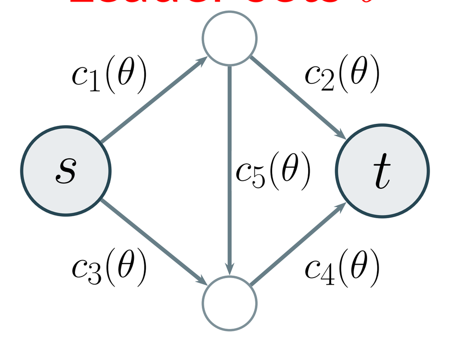
Zeroth-Order Stackelberg Control in Combinatorial Congestion Games

Saeed Masiha **Sepehr Elahi** Negar Kiyavash Patrick Thiran

EPFL, Switzerland

Stackelberg Control in Combinatorial Congestion Games

Leader sets θ



Followers choose routes

Problem

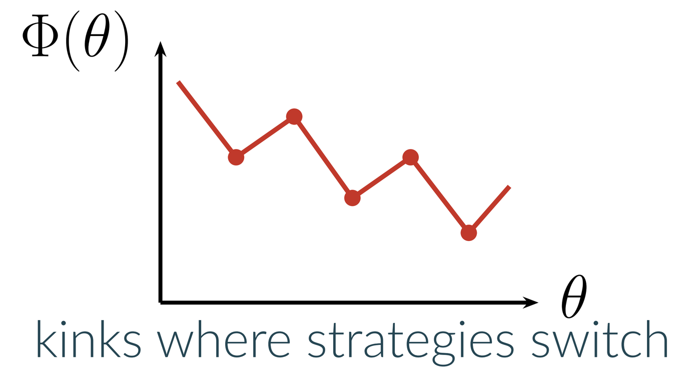
$$\min_{\theta\in\Theta} \underbrace{F\big(\theta,\ y^\star(\theta)\big)}_{\text{social cost } \Phi(\theta)}$$

subject to

$$y^\star(\theta) \in \arg\min_{y\in\mathcal{C}} f(\theta, y)$$

Wardrop equilibrium

True objective is nonsmooth!



kinks where strategies switch

A **leader** sets θ. **Followers** choose discrete routes and reach equilibrium $y^\star(\theta)$. As θ changes, the used routes can switch abruptly, creating **kinks** in $\Phi(\theta)$.

Our Method: Frank–Wolfe Inside, Zeroth-Order Outside

Inner loop / Outer loop

Leader parameter θ_t → FW-Equilibrium $y_T(\theta) \approx y^\star(\theta)$ → Evaluate $\hat\Phi_T(\theta \pm \rho u)$ → ZO update θ_{t+1}

Inner: FW-Equilibrium

1. Compute gradient g_t
2. Oracle returns cheapest strategy
3. Line search $\to y_{t+1}$

Never lists all strategies—just oracle calls.

Outer: Zeroth-Order

1. Sample direction u
2. Evaluate two nearby points
3. Update θ

No backprop.

Why Not Differentiate Through the Equilibrium?

Differentiate [1]		Ours (ZO-Stackelberg)
Smooth LMO	vs.	Exact / sub-sampled LMO
Backprop through T steps		Black-box equilibrium solver
Optimize a surrogate		Optimize the true objective Φ

× Surrogate $\neq$ true objective ✓ Nonsmoothness handled naturally

One Framework, Three Oracle Regimes

The **same outer method** works with all three oracle choices:

