Zeroth-Order Stackelberg Control in Combinatorial Congestion Games

Saeed Masiha* **Sepehr Elahi*** **Negar Kiyavash** **Patrick Thiran**
EPFL, Switzerland

Abstract

We study Stackelberg (leader–follower) control of network parameters (tolls, capacity parameters in costs, incentives) in combinatorial congestion games, where selfish users choose *discrete* routes (or other combinatorial strategies) and settle at a congestion equilibrium. The leader minimizes a system-level objective (e.g., total travel time) evaluated at equilibrium, but this objective can be nonsmooth because the set of used strategies can change abruptly. We propose ZEROTH-ORDER STACKELBERG (ZOS), which couples a projection-free Frank–Wolfe equilibrium solver with a zeroth-order outer update, avoiding differentiation through equilibria. We prove an expected approximate generalized Goldstein-stationarity guarantee for the true equilibrium objective, accounting for finite equilibrium error. For subsampled oracles with uniform optimizer-hit probability at least κ_m, the expected Frank–Wolfe error decays as $\mathcal{O}(1/(\kappa_m T))$. We derive conditions under which sampling by strategy length increases the probability of finding an exact LMO optimizer compared with uniform sampling. In public road-network experiments covering large strategy spaces requiring subsampled oracles, ZOS reaches small follower-equilibrium gaps and final social costs comparable to differentiation-based methods, while reducing runtime per outer iteration by 20–1000$\times$ and using much less peak memory.

1 Introduction

Transportation, communication, and logistics systems can be steered by a Stackelberg leader who modifies resource costs through tolls, capacity parameters, or incentives [Migdalas, 1995, Korilis et al., 1997]. Followers then choose routes or other combinatorial strategies in a combinatorial congestion game (CCG) and reach a Wardrop equilibrium [Wardrop, 1952]. Applications include traffic pricing, EV routing and charging, multicast routing, and service tours [Yang and Bell, 1997, Zeng et al., 2024, Kim et al., 2013, Interian and Ribeiro, 2017]; Appendix E details their modeling requirements.

A structured bilevel problem. The leader chooses $\theta \in \Theta \subseteq \mathbb{R}^{d_\theta}$, and a unit-mass nonatomic population chooses strategies $S \in \mathcal{S}$ built from n resources. Its aggregate load belongs to $\mathcal{C} = \mathrm{conv}\{\mathbf{1}_S : S \in \mathcal{S}\} \subseteq [0,1]^n$. The Wardrop load $y^\star(\theta)$ uniquely minimizes the convex Beckmann potential $f(\theta, \cdot)$ over $\mathcal{C}$ (Proposition 2.3), equivalently solving a potential variational inequality (VI). The leader minimizes $\Phi(\theta) := F(\theta, y^\star(\theta))$, where F measures system-level performance. Each strategy $S \in \mathcal{S}$ is a discrete subset of resources, but the aggregate load $y \in \mathcal{C}$ is continuous. The combinatorial structure therefore lies in the potentially exponential strategy family and the representation of its load polytope; the lower-level equilibrium problem is a convex optimization problem, not an integer program. Capacity parameters modify resource costs here; they do not impose hard aggregate load constraints.

*Equal contribution.

40th Conference on Neural Information Processing Systems (NeurIPS 2026).

Three distinct challenges. First, our access to $\mathcal{C}$ is through exact or subsampled linear minimization oracles (LMOs), which do not directly supply an explicit constraint system, projections, equilibrium multipliers, or active-face information. We also use exact lower gradients, values of F, and projection onto the leader set Θ. Second, changes in the leader parameter can change which constraints are binding at equilibrium, an active-set change. Such transitions can make Φ nondifferentiable even when the resource costs and upper-level objective are smooth (Proposition 2.7). This is a feature of parameterized constrained optimization, not a new phenomenon specific to combinatorial representation [Bonnans and Shapiro, 2000, 1998]. Third, finite, possibly randomized equilibrium solves introduce errors that must be propagated to the outer optimization.

Relation to bilevel optimization. Classical KKT, value-function, and VI formulations remain valid and can be preferable when their derivatives and constraint-handling operations are efficiently available [Zemkoho and Zhou, 2021, Facchinei and Pang, 2003]. These operations are not assumed in our LMO access model. Projection-free alternatives also exist: Giang-Tran et al. [2025] minimize a convex objective over the minimizers of a fixed convex problem using first-order information and a base-set LMO. Here θ changes the lower problem, and the outer oracle supplies values of a potentially nonsmooth, nonconvex Φ. The mixed-integer linear decomposition methods of Yue et al. [2019], Zeng and An [2014] instead exploit explicit linear formulations and integer-assignment decomposition to obtain global guarantees under their assumptions. They do not apply as written to a generally nonlinear Beckmann potential with LMO access. Appendix D expands these comparisons.

Oracle-based equilibrium optimization. Our ZEROTH-ORDER STACKELBERG (ZOS) framework couples Frank–Wolfe (FW) equilibrium computation with mini-batched two-point outer updates. Each inner step minimizes additive strategy costs using a combinatorial routine, such as shortest path, or a zero-suppressed binary decision diagram (ZDD) [Jaggi, 2013, Dijkstra, 1959, Minato, 1993]. When exact minimization is expensive, the LMO selects the best of sampled strategies. Unlike differentiable CCG solvers that smooth the LMO and backpropagate through FW [Sakaue and Nakamura, 2021], ZOS uses values at approximate equilibria and controls their error. Derivative-free bilevel methods provide related approaches [Ehrhardt and Roberts, 2021, Aghasi and Ghadimi, 2025, Chen et al., 2025]; our analysis connects the outer guarantee specifically to exact or subsampled combinatorial FW solves.

Contributions. Our contributions are:

- An oracle-based framework coupling FW over an implicit combinatorial load polytope with zeroth-order optimization of the true hyper-objective, with an end-to-end expected approximate generalized Goldstein-stationarity guarantee accounting for finite inner error (Theorem 5.6).
- We prove that the expected FW potential error decreases at rate $\mathcal{O}(1/(\kappa_m T))$ after an initial phase of T_0 FW iterations, provided every inner step includes an exact LMO optimizer among its m sampled strategies with probability at least κ_m (Theorem 5.4). This probability bound must hold uniformly across the queried inner problems. We also derive conditions under which sampling by strategy length increases this probability compared with uniform sampling (Proposition 5.5).
- We evaluate ZOS on public transportation-network case studies spanning shortest-path, compact-ZDD, and large-scale subsampled oracle regimes. In each regime, the recommended oracle is competitive with the differentiable baseline in social cost and FW-gap accuracy while reducing per-iteration wall-clock time by 20–1000$\times$ and peak memory by an order of magnitude.

Section 2 defines the model, Sections 3 and 4 present the algorithm and oracles, and Sections 5 and 6 give theory and experiments. Section 7 summarizes the scope, limitations, and future directions.

2 Problem Setting

We formalize the Stackelberg CCG model and state the assumptions used in our analysis. A summary of notation appears in Appendix A.

2.1 Combinatorial Congestion Games (CCGs)

Let $[n] := \{1, \dots, n\}$ index the resources (e.g., edges) and let $\mathcal{S} \subseteq 2^{[n]}$ be the set of feasible strategies. We represent a nonatomic population by a distribution $z = (z_S)_{S \in \mathcal{S}} \in \Delta^d$ over strategies, where $d := |\mathcal{S}|$ and $\Delta^d := \{z \in \mathbb{R}^d : z_S \geq 0,\ \sum_{S \in \mathcal{S}} z_S = 1\}$. We encode S by its incidence

vector $\mathbf{1}_S \in \{0,1\}^n$ with $[\mathbf{1}_S]_i = 1$ if resource i is used. Then the induced resource-load vector is $y(z) := \sum_{S \in \mathcal{S}} z_S \, \mathbf{1}_S$, and the feasible set of loads is the polytope $\mathcal{C} := \operatorname{conv}\{\mathbf{1}_S : S \in \mathcal{S}\} \subseteq [0,1]^n$. Each resource $i \in [n]$ has a congestion-dependent cost function $c_i : [0,1] \to \mathbb{R}$, and the cost of a strategy S under load $y \in \mathcal{C}$ is $c_S(y) := \sum_{i \in S} c_i(y_i)$.

Assumption 2.1 (Resource costs)**.** *For each $i \in [n]$, the function c_i is continuous and strictly increasing on $[0,1]$.*

2.2 Wardrop Equilibrium and Potential Minimization

We adopt the notion of Wardrop equilibrium [Wardrop, 1952], which is the standard equilibrium concept for nonatomic congestion games, under which, for a given induced load, every strategy used with positive mass has minimum cost. For ease of notation, we temporarily omit the parameter θ; the parameterized Stackelberg model is introduced in Section 2.3.

Definition 2.2 (Wardrop equilibrium)**.** *A population profile $z \in \Delta^d$ with induced load $y = y(z) \in \mathcal{C}$ is a* Wardrop equilibrium *if the following holds: for every strategy $S \in \mathcal{S}$ with $z_S > 0$, we have $c_S(y) \le c_{S'}(y)$ for all $S' \in \mathcal{S}$.*

We define the standard convex potential of the game [Beckmann et al., 1956] as $f(y) := \sum_{i \in [n]} \int_0^{y_i} c_i(u)\,\mathrm{d}u$. Under Assumption 2.1, f is continuously differentiable and strictly convex on $[0,1]^n$, with $\nabla f(y)_i = c_i(y_i)$. As a result, f admits a unique minimizer over the convex set $\mathcal{C}$. The next proposition recalls the classical equivalence between Wardrop equilibria and minimizers of this problem, originating with the Beckmann transformation [Beckmann et al., 1956]; see also standard references such as Sheffi [1985, Section 3.2]. We provide a proof in Appendix G.2 for completeness.

Proposition 2.3 (Equilibrium $\iff$ potential minimizer)**.** *Let $z \in \Delta^d$ and $y = y(z)$. Under Assumption 2.1, z is a Wardrop equilibrium if and only if $y = \arg\min_{y' \in \mathcal{C}} f(y')$.*

Uniqueness concerns the aggregate equilibrium load $y^\star$, not the distribution z: different strategy distributions can induce the same load. The leader objective below depends only on this load.

2.3 Bilevel Model of CCG

The leader controls a vector $\theta \in \Theta \subseteq \mathbb{R}^{d_\theta}$ (closed and convex) that parametrizes the resource cost functions. For each resource $i \in [n]$, let $c_i(\cdot;\theta) : [0,1] \to \mathbb{R}$ satisfy Assumption 2.1 for every fixed $\theta \in \Theta$, and assume $\theta \mapsto c_i(y_i;\theta)$ is continuously differentiable for every fixed $y_i \in [0,1]$. The associated potential is $f(\theta, y) := \sum_{i \in [n]} \int_0^{y_i} c_i(u;\theta)\,du$, where $y \in \mathcal{C}$. For each $\theta \in \Theta$, define the equilibrium load $y^\star(\theta) := \arg\min_{y \in \mathcal{C}} f(\theta, y)$. The leader evaluates an equilibrium via an upper-level objective $F : \Theta \times \mathcal{C} \to \mathbb{R}$ and solves the Stackelberg problem

$$\min_{\theta \in \Theta} F(\theta, y) \quad \text{s.t.} \quad y \in \arg\min_{u \in \mathcal{C}} f(\theta, u). \tag{1}$$

Since the lower-level minimizer is unique, we write $y^\star(\theta)$ and optimize the hyper-objective $\Phi(\theta) := F(\theta, y^\star(\theta))$ over Θ.

Lipschitz stability of the equilibrium map. As discussed in Section 1, Φ can be nonsmooth because $y^\star(\theta)$ may switch between faces of the polytope $\mathcal{C}$. Even in this nonsmooth regime, the equilibrium map $\theta \mapsto y^\star(\theta)$ (and thus the hyper-objective Φ) can still be stable in a weaker sense: under mild conditions, both are globally Lipschitz. To formalize this, we assume a local quadratic growth property around each equilibrium, together with standard Lipschitz regularity of $\nabla_y f$ and of F.

Assumption 2.4 (Local quadratic growth around equilibrium)**.** *There exist constants $\alpha > 0$ and $r > 0$ such that for all $\theta \in \Theta$ and all $y \in \mathcal{C}$ with $\|y - y^\star(\theta)\| \le r$, we have $f(\theta, y) - f(\theta, y^\star(\theta)) \ge (\alpha/2)\|y - y^\star(\theta)\|^2$.*

Assumption 2.5 (Regularity of the lower- and upper-level problems)**.** *There exists $L_{f,2} \ge 0$ such that for all $\theta, \theta' \in \Theta$ and all $y, y' \in \mathcal{C}$, $\big\|\nabla_y f(\theta, y) - \nabla_y f(\theta', y')\big\| \le L_{f,2}\big(\|\theta - \theta'\| + \|y - y'\|\big)$. Moreover, $F : \Theta \times \mathcal{C} \to \mathbb{R}$ is Lipschitz: there exists $L_{F,1} \ge 0$ such that for all $\theta, \theta' \in \Theta$ and all $y, y' \in \mathcal{C}$, $|F(\theta, y) - F(\theta', y')| \le L_{F,1}[\|\theta - \theta'\| + \|y - y'\|]$.*

See Appendix K for simple bounds and typical scaling of $L_{f,2}$ and $L_{F,1}$ in our setting. Under Assumptions 2.4 and 2.5, the equilibrium map is stable to perturbations of θ; Lemma 2.6 makes this precise (proof in Appendix G.3).

Lemma 2.6 (Lipschitzness of the equilibrium map and hyper-objective)**.** *Let Assumptions 2.4 and 2.5 hold and assume that Θ is convex. Then the equilibrium map $\theta \mapsto y^\star(\theta)$ and the hyper-objective $\Phi(\theta) = F(\theta, y^\star(\theta))$ are globally Lipschitz on Θ: for all $\theta, \theta' \in \Theta$, $\|y^\star(\theta) - y^\star(\theta')\| \le (2L_{f,2}/\alpha)\|\theta - \theta'\|$ and $|\Phi(\theta) - \Phi(\theta')| \le L_\Phi \|\theta - \theta'\|$, where $L_\Phi := L_{F,1}\big(1 + 2L_{f,2}/\alpha\big)$.*

When does quadratic growth hold? Assumption 2.4 is implied by (local) strong convexity of $y \mapsto f(\theta, y)$ over $\mathcal{C}$. For the separable potential $f(\theta, y) = \sum_i \int_0^{y_i} c_i(u; \theta)\, du$, a simple sufficient condition is a uniform slope lower bound: if each $c_i(\cdot; \theta)$ is differentiable in its first argument and $\partial_{y_i} c_i(y_i; \theta) \ge \mu > 0$ for all $y_i \in [0, 1]$ and all $\theta \in \Theta$, then $f(\theta, \cdot)$ is μ-strongly convex, so the inequality in Assumption 2.4 holds globally with $r = \infty$ and $\alpha = \mu$.

For the fractional costs $c_i(y_i; \theta_i) = d_i(1 + Cy_i/(\theta_i + 1))$ with $C > 0$, $d_{\min} := \min_i d_i > 0$, and $\Theta = \{\theta \in \mathbb{R}^n_+ : \mathbf{1}^\top \theta = n\}$, we have $\partial_{y_i} c_i(y_i; \theta_i) = Cd_i/(\theta_i + 1) \ge Cd_{\min}/(n + 1) > 0$. This certifies global strong convexity and quadratic growth for these costs. The certificate is sufficient, not necessary. For Bureau of Public Roads (BPR) travel-time costs with exponent $\beta > 1$, the derivative vanishes at zero load, so this certificate does not establish Assumption 2.4 globally. Appendix C defines the BPR model and discusses the separate growth and stability conditions needed.

Nonsmoothness via active-set changes. The active-set nonsmoothness noted in Section 1 is not an artifact of a single example: it already appears in a simple parametric class satisfying our standing assumptions.

Proposition 2.7 (A class of Stackelberg control problems with nonsmooth social-cost hyper-objective)**.** *Let $\mathcal{G} := \{g \in C^1([0, 1]) : \|g\|_\infty \le 1\}$. For any $g \in \mathcal{G}$, consider the two-link problem with $\Theta = [0, 1]$, $\mathcal{S} = \{\{1\}, \{2\}\}$, $\mathcal{C} = \{(y_1, y_2) \in \mathbb{R}^2_{\ge 0} : y_1 + y_2 = 1\}$, costs $c_1(y_1; \theta) = y_1$ and $c_2(y_2; \theta) = y_2 - 1 + g(\theta)$, and social-cost upper level $F_{\mathrm{sc}}(\theta, y) := \sum_{i=1}^2 y_i\, c_i(y_i; \theta)$. Then Assumptions 2.1, 2.4 and 2.5 hold, and the induced hyper-objective $\Phi_{\mathrm{sc}}(\theta) := F_{\mathrm{sc}}(\theta, y^\star(\theta))$ is nondifferentiable at each interior simple zero $\bar\theta \in (0, 1)$ of g, meaning $g(\bar\theta) = 0$ and $g'(\bar\theta) \ne 0$.*

If g has infinitely many interior simple zeros, the hyper-objective has infinitely many nondifferentiable points. Appendix G.1 proves this claim by computing Φ_{sc} explicitly, gives an oscillatory choice of g, and presents a separate example with $\Omega(|\mathcal{S}|)$ kinks in the equilibrium map. This motivates our zeroth-order outer loop, which does not require differentiability of Φ and is analyzed using generalized Goldstein stationarity for constrained nonsmooth optimization.

3 Zeroth-Order Algorithm for Stackelberg Control

Our method ZOS combines a first-order FW equilibrium solver with a zeroth-order outer method, using the information available at each level. Under Assumptions 2.1 and 2.5, the Beckmann potential is continuously differentiable with exact gradient $\nabla_y f(\theta, y) = c(y; \theta)$. FW uses this gradient with an LMO over $\mathcal{C}$, avoiding gradient estimation in the inner loop. The hyper-objective, however, can be nonsmooth (Proposition 2.7). Values at approximate FW equilibria remain controlled: $|F(\theta, y_T(\theta)) - \Phi(\theta)| \le L_{F,1}\|y_T(\theta) - y^\star(\theta)\|$. Our outer two-point updates exploit this bound without differentiating $y^\star$.

First-order methods can also handle nonsmooth objectives. To use them here, however, we would need to compute suitable generalized gradients of Φ and bound the errors caused by approximate equilibrium solves. Our FW solver does not directly provide these quantities; constructing such an outer oracle is future work (Appendix C).

Penalty and value-function methods can avoid differentiating $y^\star(\theta)$ [Shen and Chen, 2023] and can be adapted to use FW instead of projections onto $\mathcal{C}$. The FW-based penalty variants evaluated in Section 6.2 require $\nabla_y F$ for their lower-level updates; gradient-based leader updates additionally use $\nabla_\theta F$ and $\nabla_\theta f$. These derivative requirements are specific to those implementations. In contrast, ZOS uses only function values of F, the lower gradient $\nabla_y f$, and the LMO, together with projection onto Θ.

Combining a derivative-free outer method with gradient-based inner solvers also has precedent: Ehrhardt and Roberts [2021] use this approach for image denoising and MRI reconstruction. These applications motivate the design but are not covered by our CCG guarantees.

Algorithm 1 ZOS: zeroth-order Stackelberg optimization

Input: initial $\theta_0 \in \Theta$, outer iterations K, inner iterations T, smoothing $\rho > 0$, stepsize $\eta > 0$, direction batch size $B_{\text{out}} \geq 1$
Inner solver: FW-EQUILIBRIUM(θ, T) is given in Appendix B (Algorithm 2)
for $t = 0, 1, \ldots, K-1$ **do**
  for $b = 1, \ldots, B_{\text{out}}$ **do**
    sample $u_{t,b}$ uniformly from the unit sphere in $\mathbb{R}^{d_\theta}$
    $y^+_{t,b} :=$ FW-EQUILIBRIUM$(\theta_t + \rho u_{t,b}, T)$
    $y^-_{t,b} :=$ FW-EQUILIBRIUM$(\theta_t - \rho u_{t,b}, T)$
    $\widehat{\Phi}^+_{t,b} := F(\theta_t + \rho u_{t,b}, y^+_{t,b})$
    $\widehat{\Phi}^-_{t,b} := F(\theta_t - \rho u_{t,b}, y^-_{t,b})$
    $\widehat{g}_{t,b} := \frac{d_\theta}{2\rho}\left(\widehat{\Phi}^+_{t,b} - \widehat{\Phi}^-_{t,b}\right) u_{t,b}$
  end for
  $\widehat{g}_t := B_{\text{out}}^{-1} \sum_{b=1}^{B_{\text{out}}} \widehat{g}_{t,b}$
  $\theta_{t+1} := \Pi_\Theta\big(\theta_t - \eta\, \widehat{g}_t\big)$
end for
return θ_K

- **Inner equilibrium loop.** Given θ, we approximate $\min_{y \in \mathcal{C}} f(\theta, y)$ using Frank–Wolfe (FW). At iterate y_t, FW forms the weight vector $\omega_t := \nabla_y f(\theta, y_t)$ and calls a linear minimization oracle (LMO) over $\mathcal{C}$: $\mathrm{LMO}(g) \in \arg\min_{y \in \mathcal{C}} \langle g, y \rangle = \arg\min_{S \in \mathcal{S}} \sum_{i \in S} g_i$. We use either the exact LMO or a subsampled variant LMO_m (Section 4); Appendix B gives the pseudocode for the inner loop (Algorithm 2).
- **Outer zeroth-order loop.** The outer loop performs projected zeroth-order updates on Θ using two-point finite differences. At iteration t, sample B_{out} independent directions $u_{t,1}, \ldots, u_{t,B_{\text{out}}}$ uniformly from the unit sphere[1], form the mini-batched estimator

$$\widehat{g}_t := \frac{1}{B_{\text{out}}} \sum_{b=1}^{B_{\text{out}}} \frac{d_\theta}{2\rho} \big(\widehat{\Phi}_T(\theta_t + \rho u_{t,b}) - \widehat{\Phi}_T(\theta_t - \rho u_{t,b})\big) u_{t,b},$$

and update $\theta_{t+1} := \Pi_\Theta(\theta_t - \eta\, \widehat{g}_t)$. Each outer iteration makes $2B_{\text{out}}$ calls to the inner solver (and hence $2B_{\text{out}}T$ LMO calls). Algorithm 1 returns the last iterate θ_K; for analysis we consider a uniformly random iterate from $\{\theta_0, \ldots, \theta_{K-1}\}$ (Section 5).

4 Implementing the Linear Minimization Oracle (LMO)

Each Frank–Wolfe step in Algorithm 2 requires a *linear minimization oracle* (LMO): given weights $g \in \mathbb{R}^n$, return $\mathrm{LMO}(g) \in \arg\min_{S \in \mathcal{S}} \sum_{i \in S} g_i$. Thus the computational bottleneck is entirely determined by the strategy family $\mathcal{S}$. For graph-based CCGs, where resources are edges (or nodes) and strategies are paths, trees, or tours [Roughgarden and Tardos, 2002, Anshelevich et al., 2004, Garey et al., 1976], we use three oracle regimes.

- **Polynomial-time exact LMO.** When the minimum-weight feasible strategy can be found in polynomial time, FW calls that routine directly without enumerating $\mathcal{S}$. For example, in s–t routing on a directed graph $G = (V, E)$, the LMO is a shortest-path problem $\min_{P \in \mathcal{P}_{s \to t}} \sum_{e \in P} g_e$, solvable by Dijkstra's algorithm when weights are nonnegative [Dijkstra, 1959].
- **Exact LMO via dynamic programming on the ZDD.** For some strategy families, finding the minimum-weight strategy is NP-hard in the worst case. This does not preclude exact optimization on a fixed structured instance. Since $\mathcal{S}$ is fixed across all FW and outer iterations, we can compile it once into a *zero-suppressed binary decision diagram* (ZDD), when such a compact representation exists [Minato, 1993]. Then each exact LMO query reduces to a minimum-cost root-to-terminal dynamic

[1] Boundary queries $\theta_t \pm \rho u$ may leave Θ. The analysis uses an L_Φ-Lipschitz extension of Φ to the queried ρ-neighborhood and requires the same uniform evaluation-error control at every queried point. Existence of a Lipschitz extension alone does not supply an evaluation oracle for it.

program on the diagram, with the current weights g as edge/item costs. The caveat is that the ZDD itself can be exponential, so this regime is useful only for instances whose diagrams are small enough to build and reuse. See more details in Appendix I.

- **Approximate LMO via Subsampling.** When the ZDD is too large for repeated exact minimization, we use it instead to *sample* feasible strategies. Since root-to-terminal paths correspond to strategies, sampling paths in the ZDD produces samples from $\mathcal{S}$ without enumerating $\mathcal{S}$. Given weights $g \in \mathbb{R}^n$, we sample m i.i.d. feasible strategies $S^{(1)}, \ldots, S^{(m)} \in \mathcal{S}$, score each by $\sum_{i \in S} g_i$, and return the best one(s) as a subsampled oracle LMO_m. We discuss sampling distributions and analyze the resulting subsampled FW rate in Section 5.1, with implementation details in Appendix I.3. In our experiments, LMO_m closely tracks the exact ZDD oracle while remaining practical when repeated exact minimization is not (Section 6). Appendix J summarizes preprocessing and per-iteration costs.

  Unlike Sakaue and Nakamura [2021], who smooth the LMO to backpropagate through FW iterates, we use ZDDs only as exact or sampled combinatorial oracles inside standard FW. This avoids softmin bias and the memory cost of differentiating through the inner loop; the outer loop handles the resulting nonsmooth dependence on θ using zeroth-order updates.

5 Theoretical Analysis

We analyze Algorithm 1 for minimizing $\Phi(\theta) = F(\theta, y^\star(\theta))$, where $y^\star(\theta) \in \arg\min_{y \in \mathcal{C}} f(\theta, y)$, using either an exact or subsampled LMO.

5.1 Inner Loop: Equilibrium Approximation via (Subsampled) Frank–Wolfe

For a fixed leader parameter θ, the inner equilibrium problem is the convex program $\min_{y \in \mathcal{C}} f(\theta, y)$, where $\mathcal{C} = \mathrm{conv}\{\mathbf{1}_S : S \in \mathcal{S}\}$. We solve this problem by Frank–Wolfe; the pseudocode is given in Appendix B (Algorithm 2). Each FW step linearizes $f(\theta, \cdot)$ at the current load and calls a linear minimization oracle over $\mathcal{C}$. Let $D := \max\{\|u - v\| : u, v \in \mathcal{C}\}$.

Exact and subsampled linear minimization. Given weights $g \in \mathbb{R}^n$, the exact oracle returns $\mathrm{LMO}(g) \in \arg\min_{y \in \mathcal{C}} \langle g, y \rangle = \arg\min_{S \in \mathcal{S}} \langle g, \mathbf{1}_S \rangle$; the equality uses the fact that a linear objective over $\mathcal{C}$ is minimized at a strategy vector. When exact minimization is too expensive, let q be a probability distribution over $\mathcal{S}$. We draw candidates $S^{(1)}, \ldots, S^{(m)} \overset{\text{i.i.d.}}{\sim} q$ and use the best sampled strategy, $\mathrm{LMO}_m(g) \in \arg\min_{1 \le j \le m} \langle g, \mathbf{1}_{S^{(j)}} \rangle$. The resulting inner loop is standard FW with LMO replaced by LMO_m.

Remark 5.1 (Exact LMO (standard))**.** *If the exact oracle* LMO *is available, FW admits the standard $\mathcal{O}(1/T)$ convergence rate (e.g., Frank and Wolfe [1956], Jaggi [2013]). Exact LMOs do not require Assumption 5.2; the bounds below specialize to $\kappa_m = 1$. We focus on subsampled oracles.*

The single quantity that controls the effect of subsampling. Let $\mathrm{Opt}(g) := \arg\min_{S \in \mathcal{S}} \langle g, \mathbf{1}_S \rangle$ and $p(g) := q(\mathrm{Opt}(g)) = \sum_{S \in \mathrm{Opt}(g)} q(S)$. With m i.i.d. samples, the probability of sampling an exact minimizer is $\kappa_m(g) := 1 - (1 - p(g))^m$. For a clean FW rate we assume that $p(g)$ is uniformly bounded from below along the inner run:

Assumption 5.2 (Uniform optimizer mass)**.** *There exists $\underline{p} \in (0, 1]$ such that for all FW gradients g encountered by the inner loop, $p(g) = q(\mathrm{Opt}(g)) \ge \underline{p}$.*

Under Assumption 5.2, $\kappa_m(g) \ge \kappa_m := 1 - (1 - \underline{p})^m$, and κ_m is the only dependence on q in the FW rate. In the bilevel setting, we require this bound uniformly over the inner FW runs invoked by the outer loop, so κ_m does not depend on θ. Full support on finite $\mathcal{S}$ ensures positive mass, but it can still be exponentially small; strict cost monotonicity or quadratic growth does not establish a useful optimizer-hit bound.

