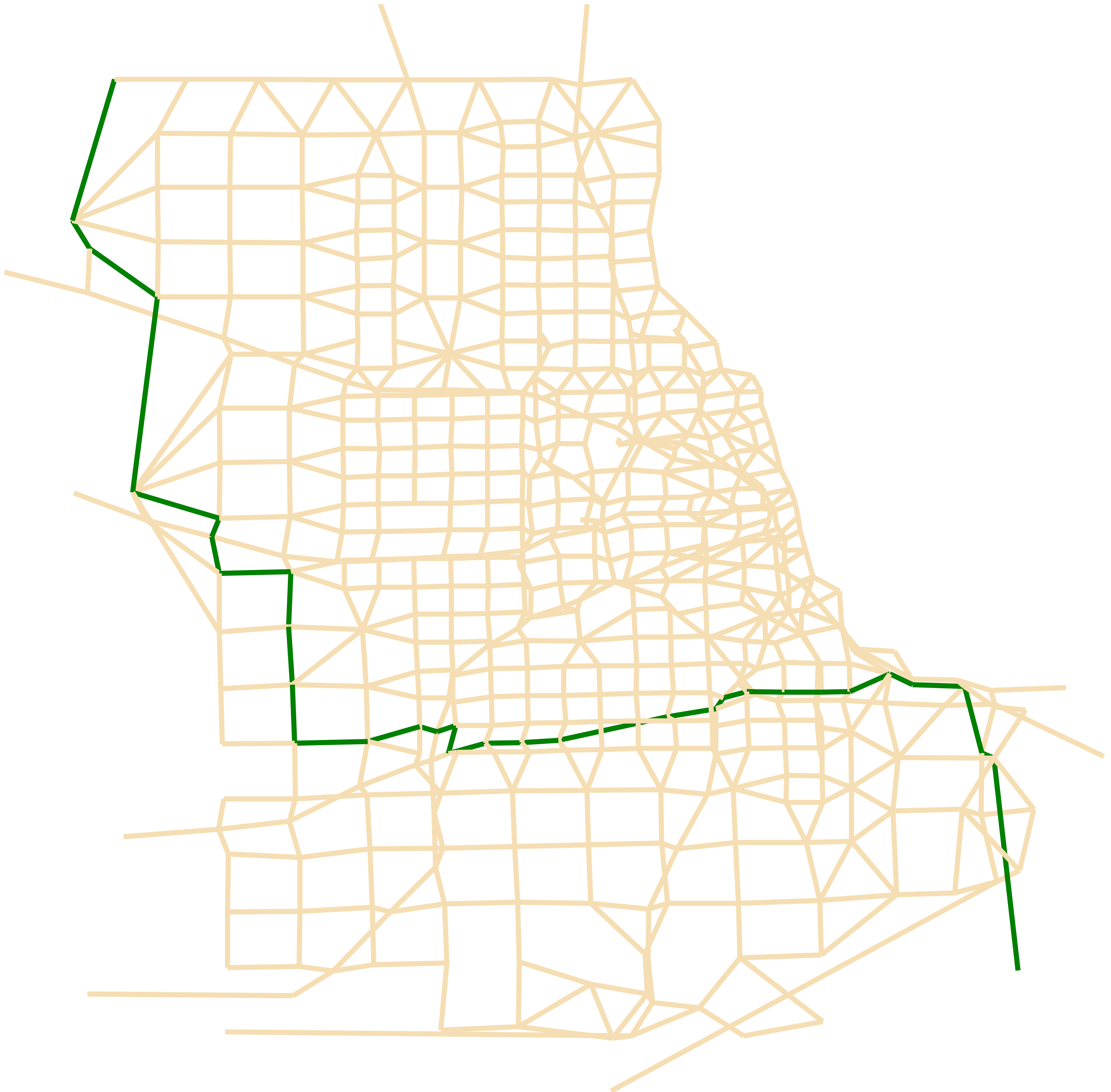
Optimal Route Choice





Optimal Route Choice

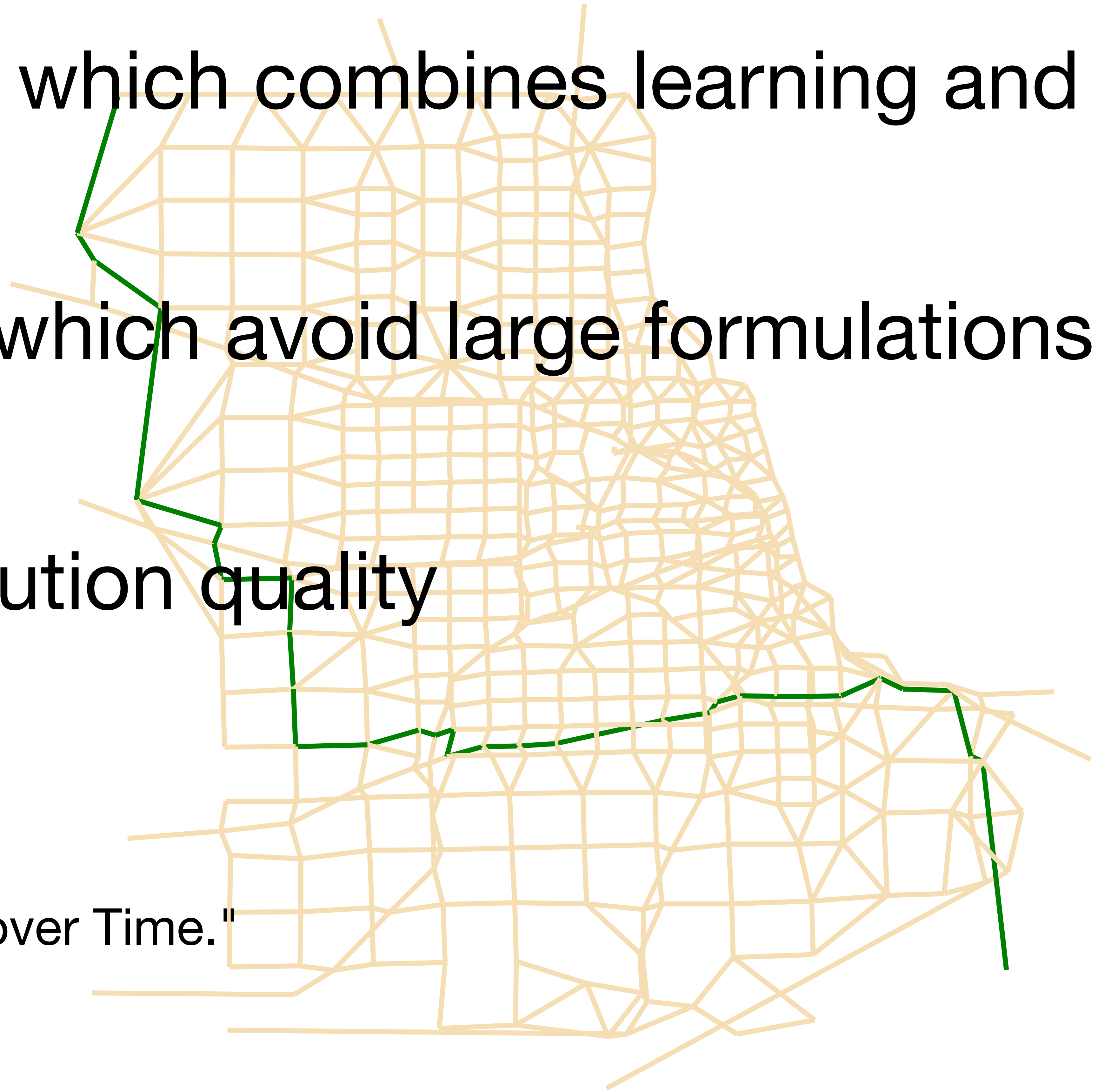




Continuous decrease in objective value but discrete jumps due to changing solution

Conclusions

- Method for optimization under uncertainty which combines learning and Distributionally Robust Optimization
- Iterative algorithm for solution of problem which avoid large formulations and maintains structure
- Theoretical proofs of convergence and solution quality
- Numerical illustrations for the results



Aigner et al. "Data-driven Distributionally Robust Optimization over Time." *INFORMS Journal on Optimization* (2023).

Email: kartikey.sharma@zib.de

thank you!