Regime 1 — Poly-Time LMO
Easy oracle
shortest path via Dijkstra

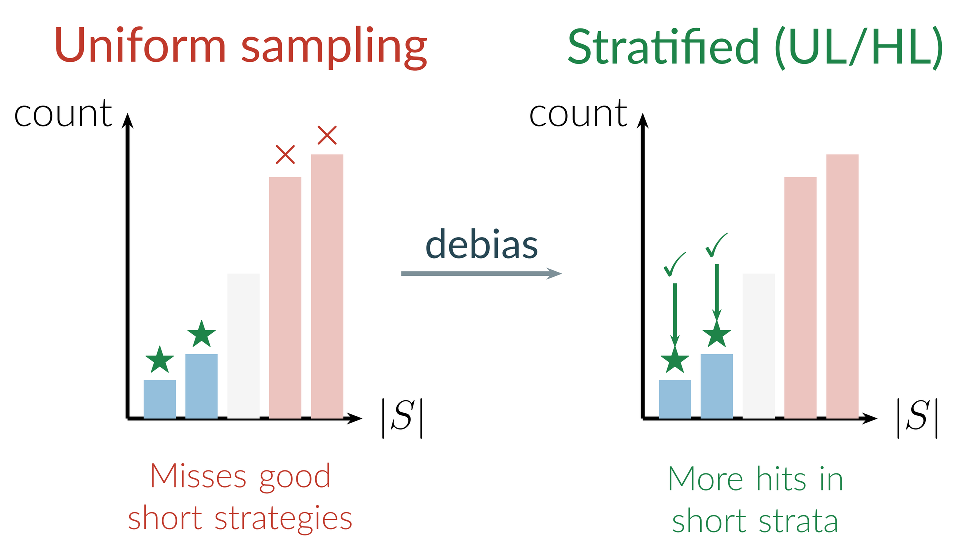
Regime 2 — Exact ZDD LMO
Hard problem, manageable ZDD
compile once, reuse

Regime 3 — Sampled ZDD LMO
ZDD too large for exact search
stratified sampling boosts κ_m

increasing hardness ⟶

In the largest regime, stratified sampling (UL/HL) raises κ_m by spreading samples across lengths.

Why Stratified Sampling?



- When $g \geq 0$, the best strategies are often **short** (★).
- Uniform sampling wastes effort on many long strategies.
- **UL/HL** spread samples across lengths and raise κ_m.

Theoretical Guarantees

Lipschitz, not smooth. $y^\star(\theta)$ and $\Phi(\theta)$ stay stable, but support switches create kinks.

Sampling quality matters. κ_m is the chance of hitting an exact LMO minimizer; the inner error decays as $\mathcal{O}(1/(\kappa_m T))$.

End-to-end guarantee. ZO-Stackelberg converges to a **Goldstein stationary point** of the true objective Φ.

Summary of Contributions

Problem
Route switches create **kinks**—the true Stackelberg objective is nonsmooth.

Method
Optimize the **true objective** directly via zeroth-order updates. No backprop.

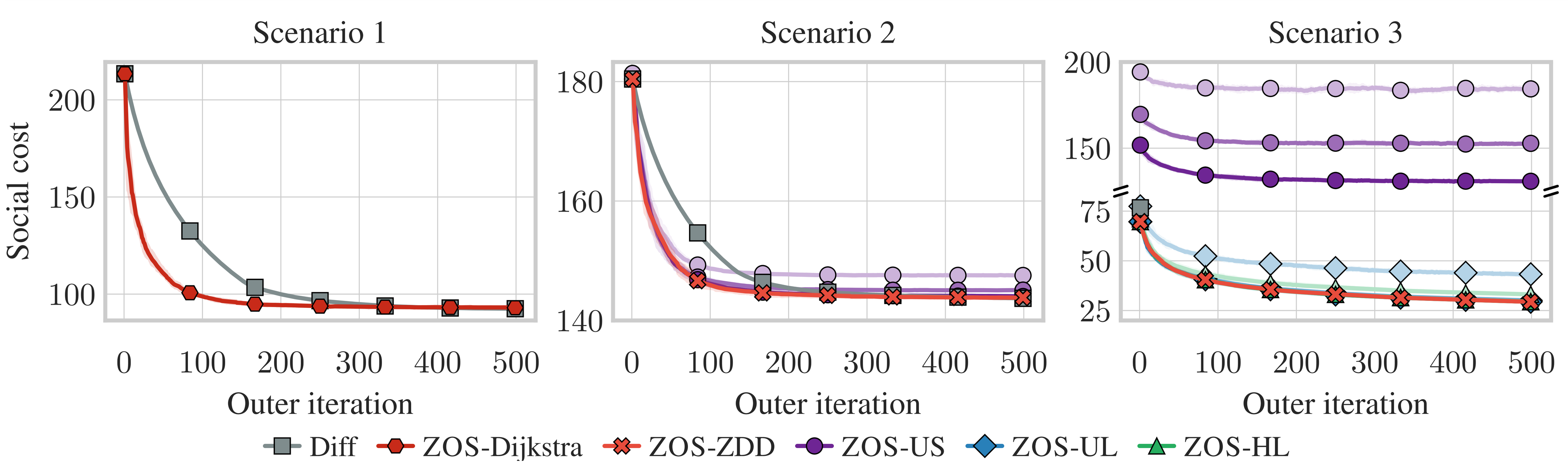
Theory
FW with stratified sampling: $\mathcal{O}(1/(\kappa_m T))$.
End-to-end: **Goldstein**