Remark 5.3 (Relation to uniform-inclusion subsampling)**.** *Kerdreux et al. [2018] obtain an $\mathcal{O}(1/(\eta T))$ rate when every atom is included in the candidate set with the same probability η. Then any exact LMO minimizer is hit with probability at least η. Assumption 5.2 keeps only this optimizer-hit requirement: for an arbitrary sampler q, the m-sample candidate set must hit an exact minimizer with probability at least κ_m. Uniform inclusion is recovered by taking $\kappa_m = \eta$, while non-uniform structure-aware samplers can have larger κ_m.*

Theorem 5.4 (Convergence of FW-EQUILIBRIUM with subsampled LMO)**.** *Let Assumptions 2.1, 2.5 and 5.2 hold. Fix $\theta \in \Theta$ and initialize $y_0 \in \mathcal{C}$. Run Algorithm 2 with the subsampled oracle LMO_m and exact line-search.*[2] *Define the initial optimality gap $\Delta_0 := f(\theta, y_0) - f(\theta, y^\star(\theta))$ and the burn-in $T_0 := \lceil \kappa_m^{-1} \log\big(1 + 2\Delta_0/(L_{f,2}D^2)\big) \rceil$. Then for all $T \geq T_0$,*

$$\mathbb{E}[f(\theta, y_T)] - f(\theta, y^\star(\theta)) \;\leq\; \frac{2L_{f,2}D^2}{\kappa_m\,(T - T_0 + 1)}. \tag{2}$$

Moreover, if Assumption 2.4 holds, then for every $\varepsilon \in (0, r]$, $\mathbb{E}[\|y_T - y^\star(\theta)\|^2] \leq \varepsilon^2$ whenever $T \geq \lceil 4L_{f,2}D^2 \max\{1, D/r\}/(\alpha\,\kappa_m\,\varepsilon^2) - 1 \rceil + T_0$, where T_0 is the burn-in above. If $\bar{\Delta}_0 := \sup_{\theta\in\Theta} \Delta_0(\theta) < \infty$ and $\bar{T}_0 := \lceil \kappa_m^{-1} \log\big(1 + 2\bar{\Delta}_0/(L_{f,2}D^2)\big) \rceil$, then the same bound holds uniformly with $\bar{T}_0$ in place of T_0.[3]

The proof is given in Appendix H.4. The rate (2) is the standard $\mathcal{O}(1/T)$ FW rate up to the logarithmic burn-in T_0, with subsampling captured by κ_m; we next discuss how to make κ_m nontrivial.

Improving κ_m via stratified sampling. If q is uniform over $\mathcal{S}$, then $p(g) = |\mathrm{Opt}(g)|/|\mathcal{S}|$. When $|\mathcal{S}|$ is huge (often exponential in n), $p(g)$ is tiny unless $\mathrm{Opt}(g)$ is enormous, and when $m\,p(g) \ll 1$, $\kappa_m(g) = 1 - (1 - p(g))^m \approx m\,p(g)$, so the bound in (2) is vacuous unless m is extremely large. Uniform sampling can be ineffective in *imbalanced* families, where the number of feasible strategies varies widely across simple strata (e.g., by cardinality $|S|$). In such cases, uniform draws are dominated by the largest strata, even though LMO minimizers often come from more structured classes (e.g., short strategies)[4]. Let $\tau : \mathcal{S} \to \mathcal{T}$ be a user-chosen stratification statistic (e.g., $\tau(S) = |S|$) and define strata $\mathcal{S}_t := \{S \in \mathcal{S} : \tau(S) = t\}$. We retain only nonempty strata. We sample in two stages: choose $t \sim w$ and then sample S uniformly from $\mathcal{S}_t$, i.e., $q(S) = w_{\tau(S)}/|\mathcal{S}_{\tau(S)}|$.

Write $\pi_t := |\mathcal{S}_t|/|\mathcal{S}|$ for the population share and $r_t(g) := |\mathrm{Opt}(g) \cap \mathcal{S}_t|/|\mathcal{S}_t|$ for the optimizer fraction. The stratified and uniform single-draw optimizer probabilities are $p_{\mathrm{strat}}(g) = \sum_t w_t r_t(g)$ and $p_{\mathrm{unif}}(g) = \sum_t \pi_t r_t(g)$, hence

$$p_{\mathrm{strat}}(g) - p_{\mathrm{unif}}(g) = \sum_{t\in\mathcal{T}} (w_t - \pi_t) r_t(g).$$

Thus, improvement requires assigning more mass than uniform sampling to strata with favorable optimizer fractions. These strata must be identifiable and efficiently sampled, with enough mass uniformly over the relevant FW gradients. There is no gain with one stratum or equal optimizer fractions; even an enriched stratum may contain too many strategies for a useful hit probability.

Two length-stratified choices. A natural choice is $\tau(S) = |S|$, which motivates length-stratified sampling. Let $\mathcal{L} := \{|S| : S \in \mathcal{S}\}$ and $N_t := |\{S \in \mathcal{S} : |S| = t\}|$. We use **uniform-over-length** (UL), $q(S) = 1/(|\mathcal{L}|\,N_{|S|})$, and **harmonic-length** (HL), $q(S) \propto 1/(|S|\,N_{|S|})$. For positive LMO weights with bounded dynamic range, any exact minimizer cannot be much longer than a shortest feasible strategy; hence only relatively short strata can contain optimizers.

Proposition 5.5 (Length-stratified sampling in imbalanced families)**.** *Let $\ell_{\min} := \min_{S\in\mathcal{S}} |S| \geq 1$ and suppose the LMO weights satisfy $0 < g_{\min} \leq g_i \leq g_{\max}$ for all i. Set $R := \lfloor (g_{\max}/g_{\min})\ell_{\min} \rfloor$ and $M_R := \max_{\ell\in\mathcal{L}:\ell\leq R} N_\ell$. Under UL sampling, with $p_{\mathrm{UL}}(g) := q_{\mathrm{UL}}(\mathrm{Opt}(g))$, $\kappa_{m,\mathrm{UL}}(g) \geq 1 - \Big(1 - \frac{1}{|\mathcal{L}|M_R}\Big)^m$. For comparison, uniform-over-strategies sampling has $p_{\mathrm{US}}(g) = |\mathrm{Opt}(g)|/|\mathcal{S}|$. Moreover, $p_{\mathrm{UL}}(g) \geq |\mathcal{S}|\,p_{\mathrm{US}}(g)/(|\mathcal{L}|M_R)$.*

The bound is useful when $|\mathcal{S}| \gg |\mathcal{L}|M_R$, i.e., most strategies lie in large strata that cannot contain LMO minimizers. Appendix H.2 gives the proof, the HL analogue, and an exponential-separation example. Appendix F.3 reports the same mechanism empirically in the large-strategy experiment: exact minimizers lie in short-length strata, making κ_{1000} negligible for uniform sampling but close to one for UL/HL.

Applicability and limits. Proposition 5.5 is pointwise in g; the bilevel guarantee requires a useful bound uniformly over all queried inner runs. For standard Hamiltonian tours, as in the traveling

[2] Appendix H.4 shows the same bound holds with the standard smoothness-based short-step (approximate line search) $\gamma_t = \min\{\max\{\langle \nabla_y f(\theta, y_t), y_t - s_t\rangle/(L_{f,2}D^2), 0\},\, 1\}$.

[3] This uniform boundedness holds when Θ is compact, since f and $y^\star(\cdot)$ are continuous on the compact set $\Theta \times \mathcal{C}$.

[4] If the LMO weights are nonnegative, adding resources cannot reduce $\sum_{i\in S} g_i$, so minimizers tend to avoid detours; longer strategies can still win if they use sufficiently cheaper resources.

salesman problem (TSP), every strategy contains $|V|$ edges; standard position–item encodings of permutations likewise have fixed cardinality. Length stratification then has one stratum, so UL and HL reduce to uniform sampling. Other TSP-specific strata can help only with additional structure and efficient conditional sampling; Appendix H.3 gives a sufficient zone-transition certificate and the limits of fixed, weight-independent sampling.

5.2 Outer Loop: Goldstein Stationarity and Convergence of Algorithm 1

Since the equilibrium map can switch combinatorially with θ, the hyper-objective $\Phi(\theta) = F(\theta, y^\star(\theta))$ is typically Lipschitz but nonsmooth. We therefore measure progress by Goldstein stationarity and relate the outer-loop accuracy to the inner equilibrium error.

Stationarity measure. Let $\varphi : \mathbb{R}^{d_\theta} \to \mathbb{R}$ be locally Lipschitz and $\Theta \subseteq \mathbb{R}^{d_\theta}$ be closed and convex. For $\rho > 0$, define $\partial_\rho \varphi(\theta) := \operatorname{conv}\big(\bigcup_{\|v-\theta\| \le \rho} \partial\varphi(v)\big)$, where $\partial\varphi(\cdot)$ is the Clarke subdifferential [Clarke, 1990]; see Goldstein [1977]. For $\eta > 0$, define $G_\Theta(\theta, g; \eta) := \eta^{-1}\big(\theta - \Pi_\Theta(\theta - \eta g)\big)$. We call $\theta \in \Theta$ an (ϵ, ρ, η)-generalized Goldstein stationary point[5] (GGSP) of φ if $\min_{g \in \partial_\rho \varphi(\theta)} \|G_\Theta(\theta, g; \eta)\| \le \epsilon$.

Inexact objective evaluations. The outer loop does not access Φ directly; instead it queries $\widehat{\Phi}_T(\theta) := F(\theta, y_T(\theta))$, where $y_T(\theta)$ is the output of running Algorithm 2 for T steps. The only inner-loop quantity that enters the outer-loop bound is $\varepsilon_y := \sup_{\theta \in \Theta} \sqrt{\mathbb{E}[\|y_T(\theta) - y^\star(\theta)\|^2]}$. By Assumption 2.5, $F(\theta, \cdot)$ is $L_{F,1}$-Lipschitz on $\mathcal{C}$, so $\mathbb{E}[|\widehat{\Phi}_T(\theta) - \Phi(\theta)|] \le L_{F,1}\, \mathbb{E}[\|y_T(\theta) - y^\star(\theta)\|] \le L_{F,1}\varepsilon_y$. This inexactness induces a bias term in the two-point estimator used by Algorithm 1.

Theorem 5.6 (GGSP bound for Algorithm 1)**.** *Let Assumptions 2.1, 2.4 and 2.5 hold. Let Φ be L_Φ-Lipschitz on Θ (Lemma 2.6). Fix $\rho > 0$, run Algorithm 1 under the query convention in Section 3 for K iterations with direction batch size $B_{\text{out}} \ge 1$ and constant stepsize $\eta \le c_0\rho/(L_\Phi\sqrt{d_\theta})$, and let $\widehat{\theta}$ be a uniformly random iterate from $\{\theta_0, \dots, \theta_{K-1}\}$. Let ε_y be as defined above and define the initial hyper-objective gap $\Delta_{\Phi,0} := \Phi(\theta_0) - \inf_{\theta\in\Theta} \Phi(\theta)$. Then there exist universal constants $C_1, C_2, C_3 > 0$ such that*

$$\mathbb{E}\left[\min_{g \in \partial_\rho \Phi(\widehat{\theta})} \|G_\Theta(\widehat{\theta}, g; \eta)\|\right] \le C_1\sqrt{\frac{\Delta_{\Phi,0} + 2\rho L_\Phi}{\eta K}} + C_2 L_\Phi \sqrt{\frac{d_\theta}{B_{\text{out}}}} + C_3 \frac{d_\theta}{\rho}\, L_{F,1}\, \varepsilon_y. \tag{3}$$

In particular, for any $\epsilon > 0$, if $\varepsilon_y \le r$ and $\varepsilon_y \le \rho\epsilon\,(3C_3 d_\theta L_{F,1})^{-1}$, $\eta \le c_0\rho/(L_\Phi\sqrt{d_\theta})$, $B_{\text{out}} \ge 9C_2^2 L_\Phi^2 d_\theta\, \epsilon^{-2}$, and $K \ge 9C_1^2\big(\Phi(\theta_0) - \inf_\Theta \Phi + 2\rho L_\Phi\big)\, \eta^{-1}\epsilon^{-2}$, then $\widehat{\theta}$ is an (ϵ, ρ, η)-GGSP of Φ in expectation.

Corollary 5.7 (Explicit oracle complexity)**.** *Under the hypotheses of Theorems 5.4 and 5.6, assume uniform inner bounds on the queried domain, finite $\bar{\Delta}_0$ and $\Delta_{\Phi,0}$, and $L_{f,2}D^2 > 0$. Let $0 < \epsilon \le L_\Phi$, $\Delta_\Phi := \Delta_{\Phi,0} + 2\rho L_\Phi$, and*

$$A := \frac{L_{f,2} D^2 \max\{1, D/r\}}{\alpha \kappa_m}, \qquad \bar{\varepsilon}_y := \min\left\{r, \frac{\rho\epsilon}{3C_3 d_\theta L_{F,1}}\right\}.$$

Choose $\eta = c_0\rho/(L_\Phi\sqrt{d_\theta})$, $B_{\text{out}} = \lceil 9C_2^2 L_\Phi^2 d_\theta/\epsilon^2\rceil$, $K = \lceil 9C_1^2\Delta_\Phi/(\eta\epsilon^2)\rceil$, and $T = \bar{T}_0 + \lceil 4A/\bar{\varepsilon}_y^2\rceil$. The random output is an (ϵ, ρ, η)-GGSP in expectation, using

$$N_F := 2B_{\text{out}}K = \mathcal{O}\Big(\Delta_\Phi L_\Phi^3 {d_\theta}^{3/2}\rho^{-1}\epsilon^{-4}\Big)$$

upper-level function evaluations. If $\rho\epsilon \le 3C_3 d_\theta L_{F,1} r$, the number of inner gradients and subsampled-LMO calls is

$$N_{\text{inner}} := N_F T = \mathcal{O}\Big(A\Delta_\Phi L_\Phi^3 L_{F,1}^2 {d_\theta}^{7/2}\rho^{-3}\epsilon^{-6} + \Delta_\Phi L_\Phi^3(1 + \bar{T}_0){d_\theta}^{3/2}\rho^{-1}\epsilon^{-4}\Big).$$

The sampled-strategy count is mN_{inner}. With exact LMOs, set $\kappa_m = 1$ and omit Assumption 5.2 and the sampling count.

Appendix H.6 derives the bounds, including the other accuracy regime and integer rounding. These counts exclude oracle preprocessing and per-call arithmetic; the problem constants can themselves depend on dimension (Appendix K).

[5] This notion is consistent with standard first-order stationarity when φ is smooth: if φ is C^1, then the Clarke subdifferential satisfies $\partial\varphi(\theta) = \{\nabla\varphi(\theta)\}$, so $\partial_\rho\varphi(\theta)$ reduces to nearby gradients and collapses to $\{\nabla\varphi(\theta)\}$ as $\rho \downarrow 0$.

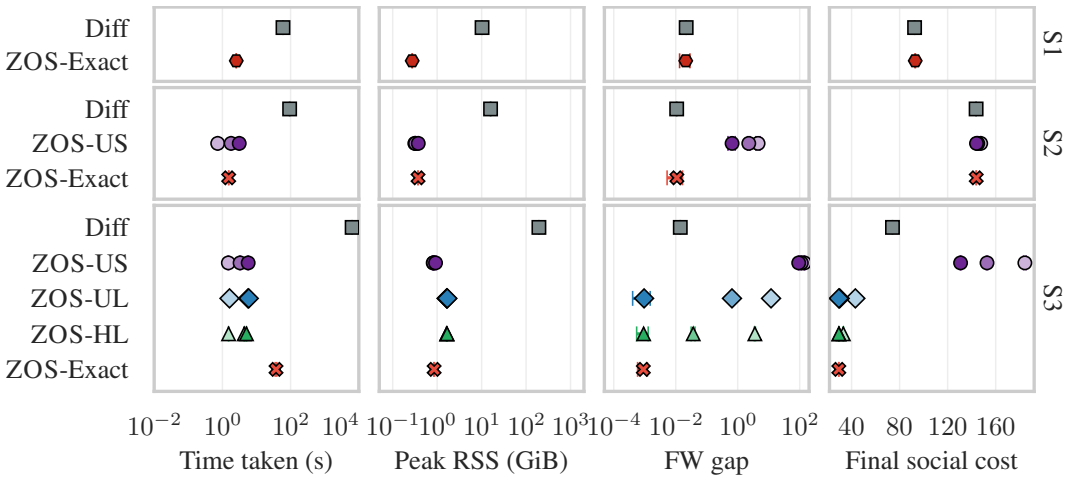


Figure 1: Runtime, memory, and final-iterate diagnostics for Scenarios 1–3: average wall-clock time per executed outer iteration, peak RSS, final FW gap, and final social cost. Exact-oracle ZOS variants are labeled ZOS-Exact. For subsampling-based variants, lighter shades denote smaller candidate budgets $m \in \{10, 100, 1000\}$ in Scenarios 2–3; markers are means and bars are 99% CIs over 10 runs. In Scenario 3 the Diff marker is a feasibility-capped diagnostic (Appendix F.1).

6 Experiments

We evaluate ZOS on three transportation networks from the Transportation Networks for Research Core Team [2016] (TNTP) benchmark suite.[6] The scenarios isolate the role of the lower-level linear minimization oracle (LMO) across three regimes: a tractable shortest-path oracle, a compact exact ZDD oracle, and a large ZDD setting where repeated exact minimization is expensive and subsampled oracles become practically important. Sections 6.1 and 6.2 compare ZOS with the differentiable baseline and other outer methods; Appendix F.4 adds ablations.

Networks and strategy families. We use three TNTP networks (Winnipeg/Chicago/Philadelphia), inducing s–t paths (Scenario 1), Hamiltonian s–t paths (Scenario 2), and Steiner cycles through a fixed terminal set (Scenario 3); Table 2 in Appendix F.1 gives their sizes.

Cost model and Stackelberg objective. The fractional edge costs $c_i(y_i; \theta_i) = d_i(1 + Cy_i/(\theta_i + 1))$ of Sakaue and Nakamura [2021], the leader set Θ of Section 2.3, and leader objective $F(\theta, y) = \sum_i c_i(y_i; \theta_i)\, y_i$ satisfy Assumptions 2.4–2.5 on the experimental domain (Appendix K); each objective query runs T=3000 FW steps.

Algorithms. (i) Diff [Sakaue and Nakamura, 2021]: Diff replaces the LMO by a differentiable softmin over the ZDD and, as presented, computes each outer hypergradient by retaining and backpropagating through the full unrolled inner FW trajectory. **(ii) ZOS (our approach):** We run Algorithm 1 with an LMO implemented via Dijkstra's algorithm (Scenario 1), an exact ZDD argmin (Scenario 2), and either an exact ZDD argmin or ZDD-guided US/UL/HL subsampling (Scenario 3; Section 5.1). Each ZOS outer step still performs the full inner FW solves; the saving comes from avoiding the backward pass through the unrolled FW/ZDD computation (Appendix F.1 reports Diff variants with truncated unrolls and adaptive inner steps).

6.1 Comparison with the Differentiable Baseline

Results. Figure 1 reports four diagnostics; small final FW gaps, as for ZOS-Exact in Scenarios 1–2 and ZOS-UL/HL at m=1000 in Scenario 3, indicate costs evaluated near a true equilibrium.

Scenario 1: Polynomial-time regime. With the exact shortest-path oracle, ZOS matches the social cost of Diff with about 23× less time per iteration and 0.28 GiB instead of 10.2 GiB peak RSS.

[6] Code is available at `https://github.com/sepehr78/ZOS-congestion`.

Table 1: Certified social cost under a matched LMO budget (mean $\pm$ 99% CI over 10 runs; 5 in Scenario 3, where all methods use HL with m=1000). ValueGap-CG is deterministic in Scenarios 1–2.

Method	Scenario 1	Scenario 2	Scenario 3
ZOS	$\mathbf{93.01} \pm 0.24$	$\mathbf{143.83} \pm 0.11$	29.40 ± 0.13
SPSA-FW	108.08 ± 0.99	150.69 ± 0.81	45.46 ± 1.19
ValueGap-ZO	128.56 ± 2.39	148.59 ± 0.81	29.77 ± 0.33
ValueGap-CG	116.33	145.96	$\mathbf{24.97} \pm <0.01$

Scenario 2: Compact-ZDD regime. When the ZDD is compact enough for exact minimization, ZOS-Exact is the best choice: it matches the social cost of Diff with roughly $61\times$ less time per iteration and $0.37\,\mathrm{GiB}$ rather than $15.9\,\mathrm{GiB}$ peak memory.

Scenario 3: Massive-ZDD regime. When the ZDD is massive, stratified subsampling is the practical regime: UL/HL at m=1000 approach ZOS-Exact in social cost and equilibrium accuracy while taking roughly 6–7$\times$ less time per iteration, whereas Diff is impractical ($194\,\mathrm{GiB}$ peak and only five outer iterations in 10 hours). Appendix F.3 verifies the mechanism: the median κ_{1000} is essentially zero for US but exceeds 0.93 for UL and HL.

Total time to quality. Since methods need different numbers of outer iterations, we also measure the total time until the *certified* social cost (F with the follower re-solved to FW gap 10^{-8}) first drops by 90% of the mean ZOS reduction: 7.2, 2.0, and 103 minutes for ZOS in Scenarios 1–3 versus 219 and 316 for Diff in Scenarios 1–2, and over 10 hours for even 50% in Scenario 3.

6.2 Comparison of Outer Methods

To isolate the outer method, we compare ZOS with SPSA [Spall, 1992] (SPSA-FW) and the penalty method of Shen and Chen [2023] with a two-point (ValueGap-ZO) or gradient (ValueGap-CG) update of θ, all with the FW solver and LMO budget of ZOS (Appendix F.4). In certified social cost (Table 1), ZOS is lowest in Scenarios 1–2; in Scenario 3, its mean is below those of SPSA-FW and ValueGap-ZO, and ValueGap-CG, which uses $\nabla_\theta F$ and $\nabla_\theta f$, is lowest.

7 Conclusion

ZOS couples FW equilibrium computation over an implicit load polytope with zeroth-order optimization of the true hyper-objective. The analysis propagates finite equilibrium error to expected approximate generalized Goldstein stationarity, not global optimality, and quantifies subsampled-FW convergence through κ_m. Its applicability depends on uniform growth and regularity, tractable oracle operations, and, for subsampling, useful uniform optimizer mass. Length stratification is not universal, ZDDs can be exponential, and the global positive-slope certificate does not cover BPR costs with exponent greater than one at zero load.

Scope and extensions. Restrictions that do not depend on θ can be incorporated into $\mathcal{S}$. For example, $\mathcal{S}$ may contain only paths that avoid forbidden roads or satisfy a prescribed resource budget. The framework then requires an LMO or sampler for this restricted family, together with the same regularity and sampling assumptions. Finding or sampling suitable strategies under these restrictions may itself be computationally difficult.

Capacity constraints on aggregate loads restrict $\mathcal{C}$ itself and may require a different LMO. Constraints that change with θ need additional feasibility and stability analysis. Atomic followers and equilibria without a potential function require new inner formulations or solvers, together with bounds on the resulting equilibrium errors.

Beyond CCGs, the proof can be reused for a smooth convex lower problem on a fixed compact convex set, under the same uniqueness, uniform growth, and regularity conditions. This also requires lower gradients and an LMO, Lipschitz F, projection onto Θ, and controlled function-value errors at every outer query. Appendix E.1 details these requirements and extensions. Weaker growth, generalized-gradient outer methods, and improved problem-specific sampling guarantees remain future directions.

Acknowledgements and Disclosures

This work was supported as a part of NCCR Automation, a National Centre of Competence in Research, funded by the Swiss National Science Foundation (grant number 51NF40_225155). The authors declare no competing interests.

References

Alireza Aghasi and Saeed Ghadimi. Fully zeroth-order bilevel programming via Gaussian smoothing. *Journal of Optimization Theory and Applications*, 205(2):31, 2025. doi: 10.1007/s10957-025-02647-y.

Akshay Agrawal, Brandon Amos, Shane Barratt, Stephen Boyd, Steven Diamond, and Zico Kolter. Differentiable convex optimization layers. In *Advances in Neural Information Processing Systems*, volume 32, pages 9558–9570, 2019.

Mahnoosh Alizadeh, Hoi-To Wai, Mainak Chowdhury, Andrea Goldsmith, Anna Scaglione, and Tara Javidi. Optimal pricing to manage electric vehicles in coupled power and transportation networks. *IEEE Transactions on Control of Network Systems*, 4(4):863–875, 2017. doi: 10.1109/TCNS.2016.2590259.

Brandon Amos and J. Zico Kolter. Optnet: Differentiable optimization as a layer in neural networks. In Doina Precup and Yee Whye Teh, editors, *Proceedings of the 34th International Conference on Machine Learning*, volume 70 of *Proceedings of Machine Learning Research*, pages 136–145. PMLR, 2017.

Elliot Anshelevich, Anirban Dasgupta, Jon Kleinberg, Éva Tardos, Tom Wexler, and Tim Roughgarden. The price of stability for network design with fair cost allocation. In *45th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 295–304, 2004. doi: 10.1109/FOCS.2004.68.

Babak Ghaffarzadeh Bakhshayesh and Hamed Kebriaei. Generalized wardrop equilibrium for charging station selection and route choice of electric vehicles in joint power distribution and transportation networks. *IEEE Transactions on Control of Network Systems*, 10(3):1245–1254, 2023. doi: 10.1109/TCNS.2023.3239535.

Martin J. Beckmann, C. B. McGuire, and Christopher B. Winsten. *Studies in the Economics of Transportation*. Yale University Press, 1956. Published for the Cowles Commission for Research in Economics.

David Bergman, André A. Cire, Willem-Jan van Hoeve, and John Hooker. *Decision Diagrams for Optimization*. Springer, 2016. ISBN 9783319428499. doi: 10.1007/978-3-319-42849-9.

J. Frédéric Bonnans and Alexander Shapiro. Optimization problems with perturbations: A guided tour. *SIAM Review*, 40(2):228–264, 1998. doi: 10.1137/S0036144596302644.

J. Frédéric Bonnans and Alexander Shapiro. *Perturbation Analysis of Optimization Problems*. Springer Series in Operations Research and Financial Engineering. Springer, New York, 2000. ISBN 978-0-387-98705-7. doi: 10.1007/978-1-4612-1394-9.

Randal E. Bryant. Graph-based algorithms for Boolean function manipulation. *IEEE Transactions on Computers*, 35(8):677–691, 1986. doi: 10.1109/TC.1986.1676819.

James V. Burke, Adrian S. Lewis, and Michael L. Overton. A robust gradient sampling algorithm for nonsmooth, nonconvex optimization. *SIAM Journal on Optimization*, 15(3):751–779, 2005. doi: 10.1137/030601296.

He Chen, Jiajin Li, and Anthony Man-Cho So. Set smoothness unlocks Clarke hyper-stationarity in bilevel optimization. In *Advances in Neural Information Processing Systems*, volume 38, 2025.

Frank H. Clarke. *Optimization and Nonsmooth Analysis*, volume 5 of *Classics in Applied Mathematics*. SIAM, Philadelphia, PA, 1990. doi: 10.1137/1.9781611971309.

Andrew R. Conn, Katya Scheinberg, and Luis N. Vicente. *Introduction to Derivative-Free Optimization*. Society for Industrial and Applied Mathematics, 2009. ISBN 9780898716689. doi: 10.1137/1.9780898718768.

G. B. Dantzig and J. H. Ramser. The truck dispatching problem. *Management Science*, 6(1):80–91, 1959. doi: 10.1287/mnsc.6.1.80.

Damek Davis, Dmitriy Drusvyatskiy, Sham Kakade, and Jason D. Lee. Stochastic subgradient method converges on tame functions. *Foundations of Computational Mathematics*, 20(1):119–154, 2020. doi: 10.1007/s10208-018-09409-5.

Robert B. Dial. Minimal-revenue congestion pricing part i: A fast algorithm for the single-origin case. *Transportation Research Part B: Methodological*, 33(3):189–202, 1999. doi: 10.1016/S0191-2615(98)00026-5.

E. W. Dijkstra. A note on two problems in connexion with graphs. *Numerische Mathematik*, 1(1): 269–271, 1959. doi: 10.1007/BF01386390.

Dmitriy Drusvyatskiy, Maryam Fazel, and Lillian J. Ratliff. Improved rates for derivative free gradient play in strongly monotone games. In *2022 IEEE 61st Conference on Decision and Control (CDC)*, pages 3403–3408, 2022. doi: 10.1109/cdc51059.2022.9992950.

John C. Duchi, Michael I. Jordan, Martin J. Wainwright, and Andre Wibisono. Optimal rates for zero-order convex optimization: The power of two function evaluations. *IEEE Transactions on Information Theory*, 61(5):2788–2806, 2015. doi: 10.1109/TIT.2015.2409256.

Matthias J. Ehrhardt and Lindon Roberts. Inexact derivative-free optimization for bilevel learning. *Journal of Mathematical Imaging and Vision*, 63:580–600, 2021. doi: 10.1007/s10851-021-01020-8.

Jonas Eliasson, Lars Hultkrantz, Lena Nerhagen, and Lena Smidfelt Rosqvist. The Stockholm congestion-charging trial 2006: Overview of effects. *Transportation Research Part A: Policy and Practice*, 43(3):240–250, 2009. doi: 10.1016/j.tra.2008.09.007.

Francisco Facchinei and Jong-Shi Pang. *Finite-Dimensional Variational Inequalities and Complementarity Problems*, volume I. Springer, New York, 2003. doi: 10.1007/b97543.

Abraham D. Flaxman, Adam Tauman Kalai, and H. Brendan McMahan. Online convex optimization in the bandit setting: Gradient descent without a gradient. In *Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 385–394. SIAM, 2005.

Marguerite Frank and Philip Wolfe. An algorithm for quadratic programming. *Naval Research Logistics Quarterly*, 3(1–2):95–110, 1956. doi: 10.1002/nav.3800030109.

M. R. Garey, D. S. Johnson, and R. Endre Tarjan. The planar Hamiltonian circuit problem is NP-complete. *SIAM Journal on Computing*, 5(4):704–714, 1976. doi: 10.1137/0205049.

Saeed Ghadimi and Guanghui Lan. Stochastic first- and zeroth-order methods for nonconvex stochastic programming. *SIAM Journal on Optimization*, 23(4):2341–2368, 2013. doi: 10.1137/120880811.

Khanh-Hung Giang-Tran, Nam Ho-Nguyen, and Dabeen Lee. A projection-free method for solving convex bilevel optimization problems. *Mathematical Programming*, 213(1–2):907–940, 2025. doi: 10.1007/s10107-024-02157-1.

Allen A. Goldstein. Optimization of lipschitz continuous functions. *Mathematical Programming*, 13 (1):14–22, 1977. doi: 10.1007/BF01584320.

Aric A. Hagberg, Daniel A. Schult, and Pieter J. Swart. Exploring network structure, dynamics, and function using NetworkX. In *Proceedings of the 7th Python in Science Conference*, pages 11–15, 2008.

Elad Hazan and Haipeng Luo. Variance-reduced and projection-free stochastic optimization. In *Proceedings of the 33rd International Conference on Machine Learning*, volume 48 of *Proceedings of Machine Learning Research*, pages 1263–1271. PMLR, 2016.

D. W. Hearn, M. B. Yildirim, M. V. Ramana, and L. H. Bai. Computational methods for congestion toll pricing models. In *Proceedings of the 2001 IEEE Intelligent Transportation Systems Conference (ITSC)*, pages 257–262, 2001. doi: 10.1109/ITSC.2001.948665.

Haoqiang Huang, Zihe Wang, Zhide Wei, and Jie Zhang. Cost minimization for equilibrium transition. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pages 9765–9772, 2024. doi: 10.1609/aaai.v38i9.28835.

Takeru Inoue, Hiroaki Iwashita, Jun Kawahara, and Shin-ichi Minato. Graphillion: Software library for very large sets of labeled graphs. *International Journal on Software Tools for Technology Transfer*, 18(1):57–66, 2016.

Ruben Interian and Celso C. Ribeiro. A GRASP heuristic using path-relinking and restarts for the Steiner traveling salesman problem. *International Transactions in Operational Research*, 24(6): 1307–1323, 2017. doi: 10.1111/itor.12419.

Martin Jaggi. Revisiting Frank-Wolfe: Projection-free sparse convex optimization. In *Proceedings of the 30th International Conference on Machine Learning*, volume 28 of *Proceedings of Machine Learning Research*, pages 427–435. PMLR, 2013.

George Karakostas and Stavros G. Kolliopoulos. Stackelberg strategies for selfish routing in general multicommodity networks. *Algorithmica*, 53(1):132–153, 2009. doi: 10.1007/s00453-007-9018-5.

Thomas Kerdreux, Fabian Pedregosa, and Alexandre d'Aspremont. Frank-Wolfe with subsampling oracle. In Jennifer Dy and Andreas Krause, editors, *Proceedings of the 35th International Conference on Machine Learning*, volume 80 of *Proceedings of Machine Learning Research*, pages 2591–2600. PMLR, 2018.

Moonseong Kim, Hyunseung Choo, Matt W. Mutka, Hyung-Jin Lim, and Kwangjin Park. On QoS multicast routing algorithms using k-minimum Steiner trees. *Information Sciences*, 238:190–204, 2013. doi: 10.1016/j.ins.2013.03.006.

Y. A. Korilis, A. A. Lazar, and A. Orda. Achieving network optima using Stackelberg routing strategies. *IEEE/ACM Transactions on Networking*, 5(1):161–173, 1997. doi: 10.1109/90.554730.

Larry J. LeBlanc, Edward K. Morlok, and William P. Pierskalla. An efficient approach to solving the road network equilibrium traffic assignment problem. *Transportation Research*, 9(5):309–318, 1975. doi: 10.1016/0041-1647(75)90030-1.

Jiaoyang Li, Andrew Tinka, Scott Kiesel, Joseph W. Durham, T. K. Satish Kumar, and Sven Koenig. Lifelong multi-agent path finding in large-scale warehouses. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pages 11272–11281, 2021. doi: 10.1609/aaai.v35i13.17344.

Sarah H. Q. Li, Daniel Calderone, and Behçet Açıkmese. Congestion-aware path coordination game with markov decision process dynamics. *IEEE Control Systems Letters*, 7:431–436, 2023. doi: 10.1109/LCSYS.2022.3189323.

Zhuanghua Liu, Cheng Chen, Luo Luo, and Bryan Kian Hsiang Low. Zeroth-order methods for constrained nonconvex nonsmooth stochastic optimization. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pages 30842–30872. PMLR, 2024.

Chinmay Maheshwari, Chih-Yuan Chiu, Eric Mazumdar, Shankar Sastry, and Lillian Ratliff. Zeroth-order methods for convex-concave min-max problems: Applications to decision-dependent risk minimization. In *Proceedings of The 25th International Conference on Artificial Intelligence and Statistics*, volume 151 of *Proceedings of Machine Learning Research*, pages 6702–6734. PMLR, 2022.

Alberto Marchesi, Matteo Castiglioni, and Nicola Gatti. Leadership in congestion games: Multiple user classes and non-singleton actions. In *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence (IJCAI-19)*, pages 485–491, 2019. doi: 10.24963/ijcai.2019/69.

Saeed Masiha, Zebang Shen, Negar Kiyavash, and Niao He. Superquantile-Gibbs relaxation for minima-selection in bilevel optimization. *arXiv preprint arXiv:2505.05991*, 2025. doi: 10.48550/arXiv.2505.05991. Version 4, revised March 2026.

Athanasios Migdalas. Bilevel programming in traffic planning: Models, methods and challenge. *Journal of Global Optimization*, 7(4):381–405, 1995. doi: 10.1007/BF01099649.

Shin-ichi Minato. Zero-suppressed BDDs for set manipulation in combinatorial problems. In *Proceedings of the 30th ACM/IEEE Design Automation Conference (DAC)*, pages 272–277, 1993. doi: 10.1145/157485.164890.

Aryan Mokhtari, Hamed Hassani, and Amin Karbasi. Stochastic conditional gradient methods: From convex minimization to submodular maximization. *Journal of Machine Learning Research*, 21 (105):1–49, 2020.

Yurii Nesterov and Vladimir Spokoiny. Random gradient-free minimization of convex functions. *Foundations of Computational Mathematics*, 17(2):527–566, 2017. doi: 10.1007/s10208-015-9296-2.

Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, Alban Desmaison, Andreas Köpf, Edward Yang, Zachary DeVito, Martin Raison, Alykhan Tejani, Sasank Chilamkurthy, Benoit Steiner, Lu Fang, Junjie Bai, and Soumith Chintala. PyTorch: An imperative style, high-performance deep learning library. In *Advances in Neural Information Processing Systems*, volume 32, 2019.

M. Patriksson. *Traffic Assignment Problems: Models and Methods*. V.S.P. Intl Science, 1994. ISBN 9789067641814.

Marco Rando and Samuel Vaiter. ZOBA: An Efficient Single-loop Zeroth-order Bilevel Optimization Algorithm, 2026. URL `https://arxiv.org/abs/2601.21836`.

Tim Roughgarden and Éva Tardos. How bad is selfish routing? *Journal of the ACM*, 49(2):236–259, 2002. doi: 10.1145/506147.506153.

Shinsaku Sakaue and Kengo Nakamura. Differentiable equilibrium computation with decision diagrams for Stackelberg models of combinatorial congestion games. In *Advances in Neural Information Processing Systems*, volume 34, pages 9416–9428, 2021.

Juan-Jose Salazar-Gonzalez. The Steiner cycle polytope. *European Journal of Operational Research*, 147(3):671–679, 2003. doi: 10.1016/S0377-2217(02)00359-4.

William H. Sandholm. *Population Games and Evolutionary Dynamics*. Economic Learning and Social Evolution. MIT Press, Cambridge, MA, 2010. ISBN 978-0-262-19587-8.

Yosef Sheffi. *Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods*. Prentice-Hall, Englewood Cliffs, NJ, 1985. ISBN 0139397299.

Han Shen and Tianyi Chen. On penalty-based bilevel gradient descent method. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pages 30992–31015. PMLR, 2023. URL `https://proceedings.mlr.press/v202/shen23c.html`.

J. C. Spall. Multivariate stochastic approximation using a simultaneous perturbation gradient approximation. *IEEE Transactions on Automatic Control*, 37(3):332–341, 1992. doi: 10.1109/9.119632.

Mathias Staudigl, Simon Weissmann, and Tristan van Leeuwen. Derivative-free stochastic bilevel optimization for inverse problems. *Computational Optimization and Applications*, 93:967–1020, 2026. doi: 10.1007/s10589-025-00745-1.

Transportation Networks for Research Core Team. Transportation networks for research, 2016. URL `https://github.com/bstabler/TransportationNetworks`. Accessed May 7, 2026.

J. G. Wardrop. Some theoretical aspects of road traffic research. *Proceedings of the Institution of Civil Engineers*, 1(3):325–362, 1952. doi: 10.1680/ipeds.1952.11259.

H. Yaiche, R. R. Mazumdar, and C. Rosenberg. A game theoretic framework for bandwidth allocation and pricing in broadband networks. *IEEE/ACM Transactions on Networking*, 8(5):667–678, 2000. doi: 10.1109/90.879352.

Hai Yang and Michael G. H. Bell. Traffic restraint, road pricing and network equilibrium. *Transportation Research Part B: Methodological*, 31(4):303–314, 1997. doi: 10.1016/S0191-2615(96)00030-6.

Dajun Yue, Jiyao Gao, Bo Zeng, and Fengqi You. A projection-based reformulation and decomposition algorithm for global optimization of a class of mixed integer bilevel linear programs. *Journal of Global Optimization*, 73(1):27–57, 2019. doi: 10.1007/s10898-018-0679-1.

Alain B. Zemkoho and Shenglong Zhou. Theoretical and numerical comparison of the Karush–Kuhn–Tucker and value function reformulations in bilevel optimization. *Computational Optimization and Applications*, 78(2):625–674, 2021. doi: 10.1007/s10589-020-00250-7.

Bo Zeng and Yu An. Solving bilevel mixed integer program by reformulations and decomposition. Technical report, Department of Industrial and Management Systems Engineering, University of South Florida, June 2014. URL `https://optimization-online.org/wp-content/uploads/2014/07/4455.pdf`.

Hongtai Zeng, Yujie Sheng, Hongbin Sun, Yanzhen Zhou, Yixun Xue, and Qinglai Guo. A conic relaxation approach for solving Stackelberg pricing game of electric vehicle charging station considering traffic equilibrium. *IEEE Transactions on Smart Grid*, 15(3):3080–3097, 2024. doi: 10.1109/TSG.2023.3329651.

A Notation

The table below summarizes the main notation used throughout the paper.

Symbol	Meaning
n; $[n]$	Number of resources; index set $\{1, \dots, n\}$.
$\mathcal{S}$; $d := \lvert\mathcal{S}\rvert$	Set of feasible (combinatorial) strategies; number of strategies.
$S \in \mathcal{S}$; $\mathbf{1}_S \in \{0,1\}^n$	A strategy (subset of resources); its incidence vector.
Δ^d; $z \in \Delta^d$	Probability simplex; population distribution over strategies.
$y(z) \in \mathbb{R}^n$; $\mathcal{C}$	Induced load $y(z) = \sum_S z_S \mathbf{1}_S$; feasible load polytope $\mathcal{C} = \mathrm{conv}\{\mathbf{1}_S : S \in \mathcal{S}\} \subseteq [0,1]^n$.
$c_i(\cdot;\theta)$; $c_S(y)$	Cost of resource i under parameter θ; strategy cost $c_S(y) = \sum_{i \in S} c_i(y_i;\theta)$.
$\theta \in \Theta \subseteq \mathbb{R}^{d_\theta}$; d_θ	Leader parameter and its feasible set; dimension.
$f(\theta, y)$	Lower-level (Beckmann) potential; Wardrop equilibria are minimizers over $\mathcal{C}$.
$y^\star(\theta)$	Wardrop equilibrium load: $y^\star(\theta) \in \arg\min_{y \in \mathcal{C}} f(\theta, y)$.
$F(\theta, y)$; $\Phi(\theta)$	Upper-level objective; hyper-objective $\Phi(\theta) = F(\theta, y^\star(\theta))$.
y_t; T	Inner Frank–Wolfe iterates; number of inner iterations.
$\omega_t = \nabla_y f(\theta, y_t)$	FW weight vector used in the linear minimization oracle (LMO).
$\mathrm{LMO}(g)$; $\mathrm{LMO}_m(g)$	Exact LMO $\arg\min_{y \in \mathcal{C}} \langle g, y\rangle$; subsampled LMO over m sampled strategies.
q; $S^{(1:m)} \sim q$; m	Sampling distribution over $\mathcal{S}$; sampled candidates; sample size.
$\mathrm{Opt}(g)$; $p(g)$; $\kappa_m(g)$	Exact LMO minimizers; optimizer mass $p(g) = q(\mathrm{Opt}(g))$; optimizer-hit probability $\kappa_m(g) = 1 - (1 - p(g))^m$.
$\underline{p}$; $\underline{\kappa}_m$	Uniform lower bound $p(g) \geq \underline{p}$ along the inner run; corresponding uniform hit probability $\underline{\kappa}_m = 1 - (1 - \underline{p})^m$.
$\tau : \mathcal{S} \to \mathcal{T}$; $\mathcal{S}_t$; w	Stratification map; stratum $\mathcal{S}_t := \{S : \tau(S) = t\}$; stratum weights $w = (w_t)_{t \in \mathcal{T}}$.
$\mathcal{L}$; N_t	Set of attainable lengths $\mathcal{L} = \{\lvert S\rvert : S \in \mathcal{S}\}$; stratum size $N_t := \lvert\mathcal{S}_t\rvert$.
Δ_0	Initial inner optimality gap: $\Delta_0 = f(\theta, y_0) - f(\theta, y^\star(\theta))$.
$\bar{\Delta}_0$; $\bar{T}_0$	Uniform bound $\bar{\Delta}_0 = \sup_{\theta \in \Theta} \Delta_0(\theta)$; corresponding uniform burn-in bound $\bar{T}_0$ in Theorem 5.4.
T_0	Logarithmic burn-in length in Theorem 5.4.
D	Diameter of $\mathcal{C}$: $D = \max\{\lVert u - v\rVert : u, v \in \mathcal{C}\}$.
$\lVert\cdot\rVert$; $\langle\cdot,\cdot\rangle$; $\lvert\cdot\rvert$	Euclidean norm ($\lVert\cdot\rVert_2$) and inner product; absolute value for scalars.
θ_t; K	Outer (leader) iterates; number of outer iterations.
$\Pi_\Theta(\cdot)$	Euclidean projection onto the leader feasible set Θ.
$\widehat{\Phi}_T(\theta)$	Approximate hyper-objective value using T inner steps: $\widehat{\Phi}_T(\theta) = F(\theta, y_T(\theta))$.
ε_y	Uniform mean-square inner accuracy: $\varepsilon_y := \sup_{\theta \in \Theta} \sqrt{\mathbb{E}[\lVert y_T(\theta) - y^\star(\theta)\rVert^2]}$.
$\widehat{g}_t$; B_{out}	Mini-batched two-point gradient estimator; outer direction batch size in Algorithm 1.
ρ; $\mathbb{B}$; $\mathbb{S}^{d_\theta - 1}$	Smoothing radius; unit ball; unit sphere in $\mathbb{R}^{d_\theta}$.
$u \sim \mathrm{Unif}(\mathbb{S}^{d_\theta - 1})$	Random direction used in two-point finite differences.
η	Outer stepsize.
$\partial_\rho \Phi(\theta)$; $G_\Theta(\theta, g; \eta)$	Goldstein ρ-subdifferential; projected gradient mapping.
$L_{f,2}, L_{F,1}, L_\Phi$; α, r	Smoothness/Lipschitz constants; quadratic-growth parameters.

B Algorithm Details

In our experiments we use exact line search along the FW segment, but the theory in Section 5 also covers the standard smoothness-based short-step rule, so exact line search is not required by the method. The output $y_T(\theta)$ defines the approximate hyper-objective value $\widehat{\Phi}_T(\theta) := F(\theta, y_T(\theta))$. We quantify the inner accuracy as a function of T in Section 5.

Algorithm 2 FW-EQUILIBRIUM(θ, T): Frank–Wolfe equilibrium solver

Input: parameter $\theta \in \Theta$, iterations $T \geq 1$
initialize $y_0 \in \mathcal{C}$ (e.g., $y_0 = \mathbf{1}_{S_0}$ for some $S_0 \in \mathcal{S}$)
for $t = 0, 1, \ldots, T-1$ **do**
  compute $\omega_t := \nabla_y f(\theta, y_t)$
  oracle call $s_t := \mathrm{LMO}(\omega_t)$ (or $s_t := \mathrm{LMO}_m(\omega_t)$)
  choose $\gamma_t \in [0, 1]$ (exact line search or short-step rule)
  update $y_{t+1} := (1 - \gamma_t) y_t + \gamma_t s_t$
end for
return $y_T(\theta) := y_T$

C Discussion, Limitations, and Future Directions

The main conclusion and scope discussion appear in Section 7. The guarantee is approximate generalized Goldstein stationarity, not global optimality. Uniform quadratic growth controls equilibrium error and stability; useful optimizer mass is a separate condition needed only for subsampled LMOs. Strongly increasing resource costs do not, by themselves, verify that sampling condition. Sharper problem-specific hit bounds and weaker-growth theory remain future work.

Our experiments cover three TNTP-derived road-network cases spanning the polynomial-time, exact-ZDD, and subsampled oracle regimes; a natural next empirical step is to instantiate the same outer algorithm with other domain oracles, such as communication-network routing, EV charging, and multi-agent path coordination.

Clarke-based alternatives. A nonsmooth first-order outer method is possible when it has a valid generalized-gradient oracle for Φ and satisfies the chosen method's regularity conditions. At differentiability points, the chain rule uses $Dy^\star(\theta)$ in addition to derivatives of F. At a kink, derivatives of F and f alone need not identify a Clarke subgradient, i.e., an element of the closed convex hull of limiting nearby gradients of Φ [Clarke, 1990, Bonnans and Shapiro, 2000]. An explicit constraint representation with suitable regularity can permit differentiation of KKT systems on neighboring active-set regimes; this is one route, not a requirement of every generalized-gradient method. Gradient sampling instead uses gradients at nearby differentiability points [Burke et al., 2005]. Either route must justify the computed generalized gradients and control errors from finite equilibrium solves; an unverified approximation to $Dy^\star$ in the smooth chain rule does not provide that guarantee. Developing such an oracle for implicit load polytopes is a promising extension. Nonsmoothness of $y^\star$ alone does not imply nonsmoothness of Φ; Proposition 2.7 proves the latter for its specified upper objective.

Growth and BPR costs. For BPR travel times $c_i(y_i) = t_i^0[1 + a(y_i/u_i)^\beta]$ with $t_i^0, a, u_i > 0$ and $\beta > 1$, the derivative at zero vanishes. Hence the global positive-slope certificate in Section 2.3 does not apply on all of $[0, 1]$. A particular network and parameter range can still satisfy uniform local quadratic growth, but this and the required stability must be checked separately. Replacing quadratic growth by $f - f^\star \geq c\|y - y^\star\|^s$ would bound load error by $[(f - f^\star)/c]^{1/s}$; it does not automatically give the Lipschitz response used in the present outer analysis. A general weak-growth extension therefore requires new analysis.

Representation and changing networks. ZDD preprocessing is reusable when $\mathcal{S}$ is fixed and only additive weights change. It can be exponential even then. Road closures, changing terminals, or moving feasibility constraints require updating the representation and establishing uniform feasibility and equilibrium stability, not merely reweighting a fixed diagram. Incremental diagrams and alternative domain-specific oracles are possible algorithmic directions; Appendix E.1 details the modeling boundaries.

D Related Work

Convex lower levels, reformulations, and VIs. The Beckmann lower level is a convex program, equivalently the potential VI $\langle c(y; \theta), u - y \rangle \geq 0$ for every $u \in \mathcal{C}$; VI formulations and their

stability theory are classical [Facchinei and Pang, 2003, Bonnans and Shapiro, 2000]. With an explicit constraint representation and the relevant constraint qualifications, its KKT conditions give a single-level reformulation. Alternatively, the value-function constraint $f(\theta, y) \leq \min_{u \in \mathcal{C}} f(\theta, u)$ enforces lower optimality; Zemkoho and Zhou [2021] compare these two approaches. Neither formulation requires a differentiable equilibrium map merely to be stated. For example, the projected penalty-gradient method of Shen and Chen [2023] avoids differentiating that map, but uses objective derivatives and projections. These methods may be preferable when their required operations are efficiently available. Our access to $\mathcal{C}$ through an LMO does not directly supply projections, an inequality system, multipliers, or active-face information; we do not claim that such information is mathematically impossible to recover.

Projection-free convex bilevel optimization. Giang-Tran et al. [2025] minimize a smooth convex objective over the minimizers of a fixed smooth convex problem on a compact convex set. Their method uses gradients of both objectives and an LMO over the base set, with bounds on both objective gaps. Here the leader parameter changes the lower problem, and the resulting hyper-objective may be nonsmooth and nonconvex. Their guarantee therefore does not directly cover our function-value outer oracle or finite, subsampled equilibrium errors. This comparison concerns that formulation, not all projection-free bilevel or VI methods. For fixed θ, our potential VI is solved by FW through Proposition 2.3; extensions to non-potential VIs require the additional analysis discussed in Appendix E.

Mixed-integer reformulation and decomposition. Zeng and An [2014] and Yue et al. [2019] develop column-and-constraint-generation methods for explicitly represented mixed-integer bilevel linear programs. Their reformulations exploit lower-level integer assignments and LP KKT or duality conditions, with finite global-convergence guarantees under their respective assumptions. The latter work extends the decomposition to upper-level constraints involving lower-level variables and uses a projection-based formulation to handle failures of the relatively complete response property. Our follower variables are continuous aggregate loads, the Beckmann objective is generally nonlinear, and the load polytope is accessed through an LMO. Applying those algorithms here would require a suitable explicit reformulation and new analysis; their global guarantees do not transfer directly to our stationarity result.

Stackelberg control and congestion games. Congestion games and selfish routing are classical models for large-scale transportation and communication systems, with Wardrop equilibria as the standard notion of flow-level equilibrium [Wardrop, 1952]. The potential formulation of Wardrop equilibria and its connection to convex optimization dates back to the Beckmann transformation [Beckmann et al., 1956] and has since been widely used in algorithmic and economic analyses of routing [Roughgarden and Tardos, 2002, Sandholm, 2010]. Stackelberg control in these models includes leader decisions such as tolling, pricing, or capacity design that influence equilibrium behavior; see, e.g., Karakostas and Kolliopoulos [2009] and references therein. More general Stackelberg formulations in congestion games allow richer action sets (e.g., multiple user classes and non-singleton actions), further motivating combinatorial strategy spaces beyond simple path choices [Marchesi et al., 2019]. Related work also studies how monetary interventions can steer the system between equilibria [Huang et al., 2024]. For large networks, equilibrium computation is typically performed using methods that exploit shortest-path computations and other combinatorial subroutines. Classic traffic-assignment algorithms include Frank–Wolfe-type schemes specialized to the Beckmann potential [LeBlanc et al., 1975]. Since Wardrop equilibria coincide with minimizers of the (convex) Beckmann potential, each Frank–Wolfe step solves the linearization of the potential over the feasible flow polytope; in network models this reduces to an all-or-nothing assignment subproblem (typically implemented via shortest-path computations). This yields a projection-free method whose per-iteration cost is dominated by combinatorial primitives rather than Euclidean projections, which is one reason Frank–Wolfe is widely used in practice. Comprehensive discussions of step-size rules (e.g., exact line search vs. method-of-successive-averages heuristics), convergence, and practical variants can be found in standard references such as Sheffi [1985], Patriksson [1994].

Compact representations of combinatorial strategy sets. When the strategy family $\mathcal{S}$ is exponentially large, one approach is to represent $\mathcal{S}$ implicitly via compact data structures that share substructure across many feasible strategies. Binary decision diagrams (BDDs) encode a Boolean function as a reduced, ordered decision DAG [Bryant, 1986]; in our setting, the function is simply

the indicator of feasibility over subsets of $[n]$. Zero-suppressed BDDs (ZDDs) modify the reduction rules to better match *families of sparse sets*, roughly, instances where most sets omit most items, which often leads to dramatically smaller diagrams in practice [Minato, 1993]. Once constructed, the ZDD supports efficient dynamic programming on the DAG to perform common set-family operations (e.g., union/intersection), to count feasible strategies, to sample feasible strategies, and to optimize additive costs over $\mathcal{S}$ without enumerating it [Minato, 1993, Bergman et al., 2016]. Although ZDD size is problem- and ordering-dependent (and can be exponential in the worst case), decision-diagram techniques have become a practical tool for discrete optimization and for implementing combinatorial oracles over large strategy spaces [Bergman et al., 2016]. Conceptually, while our use of ZDDs is similar to the equilibrium computations in Sakaue and Nakamura [2021], our role for them is different. We use ZDDs purely as a *combinatorial oracle* for linear minimization inside Frank–Wolfe; the inner loop remains the standard (non-smoothed) FW method, and we never require its iterations to be differentiable with respect to θ. This stands in contrast to differentiable approaches that modify the FW steps (e.g., via softmin relaxations) to enable backpropagation through the equilibrium map and rely on gradient-based updates for the leader's problem. Our algorithm instead keeps the equilibrium computation as a black box and handles the non-smooth dependence on θ entirely at the outer level via zeroth-order methods.

Differentiating through equilibria and decision diagrams. A contrasting line of work makes the equilibrium computation itself *differentiable* and then applies gradient-based bilevel optimization by differentiating through the lower-level computation. For Stackelberg CCGs, Sakaue and Nakamura [2021] is the closest prior work to ours at the modeling level: it studies the same leader–follower CCG framework and also uses ZDDs to cope with exponentially large strategy families. The crucial difference is algorithmic and analytical. Sakaue and Nakamura [2021] make the equilibrium solver differentiable by replacing the exact LMO with a softmin relaxation, then backpropagate through a finite number of inner iterations and optimize the resulting surrogate map $\theta \mapsto y_T(\theta)$ and surrogate objective value $F(\theta, y_T(\theta))$. By contrast, we keep the inner Frank–Wolfe solver in its standard projection-free form, use ZDDs only as exact or sampled combinatorial oracles, and place the nonsmoothness entirely in the outer loop. More broadly, differentiable optimization layers provide derivatives through suitably regular optimization problems [Amos and Kolter, 2017, Agrawal et al., 2019]. For smoothed or finitely unrolled solvers, these derivatives instead concern the resulting approximation to $\Phi(\theta) = F(\theta, y^\star(\theta))$; this qualification does not apply to every implicit-differentiation method. Establishing end-to-end guarantees for Φ requires quantifying both the equilibrium approximation error $y_T(\theta) \approx y^\star(\theta)$ and the mismatch between surrogate gradients and suitable generalized derivatives of Φ; this difficulty is explicitly noted in Sakaue and Nakamura [2021]. In contrast, our method avoids differentiating through equilibria and provides convergence to generalized Goldstein stationary points of the true, typically Lipschitz but nonsmooth, hyper-objective, with explicit dependence on the inner equilibrium error.