Results
23–61× speedup on three TNTP networks with 50× less memory.

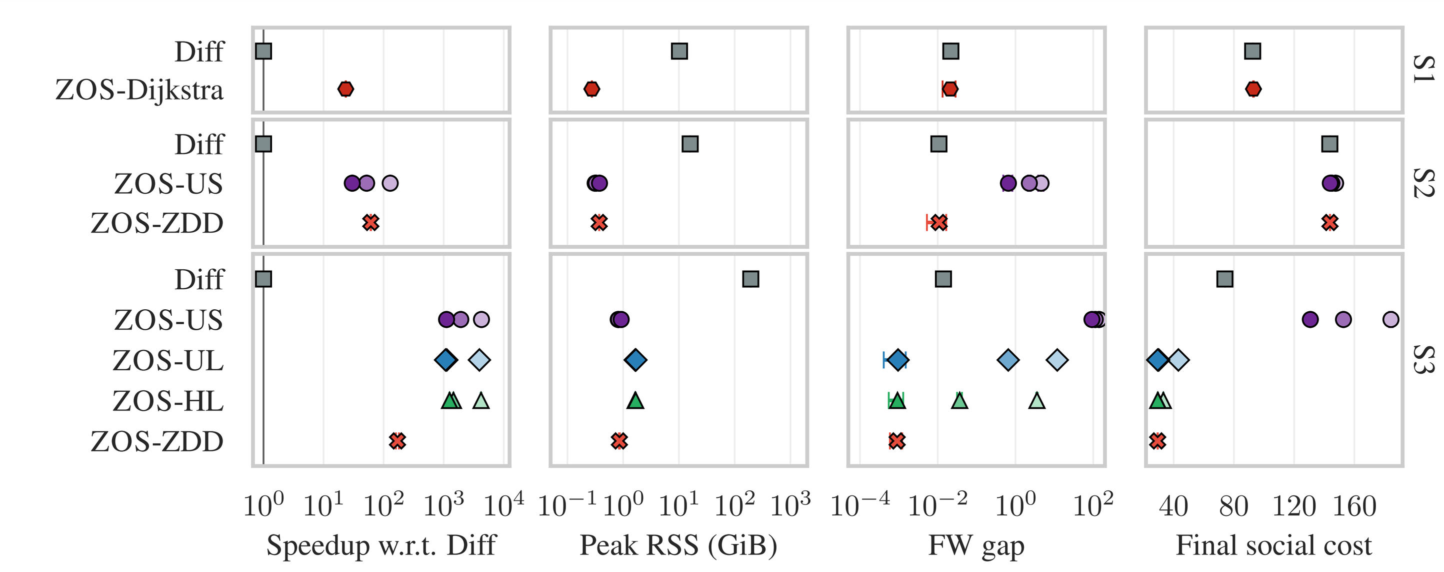
References

[1] Sakaue & Nakamura, *Differentiable Equilibrium Computation with Decision Diagrams for Stackelberg Models of Combinatorial Congestion Games*, NeurIPS 2021.

Experiments: Same Quality, Much Cheaper



Leader objective vs. outer iterations on three TNTP networks. **ZOS** matches or beats Diff. In Scenario 3, **uniform sampling stalls** while **UL/HL recovers**.



Final diagnostics: runtime, memory, FW gap, and social cost.

23× faster	61× faster	Diff fails
S1: 0.28 vs 10.2 GiB	S2: 0.37 vs 15.9 GiB	S3: 194 GiB peak

Takeaway: Left plot shows convergence quality; right plot summarizes cost, speed, and memory.