Zeroth-order (derivative-free) optimization. Because the hyper-objective $\Phi(\theta) = F(\theta, y^\star(\theta))$ is typically *Lipschitz but nonsmooth* due to active-set changes in $y^\star(\theta)$, we adopt a weak first-order stationarity notion based on Clarke and Goldstein subdifferentials [Clarke, 1990, Goldstein, 1977]. These generalized derivatives and stationarity measures also underpin classic algorithms for nonsmooth nonconvex optimization, most notably gradient sampling [Burke et al., 2005], as well as modern analyses of stochastic methods on tame/nonsmooth landscapes [Davis et al., 2020, Liu et al., 2024]. Our outer loop is a zeroth-order method based on *randomized smoothing* and *two-point* finite-difference estimators, a standard approach in derivative-free optimization and bandit/zeroth-order learning: foundational results include Flaxman et al. [2005], with refined rates and smoothing-based analyses developed in, e.g., Duchi et al. [2015], Nesterov and Spokoiny [2017], Ghadimi and Lan [2013]. Zeroth-order methods have also been analyzed in game-theoretic settings, including convex–concave minmax formulations and (strongly) monotone games [Maheshwari et al., 2022, Drusvyatskiy et al., 2022]. More classical derivative-free schemes include stochastic approximation methods such as SPSA [Spall, 1992] and general-purpose trust-region frameworks for derivative-free optimization [Conn et al., 2009]. Zeroth-order approaches are also useful in bilevel problems when the outer objective is evaluated through an inner solve and may be nonsmooth or otherwise unsuitable for backpropagation. Several recent papers share part of this perspective, especially the use of derivative-free outer updates, inexact lower-level solves, or nonsmooth hyper-objectives. The main distinction is that our guarantee is tied to the particular access model of Stackelberg CCGs: a true equilibrium objective induced by a unique Wardrop equilibrium over a combinatorial load polytope,

where each function evaluation is produced by a Frank–Wolfe equilibrium solver using exact or subsampled weighted minimum-cost strategy oracles.

- **Ehrhardt and Roberts [2021].** Their inexact derivative-free trust-region method learns image-denoising parameters and MRI sampling patterns using gradient descent or FISTA for lower-level reconstruction (Sections 2.3 and 4 of that work). This provides a concrete precedent for derivative-free outer optimization with derivative-based inner solvers. Their analysis controls inexact reconstruction error; ours propagates FW equilibrium error, including subsampled-LMO error, to generalized Goldstein stationarity. The imaging applications motivate the oracle design but are not asserted to satisfy the present CCG theorem.
- **Staudigl et al. [2026].** This paper also uses derivative-free ideas for bilevel problems in which standard gradients are difficult to obtain, and it gives complexity guarantees for stochastic bilevel inverse problems. The overlap is the black-box treatment of the outer objective and the need to handle inexact lower-level information. The model class is different: their lower level is a sample-dependent convex nonsmooth variational reconstruction problem, so the nonsmoothness and stochasticity come from Euclidean inverse-problem structure and objective-value samples. In our setting, the lower-level potential is smooth and the Wardrop equilibrium is unique, but the resulting hyper-objective is nonsmooth because the active combinatorial strategies can change as the leader parameter changes. Our guarantee is tailored to this equilibrium-switching mechanism and to the FW/ZDD oracle used to evaluate the objective.
- **Aghasi and Ghadimi [2025].** Like our outer loop, this work relies on zeroth-order information and Gaussian/randomized smoothing, and it provides non-asymptotic sample-complexity guarantees for stochastic bilevel optimization when first-order information is unavailable. The role of zeroth-order information is nevertheless different. Their method estimates first- and second-order partial derivatives of a smooth Euclidean bilevel problem using function values, while our lower-level equilibrium is not differentiated or estimated through zeroth-order partial derivatives. We solve the lower level approximately by FW, treat the resulting equilibrium computation as a black-box objective query, and prove stationarity for the Stackelberg hyper-objective with the FW equilibrium error made explicit.
- **Rando and Vaiter [2026].** This work is close on the algorithmic axis of finite-difference bilevel optimization: it studies how to avoid expensive nested finite-difference loops by using delayed information in a single-loop method and proves rates for a Euclidean value-function model. We do not improve on that loop-complexity result, and our algorithm intentionally keeps a nested structure because each outer evaluation requires an equilibrium solve. The contribution here is instead complementary: we analyze the regime in which the inner solve is a combinatorial Wardrop equilibrium computation with exact or sampled LMOs, and we show how the resulting inner approximation error affects stationarity of the true Stackelberg objective.
- **Chen et al. [2025].** This paper is the closest comparison on the nonsmooth-hyperobjective axis. It obtains approximate Clarke hyper-stationarity for optimistic/pessimistic BLO under a set-smoothness condition, and in that sense its stationarity target is stronger than the Goldstein stationarity used here. The source of nonsmoothness is different, however. Their hyper-objective is built by minimizing or maximizing over a possibly set-valued lower-level solution map. Our lower-level Wardrop equilibrium load is unique under our assumptions; nonsmoothness arises because this unique equilibrium can move across faces of a combinatorial load polytope. Our result therefore does not address set-valued minima selection, but it does cover the oracle-based equilibrium computation and its approximation error, which are not part of their framework.
- **Masiha et al. [2025].** This work is also close on nonsmooth BLO and Goldstein stationarity. It treats the non-unique-minimizer regime, where defining and optimizing the pessimistic hyper-objective requires a minima-selection mechanism; under a PL$^\circ$-type condition, the Superquantile–Gibbs relaxation gives a way to handle this selection problem. Our paper shares the use of Goldstein-type stationarity, but the lower-level pathology is different. We do not have multiple lower-level equilibria to select among; the equilibrium load is unique. The difficulty is instead that the unique response can be nonsmooth in θ due to active-set changes, and the objective can only be evaluated after running an approximate combinatorial equilibrium solver.
- **Maheshwari et al. [2022] and Drusvyatskiy et al. [2022].** These works show that derivative-free methods can be analyzed in equilibrium-style problems, respectively convex–concave minmax formulations and strongly monotone games. They overlap with us in using zeroth-order estimators to avoid direct gradient access in game-theoretic settings. They do not, however, study a bilevel

Stackelberg objective in which the follower response is itself a Wardrop equilibrium over an exponentially large combinatorial strategy family. Our analysis is therefore not a replacement for their game results; it is a bilevel stationarity result for a different equilibrium interface.

- **Liu et al. [2024].** This paper is very relevant for the outer optimization language: it develops non-asymptotic zeroth-order stationarity guarantees for constrained nonsmooth nonconvex optimization using generalized gradient-mapping and Frank–Wolfe-gap notions. If our hyper-objective were given directly by a zeroth-order stochastic oracle, that framework would be much closer. The missing piece in our setting is that each value of Φ is induced by a lower-level equilibrium computation, and the evaluation error is structured by the FW solver and its LMO. Our result can be viewed as adapting this constrained nonsmooth ZO viewpoint to a bilevel equilibrium objective with an explicit inner-solver error term.

These comparisons locate our contribution in the oracle model and error analysis: Theorem 5.6 gives an end-to-end expected approximate generalized Goldstein-stationarity guarantee for the true hyper-objective, accounting for equilibrium error from exact or subsampled FW solves. We do not claim novelty for nonsmooth solution mappings or FW equilibrium computation.

Randomized and approximate Frank–Wolfe oracles. To reduce the cost of LMO calls when $\mathcal{S}$ is very large, several works study conditional-gradient methods with randomized or approximate oracles. In particular, Kerdreux et al. [2018] analyze a subsampling-oracle Frank–Wolfe method under a uniform-inclusion model, obtaining an $\mathcal{O}(1/(\eta T))$ rate. Our lower-level analysis starts from the same conditional-gradient viewpoint but generalizes the sampling model: instead of assuming uniform inclusion, we isolate a single optimizer-hit probability κ_m that covers arbitrary non-uniform samplers and therefore permits structure-aware strategies such as UL/HL. Related variants of stochastic/online Frank–Wolfe have been studied in different settings, including projection-free stochastic optimization [Hazan and Luo, 2016, Mokhtari et al., 2020].

E Applications of Stackelberg Control in Congestion Games

The experiments in the main text use transportation-derived graphs, but the modeling and algorithmic interface are not road-specific. The common structure is a Stackelberg control problem in a nonatomic potential congestion game: the leader chooses parameters θ, followers choose combinatorial strategies $S \in \mathcal{S}$, the induced resource load is $y \in \mathcal{C} = \mathrm{conv}\{\mathbf{1}_S : S \in \mathcal{S}\}$, and the Wardrop equilibrium is the minimizer of the Beckmann potential

$$f(\theta, y) = \sum_{i \in [n]} \int_0^{y_i} c_i(u; \theta)\, du.$$

Once an application supplies the weighted minimum-cost strategy oracle

$$\arg\min_{S \in \mathcal{S}} \sum_{i \in S} g_i,$$

the same algorithm can be used provided the cost, growth, regularity, and query assumptions also hold. The oracle is the domain-specific computational component, but its availability alone does not verify these assumptions.

A reusable modeling template. For each domain, one specifies five objects. First, resources $i \in [n]$ are the congestible units, such as road links, charging stations, communication links, warehouse space–time cells, or freight corridors. Second, strategies $S \in \mathcal{S}$ are feasible combinatorial choices, such as paths, route-and-charge plans, schedules, trees, or tours. Third, the leader parameter $\theta \in \Theta$ modifies perceived resource costs through tolls, access penalties, prices, reservations, or incentives. Fourth, resource costs $c_i(y_i; \theta)$ are separable and increasing in load. For example, additive controls give costs of the form $c_i(y_i; \theta) = b_i + h_i(y_i) + \theta_i$, where h_i is increasing; if h_i' is uniformly bounded below on the relevant load range, then the Beckmann potential satisfies the quadratic-growth condition used in Section 2.3. Finally, the leader objective $F(\theta, y)$ evaluates the resulting equilibrium, e.g., total delay, congestion exposure, energy or bandwidth usage, revenue-adjusted welfare, or a regularized intervention cost.

Road pricing and evacuation routing. In road networks, resources are road segments and strategies are feasible routes. For a fixed origin–destination pair, $\mathcal{S}$ is the set of s–t paths, and the LMO is a shortest-path computation. The leader controls link tolls, capacity parameters in the costs, or soft access penalties; the equilibrium load $y^\star(\theta)$ is the user-equilibrium link-flow vector. Hard restrictions are fixed in $\mathcal{S}$ in the present model. This is the classical setting of equilibrium-based road pricing and traffic-restraint models [Yang and Bell, 1997, Dial, 1999, Hearn et al., 2001], with real deployments such as congestion charging providing practical motivation [Eliasson et al., 2009]. Our Scenario 1 is a direct instance of this tractable-oracle regime.

EV routing and charging. EV routing couples road congestion with charging-station congestion. Resources include both road links and charging-station resources; a strategy is a trip plan that specifies a route and charging choices. Using an expanded road–charging network, charging choices can be encoded by virtual arcs or station resources, so the LMO remains a shortest-path oracle when feasibility is captured by the expansion; with explicit battery constraints it becomes a resource-constrained shortest-path oracle. The leader can set charging prices, station incentives, or road tolls to steer the joint traffic–charging equilibrium. This abstraction captures the route-choice and station-pricing component of EV equilibrium models [Alizadeh et al., 2017, Bakhshayesh and Kebriaei, 2023, Zeng et al., 2024], while full power-flow constraints would require additional modeling beyond the separable CCG used here.

Multi-robot warehouse path coordination. In large warehouses, many robots repeatedly choose paths or schedules through shared corridors, intersections, workstations, and charging docks. Under a nonatomic or repeated-job approximation, a strategy is a path in a time-expanded warehouse graph, and resources are space–time cells such as (e, t) corridor traversals, (v, t) node occupancies, or dock reservations. The leader can impose priority prices, soft access penalties, charging-dock prices, or time-window incentives. The LMO is a shortest-path or constrained-schedule oracle in the time-expanded graph. This connects the framework to multi-agent path-finding and warehouse coordination models [Li et al., 2021, 2023]. The present nonatomic CCG model captures aggregate congestion effects; exact finite-agent collision avoidance is a stricter MAPF constraint and is a natural extension.

Telecommunication traffic engineering and multicast. In communication networks, resources are links or routers and strategies are paths or multicast trees. For unicast traffic, the LMO is a shortest-path computation. For multicast traffic, each strategy can be a Steiner tree connecting a source to receivers, and the LMO becomes a minimum-cost multicast-tree oracle. The leader can set per-link prices, bandwidth charges, or routing incentives to influence equilibrium traffic allocation [Korilis et al., 1997, Yaiche et al., 2000]. Multicast routing via Steiner-tree structure is standard in communication-network design [Kim et al., 2013]. This is an example where the same Stackelberg CCG objective is meaningful, but the oracle may be substantially harder than shortest path and may benefit from ZDD-based exact or sampled implementations.

Freight logistics and service tours. In freight or service operations, each follower strategy may be a tour connecting a required set of stops. On a road network, this can be modeled as a Steiner-cycle or Steiner traveling-salesman-type family [Dantzig and Ramser, 1959, Interian and Ribeiro, 2017, Salazar-Gonzalez, 2003]. Resources are road segments or curb/zone resources, and y_i is the fraction of tours using resource i. The leader can impose segment access fees, curb-use prices, or soft zone penalties to reduce congestion or externalities. The corresponding LMO is a minimum-cost tour or Steiner-cycle oracle, which is generally combinatorial and motivates compact representations such as ZDDs. Our Scenarios 2–3 are not freight models per se, but they instantiate the same exact-ZDD and subsampled-ZDD oracle regimes needed for such large strategy families.

E.1 Scope of Supported Models and Extensions

Fixed resource-selection restrictions. A fixed family $\mathcal{S}$ can exclude forbidden resources, enforce path requirements, or restrict each strategy's resource budget. The load polytope is then the convex hull of this restricted family. The analysis applies if separable costs, uniform growth and regularity, and the relevant oracle assumptions hold. Representing feasibility in $\mathcal{S}$ does not make its weighted-strategy LMO or conditional sampler efficient.

Hard aggregate capacities and moving feasible sets. The capacity parameters appearing in resource costs are not constraints on aggregate loads. Adding $Hy \leq b$ replaces $\mathcal{C}$ by $\mathcal{C} \cap \{y : Hy \leq b\}$. New extreme points can be fractional mixtures of strategy vectors, so the new LMO need not reduce to one minimum-cost strategy. A suitable capacitated LMO could preserve the potential-minimization approach on a fixed feasible set, subject to the same regularity and error bounds; otherwise a different inner method and its error analysis are needed. Parameter-dependent restrictions additionally require uniform feasibility and stability of a moving feasible set. These conclusions are conditional, not consequences for arbitrary capacity-constrained games.

Atomic followers and non-potential VIs. Atomic or integer follower choices are outside the nonatomic load formulation and need a discrete, potentially nonconvex equilibrium model and new outer analysis. For a non-potential VI there is no Beckmann objective to minimize, so Algorithm 2 and Theorem 5.4 do not apply directly. An extension needs an appropriate VI solver and a uniform norm-error bound for its equilibrium approximation, which Section 5.2 uses to control outer function-value error. On a fixed convex set, uniform strong monotonicity and suitable parameter regularity give one route to a unique Lipschitz response [Facchinei and Pang, 2003]; monotonicity alone may require an equilibrium-selection rule. Obtaining compatible exact- or subsampled-LMO error bounds remains future work.

Beyond congestion games. The proof mechanism uses a unique minimizer of a smooth convex lower objective on a fixed compact convex set, uniform quadratic growth and joint lower-gradient regularity, and a Lipschitz upper objective. Thus it can be reused outside CCGs when these conditions, lower-gradient and LMO access, projection onto a closed convex leader domain, finite initial gaps, and the outer query convention hold. A uniformly strongly convex lower objective supplies the growth condition, but does not by itself supply the other requirements or differentiability of the hyper-objective. Subsampling additionally needs a suitable atom representation and uniform optimizer-hit control. Coupled constraints, moving feasible sets, general non-potential games, and Clarke-based outer methods need further analysis.

F Additional Experimental Details & Results

F.1 Experimental Details

Hardware and software. All TNTP experiments were run on an HPC cluster node with an AMD EPYC 9334 CPU (32 cores / 64 threads), running Red Hat Enterprise Linux 9.4, and equipped with 371 GB of RAM. All methods were implemented in Python; our ZOS implementation is available at `https://github.com/sepehr78/ZOS-congestion`. We use NetworkX [Hagberg et al., 2008] for graph preprocessing and shortest-path computations. ZDD construction uses the Graphillion library [Inoue et al., 2016],[7] but we implemented our own shortest-path and sampling routines on ZDDs to support weighted sampling and to avoid unnecessary dependencies. We implemented the differentiable baseline of Sakaue and Nakamura [2021] ourselves in Python (PyTorch [Paszke et al., 2019]), since the released implementation is not in Python and was not straightforward to integrate into our pipeline.

TNTP networks and scenario construction. We start from three networks in the TNTP benchmark suite [Transportation Networks for Research Core Team, 2016] (Winnipeg, Chicago-Sketch, and Philadelphia). TNTP provides directed networks; since Graphillion represents undirected edge sets, we form an undirected graph by merging antiparallel arcs. We work on the largest connected component and then select one fixed connected subgraph (and endpoints/terminals) per benchmark by an offline randomized search aimed at realizing the three intended oracle regimes of the paper: a tractable shortest-path case, a compact-ZDD exact-oracle case, and a large-ZDD subsampling case. The resulting instances are fixed across all runs. Figure 2 visualizes the three subgraphs used in Scenarios 1–3.

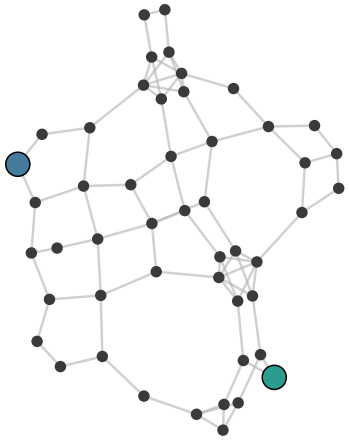
Cost model and parameters. All experiments use the fractional cost family

$$c_i(y_i; \theta_i) = d_i\Big(1 + C\,\frac{y_i}{\theta_i + 1}\Big),$$

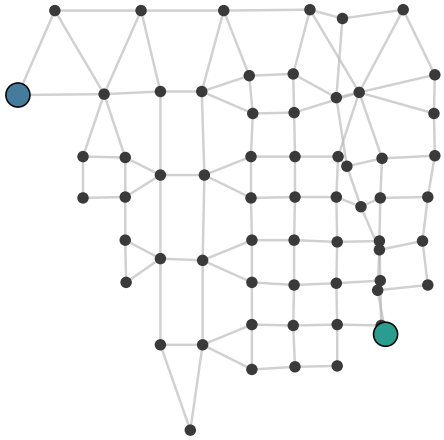
[7] `https://github.com/takemaru/graphillion`

Table 2: TNTP scenarios used in the experiments.

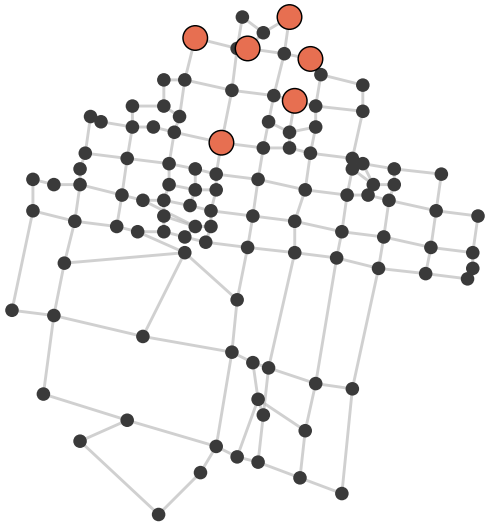
Scenario	Family	$\lvert V\rvert$	$\lvert E\rvert$	$\lvert\mathcal{S}\rvert$	ZDD nodes	ZDD build (s)
1	s–t paths	49	82	3.12×10^7	1.36×10^4	9.12 ± 2.10
2	Hamiltonian s–t paths	63	118	2.53×10^6	2.88×10^4	4.23 ± 1.32
3	Steiner cycles	110	176	3.81×10^{14}	1.08×10^6	301.65 ± 15.47



(a) Scenario 1 (Winnipeg): s–t paths (source s in green; target t in blue).



(b) Scenario 2 (Chicago-Sketch): Hamiltonian s–t paths (source s in green; target t in blue).



(c) Scenario 3 (Philadelphia): Steiner cycles (terminals highlighted).

Figure 2: TNTP-derived subgraphs used in Scenarios 1–3.

with leader feasible set $\Theta = \{\theta \in \mathbb{R}^n_+ : \mathbf{1}^\top \theta = n\}$. For Scenarios 1–2, d_i is the normalized free-flow travel time extracted from the TNTP network files (after symmetrization), scaled so that $\max_i d_i = 1$. For Scenario 3, d_i is based on Euclidean edge lengths computed from the provided node coordinates. The scaling parameter is $C = 500$ (Scenario 1), $C = 20$ (Scenario 2), and $C = 10$ (Scenario 3).

Frank–Wolfe gap. For a fixed leader parameter θ and inner iterate y_t, we report the (exact) Frank–Wolfe duality gap

$$g_t := \langle \nabla_y f(\theta, y_t),\, y_t - s_t^\star \rangle, \qquad s_t^\star \in \arg\min_{s \in \mathcal{C}} \langle \nabla_y f(\theta, y_t),\, s \rangle.$$

Since $\nabla_y f(\theta, y_t) = c(y_t; \theta)$ in our congestion model, g_t is computed by a single exact oracle call (shortest-path for Scenario 1; exact ZDD minimization for Scenarios 2–3). The reported "final FW gap" corresponds to g_T at the last inner iteration at the final outer iterate.

Algorithm details. For ZOS, the final experiments use the standard unit-sphere two-point estimator, matching Algorithm 1. For a batch of B_{out} directions, $u_j \sim \text{Unif}(\mathbb{S}^{d_\theta - 1})$, the implemented estimator is

$$\widehat{g} = \frac{1}{B_{\text{out}}} \sum_{j=1}^{B_{\text{out}}} \frac{d_\theta}{2\rho} \Big(\widehat{\Phi}(\Pi_\Theta(\theta + \rho u_j)) - \widehat{\Phi}(\Pi_\Theta(\theta - \rho u_j)) \Big) u_j.$$

Batching only averages independent unit-sphere two-point estimates and does not change the smoothing convention or the factor $d_\theta/(2\rho)$. We use smoothing radius $\rho = 0.05$ and constant outer stepsize $\eta = 0.05$. We use 4 threads to parallelize these outer-loop direction queries, and 16 threads parallelize sampling m strategies inside the inner loop.

For the differentiable baseline (Diff), we use outer gradient descent with learning rate 0.1 and the full-unroll differentiable Frank–Wolfe routine of Sakaue and Nakamura [2021]. Diff replaces the exact LMO with a softmin over the full compiled ZDD and backpropagates through the full unrolled differentiable-FW trajectory used in the run; for an inner budget T, the hypergradient retains the T softmin/FW states. There is no separate softmin temperature: softmin probabilities are proportional to $\exp(-\langle c, S \rangle)$, and the accumulated differentiable-FW cost vector in the reported full-unroll runs uses $\eta_{\text{Diff}} = 10^{-3}$. We also tested reviewer-suggested non-fixed inner updates, including the classical FW schedule $\gamma_t = 2/(t+2)$ and line-search/short-step variants; these did not materially change the final-iterate social cost or final FW gap diagnostics, so we kept the simpler fixed full-unroll Diff protocol. We report full-unroll Diff rather than a truncated 10–50-step variant because truncation made the runs cheaper but performance dropped considerably under the same final-iterate diagnostics; the plotted Diff row is therefore the strongest Diff protocol we found, not an intentionally under-converged one. Although Diff is implemented in PyTorch, in our setup running it on GPU did not yield a wall-clock speedup (and was often slower). We attribute this to the irregular, memory-bound dynamic programming passes over the ZDD (bottom-up and top-down), together with kernel-launch and host–device overheads.

Inner-loop stepsizes and reproducibility. All Wardrop equilibria are computed using T=3000 steps of Frank–Wolfe. ZOS and Diff are run for up to 500 outer iterations (Appendix F.4 gives the budgets of the other outer methods), with a wall-clock time limit of 10 hours per run; in Scenario 3, matching the ZOS budget with full-unroll Diff timed out because every Diff outer step backpropagates through the full inner FW/ZDD unroll, and Diff only completes 5 outer iterations within this limit. The Scenario 3 Diff point is therefore a feasibility-capped diagnostic, not a converged equal-budget final iterate. For methods with algorithmic randomness, we report averages over 10 runs with different random seeds; Diff is deterministic. For scenarios requiring ZDDs (Scenarios 2–3, and the Diff baseline in Scenario 1), we construct each ZDD once and cache it for reuse across runs. The reported "time taken" is the average wall-clock time per executed outer iteration, computed from the solver wall clock divided by the number of completed outer iterations. Diff uses the same CPU allocation through Torch intra-op/inter-op execution and threaded BLAS, so the ZOS speedups are not measured against a single-threaded Diff baseline.

F.2 Additional Scenario-Level Interpretation

Scenario 1: polynomial-time regime. Since ZOS treats FW as a black box, Scenario 1 simply plugs in the exact shortest-path oracle. The resulting small FW gaps show that the reported social cost is evaluated near a true follower equilibrium. The speed and memory advantage over Diff comes from avoiding backpropagation through the unrolled inner FW computation rather than from weakening the lower-level solve.

Scenario 2: compact-ZDD regime. Scenario 2 is NP-hard as a strategy family, so both methods rely on a ZDD; however, the diagram is small enough that exact LMO minimization is inexpensive. The main gap is computational: Diff backpropagates through the ZDD across all inner iterations, whereas ZOS treats the equilibrium solver as a black box and never differentiates through it. Auxiliary subsampling checks did not improve the plotted ZOS-Exact result in this compact-ZDD regime. Thus, when the ZDD is tractable (e.g., fewer than 10^5 nodes), ZOS-Exact is the recommended choice.

Scenario 3: massive-ZDD regime. This setting yields a massive ZDD, making repeated exact minimization costly and motivating subsampling. Uniform-strategy sampling (US) essentially stalls in both social cost and FW gap because the probability of hitting an exact LMO minimizer is vanishingly small (indeed, $\kappa_m \approx m/|\mathcal{S}|$ when the minimizer is unique). In contrast, UL/HL exploit the structure induced by nonnegative weights: when $g \geq 0$, adding resources cannot decrease the linear cost, so LMO minimizers are typically concentrated in smaller-cardinality strata. Appendix F.3 verifies this mechanism directly: along an exact-FW trajectory, the median candidate-set hit probability κ_{1000} is 2.62×10^{-12} for US, compared with 0.934 for UL and 0.9997 for HL. By over-sampling these strata, UL/HL dramatically increase the effective optimizer-hit probability and, once m is large enough, match or outperform ZOS-Exact in final-iterate social cost and FW gap while being much faster.

F.3 Empirical Optimizer-Hit Probabilities in Scenario 3

The subsampled-FW rate in Theorem 5.4 depends on $\kappa_m(g) = 1 - (1 - p(g))^m$, where $p(g) = q(\mathrm{Opt}(g))$ is the single-draw probability of sampling an exact LMO minimizer. To check that the length-stratified samplers used in Scenario 3 actually improve this quantity, we compute it exactly along a representative inner FW trajectory. This diagnostic runs exact ZDD-FW for 3000 steps at a fixed θ (obtained by projecting an i.i.d. uniform vector onto Θ with random seed 0), saving one FW gradient every 100 steps and yielding 30 gradients along a representative inner trajectory. For each saved gradient g we compute $\mathrm{Opt}(g)$ and $|\mathrm{Opt}(g)|$ by dynamic programming over the ZDD; combined with the exact length counts N_ℓ, this yields closed-form optimizer-hit probabilities for US, UL, and HL. We report κ_{1000} because $m = 1000$ is the largest candidate budget in the experimental sweep $m \in \{10, 100, 1000\}$; this shows that US remains negligible even at the largest plotted budget, while UL/HL already have high optimizer-hit probability.

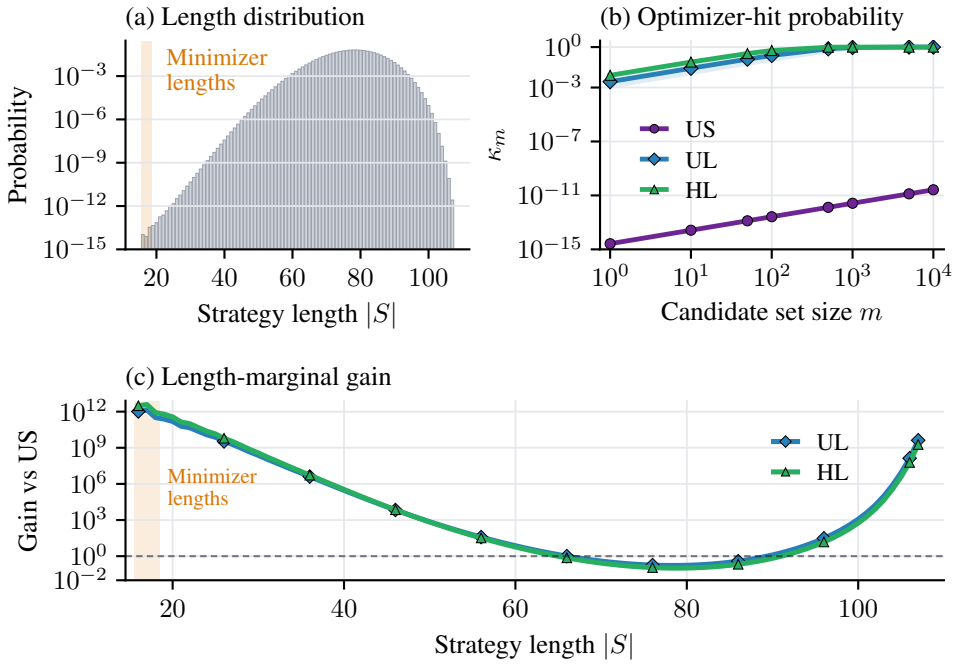


Figure 3: Empirical optimizer-hit probabilities for Scenario 3. In panel (a), the orange band marks the strategy lengths attained by exact LMO minimizers over the saved FW gradients; uniform-strategy sampling places most mass on lengths around 75–82, whereas these minimizers have length 16 or 18. Consequently, the median candidate-set hit probability κ_{1000} is 2.62×10^{-12} for US, 0.934 for UL, and 0.9997 for HL. Panel (c) shows the per-length probability gain of UL and HL relative to US; at the minimizer lengths, stratification amplifies sampling mass by roughly 10^{12}. Shaded bands in panel (b) show the 10–90% range across FW snapshots.

Table 3: Scenario 3 exact optimizer-hit summaries for the ZDD sampling schemes defined in Section 5.1 and Appendix I.3. For each saved FW gradient g, $p(g) = q(\mathrm{Opt}(g))$ is the single-draw probability of sampling an exact LMO minimizer, $\kappa_{1000}(g) = 1 - (1 - p(g))^{1000}$, and the 50% column is the smallest integer candidate-set size m with $1 - (1 - p(g))^m \geq 1/2$. Entries are medians over the 30 saved FW gradients; gain is relative to uniform-over-strategies sampling.

Sampler	Median $p(g)$	Median $\kappa_{1000}(g)$	Min. m for 50% hit	Gain vs US
Uniform-over-strategies (US)	2.62×10^{-15}	2.621×10^{-12}	2.64×10^{14}	1
Uniform-over-length (UL)	0.002 72	0.9342	255	1.04×10^{12}
Harmonic-length (HL)	0.008 07	0.9997	86	3.08×10^{12}

Figure 3 and Table 3 show that the Scenario 3 strategy family is highly imbalanced: there are 92 feasible lengths, ranging from 16 to 107, but US length mass concentrates near much longer lengths than the exact LMO minimizers. Across the 30 saved FW snapshots, the minimizer is unique and has length 16 in 23 snapshots and length 18 in 7 snapshots. Panel (c) plots relative length-marginal gains, $\Pr_{S\sim q_{\mathrm{UL}}}(|S| = \ell)/\Pr_{S\sim q_{\mathrm{US}}}(|S| = \ell)$ and $\Pr_{S\sim q_{\mathrm{HL}}}(|S| = \ell)/\Pr_{S\sim q_{\mathrm{US}}}(|S| = \ell)$, rather than the raw length priors. Thus the HL curve need not decrease monotonically even though the HL length prior is proportional to $1/\ell$: the denominator $\Pr_{S\sim q_{\mathrm{US}}}(|S| = \ell) = N_\ell/|\mathcal{S}|$ is itself highly nonuniform, with very little mass in both length tails and most mass in the combinatorial bulk. The right-tail increase in panel (c) reflects this small US denominator at very long lengths; the relevant effect for Scenario 3 is the left orange band, where the exact minimizers occur. Thus US has single-draw optimizer mass $1/|\mathcal{S}| = 2.62 \times 10^{-15}$, so even $m = 1000$ candidates give a negligible hit probability. UL and HL instead first choose a length stratum and then sample uniformly within it; because the minimizer lies in a tiny short-length stratum, this raises $p(g)$ by roughly 10^{12}, reduces the median candidate budget for a 50% exact hit from 2.64×10^{14} under US to 255 under UL and 86 under HL, and makes the effective κ_m nontrivial for the sampling budgets used in the experiments. This is exactly the regime where Proposition 5.5 is informative: $|\mathcal{S}|/(|\mathcal{L}|M_R)$ is enormous because exact LMO minimizers concentrate in two tiny short-length strata while most feasible strategies live in much larger long-length strata. The numbers above are computed at a single θ along one inner FW trajectory and so are not an additional worst-case guarantee. However, the US/UL/HL gap arises from a θ-independent length-mass mismatch—US concentrates length mass on long strategies (lengths $\approx$ 75–82) regardless of g, whereas exact minimizers occur at length 16 or 18—so the qualitative comparison is expected to persist along the outer-loop trajectories. This empirically supports structure-aware stratification when prior information indicates which strata are likely to contain LMO optimizers.

Along optimization trajectories. To check that this comparison persists away from the fixed θ above, we repeat the computation at the leader parameters of five ZOS-HL runs (m=1000) after 0, 125, 250, 375, and 499 outer iterations, at FW steps 1, 10, 100, 1000, and 3000 of an exact-LMO inner solve. This gives 125 evaluations; since the five runs share their initialization, 105 of them are distinct. We count as optimizers all strategies within relative cost 10^{-12} of the minimum; the summaries below are identical for tolerances 10^{-10} and 10^{-14}. The exact minimizers have length between 16 and 23, and length 16 in most evaluations. As Figure 4 shows, the median κ_{1000} is again 2.62×10^{-12} for US, 0.934 for UL, and 0.9997 for HL, and the median holds at every recorded outer iteration. The smallest values over all evaluations are 0.059 for UL and 0.118 for HL, both at the first FW step from the shared initialization. These evaluations come from one instance and from serially dependent iterates, so they describe our experiments rather than certify Assumption 5.2.

F.4 Outer-Method Baselines, Time to Quality, and Ablations

This subsection gives the protocols and full results for the comparisons summarized in Sections 6.1 and 6.2. All runs use the networks, cost model, and leader set of Appendix F.1.

Certified social cost. At the end of a run, each optimizer holds only an approximate follower equilibrium, whose accuracy depends on the method and the inner budget. To compare final leader parameters on equal terms, we re-solve the follower problem at each reported θ with a fully corrective FW method that uses the exact LMO, until the exact FW gap is at most 10^{-8}. We report $F(\theta, y)$ at

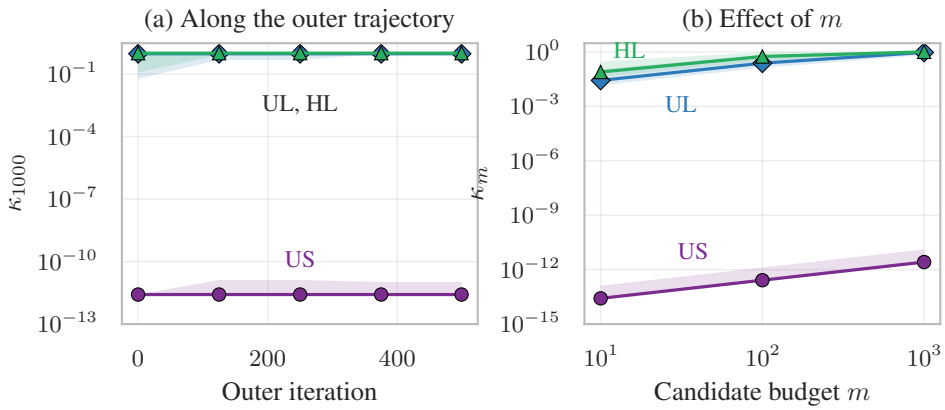


Figure 4: Optimizer-hit probabilities along five ZOS-HL runs in Scenario 3. (a) κ_{1000} at five outer iterations; lines are medians and bands the range over 25 evaluations (5 runs $\times$ 5 FW steps). (b) κ_m against m; bands are 10–90% quantiles over all 125 evaluations.

the re-solved load y as the *certified* social cost. This re-solve is a post-hoc diagnostic: it is not part of any optimizer, and its time is excluded from all reported times.

Baselines. All baselines use the FW solver and LMO of ZOS, with $T = 3000$ steps per full follower solve.

- **SPSA-FW** applies SPSA [Spall, 1992] to $\widehat{\Phi}_T$. At update t, it draws a Rademacher direction $u_t \in \{\pm 1\}^{d_\theta}$ and sets $\theta_{t+1} = \Pi_\Theta(\theta_t - \eta_t \widehat{g}_t)$, where

$$\widehat{g}_t = \frac{1}{2\rho_t}\Big(\widehat{\Phi}_T\big(\Pi_\Theta(\theta_t + \rho_t u_t)\big) - \widehat{\Phi}_T\big(\Pi_\Theta(\theta_t - \rho_t u_t)\big)\Big)u_t,$$

  with the standard SPSA gain sequences $\eta_t = \eta_0/(t+11)^{0.602}$ and $\rho_t = \rho_0/(t+1)^{0.101}$. We use $(\eta_0, \rho_0) = (0.05, 0.05)$ in Scenario 1 and $(0.1, 0.02)$ in Scenarios 2–3, where we report the running average of the iterates.
- **ValueGap-ZO** and **ValueGap-CG** follow the value-function penalty approach of Shen and Chen [2023], which replaces the equilibrium constraint by the penalty

$$L_\gamma(\theta, y) := F(\theta, y) + \gamma\big[f(\theta, y) - f(\theta, \bar{y})\big],$$

  where $\bar{y} \approx y^\star(\theta)$ is a full follower solve. Each update takes five FW steps on $y \mapsto F(\theta, y) + \gamma f(\theta, y)$ over $\mathcal{C}$, warm-started from the previous y, and then updates θ with y fixed; unlike ZOS, these FW steps require $\nabla_y F$. ValueGap-CG takes a projected step along $\nabla_\theta L_\gamma(\theta, y)$ with $\bar{y}$ also fixed, which requires $\nabla_\theta F$ and $\nabla_\theta f$. ValueGap-ZO replaces this derivative by the average of four two-point differences of L_γ along Rademacher directions, re-solving $\bar{y}$ at each perturbed point. The suffix CG stands for the conditional-gradient (FW) updates of y; both variants use them and differ only in the update of θ. Both methods then re-solve $\bar{y}$ at the new θ and use $\gamma = 1$. ValueGap-ZO uses $\rho = 0.05$, with $\eta = 0.02$ in Scenarios 1–2 and $\eta = 0.05$ in Scenario 3, and ValueGap-CG uses $\eta = 0.05$. Along the trajectory, we record $F(\theta, \bar{y})$ for both methods.

ZOS uses the settings of Appendix F.1, with $B_{\text{out}} = 4$ directions and $K = 500$ outer iterations. The hyperparameters of SPSA-FW and ValueGap-ZO were selected on pilot runs with seeds disjoint from the evaluation seeds. Each method is run with 10 seeds in Scenarios 1–2 and 5 seeds in Scenario 3, where all methods use the HL sampler with $m = 1000$. ValueGap-CG is deterministic in Scenarios 1–2 and is run once there.

Matched oracle budget. We match the total number of LMO operations: exact LMO calls in Scenarios 1–2 and scored sampled strategies in Scenario 3. Each ZOS run makes $1+(2B_{\text{out}}+1)K = 4{,}501$ follower solves: one at θ_0 and, in each outer iteration, $2B_{\text{out}}$ solves at the perturbed points and one at the new iterate, which provides the recorded cost and the warm start for the next iteration. In Scenarios 1–2, this amounts to about 1.35×10^7 exact LMO calls. SPSA-FW makes three follower

solves per update and uses 1,500 updates in Scenario 1; in Scenarios 2–3, it makes a fourth solve at the averaged iterate and uses 1,125 updates. In both cases, it makes exactly 4,501 follower solves. Each ValueGap-ZO update makes nine full follower solves (eight at perturbed points and one at the new iterate) and one five-step solve, and each ValueGap-CG update makes one of each. With $K = 500$ and $K = 4{,}493$ updates, respectively, both methods exceed the ZOS LMO budget by less than 0.02%. Since a five-step solve uses few LMO calls, the ValueGap methods make more solver calls than ZOS for the same LMO budget (Table 4); truncating them at 4,501 solver calls leaves the ordering of the four methods unchanged in every scenario.

Results. Table 4 reports certified final costs and optimizer times, and Figure 5 shows the social cost along each trajectory against the LMO work spent. ZOS attains the lowest certified cost in Scenarios 1–2, improving on SPSA-FW by 15.1 (99% bootstrap CI $[14.3, 15.8]$) in Scenario 1 and on the runner-up, ValueGap-CG, by 2.1 in Scenario 2. In Scenario 3, ZOS attains a lower mean cost than SPSA-FW and ValueGap-ZO, and ValueGap-CG attains the lowest cost using $\nabla_\theta F$ and $\nabla_\theta f$, which the oracle model of ZOS does not assume. Optimizer times are solver wall-clock times, excluding ZDD construction and certification.

Table 4: Outer methods that share the FW solver and LMO of ZOS under a matched LMO budget: certified final social cost (mean ± 99% CI), mean optimizer wall-clock time in minutes, and follower solves per run, counting each five-step ValueGap solve as one. ValueGap-CG is deterministic in Scenarios 1–2 and is reported from one run.

	Scenario 1		Scenario 2		Scenario 3 (HL, m=1000)		
Method	Cost	Time	Cost	Time	Cost	Time	Solves
ZOS	**93.01** ± 0.24	49.2	**143.83** ± 0.11	13.4	29.40 ± 0.13	37.3	4,501
SPSA-FW	108.08 ± 0.99	34.9	150.69 ± 0.81	33.1	45.46 ± 1.19	70.0	4,501
ValueGap-ZO	128.56 ± 2.39	59.5	148.59 ± 0.81	13.2	29.77 ± 0.33	169.8	5,001
ValueGap-CG	116.33	32.2	145.96	32.9	**24.97** ± <0.01	52.5	8,987

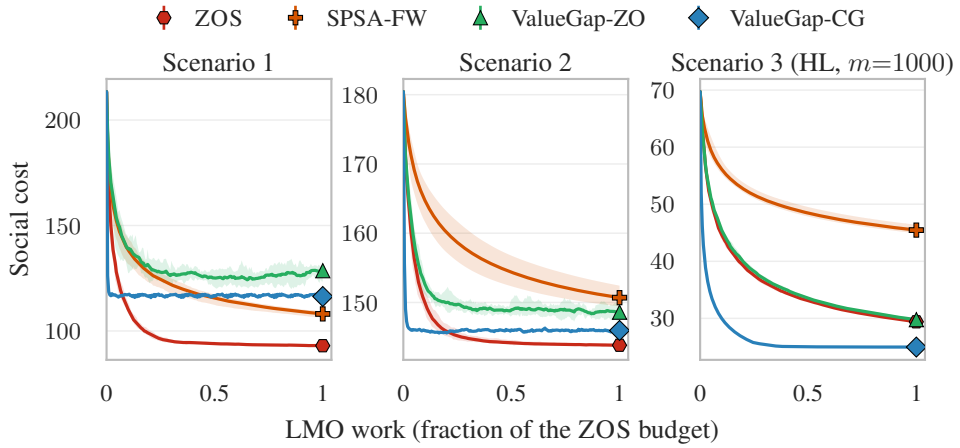


Figure 5: Social cost along the optimizer trajectory versus cumulative LMO work, normalized by the ZOS budget (exact LMO calls in Scenarios 1–2; scored sampled strategies in Scenario 3). Lines are means over runs and shaded bands span the minimum to the maximum; along the trajectory, the cost is evaluated at each method's own follower iterate. End markers are certified final costs with 99% CIs.

Total time to quality. The per-iteration times in Figure 1 do not account for the number of outer iterations each method needs, so we also measure the total wall-clock time to reach a fixed solution quality. For each scenario, let C_0 be the common initial cost and $\bar{C}$ the mean final certified cost of ZOS in Table 4; the target at level $p \in \{0.5, 0.75, 0.9\}$ is $C_0 - p\,(C_0 - \bar{C})$. The time to quality is the solver wall-clock time from initialization, including the initial follower solve, until the first certified

iterate at or below the target; one-time setup and certification are excluded. For this comparison, we reran ZOS with certified checkpoints at a fixed grid of outer iterations that is denser early $(0, 1, 2, 5, 10, \ldots, 500)$, using 10 runs in Scenarios 1–2 and 5 runs in Scenario 3. Since we do not interpolate between checkpoints, this spacing can only overstate the time of ZOS. Since every checkpoint is re-solved to an exact FW gap of at most 10^{-8}, each plotted cost is evaluated at an accurate follower equilibrium, so we do not plot the FW gap separately. We reran Diff with the protocol of Appendix F.1 in Scenarios 1–2, certifying the iterates on either side of each target crossing. In Scenario 3, we use the original Diff run, which completed five outer iterations within the 10-hour limit. Diff is deterministic, so we report one run per scenario.

Table 5 and Figure 6 show that ZOS reaches every target 30–238$\times$ faster than Diff in Scenarios 1–2. In Scenario 3, none of the five certified Diff iterates reaches the 50% target, whereas ZOS reaches the 90% target after 103 minutes on average. Every ZOS run in Scenarios 1–2 also reaches each target in fewer outer iterations than Diff (e.g., at most 75 versus 157 for the 90% target in Scenario 1), so the per-iteration speedups in Figure 1 understate the speedups in total time.

Table 5: Total wall-clock time in minutes until the certified cost first reaches the target at level p, i.e., a fraction p of the mean ZOS cost reduction. ZOS: mean and 99% bootstrap CI over runs; Diff: one deterministic run; NR: not reached within the 10-hour limit.

	Scenario 1		Scenario 2		Scenario 3 (HL, m=1000)	
p	ZOS	Diff	ZOS	Diff	ZOS	Diff
50%	1.05 [0.81, 1.30]	84.3	0.45 [0.41, 0.51]	108.1	18.2 [17.2, 20.4]	NR
75%	3.04 [2.50, 3.73]	154.0	1.03 [0.94, 1.12]	209.2	41.6 [34.0, 57.8]	NR
90%	7.21 [6.34, 8.01]	219.4	2.05 [1.92, 2.26]	316.2	103.3 [97.1, 116.8]	NR

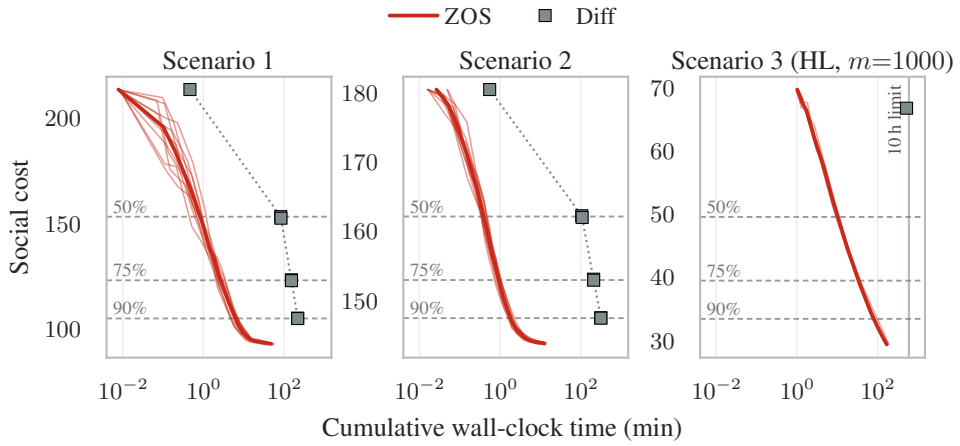


Figure 6: Certified social cost versus cumulative wall-clock time. Thin lines are individual ZOS runs and the thick line is their mean; Diff markers are certified iterates of its single deterministic run. Dashed lines mark the 50%, 75%, and 90% targets. In Scenario 3, only the last of the five Diff iterates has a recorded time, and the vertical line marks the 10-hour limit.

Ablations. We vary one ZOS parameter at a time from the defaults $T = 3000$, $\rho = 0.05$, and exact LMO, keeping the other settings and the seeds fixed so that runs are paired across values. Table 6 reports $T \in \{500, 1000, 3000\}$ and $\rho \in \{0.025, 0.05, 0.1\}$ over 10 runs in Scenarios 1–2, together with the FW gap of the optimizer's final follower iterate before re-solving. Larger T reduces this FW gap, while the mean certified cost changes by less than 0.002 across the tested T in both scenarios. Over the fourfold range of ρ, the certified cost varies by less than 0.5% in Scenario 1 and 0.2% in Scenario 2. This indicates robustness over the tested range, not that $\rho = 0.05$ is optimal; $\rho = 0.1$ gives a slightly lower cost in both scenarios.

Table 6: One-factor ablations of the inner budget T and smoothing radius ρ of ZOS with the exact LMO (mean over 10 paired runs). FW gap: exact FW gap of the optimizer's final follower iterate before re-solving.

		T			ρ		
		500	1000	3000	0.025	0.05	0.1
Scenario 1	Certified cost	93.009	93.009	93.008	93.137	93.008	92.753
	FW gap	0.358	0.177	0.059	0.059	0.059	0.057
Scenario 2	Certified cost	143.827	143.827	143.827	143.953	143.827	143.702
	FW gap	0.095	0.044	0.017	0.014	0.017	0.016

Table 7 varies the candidate budget m of a uniformly subsampled LMO in Scenario 2. All $2{,}526{,}940$ strategies in this scenario contain 62 edges, so UL and HL coincide with uniform sampling, and the only available guarantee is $\kappa_m \geq 1 - (1 - 1/|\mathcal{S}|)^m$. Scenario 2 therefore isolates the effect of m without any gain from length stratification. We report the five runs per m underlying Figure 1 and five additional runs with new seeds as separate sets, since the additional runs were executed in a separate batch and give higher subsampled costs at every m for reasons we could not isolate. In both sets, the certified cost decreases monotonically toward the exact-LMO value as m grows; at $m = 10$, it remains about 2.0 above this value in the original runs and 2.7 in the additional runs. This improvement is larger than κ_m alone suggests: the bound counts only exact optimizer hits, whereas near-optimal candidates and fresh samples at each of the T FW steps also reduce the gap.

Table 7: Candidate budget m of the uniformly subsampled LMO in Scenario 2 (mean certified cost over 5 runs per set). Original: the runs underlying Figure 1; Additional: runs with new seeds. κ_m bound: $1 - (1 - 1/|\mathcal{S}|)^m$. FW gap: exact FW gap of the optimizer's final follower iterate before re-solving, for the original runs.

		Certified cost		
LMO	κ_m bound	Original	Additional	FW gap
$m = 10$	3.96×10^{-6}	145.85	146.54	4.54
$m = 100$	3.96×10^{-5}	144.60	146.10	2.27
$m = 1000$	3.96×10^{-4}	143.94	145.63	0.66
Exact	1	143.82	143.84	0.015

G Proofs for Section 2

G.1 Nonsmoothness of the Social-Cost Hyper-Objective

We prove Proposition 2.7 by evaluating the social cost at equilibrium and showing that its one-sided derivatives differ. Nonsmoothness of the equilibrium map alone would not establish this conclusion.

Proof of Proposition 2.7.

Proof. Step 1: verify the assumptions. Both resource costs are continuous and strictly increasing in their load variables, so Assumption 2.1 holds. Since $g \in C^1([0,1])$, its derivative is bounded and g is Lipschitz. Thus the lower gradient $\nabla_y f(\theta, y) = (y_1, y_2 - 1 + g(\theta))$ is jointly Lipschitz, and each cost is continuously differentiable in θ. The Hessian $\nabla^2_{yy} f$ is the identity matrix, giving uniform 1-strong convexity and global quadratic growth with $\alpha = 1$. Moreover, $F_{\mathrm{sc}}(\theta, y) = y_1^2 + y_2(y_2 - 1 + g(\theta))$ is continuously differentiable with bounded derivatives on the compact convex set $\Theta \times \mathcal{C}$, hence Lipschitz. This verifies Assumptions 2.4 and 2.5.

Step 2: compute the hyper-objective. Substituting $y_2 = 1 - y_1$ into the Beckmann potential gives

$$f(\theta, y_1, 1 - y_1) = y_1^2 - g(\theta) y_1 + g(\theta) - \frac{1}{2}, \qquad y_1 \in [0,1].$$

Its unconstrained minimizer is $g(\theta)/2 \le 1/2$ because $\|g\|_\infty \le 1$. Therefore

$$y_1^\star(\theta) = \max\{0, g(\theta)/2\}, \qquad y_2^\star(\theta) = 1 - y_1^\star(\theta). \tag{4}$$

If $g(\theta) \le 0$, all flow uses link 2, whose cost at load 1 is $g(\theta)$. If $g(\theta) > 0$, both links are used and each has cost $g(\theta)/2$. Since total flow is one, the social-cost hyper-objective is

$$\Phi_{\mathrm{sc}}(\theta) = F_{\mathrm{sc}}(\theta, y^\star(\theta)) = \begin{cases} g(\theta), & g(\theta) \le 0, \\ g(\theta)/2, & g(\theta) > 0. \end{cases} \tag{5}$$

Step 3: prove nondifferentiability of Φ_{sc}. Let $\bar\theta \in (0,1)$ satisfy $g(\bar\theta) = 0$ and set $a := g'(\bar\theta) \ne 0$. The expansion $g(\bar\theta + h) = ah + o(h)$ shows that g has opposite signs on the two sides of $\bar\theta$. Consequently, (5) gives the left and right derivatives

$$\big(\Phi'_{\mathrm{sc},-}(\bar\theta), \Phi'_{\mathrm{sc},+}(\bar\theta)\big) = \begin{cases} (a, a/2), & a > 0, \\ (a/2, a), & a < 0. \end{cases}$$

These derivatives differ because $a \ne 0$, proving the claimed nondifferentiability of the hyper-objective. The same sign change moves the equilibrium between the boundary face $y_1 = 0$ and the relative interior of $\mathcal{C}$; the explicit social-cost calculation above shows that this switch produces a kink in Φ_{sc}. If g has infinitely many interior simple zeros, the argument applies at each of them. □

Remark G.1 (A concrete oscillatory member of the class)**.** *Let*

$$g(\theta) = \begin{cases} \theta^4 \sin(1/\theta^2), & \theta > 0, \\ 0, & \theta = 0. \end{cases}$$

Here $|g(\theta)| \le \theta^4 \le 1$, $g'(0) = 0$, *and* $g'(\theta) = 4\theta^3 \sin(1/\theta^2) - 2\theta\cos(1/\theta^2)$ *tends to zero as* $\theta \downarrow 0$. *Thus* $g \in \mathcal{G}$. *Its zeros* $\theta_n = 1/\sqrt{n\pi}$, $n = 1, 2, \ldots$, *lie in* $(0,1)$ *and satisfy*

$$g'(\theta_n) = -2\theta_n(-1)^n \ne 0.$$

Proposition 2.7 therefore gives infinitely many nondifferentiable points of Φ_{sc}.

Additional example: kinks in the equilibrium map. The following example concerns $y^\star$ rather than a specified upper objective. It shows that the number of equilibrium-map kinks can grow with $|\mathcal{S}|$; it is separate from the proof of Proposition 2.7.

Example G.2 (Many kinks scaling with the number of strategies)**.** *Fix* $n \ge 3$ *and consider a parallel network with* n *resources and unit demand. The strategy set is* $\mathcal{S} = \{\{i\} : i \in [n]\}$ *(so* $|\mathcal{S}| = n$*) and the feasible load set is*

$$\mathcal{C} = \Big\{ y \in \mathbb{R}^n_{\ge 0} : \sum_{i=1}^n y_i = 1 \Big\}.$$

Let $M > 2$ *and define smooth costs*

$$c_i(y_i; \theta) = y_i + M\,\phi_i(\theta), \qquad \phi_i(\theta) := -i\theta + \frac{i(i-1)}{2}, \qquad i \in [n],$$

where $\theta \in \mathbb{R}$ *is a scalar parameter. The Beckmann potential is*

$$f(\theta, y) = \sum_{i=1}^n \Big(\tfrac{1}{2} y_i^2 + M\,\phi_i(\theta)\, y_i \Big), \qquad y \in \mathcal{C}.$$

For each θ, *the Wardrop equilibrium* $y^\star(\theta)$ *is the unique minimizer of* $\min_{y \in \mathcal{C}} f(\theta, y)$ *(uniqueness holds since* $f(\theta, \cdot)$ *is* 1*-strongly convex).*

KKT conditions (necessary and sufficient here) imply that there exists a scalar $\tau(\theta)$ *such that*

$$y_i^\star(\theta) = \big[\tau(\theta) - M\phi_i(\theta)\big]_+, \qquad \sum_{i=1}^n \big[\tau(\theta) - M\phi_i(\theta)\big]_+ = 1,$$

where $[t]_+ = \max\{t, 0\}$. *Hence* $y^\star(\theta)$ *is piecewise affine in* θ, *and kinks occur when some coordinate hits* 0 *(i.e., the active set changes).*

Which ϕ_i is smallest. *For $\theta \in [i-1, i]$, the function $\phi_i(\theta)$ is smaller than both neighbors:*

$$\phi_{i-1}(\theta) - \phi_i(\theta) = \theta - (i-1) \geq 0, \qquad \phi_{i+1}(\theta) - \phi_i(\theta) = i - \theta \geq 0.$$

By chaining these inequalities, we get $\phi_i(\theta) \leq \phi_j(\theta)$ for all *$j \in [n]$ whenever $\theta \in [i-1, i]$. So as θ increases from 0 to $n-1$, the smallest $\phi_i(\theta)$ moves from strategy 1 to 2 to $\cdots$ to n.*

Equilibrium uses only strategies within $1/M$ of the minimum. *If two strategies i and j are both used (i.e., $y_i^\star(\theta), y_j^\star(\theta) > 0$), Wardrop conditions give equal costs:*

$$y_i^\star(\theta) + M\phi_i(\theta) = y_j^\star(\theta) + M\phi_j(\theta) \quad \Longrightarrow \quad y_i^\star(\theta) - y_j^\star(\theta) = M\big(\phi_j(\theta) - \phi_i(\theta)\big).$$

Since $y_i^\star, y_j^\star \in [0,1]$, the left-hand side lies in $[-1,1]$, hence any used j must satisfy

$$\big|\phi_j(\theta) - \phi_i(\theta)\big| \leq \frac{1}{M}.$$

In other words, only strategies whose offsets are within $1/M$ of the minimum can be active.

An explicit "switching window" around each integer. *Fix $i \in \{1, \ldots, n-2\}$. Around $\theta = i$, the two closest offsets are ϕ_i and ϕ_{i+1}, and*

$$\phi_{i+1}(\theta) - \phi_i(\theta) = i - \theta.$$

For $\theta \in [i - 1/M,\ i + 1/M]$, this difference has magnitude at most $1/M$, so only strategies i and $i+1$ can be active. Solving Wardrop with $y_{i+1} = 1 - y_i$ gives

$$y_i^\star(\theta) = \frac{1}{2} + \frac{M}{2}(i - \theta), \qquad y_{i+1}^\star(\theta) = \frac{1}{2} - \frac{M}{2}(i - \theta), \qquad \theta \in \big[i - \tfrac{1}{M},\ i + \tfrac{1}{M}\big],$$

and all other coordinates are 0. Outside this window, the equilibrium is pure:

$$y^\star(\theta) = e_i \text{ for } \theta \in [i - 1 + \tfrac{1}{M},\ i - \tfrac{1}{M}], \qquad y^\star(\theta) = e_{i+1} \text{ for } \theta \in [i + \tfrac{1}{M},\ i + 1 - \tfrac{1}{M}],$$

where e_i is the ith standard basis vector in $\mathbb{R}^n$.

Thus $y^\star(\theta)$ is continuous *and piecewise linear, but it has kinks at the two endpoints $\theta = i \pm \frac{1}{M}$ for each $i = 1, \ldots, n-2$. Therefore, the equilibrium map has at least $2(n-2) = \Omega(|\mathcal{S}|)$ nondifferentiable points on $[0, n-1]$.*

G.2 Proof of the Equilibrium–Potential Proposition

Proof. We first recall the first-order optimality condition for $\min_{y \in \mathcal{C}} f(y)$. Since f is differentiable and convex, a point $\bar{y} \in \mathcal{C}$ minimizes f over $\mathcal{C}$ if and only if

$$\langle \nabla f(\bar{y}),\, y - \bar{y} \rangle \ \geq\ 0 \qquad \text{for all } y \in \mathcal{C}. \tag{6}$$

(Equilibrium $\Rightarrow$ minimizer). Let $z \in \Delta^d$ be a Wardrop equilibrium and $y = y(z)$. By definition of f we have $\nabla f(y)_i = c_i(y_i)$, so for any $y' \in \mathcal{C}$,

$$\langle \nabla f(y),\, y' - y \rangle = \sum_{i \in [n]} c_i(y_i)\,(y_i' - y_i).$$

Every $y' \in \mathcal{C}$ can be written as $y' = \sum_{S \in \mathcal{S}} \xi_S\, \mathbf{1}_S$ for some $\xi \in \Delta^d$. Using $y = \sum_S z_S \mathbf{1}_S$ we obtain

$$\begin{aligned}
\langle \nabla f(y),\, y' - y \rangle &= \sum_{i \in [n]} c_i(y_i) \Big(\sum_{S \in \mathcal{S}} \xi_S [\mathbf{1}_S]_i - \sum_{S \in \mathcal{S}} z_S [\mathbf{1}_S]_i \Big) \\
&= \sum_{S \in \mathcal{S}} \xi_S \sum_{i \in S} c_i(y_i) \ -\ \sum_{S \in \mathcal{S}} z_S \sum_{i \in S} c_i(y_i) \\
&= \sum_{S \in \mathcal{S}} \xi_S\, c_S(y) \ -\ \sum_{S \in \mathcal{S}} z_S\, c_S(y).
\end{aligned}$$

Let $C_{\min} := \min_{S \in \mathcal{S}} c_S(y)$. By the Wardrop condition, $c_S(y) = C_{\min}$ whenever $z_S > 0$, and therefore $\sum_S z_S c_S(y) = C_{\min}$. On the other hand, $\sum_S \xi_S c_S(y) \geq C_{\min}$ for any $\xi \in \Delta^d$. Hence

$$\langle \nabla f(y),\, y' - y \rangle \ =\ \sum_S \xi_S c_S(y) - C_{\min} \ \geq\ 0 \qquad \text{for all } y' \in \mathcal{C}.$$

Thus y satisfies (6) and so y minimizes f over $\mathcal{C}$. By uniqueness of the minimizer, $y = y^\star$.

(Minimizer ⇒ equilibrium). Conversely, let $y \in \mathcal{C}$ minimize f over $\mathcal{C}$, and let $z \in \Delta^d$ be any population profile with $y = y(z)$. Assume for contradiction that z is not a Wardrop equilibrium. Then there exist strategies $S, S' \in \mathcal{S}$ such that $z_S > 0$ and $c_{S'}(y) < c_S(y)$.

For sufficiently small $\varepsilon > 0$ we can define a perturbed profile $z^\varepsilon \in \Delta^d$ by moving mass ε from S to S':

$$z^\varepsilon_S = z_S - \varepsilon, \qquad z^\varepsilon_{S'} = z_{S'} + \varepsilon, \qquad z^\varepsilon_T = z_T \text{ for all } T \notin \{S, S'\},$$

and let $y^\varepsilon := y(z^\varepsilon) \in \mathcal{C}$ be the induced load. Then

$$y^\varepsilon - y = \varepsilon(\mathbf{1}_{S'} - \mathbf{1}_S).$$

Using differentiability of f we compute the directional derivative of f at y in the direction $y^\varepsilon - y$:

$$\begin{aligned} \frac{d}{d\varepsilon} f(y^\varepsilon)\Big|_{\varepsilon=0} &= \langle \nabla f(y),\, \mathbf{1}_{S'} - \mathbf{1}_S \rangle \\ &= \sum_{i \in S'} c_i(y_i) - \sum_{i \in S} c_i(y_i) \;=\; c_{S'}(y) - c_S(y) \;<\; 0. \end{aligned}$$

Therefore, for all sufficiently small $\varepsilon > 0$ we have $f(y^\varepsilon) < f(y)$, contradicting the fact that y minimizes f over $\mathcal{C}$. Hence no such pair (S, S') can exist, and z must be a Wardrop equilibrium.

Combining the two directions, we obtain the claimed equivalence. □

G.3 Proof of the Lipschitz Stability Lemma

Proof. **Step 1 (a uniform local Lipschitz step).** If $L_{f,2} = 0$, the first-order condition for $y^\star(\theta')$ is also the first-order condition for the problem at θ, so uniqueness gives $y^\star(\theta) = y^\star(\theta')$. We therefore assume $L_{f,2} > 0$. Let $\delta := \alpha r/(4L_{f,2})$. We first show that whenever $\|\theta - \theta'\| \le \delta$,

$$\|y^\star(\theta) - y^\star(\theta')\| \le \frac{2L_{f,2}}{\alpha} \|\theta - \theta'\|. \tag{7}$$

Fix such θ, θ' and denote $y_\theta := y^\star(\theta)$, $y_{\theta'} := y^\star(\theta')$, and $s := \|y_{\theta'} - y_\theta\|$. Since $y_{\theta'}$ minimizes $f(\theta', \cdot)$ over $\mathcal{C}$, the first-order optimality condition can be written as

$$0 \in \nabla_y f(\theta', y_{\theta'}) + N_{\mathcal{C}}(y_{\theta'}), \tag{8}$$

where $N_{\mathcal{C}}(y_{\theta'}) := \{v \in \mathbb{R}^n : \langle v, u - y_{\theta'} \rangle \le 0 \ \forall u \in \mathcal{C}\}$ is the normal cone of $\mathcal{C}$ at $y_{\theta'}$. Equivalently, $-\nabla_y f(\theta', y_{\theta'}) \in N_{\mathcal{C}}(y_{\theta'})$.

Consider the stationarity residual of $y_{\theta'}$ for the problem at parameter θ:

$$\mathrm{res} := \mathrm{dist}\big(0, \nabla_y f(\theta, y_{\theta'}) + N_{\mathcal{C}}(y_{\theta'})\big) = \mathrm{dist}\big(-\nabla_y f(\theta, y_{\theta'}), N_{\mathcal{C}}(y_{\theta'})\big).$$

Using $-\nabla_y f(\theta', y_{\theta'}) \in N_{\mathcal{C}}(y_{\theta'})$ and Assumption 2.5 (with $y = y_{\theta'}$) gives

$$\mathrm{res} \le \big\|\nabla_y f(\theta, y_{\theta'}) - \nabla_y f(\theta', y_{\theta'})\big\| \le L_{f,2}\|\theta - \theta'\|.$$

Choose $v \in N_{\mathcal{C}}(y_{\theta'})$ such that $\mathrm{res} = \|\nabla_y f(\theta, y_{\theta'}) + v\|$. By convexity of $y \mapsto f(\theta, y)$ and the normal-cone inequality $\langle v, y_\theta - y_{\theta'} \rangle \le 0$, we have

$$f(\theta, y_{\theta'}) - f(\theta, y_\theta) \le \langle \nabla_y f(\theta, y_{\theta'}),\, y_{\theta'} - y_\theta \rangle \le \mathrm{res}\, s.$$

On the other hand, the global consequence of local quadratic growth in Lemma H.1, applied to $f(\theta, \cdot)$, gives

$$\frac{\alpha}{2} \min\{s^2, rs\} \le f(\theta, y_{\theta'}) - f(\theta, y_\theta) \le \mathrm{res}\, s.$$

If $s = 0$, (7) is immediate. Otherwise, since $\|\theta - \theta'\| \le \delta$, we have $\mathrm{res} \le \alpha r/4$. The case $s > r$ would imply $(\alpha/2)rs \le \mathrm{res}\, s \le (\alpha/4)rs$, a contradiction. Hence $s \le r$, and the quadratic branch applies:

$$\frac{\alpha}{2} s^2 \le \mathrm{res}\, s.$$

Dividing by $s > 0$ and using the residual bound yields

$$s \le \frac{2\mathrm{res}}{\alpha} \le \frac{2L_{f,2}}{\alpha} \|\theta - \theta'\|,$$

which proves (7).

Step 2 (global Lipschitzness by chaining). Fix arbitrary $\theta, \theta' \in \Theta$. If $\theta = \theta'$ the claim is trivial. Otherwise set

$$m := \left\lceil \frac{\|\theta - \theta'\|}{\delta} \right\rceil \quad \text{and} \quad \theta_i := \theta + \frac{i}{m}(\theta' - \theta), \;\; i = 0, 1, \ldots, m.$$

Since Θ is convex, $\theta_i \in \Theta$ for all i, and $\|\theta_{i+1} - \theta_i\| = \|\theta' - \theta\|/m \le \delta$. Applying the local step (7) to each consecutive pair and summing gives

$$\|y^\star(\theta) - y^\star(\theta')\| \le \sum_{i=0}^{m-1} \|y^\star(\theta_{i+1}) - y^\star(\theta_i)\| \le \frac{2L_{f,2}}{\alpha} \sum_{i=0}^{m-1} \|\theta_{i+1} - \theta_i\| = \frac{2L_{f,2}}{\alpha}\|\theta - \theta'\|,$$

which proves the equilibrium-map Lipschitz bound in Lemma 2.6.

Step 3 (Lipschitzness of Φ). Using Assumption 2.5 for F,

$$|\Phi(\theta) - \Phi(\theta')| = |F(\theta, y^\star(\theta)) - F(\theta', y^\star(\theta'))| \le L_{F,1}\|\theta - \theta'\| + L_{F,1}\|y^\star(\theta) - y^\star(\theta')\|.$$

Plugging the equilibrium-map bound into the last term gives

$$|\Phi(\theta) - \Phi(\theta')| \le L_{F,1}\Big(1 + \frac{2L_{f,2}}{\alpha}\Big)\|\theta - \theta'\|,$$

which proves the claimed Lipschitz bound for the hyper-objective. □

H Proofs for Section 5

H.1 Auxiliary Lemmas

Throughout this appendix we fix a parameter $\theta \in \Theta$ and write $f(y) := f(\theta, y)$, $y^\star := y^\star(\theta) \in \arg\min_{y \in \mathcal{C}} f(y)$. Let $D := \max\{\|u - v\| : u, v \in \mathcal{C}\}$.

Lemma H.1 (From local quadratic growth to a global error bound)**.** *Assume Assumption 2.4: there exist $\alpha > 0$ and $r > 0$ such that for all $y \in \mathcal{C}$ with $\|y - y^\star\| \le r$,*

$$f(y) - f(y^\star) \ge \frac{\alpha}{2}\|y - y^\star\|^2.$$

If f is convex on $\mathcal{C}$, then for every $y \in \mathcal{C}$,

$$f(y) - f(y^\star) \;\ge\; \frac{\alpha}{2}\min\{\|y - y^\star\|^2,\; r\|y - y^\star\|\}.$$

In particular, if $f(y) - f(y^\star) \le \frac{\alpha}{2}r^2$, then $\|y - y^\star\| \le r$ and $\|y - y^\star\|^2 \le \frac{2}{\alpha}(f(y) - f(y^\star))$.

Proof. If $\|y - y^\star\| \le r$ the claim is exactly Assumption 2.4. Otherwise let $d := \|y - y^\star\| > r$ and define the point on the segment $y_r := y^\star + \frac{r}{d}(y - y^\star)$ so that $\|y_r - y^\star\| = r$. By convexity,

$$f(y_r) \le \Big(1 - \frac{r}{d}\Big)f(y^\star) + \frac{r}{d}f(y) \quad \Rightarrow \quad f(y) - f(y^\star) \ge \frac{d}{r}\big(f(y_r) - f(y^\star)\big).$$

Applying Assumption 2.4 to y_r yields $f(y_r) - f(y^\star) \ge \frac{\alpha}{2}r^2$, hence $f(y) - f(y^\star) \ge \frac{d}{r} \cdot \frac{\alpha}{2}r^2 = \frac{\alpha r}{2}d$. This gives the stated $\min\{\cdot, \cdot\}$ lower bound and the corollaries. □

H.2 Length-Stratified Sampling and Optimizer-Hit Probability

We first prove Proposition 5.5.

Proof. Fix g and let $S^\star \in \mathrm{Opt}(g)$. Let $S_0 \in \mathcal{S}$ be any shortest strategy, so $|S_0| = \ell_{\min}$. Since $S^\star$ is an exact LMO minimizer,

$$\sum_{i \in S^\star} g_i \le \sum_{i \in S_0} g_i.$$

Using $g_{\min} \le g_i \le g_{\max}$ gives

$$g_{\min}|S^\star| \le \sum_{i \in S^\star} g_i \le \sum_{i \in S_0} g_i \le g_{\max}\ell_{\min}.$$

Hence $|S^\star| \le (g_{\max}/g_{\min})\ell_{\min}$ and, since $|S^\star|$ is integer, $|S^\star| \le R$. Thus every exact minimizer lies in a length stratum $\ell \le R$.

Under UL sampling, each $S \in \mathcal{S}_\ell$ has probability

$$q_{\mathrm{UL}}(S) = \frac{1}{|\mathcal{L}|N_\ell}.$$

For every optimizer $S^\star$, we have $|S^\star| \le R$ and $N_{|S^\star|} \le M_R$, so

$$q_{\mathrm{UL}}(S^\star) \ge \frac{1}{|\mathcal{L}|M_R}.$$

Since $\mathrm{Opt}(g)$ is nonempty,

$$p_{\mathrm{UL}}(g) = \sum_{S \in \mathrm{Opt}(g)} q_{\mathrm{UL}}(S) \ge \frac{1}{|\mathcal{L}|M_R}.$$

Therefore

$$\kappa_{m,\mathrm{UL}}(g) = 1 - \big(1 - p_{\mathrm{UL}}(g)\big)^m \ge 1 - \left(1 - \frac{1}{|\mathcal{L}|M_R}\right)^m.$$

For the comparison with uniform-over-strategies sampling, write $r_\ell(g) := |\mathrm{Opt}(g) \cap \mathcal{S}_\ell|$ and $r(g) := |\mathrm{Opt}(g)|$. Because all optimizers have length at most R,

$$p_{\mathrm{UL}}(g) = \sum_{\ell \in \mathcal{L}:\ell \le R} \frac{r_\ell(g)}{|\mathcal{L}|N_\ell} \ge \sum_{\ell \in \mathcal{L}:\ell \le R} \frac{r_\ell(g)}{|\mathcal{L}|M_R} = \frac{r(g)}{|\mathcal{L}|M_R}.$$

Since $p_{\mathrm{US}}(g) = r(g)/|\mathcal{S}|$, this gives

$$p_{\mathrm{UL}}(g) \ge \frac{|\mathcal{S}|}{|\mathcal{L}|M_R}\, p_{\mathrm{US}}(g).$$

□

Harmonic-length sampling. The same argument gives a corresponding bound for HL sampling. Let

$$Z_{\mathcal{L}} := \sum_{\ell \in \mathcal{L}} \frac{1}{\ell}, \qquad q_{\mathrm{HL}}(S) := \frac{1}{Z_{\mathcal{L}}|S|N_{|S|}}.$$

For every optimizer $S^\star$, the proof above gives $|S^\star| \le R$ and $N_{|S^\star|} \le M_R$, hence

$$q_{\mathrm{HL}}(S^\star) \ge \frac{1}{Z_{\mathcal{L}}RM_R}.$$

Consequently,

$$p_{\mathrm{HL}}(g) \ge \frac{1}{Z_{\mathcal{L}}RM_R}, \qquad \kappa_{m,\mathrm{HL}}(g) \ge 1 - \left(1 - \frac{1}{Z_{\mathcal{L}}RM_R}\right)^m.$$

Moreover,

$$p_{\mathrm{HL}}(g) = \sum_{\ell \in \mathcal{L}:\ell \le R} \frac{r_\ell(g)}{Z_{\mathcal{L}}\ell N_\ell} \ge \frac{r(g)}{Z_{\mathcal{L}}RM_R} = \frac{|\mathcal{S}|}{Z_{\mathcal{L}}RM_R}\, p_{\mathrm{US}}(g).$$

Exponential separation example. For $r \geq 2$, consider an s–t path family with one direct edge from s to t and a layered subgraph with vertices $v_0 = s, v_1, \ldots, v_r = t$, where each layer transition (v_{j-1}, v_j) has two parallel edges. The feasible strategy family consists of all simple s–t paths. There is one path of length 1 and 2^r paths of length r, so

$$|\mathcal{S}| = 2^r + 1, \qquad \mathcal{L} = \{1, r\}, \qquad N_1 = 1, \qquad N_r = 2^r.$$

Set all LMO weights equal to one. Then the direct path is the unique exact LMO minimizer. Uniform-over-strategies sampling has

$$p_{\mathrm{US}}(g) = \frac{1}{2^r + 1}, \qquad \kappa_{m,\mathrm{US}}(g) = 1 - \left(1 - \frac{1}{2^r + 1}\right)^m \leq \frac{m}{2^r + 1}.$$

In contrast, UL sampling hits the length-one stratum with probability $1/2$, hence

$$p_{\mathrm{UL}}(g) = \frac{1}{2}, \qquad \kappa_{m,\mathrm{UL}}(g) = 1 - 2^{-m}.$$

HL sampling assigns probability

$$p_{\mathrm{HL}}(g) = \frac{1}{1 + 1/r} = \frac{r}{r+1}$$

to the length-one stratum, hence

$$\kappa_{m,\mathrm{HL}}(g) = 1 - \left(\frac{1}{r+1}\right)^m.$$

Thus uniform-over-strategies sampling requires $m = \Omega(2^r)$ samples to obtain constant optimizer-hit probability, whereas UL and HL achieve constant hit probability with $m = O(1)$. Since the subsampled-FW rate in Theorem 5.4 scales as $\mathcal{O}(1/(\kappa_m T))$, this yields an exponential separation in the effective convergence constant.

H.3 Fixed-Cardinality Limits and Structured TSP Strata

No universal gain from a fixed, weight-independent sampler. For the complete undirected graph on $N \geq 3$ labeled vertices, let $\mathcal{S}$ contain undirected Hamiltonian tours, identifying rotations and reversal. Then $|\mathcal{S}| = (N-1)!/2$. Any fixed distribution q assigns some tour S probability at most $1/|\mathcal{S}|$. Giving its edges weight 1 and every other edge weight $1+\delta$, $\delta > 0$, makes S uniquely optimal. Thus uniform sampling maximizes the worst-case optimizer mass among fixed, weight-independent samplers, but its worst-case hit probability $1-(1-1/|\mathcal{S}|)^m \leq m/|\mathcal{S}|$ can still be tiny. This statement does not restrict weight-dependent samplers or additional structural information. For arbitrary costs over permutations of N objects, the analogous cardinality is $N!$, not the undirected-tour count. A fixed sampler can assign a possible unique optimizer mass at most $1/N!$, giving hit probability at most $m/N!$. Both families have fixed cardinality under their standard edge or position–item incidence encoding, so cardinality stratification alone cannot improve uniform sampling.

A sufficient zone-transition certificate. Consider instead a known partition of the N vertices into J nonempty zones, with $3 \leq J < N$. Every Hamiltonian tour remains feasible: zones provide sampling information, not feasibility constraints. Let $\tau(S)$ count edges joining different zones and set $\mathcal{A} := \{S : \tau(S) = J\}$. Each zone has an even, positive number of boundary crossings; summing them counts each between-zone edge twice. Hence $\tau(S) \geq J$, with equality exactly when every zone is visited as a contiguous block. For nonnegative LMO weights g_e, define

$$D^{\times}_{\min} := \min_{e \text{ joins different zones}} g_e, \qquad D^{\times}_{\max} := \max_{e \text{ joins different zones}} g_e, \qquad D^{\circ}_{\max} := \max_{e \text{ lies within one zone}} g_e.$$

The restriction $J < N$ ensures that within-zone edges exist. Every tour outside $\mathcal{A}$ has cost at least $(J+1)D^{\times}_{\min}$, whereas every tour in $\mathcal{A}$ has cost at most $JD^{\times}_{\max} + (N-J)D^{\circ}_{\max}$. Therefore

$$(J+1)D^{\times}_{\min} > JD^{\times}_{\max} + (N-J)D^{\circ}_{\max} \tag{9}$$

places every exact LMO minimizer in $\mathcal{A}$. Nonnegativity is needed when discarding the remaining edge costs in the lower bound. For uniform optimizer-hit control, the certificate must hold throughout the relevant FW runs.

Counting and sampling. If zone j has h_j vertices, then

$$|\mathcal{A}| = \frac{(J-1)!}{2} \prod_{j=1}^{J} h_j!.$$

To sample uniformly from $\mathcal{A}$, fix one zone first, uniformly permute the other zones and the vertices within each zone, concatenate the blocks, and close the tour. Each undirected tour in $\mathcal{A}$ has exactly two such oriented representations, proving uniformity without enumerating tours. If a two-stratum sampler puts mass ω on $\mathcal{A}$ and is uniform within each stratum, (9) gives

$$p(g) = \omega \frac{|\mathrm{Opt}(g)|}{|\mathcal{A}|}, \qquad \frac{p(g)}{p_{\mathrm{unif}}(g)} = \omega \frac{|\mathcal{S}|}{|\mathcal{A}|}.$$

Under the certificate, $\omega = 1$ maximizes this hit probability; reserving mass outside $\mathcal{A}$ is a hedge against incorrect structural information, not a theoretical necessity. Even this gain need not make $p(g)$ large when $\mathcal{A}$ is enormous.

For a ZDD, mark each edge by $b_e = \mathbf{1}\{e \text{ joins different zones}\}$. Counts of completions with r further transitions satisfy

$$N_r(v) = N_r(\mathrm{lo}(v)) + N_{r-b_e}(\mathrm{hi}(v)),$$

where e labels node v, $N_r(\top) = \mathbf{1}\{r = 0\}$, and all negative-index or $\bot$ counts are zero. Count-weighted traversal with target $r = J$ samples uniformly from $\mathcal{A}$; targets $r > J$ chosen proportionally to their counts sample uniformly from its complement (cf. Appendix I.3). This requires a tractable diagram and count preprocessing. If there is one zone or every zone is a singleton, the statistic is constant and gives no gain. For arbitrary TSP weights, neither (9) nor any useful alternative partition is guaranteed. A stratum defined by a low-cost threshold can be optimizer-enriched yet difficult to sample: rejection from uniform tours is costly when qualifying tours are rare, and direct conditional sampling can require difficult completion counts.

Numerical illustration. We instantiate the certificate on a synthetic follower problem with $N = 12$ vertices in $J = 6$ zones of two vertices each. The zone centers lie on a circle of radius 10, the two vertices of each zone lie close to its center, and each edge $e = \{u, v\}$ has base cost $d_e = \|x_u - x_v\|_2 + 200 \cdot \mathbf{1}\{e \text{ joins different zones}\}$, where $x_u \in \mathbb{R}^2$ is the position of vertex u. The follower uses the cost model of Section 6 with fixed $\theta_e = 1$ and $C = 0.1$, so $c_e(y_e) = d_e(1 + 0.05\, y_e)$. Since $0 \le y_e \le 1$, (9) holds at every feasible load with margin 10.97, so every exact LMO minimizer lies in $\mathcal{A}$ throughout every FW run. Here $|\mathcal{S}| = 11!/2 = 19{,}958{,}400$ and $|\mathcal{A}| = (5!/2)\, 2^6 = 3{,}840$. We compare uniform sampling with the two-stratum sampler that puts mass $\omega = 0.9$ on $\mathcal{A}$; we fix $\omega < 1$ without tuning, as a hedge against misspecified zones. Each run starts from the same tour in $\mathcal{A}$ and performs 300 subsampled FW steps with exact line search; we use 100 runs for each sampler and each $m \in \{10, 100, 1000\}$. We report no timings for this experiment, which was run on a laptop rather than on the cluster node used for the TNTP experiments.

Figure 7 shows the single-optimizer lower bound $\kappa_m \ge 1 - (1-p)^m$, with $p = 1/|\mathcal{S}|$ for uniform sampling and $p = \omega/|\mathcal{A}|$ for the stratified sampler, together with the final exact FW gap. Stratification raises the single-draw optimizer mass by the factor $\omega|\mathcal{S}|/|\mathcal{A}| \approx 4.7 \times 10^3$, and at $m = 1000$ the bound increases from 5.0×10^{-5} to 0.209. For $m = 10, 100, 1000$, the mean final FW gap is 55.2, 35.7, and 20.3 under uniform sampling, compared with 3.08, 0.29, and 3.0×10^{-7} under stratified sampling. This synthetic study illustrates the mechanism under a verified certificate; it does not show that arbitrary TSP instances admit useful strata.

H.4 Proof of the Subsampled Frank–Wolfe Theorem

Proof. Fix θ and abbreviate $f(y) := f(\theta, y)$ and $y^\star := y^\star(\theta)$. Let $h_t := f(y_t) - f(y^\star)$. By $L_{f,2}$-smoothness and for all $y, s \in \mathcal{C}$ with $\|s - y\| \le D$ and any $\gamma \in [0, 1]$,

$$f\big((1-\gamma)y + \gamma s\big) \le f(y) + \gamma\langle \nabla f(y), s - y\rangle + \frac{C_f}{2}\gamma^2, \qquad \text{where } C_f := L_{f,2}D^2. \tag{10}$$

Let $s_t^\star \in \arg\min_{s \in \mathcal{C}} \langle \nabla f(y_t), s\rangle$ and define the (exact) FW gap

$$\mathrm{gap}_t := \langle \nabla f(y_t), y_t - s_t^\star\rangle.$$

By convexity, $h_t \le \mathrm{gap}_t$ for all t. Let $\mathcal{F}_t$ denote the sigma-field generated by all randomness before the tth candidate set is sampled.

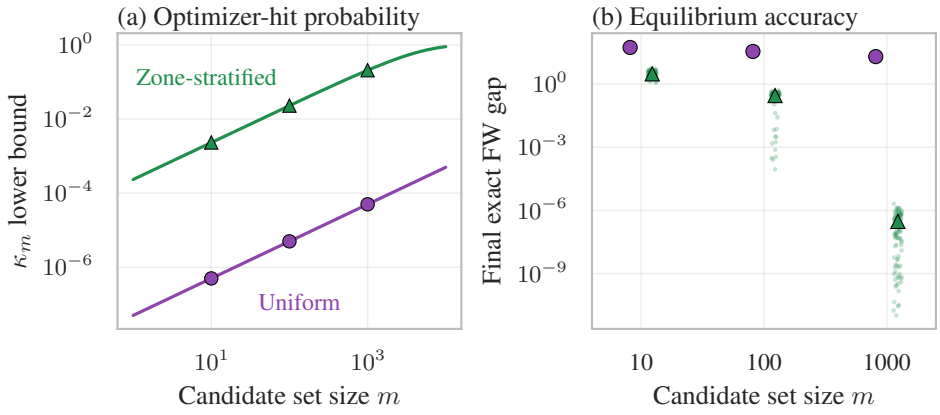


Figure 7: Zone-transition stratification on the synthetic 12-vertex TSP follower problem. (a) Single-optimizer lower bound on κ_m for uniform and zone-stratified sampling; markers indicate the budgets used in (b). (b) Exact FW gap after 300 subsampled FW steps; dots are individual runs (100 per setting) and markers are means.

Step 1: the hit event. Let $\mathcal{H}_t$ be the event that the m i.i.d. samples used by LMO_m contain an exact minimizer of $s \mapsto \langle \nabla f(y_t), s\rangle$. On $\mathcal{H}_t$, the best-of-m rule returns an exact minimizer, so $\langle \nabla f(y_t), s_t\rangle = \langle \nabla f(y_t), s_t^\star\rangle$. Since y_t is $\mathcal{F}_t$-measurable, Assumption 5.2 gives

$$\Pr(\mathcal{H}_t \mid \mathcal{F}_t) \geq \kappa_m. \tag{11}$$

Comparison to Kerdreux et al. [2018]. At a high level, the argument mirrors the subsampling-oracle FW analysis of Kerdreux et al. [2018] (see their proof of Theorem 2.1): identify a per-iteration event on which the update coincides with exact FW, and then average the resulting decrease over this event. The main additional challenge in our setting is that we do *not* assume a uniform-inclusion model. In Kerdreux et al. [2018], every atom is included with the same probability η, so the probability of hitting an exact LMO minimizer is controlled in a way that is essentially independent of the current iterate. Here, we sample candidates from an arbitrary distribution q over $\mathcal{S}$, so the hit probability depends on the current FW gradient through $\mathrm{Opt}(g)$; in particular, it is coupled with the algorithmic trajectory. To make this dependence explicit (and keep the proof clean), we isolate $p(g) = q(\mathrm{Opt}(g))$ and its m-sample amplification $\kappa_m(g) = 1 - (1 - p(g))^m$, and we assume a uniform lower bound along the run (Assumption 5.2). Given this, the proof proceeds by conditioning on the algorithm history and combining (i) an exact-FW decrease on $\mathcal{H}_t$ with (ii) monotonicity on $\mathcal{H}_t^c$ (via short-step or exact line search), which yields the conditional recursion (16) with the constant κ_m playing the role of η. This pinpoints precisely where general, structure-aware sampling enters the analysis and why improving κ_m (e.g., via stratification) improves the effective rate.

Step 2: approximate line search and definition of ψ. We analyze a smoothness-based *short-step* rule (an approximate line search): define the directional gap for the chosen atom s_t,

$$\widetilde{\mathrm{gap}}_t := \langle \nabla f(y_t), y_t - s_t\rangle,$$

and take

$$\gamma_t \in \operatorname*{arg\,max}_{\gamma\in[0,1]} \Big\{\gamma\, \widetilde{\mathrm{gap}}_t - \frac{C_f}{2}\gamma^2\Big\} \quad\Longleftrightarrow\quad \gamma_t = \min\Big\{\max\Big\{\frac{\widetilde{\mathrm{gap}}_t}{C_f}, 0\Big\}, 1\Big\}. \tag{12}$$

Applying (10) with $y = y_t$, $s = s_t$, and $\gamma = \gamma_t$ yields

$$f(y_{t+1}) \leq f(y_t) - \gamma_t\, \widetilde{\mathrm{gap}}_t + \frac{C_f}{2}\gamma_t^2 \leq f(y_t),$$

and hence $h_{t+1} \leq h_t$ always.

On the hit event $\mathcal{H}_t$ we have $\langle \nabla f(y_t), s_t\rangle = \langle \nabla f(y_t), s_t^\star\rangle$ and thus $\widetilde{\mathrm{gap}}_t = \mathrm{gap}_t$. Define

$$\psi(g) := \max_{\gamma\in[0,1]} \Big\{\gamma g - \frac{C_f}{2}\gamma^2\Big\}. \tag{13}$$

By the choice of γ_t in (12), we have $\gamma_t \widetilde{\mathrm{gap}}_t - \frac{C_f}{2}\gamma_t^2 = \psi(\widetilde{\mathrm{gap}}_t)$. Then, on $\mathcal{H}_t$,

$$h_{t+1} \le h_t - \psi(\mathrm{gap}_t). \tag{14}$$

If instead we use *exact* line search, the same bound holds since the exact minimizer over $\gamma \in [0,1]$ achieves an objective value no larger than the short-step choice in (12). Moreover, (13) has the closed form

$$\psi(g) = \begin{cases} \frac{g^2}{2C_f}, & 0 \le g \le C_f, \\ g - \frac{C_f}{2}, & g \ge C_f, \end{cases} \tag{15}$$

since the maximizer is $\gamma^\star(g) = \min\{g/C_f,\, 1\}$. In particular, $\psi(g) \ge 0$ for all $g \ge 0$.

Step 3: conditional recursion. From (13), for each fixed $\gamma \in [0,1]$ the map $g \mapsto \gamma g - \frac{C_f}{2}\gamma^2$ is affine and nondecreasing in g. Hence ψ, being the pointwise supremum over $\gamma \in [0,1]$ of such functions, is (i) *nondecreasing* and (ii) *convex* on $[0,\infty)$.

On $\mathcal{H}_t^c$ we have $h_{t+1} \le h_t$. Using (14) on $\mathcal{H}_t$ and conditioning on $\mathcal{F}_t$,

$$\begin{aligned} \mathbb{E}[h_{t+1} \mid \mathcal{F}_t] &\le \Pr(\mathcal{H}_t \mid \mathcal{F}_t)\,(h_t - \psi(\mathrm{gap}_t)) + \Pr(\mathcal{H}_t^c \mid \mathcal{F}_t)\, h_t \\ &= h_t - \Pr(\mathcal{H}_t \mid \mathcal{F}_t)\,\psi(\mathrm{gap}_t) \;\le\; h_t - \kappa_m\,\psi(\mathrm{gap}_t), \end{aligned} \tag{16}$$

where we used (11). Since $\mathrm{gap}_t \ge h_t$ and ψ is nondecreasing, $\psi(\mathrm{gap}_t) \ge \psi(h_t)$, so

$$\mathbb{E}[h_{t+1} \mid \mathcal{F}_t] \le h_t - \kappa_m\,\psi(h_t). \tag{17}$$

Take total expectation in (17):

$$\mathbb{E}[h_{t+1}] \le \mathbb{E}[h_t] - \kappa_m\,\mathbb{E}[\psi(h_t)].$$

Since ψ is convex, Jensen gives $\mathbb{E}[\psi(h_t)] \ge \psi(\mathbb{E}[h_t])$. Let $a_t := \mathbb{E}[h_t]$. Then

$$a_{t+1} \le a_t - \kappa_m\,\psi(a_t). \tag{18}$$

Step 4: solve the recursion. Using the closed form (15):

Geometric burn-in. The closed form (15) implies the global lower bound $\psi(a) \ge a - \frac{C_f}{2}$ for all $a \ge 0$. Therefore (18) yields, for every t,

$$a_{t+1} \le (1-\kappa_m)a_t + \kappa_m \frac{C_f}{2}.$$

Iterating gives

$$a_t \le (1-\kappa_m)^t a_0 + \frac{C_f}{2} \qquad \forall t \ge 0. \tag{19}$$

Sublinear regime: $a_t \le C_f$. Then $\psi(a_t) = \frac{a_t^2}{2C_f}$ and (18) becomes

$$a_{t+1} \le a_t - \frac{\kappa_m}{2C_f} a_t^2. \tag{20}$$

Let $c := \kappa_m/(2C_f)$. Since $a_{t+1} \le a_t$, we may assume $a_t > 0$. From (20) we have $a_{t+1} \le a_t(1 - ca_t)$, hence

$$\frac{1}{a_{t+1}} \ge \frac{1}{a_t(1-ca_t)}.$$

Using $a_t \le C_f$ gives $ca_t \le \kappa_m/2 < 1$, and thus $\frac{1}{1-x} \ge 1 + x$ for $x \in [0,1)$ implies

$$\frac{1}{a_{t+1}} \ge \frac{1}{a_t}\Big(1 + ca_t\Big) = \frac{1}{a_t} + c.$$

Summing from $t = 0$ to $T-1$ yields $\frac{1}{a_T} \ge \frac{1}{a_0} + cT \ge cT$, hence

$$a_T \le \frac{1}{cT} = \frac{2C_f}{\kappa_m T}. \tag{21}$$

From Regime 1 we have for all $t \geq 0$ the geometric bound

$$a_t \leq (1-\kappa_m)^t a_0 + \frac{C_f}{2}. \tag{19}$$

Choose $T_0 := \lceil \kappa_m^{-1} \log\bigl(1 + 2a_0/C_f\bigr) \rceil$. Using $1 - x \leq e^{-x}$ for $x \in [0,1]$ gives

$$(1-\kappa_m)^{T_0} a_0 \leq e^{-\kappa_m T_0} a_0 \leq \frac{C_f}{2},$$

and therefore

$$a_{T_0} \leq (1-\kappa_m)^{T_0} a_0 + \frac{C_f}{2} \leq C_f.$$

Hence, for all $t \geq T_0$ the iterates satisfy $a_t \leq C_f$, so Regime 2 applies from time T_0 onward. Applying the sublinear estimate (21) to the shifted sequence $\{a_{T_0}, a_{T_0+1}, \dots\}$ yields, for all $T \geq T_0$,

$$a_T \leq \frac{2C_f}{\kappa_m\,(T - T_0 + 1)}.$$

Recalling $a_T = \mathbb{E}[h_T] = \mathbb{E}[f(\theta, y_T) - f(\theta, y^\star(\theta))]$ proves (2).

Step 5: *mean-square* distance bound under Assumption 2.4. Let Assumption 2.4 hold. Fix any $\varepsilon \in (0, r]$ and $T \geq T_0$. Let $d_T := \|y_T - y^\star\|$. Since $y_T, y^\star \in \mathcal{C}$, we have $d_T \leq D$ almost surely. By Lemma H.1, $h_T := f(y_T) - f(y^\star) \geq \frac{\alpha}{2} \min\{d_T^2,\, r d_T\}$. If $d_T \leq r$, then $d_T^2 \leq \frac{2}{\alpha} h_T$. If $d_T > r$, then $h_T \geq \frac{\alpha r}{2} d_T$, so $d_T \leq \frac{2}{\alpha r} h_T$ and hence $d_T^2 \leq D d_T \leq \frac{2D}{\alpha r} h_T$. Thus,

$$\|y_T - y^\star\|^2 \;=\; d_T^2 \;\leq\; \frac{2}{\alpha} \max\Bigl\{1, \frac{D}{r}\Bigr\} h_T \qquad \text{almost surely.}$$

Taking expectations gives

$$\mathbb{E}\|y_T - y^\star\|^2 \;\leq\; \frac{2}{\alpha} \max\Bigl\{1, \frac{D}{r}\Bigr\} \mathbb{E}[h_T]. \tag{22}$$

Using (2) and $C_f = L_{f,2} D^2$, for all $T \geq T_0$ we have

$$\mathbb{E}[h_T] = \mathbb{E}[f(y_T)] - f(y^\star) \leq \frac{2C_f}{\kappa_m\,(T - T_0 + 1)}.$$

Combining with (22) gives

$$\mathbb{E}\|y_T - y^\star\|^2 \leq \frac{4C_f}{\alpha\,\kappa_m\,(T - T_0 + 1)} \max\Bigl\{1, \frac{D}{r}\Bigr\} = \frac{4L_{f,2} D^2}{\alpha\,\kappa_m\,(T - T_0 + 1)} \max\Bigl\{1, \frac{D}{r}\Bigr\}. \tag{23}$$

Therefore, it suffices to choose T so that the right-hand side of (23) is at most ε^2, i.e.,

$$T - T_0 + 1 \geq \frac{4L_{f,2} D^2}{\alpha\,\kappa_m\,\varepsilon^2} \max\Bigl\{1, \frac{D}{r}\Bigr\},$$

$$\text{equivalently} \qquad T \geq \left\lceil \frac{4L_{f,2} D^2}{\alpha\,\kappa_m\,\varepsilon^2} \max\Bigl\{1, \frac{D}{r}\Bigr\} - 1 \right\rceil + T_0.$$

If α, r are uniform over Θ, then the bound holds uniformly in θ, hence

$$\sup_{\theta \in \Theta} \mathbb{E}\|y_T(\theta) - y^\star(\theta)\|^2 \leq \varepsilon^2,$$

as claimed.

□

H.5 Proof of the Zeroth-Order GGSP Theorem (Theorem 5.6)

Proof. This proof is a standard "randomized smoothing + projected stochastic-gradient mapping" argument; see, e.g., [Nesterov and Spokoiny, 2017, Ghadimi and Lan, 2013, Liu et al., 2024]. The main point is to use the unbiasedness of the ideal mini-batched estimator while keeping the projection-induced variance floor explicit; the direction batch controls this floor, and the inexact inner solve contributes a separate bias term.

Step 1: smoothing and Goldstein subdifferentials. Define the ball-smoothed function

$$\Phi_\rho(\theta) := \mathbb{E}_{v\sim \mathrm{Unif}(\mathbb{B})}[\Phi(\theta+\rho v)].$$

Here Φ denotes an L_Φ-Lipschitz extension to the queried ρ-neighborhood of Θ, with the same uniform evaluation-error bound at every query. The proof below is conditional on this query convention; existence of the extension does not establish access to its values. This is the standard uniform-ball smoothing used in gradient-free methods [Nesterov and Spokoiny, 2017, Ghadimi and Lan, 2013]. The following facts are standard for Lipschitz Φ:

1. (Approximation) $|\Phi_\rho(\theta) - \Phi(\theta)| \le \rho L_\Phi$ for all θ [Nesterov and Spokoiny, 2017].
2. (Differentiability) Φ_ρ is continuously differentiable, and $\nabla\Phi_\rho$ is Lipschitz with constant $L_\rho \le c\,\frac{L_\Phi\sqrt{d_\theta}}{\rho}$ for a universal constant $c>0$ [Nesterov and Spokoiny, 2017].
3. (Goldstein inclusion) $\nabla\Phi_\rho(\theta) \in \partial_\rho\Phi(\theta)$ for all θ [Goldstein, 1977].

We use these as black-box smoothing properties. By decreasing the universal constant c_0 in the theorem if necessary, the stated stepsize bound $\eta \le c_0\rho/(L_\Phi\sqrt{d_\theta})$ implies the condition $\eta \le c_0/L_\rho$ required below.

Step 2: direction-batch variance and inner-solve error. At iteration t, let $\mathcal{F}_t$ denote the history up to θ_t, and sample $u_{t,1},\dots,u_{t,B_{\mathrm{out}}} \sim \mathrm{Unif}(\mathbb{S}^{d_\theta-1})$ independently of $\mathcal{F}_t$. For each direction, define the symmetric two-point estimator (the algorithm uses $\widehat{\Phi}_T(\theta) = F(\theta, y_T(\theta))$)

$$\widehat{g}_{t,b} := \frac{d_\theta}{2\rho}\Big(\widehat{\Phi}_T(\theta_t+\rho u_{t,b}) - \widehat{\Phi}_T(\theta_t-\rho u_{t,b})\Big)u_{t,b},$$

and its ideal exact-hyperobjective counterpart

$$g^\Phi_{t,b} := \frac{d_\theta}{2\rho}\Big(\Phi(\theta_t+\rho u_{t,b}) - \Phi(\theta_t-\rho u_{t,b})\Big)u_{t,b}.$$

The update uses $\widehat{g}_t := B_{\mathrm{out}}^{-1}\sum_{b=1}^{B_{\mathrm{out}}}\widehat{g}_{t,b}$; define $g^\Phi_t := B_{\mathrm{out}}^{-1}\sum_{b=1}^{B_{\mathrm{out}}} g^\Phi_{t,b}$. Then (standard) $\mathbb{E}[g^\Phi_t \mid \mathcal{F}_t] = \nabla\Phi_\rho(\theta_t)$ [Nesterov and Spokoiny, 2017, Ghadimi and Lan, 2013], and

$$\mathbb{E}\|g^\Phi_t - \nabla\Phi_\rho(\theta_t)\|^2 = \mathbb{E}\big[\mathrm{Var}(g^\Phi_t \mid \mathcal{F}_t)\big] \le \frac{c' d_\theta L_\Phi^2}{B_{\mathrm{out}}}, \tag{24}$$

using independence across directions and the standard single-direction second-moment bound $\mathbb{E}\|g^\Phi_{t,b}\|^2 \le c' d_\theta L_\Phi^2$ for a universal $c'>0$.

Now define the pointwise oracle errors

$$e^+_{t,b} := \widehat{\Phi}_T(\theta_t+\rho u_{t,b}) - \Phi(\theta_t+\rho u_{t,b}), \qquad e^-_{t,b} := \widehat{\Phi}_T(\theta_t-\rho u_{t,b}) - \Phi(\theta_t-\rho u_{t,b}).$$

By $L_{F,1}$-Lipschitzness of $F(\theta,\cdot)$ and the assumed mean-square inner accuracy $\mathbb{E}\|y_T(\theta)-y^\star(\theta)\|^2 \le \varepsilon_y^2$ (for all queried θ),

$$\mathbb{E}[(e^\pm_{t,b})^2] \le L_{F,1}^2\,\mathbb{E}\|y_T(\theta_t\pm\rho u_{t,b}) - y^\star(\theta_t\pm\rho u_{t,b})\|^2 \le L_{F,1}^2\,\varepsilon_y^2. \tag{25}$$

$$d_{t,b} := \widehat{g}_{t,b} - g^\Phi_{t,b} = \frac{d_\theta}{2\rho}(e^+_{t,b} - e^-_{t,b})u_{t,b}, \qquad d_t := B_{\mathrm{out}}^{-1}\sum_{b=1}^{B_{\mathrm{out}}} d_{t,b}.$$

Thus, using (25),

$$\mathbb{E}\|d_{t,b}\|^2 \le \left(\frac{d_\theta}{\rho}\right)^2 L_{F,1}^2\varepsilon_y^2 := B_y^2. \tag{26}$$

By convexity of $\|\cdot\|^2$, $\mathbb{E}\|d_t\|^2 \le B_y^2$. The conditional bias of the implemented estimator is

$$b_t := \mathbb{E}[\widehat{g}_t \mid \mathcal{F}_t] - \nabla\Phi_\rho(\theta_t) = \mathbb{E}[d_t \mid \mathcal{F}_t],$$

so Jensen's inequality and (26) give

$$\mathbb{E}\|b_t\|^2 \le B_y^2. \tag{27}$$

Moreover, since $\widehat{g}_t = g_t^{\Phi} + d_t$ and $\mathbb{E}[\widehat{g}_t \mid \mathcal{F}_t] = \nabla\Phi_\rho(\theta_t) + \mathbb{E}[d_t \mid \mathcal{F}_t]$, we have the decomposition

$$\widehat{g}_t - \mathbb{E}[\widehat{g}_t \mid \mathcal{F}_t] = \big(g_t^{\Phi} - \nabla\Phi_\rho(\theta_t)\big) + \big(d_t - \mathbb{E}[d_t \mid \mathcal{F}_t]\big).$$

Therefore, by $\|a+b\|^2 \le 2\|a\|^2 + 2\|b\|^2$, (24), and

$$\mathbb{E}\|d_t - \mathbb{E}[d_t \mid \mathcal{F}_t]\|^2 \le \mathbb{E}\|d_t\|^2 \le B_y^2,$$
$$\mathbb{E}\|\widehat{g}_t - \mathbb{E}[\widehat{g}_t \mid \mathcal{F}_t]\|^2 \le C\big(d_\theta L_\Phi^2/B_{\mathrm{out}} + B_y^2\big) \tag{28}$$

for a universal constant $C > 0$.

Step 3: projected stochastic-gradient mapping bound. We now derive the projected stochastic-gradient mapping estimate used below. Let h be L_ρ-smooth on Θ, and consider the generic projected update

$$\theta_{t+1} = \Pi_\Theta(\theta_t - \eta v_t).$$

Write $g_t := \nabla h(\theta_t)$, $b_t := \mathbb{E}[v_t \mid \mathcal{F}_t] - g_t$, and $\sigma_t^2 := \mathbb{E}[\|v_t - \mathbb{E}[v_t \mid \mathcal{F}_t]\|^2 \mid \mathcal{F}_t]$. We assume $\eta \le c_0/L_\rho$, with $c_0 > 0$ small enough below.

First, projection optimality for $\theta_{t+1} = \Pi_\Theta(\theta_t - \eta v_t)$ gives, for every $z \in \Theta$,

$$\langle \theta_{t+1} - (\theta_t - \eta v_t),\, z - \theta_{t+1}\rangle \ge 0.$$

Taking $z = \theta_t$ and using $\theta_t - \theta_{t+1} = \eta G_\Theta(\theta_t, v_t; \eta)$ yields

$$\big\langle v_t - G_\Theta(\theta_t, v_t;\eta), G_\Theta(\theta_t, v_t;\eta)\big\rangle \ge 0, \qquad \text{hence} \qquad \big\langle v_t, G_\Theta(\theta_t, v_t;\eta)\big\rangle \ge \|G_\Theta(\theta_t, v_t;\eta)\|^2.$$

By L_ρ-smoothness and $\theta_{t+1} = \theta_t - \eta G_\Theta(\theta_t, v_t;\eta)$,

$$\begin{aligned}
h(\theta_{t+1}) &\le h(\theta_t) + \langle g_t, \theta_{t+1} - \theta_t\rangle + \frac{L_\rho}{2}\|\theta_{t+1} - \theta_t\|^2 \\
&= h(\theta_t) - \eta\big\langle g_t, G_\Theta(\theta_t, v_t;\eta)\big\rangle + \frac{L_\rho\eta^2}{2}\|G_\Theta(\theta_t, v_t;\eta)\|^2 \\
&= h(\theta_t) - \eta\big\langle v_t, G_\Theta(\theta_t, v_t;\eta)\big\rangle + \eta\big\langle v_t - g_t, G_\Theta(\theta_t, v_t;\eta)\big\rangle + \frac{L_\rho\eta^2}{2}\|G_\Theta(\theta_t, v_t;\eta)\|^2 \\
&\le h(\theta_t) - \eta\Big(1 - \frac{L_\rho\eta}{2}\Big)\|G_\Theta(\theta_t, v_t;\eta)\|^2 + \eta\big\langle v_t - g_t, G_\Theta(\theta_t, v_t;\eta)\big\rangle.
\end{aligned} \tag{29}$$

Using Young's inequality, $\big\langle v_t - g_t, G_\Theta(\theta_t, v_t;\eta)\big\rangle \le \frac{1}{4}\|G_\Theta(\theta_t, v_t;\eta)\|^2 + \|v_t - g_t\|^2$, and taking $c_0 \le 1$, (29) gives

$$h(\theta_{t+1}) \le h(\theta_t) - \frac{\eta}{4}\|G_\Theta(\theta_t, v_t;\eta)\|^2 + \eta\|v_t - g_t\|^2. \tag{30}$$

Next, nonexpansiveness of projection gives

$$\|G_\Theta(\theta_t, g_t;\eta) - G_\Theta(\theta_t, v_t;\eta)\| = \eta^{-1}\big\|\Pi_\Theta(\theta_t - \eta g_t) - \Pi_\Theta(\theta_t - \eta v_t)\big\| \le \|g_t - v_t\|.$$

Therefore,

$$\|G_\Theta(\theta_t, g_t;\eta)\|^2 \le 2\|G_\Theta(\theta_t, v_t;\eta)\|^2 + 2\|v_t - g_t\|^2. \tag{31}$$

Conditioning (30) on $\mathcal{F}_t$ and rearranging gives

$$\mathbb{E}[\|G_\Theta(\theta_t, v_t;\eta)\|^2 \mid \mathcal{F}_t] \le \frac{4}{\eta}\Big(h(\theta_t) - \mathbb{E}[h(\theta_{t+1}) \mid \mathcal{F}_t]\Big) + 4\,\mathbb{E}[\|v_t - g_t\|^2 \mid \mathcal{F}_t].$$

Combining this with (31) and taking conditional expectations,

$$\|G_\Theta(\theta_t, g_t;\eta)\|^2 \le \frac{8}{\eta}\Big(h(\theta_t) - \mathbb{E}[h(\theta_{t+1}) \mid \mathcal{F}_t]\Big) + 10\,\mathbb{E}[\|v_t - g_t\|^2 \mid \mathcal{F}_t].$$

Finally,

$$\mathbb{E}[\|v_t - g_t\|^2 \mid \mathcal{F}_t] = \mathbb{E}[\|b_t + v_t - \mathbb{E}[v_t \mid \mathcal{F}_t]\|^2 \mid \mathcal{F}_t] = \|b_t\|^2 + \sigma_t^2,$$

because b_t is $\mathcal{F}_t$-measurable and $\mathbb{E}[v_t - \mathbb{E}[v_t \mid \mathcal{F}_t] \mid \mathcal{F}_t] = 0$. Taking total expectations, summing over $t = 0, \dots, K-1$, using $h(\theta_K) \ge \inf_\Theta h$, and letting τ be uniform on $\{0, \dots, K-1\}$, we obtain

$$\mathbb{E}\|G_\Theta(\theta_\tau, \nabla h(\theta_\tau);\eta)\|^2 \le C\left(\frac{h(\theta_0) - \inf_\Theta h}{\eta K} \;+\; \mathbb{E}\|b_\tau\|^2 \;+\; \mathbb{E}[\sigma_\tau^2]\right), \tag{32}$$

where $C > 0$ is universal. This is the conservative projected-SGD mapping bound for constrained nonconvex problems: the zero-mean oracle noise remains as a variance floor in the gradient-mapping residual, which is why the direction batch B_{out} is needed.

Apply (32) with $h = \Phi_\rho$ and $v_t = \widehat{g}_t$. Using (27)–(28) gives

$$\mathbb{E}\|G_\Theta(\widehat{\theta}, \nabla\Phi_\rho(\widehat{\theta}); \eta)\|^2 \le C\left(\frac{\Phi_\rho(\theta_0) - \inf_\Theta \Phi_\rho}{\eta K} \;+\; B_y^2 \;+\; d_\theta L_\Phi^2 / B_{\mathrm{out}} \;+\; B_y^2\right).$$

Absorbing duplicate B_y^2 terms into the universal constant gives

$$\mathbb{E}\|G_\Theta(\widehat{\theta}, \nabla\Phi_\rho(\widehat{\theta}); \eta)\|^2 \le C\left(\frac{\Phi_\rho(\theta_0) - \inf_\Theta \Phi_\rho}{\eta K} \;+\; B_y^2 \;+\; \frac{d_\theta L_\Phi^2}{B_{\mathrm{out}}}\right).$$

By Jensen's inequality, the smoothing approximation $\Phi_\rho(\theta_0) - \inf_\Theta \Phi_\rho \le \Phi(\theta_0) - \inf_\Theta \Phi + 2\rho L_\Phi$, and $B_y = (d_\theta/\rho) L_{F,1} \varepsilon_y$, there exist universal constants $C_1, C_2, C_3 > 0$ such that

$$\mathbb{E}\|G_\Theta(\widehat{\theta}, \nabla\Phi_\rho(\widehat{\theta}); \eta)\| \le C_1\sqrt{\frac{\Phi(\theta_0) - \inf_\Theta \Phi + 2\rho L_\Phi}{\eta K}} \;+\; C_2 L_\Phi \sqrt{\frac{d_\theta}{B_{\mathrm{out}}}} \;+\; C_3 \frac{d_\theta}{\rho} L_{F,1} \varepsilon_y,$$

where we absorbed universal constants. Finally, $\nabla\Phi_\rho(\widehat{\theta}) \in \partial_\rho \Phi(\widehat{\theta})$, so the left-hand side upper-bounds the GGSP residual in the theorem. □

H.6 Derivation of the Explicit Oracle Complexity

Use the notation and hypotheses of Corollary 5.7, and let

$$\bar{\varepsilon}_y := \min\left\{r, \frac{\rho\epsilon}{3C_3 d_\theta L_{F,1}}\right\}.$$

Choose the theorem parameters explicitly as

$$\eta = \frac{c_0 \rho}{L_\Phi \sqrt{d_\theta}}, \qquad B_{\mathrm{out}} = \left\lceil \frac{9C_2^2 L_\Phi^2 d_\theta}{\epsilon^2} \right\rceil,$$
$$K = \left\lceil \frac{9C_1^2 \Delta_\Phi}{\eta \epsilon^2} \right\rceil, \qquad T = \bar{T}_0 + \left\lceil \frac{4A}{\bar{\varepsilon}_y^2} \right\rceil.$$

Theorem 5.4 gives uniform mean-square load error at most $\bar{\varepsilon}_y^2$. Each of the three terms in (3) is consequently at most $\epsilon/3$. There are exactly $N_F = 2B_{\mathrm{out}}K$ upper-level evaluations and $N_{\mathrm{inner}} = N_F T$ lower gradients and LMO calls; a subsampled LMO draws m strategies per call. The ceilings give, for arbitrary $\epsilon > 0$,

$$N_F = \mathcal{O}\left[\left(1 + \frac{L_\Phi^2 d_\theta}{\epsilon^2}\right)\left(1 + \frac{\Delta_\Phi L_\Phi \sqrt{d_\theta}}{\rho\epsilon^2}\right)\right],$$
$$N_{\mathrm{inner}} = \mathcal{O}\left[N_F\left(1 + \bar{T}_0 + A\max\left\{r^{-2}, \frac{9C_3^2 {d_\theta}^2 L_{F,1}^2}{\rho^2\epsilon^2}\right\}\right)\right].$$

For $0 < \epsilon \le L_\Phi$, the additive ones in the bound for N_F can be absorbed: $d_\theta \ge 1$ and $\Delta_\Phi \ge 2\rho L_\Phi$ ensure that both nonconstant factors are bounded below. If also $\rho\epsilon \le 3C_3 d_\theta L_{F,1} r$, substitute $\bar{\varepsilon}_y = \rho\epsilon/(3C_3 d_\theta L_{F,1})$ to obtain the displayed bounds in Corollary 5.7. The $1 + \bar{T}_0$ term retains integer rounding and burn-in. With an exact LMO the same argument uses $\kappa_m = 1$ and needs no optimizer-hit assumption. These are oracle counts, not bounds on preprocessing or the computational cost of an individual oracle call.

I Zero-Suppressed Decision Diagrams (ZDD)

Zero-suppressed decision diagrams (ZDDs) are a canonical data structure for representing a family of subsets of a finite universe in a compact form. They can be viewed as a specialization of reduced ordered binary decision diagrams (BDDs) [Bryant, 1986] tailored to sparse set systems [Minato, 1993]. In our setting, the universe typically consists of graph resources (e.g., edges), and a feasible strategy (path, tree, cycle) is encoded by its incidence vector, i.e., a subset of universe items.

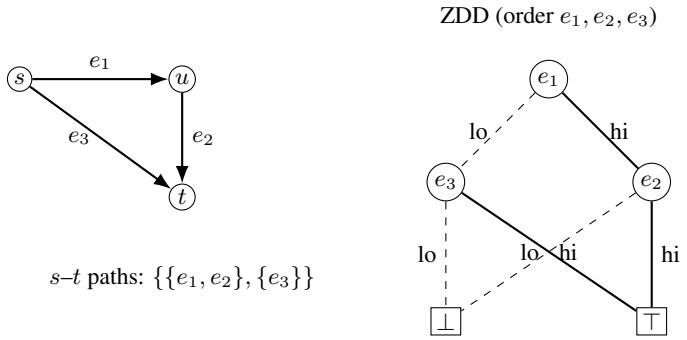


Figure 8: Left: a three-edge network with two s–t paths. Right: a ZDD encoding the corresponding strategy family. Root-to-$\top$ paths correspond to feasible strategies, with hi-arcs indicating selected edges.

I.1 Definition and Construction

ZDDs as compressed set representations. Fix an ordered universe $U = (e_1, \ldots, e_n)$ and a family $\mathcal{S} \subseteq 2^U$. A ZDD for $\mathcal{S}$ is a rooted directed acyclic graph with two terminals $\top$ and $\bot$. Each internal node v is labeled by an item index $\ell(v) \in \{1, \ldots, n\}$ and has two outgoing arcs: a *lo*-arc corresponding to excluding $e_{\ell(v)}$ and a *hi*-arc corresponding to including $e_{\ell(v)}$. A root-to-$\top$ path encodes a unique subset $S \subseteq U$ given by the labels whose hi-arcs are taken; the family represented by the diagram is exactly the collection of subsets obtained from all root-to-$\top$ paths. Figure 8 illustrates this correspondence on a toy network.

The key advantage of ZDDs is that they admit aggressive reductions. As in BDDs, isomorphic subgraphs are merged so that identical subproblems are shared [Bryant, 1986]. In addition, the *zero-suppression* rule removes any node whose hi-child is $\bot$, redirecting the incoming arc(s) to its lo-child [Minato, 1993]. Intuitively, if including an item can never lead to a feasible set, then that item is irrelevant for the family and should be skipped. After applying these reductions, the resulting diagram is canonical for a fixed variable ordering of U.

Construction and size. ZDDs can be built by incremental, constraint-driven enumeration procedures that traverse items in the chosen order and maintain a compact representation of partial feasibility. For graph families (paths, Steiner structures, cycles), such constructions often proceed by a frontier-style dynamic program that tracks only local connectivity information at the boundary between processed and unprocessed edges. The resulting preprocessing time is roughly linear in the number of ZDD nodes produced.

The practical challenge is that ZDD size is highly instance dependent. In the worst case, the number of nodes can grow exponentially with n (and it depends sensitively on the chosen ordering), which makes ZDD-based approaches unsuitable as a universal replacement for combinatorial optimization. However, when the underlying constraints have exploitable structure, ZDDs can be orders of magnitude smaller than the explicit family $\mathcal{S}$ and enable exact computations over exponentially large strategy sets [Minato, 1993].

I.2 Exact Optimization over a ZDD

A ZDD enables exact minimization of any additive weight function over the underlying family. Given weights $w \in \mathbb{R}^n$ (e.g., edge costs), consider

$$\min_{S \in \mathcal{S}} \sum_{e_i \in S} w_i.$$

Because each feasible strategy corresponds to a root-to-$\top$ path, this problem is equivalent to finding a minimum-cost root-to-$\top$ path in the ZDD DAG, where taking the hi-arc at a node labeled i incurs cost w_i and taking the lo-arc incurs cost 0. Since the ZDD is acyclic, the optimum can be computed by a single bottom-up dynamic program. Writing $\mathrm{lo}(v)$ and $\mathrm{hi}(v)$ for the children of a node v and

letting $\mathrm{cost}(\top) = 0$, $\mathrm{cost}(\bot) = +\infty$, we compute

$$\mathrm{cost}(v) = \min\{\mathrm{cost}(\mathrm{lo}(v)),\ w_{\ell(v)} + \mathrm{cost}(\mathrm{hi}(v))\},$$

in reverse topological order. Recovering an optimal strategy amounts to following the minimizing branch from the root. The runtime is linear in the number of ZDD nodes (and requires only additions and minima), so the diagram acts as an *exact linear minimization oracle* over an exponentially large family.

I.3 Sampling Strategies from a ZDD

Beyond exact optimization, ZDDs support efficient sampling of feasible strategies without enumerating $\mathcal{S}$. This is useful when exact optimization becomes costly (e.g., for very large diagrams) and one wishes to form a subsampled candidate set. Sampling is again implemented by dynamic programming followed by a randomized root-to-$\top$ traversal.

Uniform sampling over strategies. Let $N(v)$ denote the number of root-to-$\top$ completions from a node v. These counts satisfy $N(\top) = 1$, $N(\bot) = 0$, and

$$N(v) = N(\mathrm{lo}(v)) + N(\mathrm{hi}(v)).$$

After computing $N(v)$ for all nodes bottom-up, a uniformly random strategy is obtained by starting at the root and, at each node v, choosing the lo-arc with probability $N(\mathrm{lo}(v))/N(v)$ and the hi-arc with probability $N(\mathrm{hi}(v))/N(v)$. Whenever the hi-arc is chosen, we include the corresponding item $e_{\ell(v)}$. This produces an exactly uniform draw from $\mathcal{S}$.

Length-stratified sampling. In many graph families, the distribution of strategy sizes $|S|$ is highly skewed, so uniform sampling over $\mathcal{S}$ can overwhelmingly produce long, dense strategies. To control the sampled lengths, it is convenient to first choose a target length and then sample uniformly conditional on that length.

Let $N_r(v)$ denote the number of completions from v that select exactly r additional items. These length-refined counts satisfy $N_0(\top) = 1$, $N_r(\top) = 0$ for $r \geq 1$, $N_r(\bot) = 0$, and

$$N_r(v) = N_r(\mathrm{lo}(v)) + N_{r-1}(\mathrm{hi}(v)), \qquad r \geq 0,$$

with the convention $N_{-1}(\cdot) = 0$. Given a desired remaining length r, a uniformly random strategy of that length is obtained by traversing from the root and selecting the lo-arc with probability $N_r(\mathrm{lo}(v))/N_r(v)$ (keeping r unchanged) and selecting the hi-arc with probability $N_{r-1}(\mathrm{hi}(v))/N_r(v)$ (including $e_{\ell(v)}$ and decrementing r).

Choosing the length distribution. Let $\mathcal{L} := \{r \geq 0 : N_r(\mathrm{root}) > 0\}$ denote the set of feasible lengths. A general length-stratified sampler is defined by choosing a distribution π on $\mathcal{L}$ and then sampling uniformly conditional on the selected length. We use three choices:

- **Uniform strategy (US).** Choose r with probability $\pi(r) = N_r(\mathrm{root}) / \sum_{\ell \in \mathcal{L}} N_\ell(\mathrm{root})$. This recovers uniform sampling over all strategies.
- **Uniform length (UL).** Choose r uniformly over feasible lengths, i.e., $\pi(r) = 1/|\mathcal{L}|$ for $r \in \mathcal{L}$. This equalizes the mass assigned to each length class.
- **Harmonic length (HL).** Choose r with probability proportional to $1/r$ over $\mathcal{L} \cap \{1, 2, \dots\}$. This further biases sampling toward shorter strategies.

Complexity and numerical stability. Once the diagram is built, uniform-strategy sampling requires only the node counts $N(v)$ and then one root-to-terminal traversal per sample. Length-stratified sampling requires computing the table $(N_r(v))_{v,r}$ up to the maximum feasible length, after which each sample is generated by selecting a length and performing a single traversal with the appropriate conditional probabilities. In practice the counts can be astronomically large, so in our Python implementation, we store them in log-scale and form the branching probabilities via log-sum-exp arithmetic.

Table 8: Asymptotic costs of LMO implementations. Here n bounds the depth of a root-to-terminal ZDD traversal, while $\ell_{\max} \leq n$ upper-bounds the number of selected resources in any strategy. Uniform-strategy (US) sampling only requires scalar node counts, while length-stratified schemes (UL/HL) require length-refined counts up to $\ell_{\max}$ (see Appendix I.3 for more details).

LMO	Preprocessing (one-time)	Per FW iteration
Poly-time family	none	$\mathrm{poly}(n)$
ZDD exact	build ZDD: $\mathcal{O}(N_{\mathrm{ZDD}})$	min-cost DP: $\mathcal{O}(N_{\mathrm{ZDD}})$
ZDD subsampling (US)	build ZDD + counts: $\mathcal{O}(N_{\mathrm{ZDD}})$	m samples + scoring: $\mathcal{O}(mn)$
ZDD subsampling (UL/HL)	build ZDD + length counts: $\mathcal{O}(N_{\mathrm{ZDD}}\,\ell_{\max})$	m samples + scoring: $\mathcal{O}(mn)$

J Complexity of the Linear Minimization Oracle

We summarize the computational costs of the lower-level LMO used by Frank–Wolfe (Algorithm 2), separating one-time preprocessing from the per-iteration work. As in the main text, we write n for the universe size (e.g., number of edges/resources), and denote by N_{ZDD} the number of internal nodes in the ZDD representing the strategy family $\mathcal{S}$ (see Appendix I). For subsampling LMOs, m denotes the number of sampled strategies per FW step. We also define $\ell_{\max} := \max_{S\in\mathcal{S}} |S|$ as the maximum strategy length (i.e., the maximum number of selected universe items), so that $\ell_{\max} \leq n$. A ZDD sample is generated by a root-to-terminal traversal, whose worst-case length is bounded by the diagram depth, hence by n, not by the number of selected items. Table 8 summarizes the resulting preprocessing and per-iteration costs.

Poly-time strategy families. When the strategy family admits a polynomial-time combinatorial algorithm, the LMO can be implemented without preprocessing. For example, for s–t paths and additive edge weights, the LMO is a shortest-path problem, solvable in $\mathcal{O}(|E| \log |V|)$ time using Dijkstra's algorithm, where $|V|$ and $|E| = n$ are the numbers of vertices and edges in the graph.

NP-hard families with exact ZDD minimization. For NP-hard families, the ZDD provides an implicit representation of $\mathcal{S}$ that enables exact minimization of additive costs by dynamic programming on the ZDD DAG. Once the diagram is built, a min-cost oracle call runs in time linear in the ZDD size, $\mathcal{O}(N_{\mathrm{ZDD}})$. The worst-case N_{ZDD} can be exponential in n, but in many structured instances it is far smaller than $|\mathcal{S}|$; our experiments illustrate this gap across three real transportation networks.

NP-hard families with subsampling. When N_{ZDD} is too large for repeated exact minimization, we still build the ZDD once and use it as a sampler. Uniform-strategy sampling requires only node counts, while length-stratified sampling additionally uses length-refined counts up to $\ell_{\max}$. Each FW step then draws m feasible strategies and returns the best among the samples with respect to the current linearization. This reduces the per-iteration cost to $\mathcal{O}(mn)$ at the expense of replacing the exact minimizer with a randomized approximation.

K Orders of Smoothness and Lipschitz Constants

This appendix relates the abstract constants in Assumption 2.5 to the primitives $c_i(\cdot;\theta)$, and specializes the constants to the fractional cost model used in the TNTP experiments described in Appendix F.1.

K.1 Smoothness of the Beckmann Potential

Recall the Beckmann potential

$$f(\theta, y) = \sum_{i=1}^{n} \int_0^{y_i} c_i(u;\theta)\, du, \qquad \nabla_y f(\theta, y) = \big(c_i(y_i;\theta)\big)_{i=1}^{n}.$$

Hence, the constant $L_{f,2}$ in Assumption 2.5 is exactly a Lipschitz constant of the vector cost map $(\theta, y) \mapsto c(y;\theta)$ on $\Theta \times \mathcal{C}$.

A common sufficient condition is coordinatewise Lipschitzness: assume there exist constants $L_y, L_\theta \geq 0$ such that for all $i \in [n]$ and all $(\theta, y), (\theta', y') \in \Theta \times \mathcal{C}$,

$$|c_i(y_i;\theta) - c_i(y_i';\theta')| \;\leq\; L_y\, |y_i - y_i'| \;+\; L_\theta\, \|\theta - \theta'\|.$$

Then

$$\|\nabla_y f(\theta, y) - \nabla_y f(\theta', y')\| \;=\; \|c(y;\theta) - c(y';\theta')\| \;\le\; \sqrt{n}\, L_\theta \|\theta - \theta'\| + L_y \|y - y'\|,$$

so Assumption 2.5 holds with $L_{f,2} := \max\{L_y, \sqrt{n}\, L_\theta\}$. In particular, for each fixed θ, the map $y \mapsto f(\theta, y)$ is $L_{f,2}$-smooth on $\mathcal{C}$.

K.2 Lipschitzness of the Leader Objective

For the social-cost objective used in the experiments,

$$F(\theta, y) = \sum_{i=1}^{n} y_i\, c_i(y_i; \theta_i),$$

a convenient way to certify Assumption 2.5 is via gradient bounds. Assume F is differentiable on $\Theta \times \mathcal{C}$ and define the uniform coordinate bounds

$$B_y := \sup_{(\theta,y)\in\Theta\times\mathcal{C}} \max_{i\in[n]} \big|\partial_{y_i} F(\theta, y)\big|, \qquad B_\theta := \sup_{(\theta,y)\in\Theta\times\mathcal{C}} \max_{j\in[d_\theta]} \big|\partial_{\theta_j} F(\theta, y)\big|.$$

Then $\|\nabla_y F(\theta, y)\| \le \sqrt{n}\, B_y$ and $\|\nabla_\theta F(\theta, y)\| \le \sqrt{d_\theta}\, B_\theta$, hence for all $(\theta, y), (\theta', y') \in \Theta \times \mathcal{C}$,

$$|F(\theta, y) - F(\theta', y')| \;\le\; \sqrt{d_\theta}\, B_\theta\, \|\theta - \theta'\| + \sqrt{n}\, B_y\, \|y - y'\| \;\le\; L_{F,1}\big(\|\theta - \theta'\| + \|y - y'\|\big),$$

where one may take $L_{F,1} := \max\{\sqrt{d_\theta}\, B_\theta, \sqrt{n}\, B_y\}$. In particular, if the partial derivatives are uniformly $\mathcal{O}(1)$, then typically $L_{F,1} = \mathcal{O}(\sqrt{n} + \sqrt{d_\theta})$.

K.3 Representative Specialization to a Fractional Cost Model

As a concrete illustration, consider

$$c_i(y_i; \theta_i) = d_i\Big(1 + C\,\frac{y_i}{\theta_i + 1}\Big), \qquad \Theta = \{\theta \in \mathbb{R}^n_+ : \mathbf{1}^\top \theta = n\}, \qquad y \in \mathcal{C} \subseteq [0, 1]^n,$$

and we normalize $d_i \le 1$.

Bound on $L_{f,2}$. For all i and all (θ, y) in the domain,

$$\frac{\partial}{\partial y_i} c_i(y_i; \theta_i) = d_i\, \frac{C}{\theta_i + 1} \le C, \qquad \Big|\frac{\partial}{\partial \theta_i} c_i(y_i; \theta_i)\Big| = d_i\, C\, \frac{y_i}{(\theta_i + 1)^2} \le C.$$

Therefore $(\theta, y) \mapsto c(y; \theta)$ is C-Lipschitz in y and also C-Lipschitz in θ (on $\Theta \times \mathcal{C}$ under the Euclidean norm), so Assumption 2.5 holds with

$$L_{f,2} = \mathcal{O}(C).$$

Bound on $L_{F,1}$. Writing

$$F(\theta, y) = \sum_{i=1}^{n} \Big(d_i y_i + d_i C\, \frac{y_i^2}{\theta_i + 1}\Big),$$

we have, for each i,

$$\frac{\partial F}{\partial y_i} = d_i + 2 d_i C\, \frac{y_i}{\theta_i + 1} \le 1 + 2C, \qquad \Big|\frac{\partial F}{\partial \theta_i}\Big| = d_i C\, \frac{y_i^2}{(\theta_i + 1)^2} \le C.$$

Hence $\|\nabla_y F(\theta, y)\| \le \sqrt{n}\,(1 + 2C)$ and $\|\nabla_\theta F(\theta, y)\| \le \sqrt{n}\, C$, so one may take

$$L_{F,1} = \mathcal{O}\big(\sqrt{n}\,(1 + C)\big).$$

Implication for L_Φ. Combining with Lemma 2.6, a valid choice is

$$L_\Phi = L_{F,1}\Big(1 + \frac{2L_{f,2}}{\alpha}\Big) = \mathcal{O}\Big(\sqrt{n}\,(1 + C)\Big(1 + \frac{C}{\alpha}\Big)\Big).$$

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: The abstract and introduction state the scope, main assumptions, contributions, and empirical claims, including that the gains vary by oracle regime: quality and memory in the large-ZDD case, and computational savings mainly relative to the differentiable baseline; see the abstract and Sections 1 and 6. The comparisons with other outer methods and the total times to quality are reported in Sections 6.1 and 6.2 and Appendix F.4.

   Guidelines:

   - The answer [N/A] means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A [No] or [N/A] answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: Sections 2.3, 5.1, and 7 discuss the growth and sampling assumptions, scope, and limitations. Appendix C retains the longer discussion of stationarity versus global optimality, instance-dependent oracle tractability, and the empirical scope.

   Guidelines:

   - The answer [N/A] means that the paper has no limitation while the answer [No] means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
   - If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
   - While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory assumptions and proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

   Answer: [Yes]

   Justification: The assumptions are stated in Sections 2 and 5; proofs are given in Appendices G.1, G.3, and H.

   Guidelines:

   - The answer [N/A] means that the paper does not include theoretical results.
   - All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
   - All assumptions should be clearly stated or referenced in the statement of any theorems.
   - The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
   - Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
   - Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental result reproducibility**

   Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

   Answer: [Yes]

   Justification: Section 6 and Appendices F.1, F.4, and H.3 specify the benchmark instances, scenario construction, cost model, hyperparameters, runtime limits, parallelization choices, seeds, and evaluation metrics used for the reported experiments.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - If the paper includes experiments, a [No] answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
   - If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
   - Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
   - While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
     I If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
     II If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
     III If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).

IV We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

   Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

   Answer: [Yes]

   Justification: The source-code repository at `https://github.com/sepehr78/ZOS-congestion` contains our ZOS implementation, including the FW solver, the exact and subsampled LMOs, the ZDD construction and sampling utilities, the three TNTP scenarios, and scripts that run ZOS and certify its iterates; the paper uses the public TNTP benchmark suite [Transportation Networks for Research Core Team, 2016]. Appendix F.1 specifies the benchmark instances, scenario construction, cost model, hyperparameters, FW inner length, seeds, parallelization, and runtime limits required to reproduce the main experimental results, and Appendix F.4 gives the protocols of the baselines and ablations.

   Guidelines:

   - The answer [N/A] means that paper does not include experiments requiring code.
   - Please see the NeurIPS code and data submission guidelines (`https://neurips.cc/public/guides/CodeSubmissionPolicy`) for more details.
   - While we encourage the release of code and data, we understand that this might not be possible, so [No] is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
   - The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (`https://neurips.cc/public/guides/CodeSubmissionPolicy`) for more details.
   - The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
   - The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
   - At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
   - Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental setting/details**

   Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer) necessary to understand the results?

   Answer: [Yes]

   Justification: Section 6 and Appendix F.1 describe the benchmark construction, model family, objective, optimizer choices, FW inner-loop length, outer-loop hyperparameters used in the reported runs, and thread counts; Appendices F.4 and H.3 give the corresponding details for the additional experiments.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
   - The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment statistical significance**

Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

Answer: [Yes]

Justification: Figure 1, Table 1, and Appendix F.4 report 99% confidence intervals over runs with different random seeds, with the number of runs stated for each experiment: percentile bootstrap intervals (2,000 resamples) in Figure 1, Student-t intervals for mean certified costs, and bootstrap intervals for cost differences and times to quality. Diff, and ValueGap-CG in Scenarios 1–2, are deterministic.

Guidelines:

- The answer [N/A] means that the paper does not include experiments.
- The authors should answer [Yes] if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
- The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
- The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
- The assumptions made should be given (e.g., Normally distributed errors).
- It should be clear whether the error bar is the standard deviation or the standard error of the mean.
- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g., negative error rates).
- If error bars are reported in tables or plots, the authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments compute resources**

Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

Answer: [Yes]

Justification: Appendix F.1 reports the compute platform of the TNTP experiments (AMD EPYC 9334, 32 cores / 64 threads, 371 GB RAM, RHEL 9.4), thread allocation, and the per-run wall-clock cap of 10 hours; Figure 1 and Appendix F.4 report the time per outer iteration, the total optimizer time of each method, and the times to quality. The synthetic TSP study (Appendix H.3) ran on a laptop. The reported experiments comprise a few hundred runs, and we estimate that including preliminary tuning and discarded runs the total project compute is roughly an order of magnitude above the reported runs.

Guidelines:

- The answer [N/A] means that the paper does not include experiments.
- The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
- The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
- The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code of ethics**

Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics `https://neurips.cc/public/EthicsGuidelines`?

Answer: [Yes]

Justification: The work uses public benchmark network data and involves no human subjects.

Guidelines:

- The answer [N/A] means that the authors have not reviewed the NeurIPS Code of Ethics.
- If the authors answer [No], they should explain the special circumstances that require a deviation from the Code of Ethics.
- The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader impacts**

Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

Answer: [Yes]

Justification: Our methods target Stackelberg control problems such as toll pricing, EV charging coordination, and multicast routing. Improved tools for designing such mechanisms can yield positive impacts (reduced congestion, lower aggregate travel time, better-utilized infrastructure), but the same machinery could also be applied to objectives that shift cost burdens onto specific user groups or maximize operator revenue at users' expense; equity in toll incidence and access for transit-dependent populations is a relevant concern when these methods are instantiated. We view selecting socially appropriate leader objectives, and auditing realized routing outcomes for distributional effects, as essential when this line of work is deployed in practice.

Guidelines:

- The answer [N/A] means that there is no societal impact of the work performed.
- If the authors answer [N/A] or [No], they should explain why their work has no societal impact or why the paper does not address societal impact.
- Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.
- The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate Deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pre-trained language models, image generators, or scraped datasets)?

Answer: [N/A]

Justification: The paper does not release high-risk generative models, scraped datasets, or similarly misuse-prone assets; it studies optimization methods on public benchmark networks.

Guidelines:

- The answer [N/A] means that the paper poses no such risks.
- Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
- Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
- We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

    Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

    Answer: [Yes]

    Justification: We use NetworkX (BSD 3-Clause), Graphillion (MIT License), and PyTorch (BSD-style), each cited in Appendix F.1 and used within their license terms. The TNTP networks [Transportation Networks for Research Core Team, 2016] are made available for academic research with attribution required by the original authors; we use them only for academic research and credit the original sources accordingly.

    Guidelines:

    - The answer [N/A] means that the paper does not use existing assets.
    - The authors should cite the original paper that produced the code package or dataset.
    - The authors should state which version of the asset is used and, if possible, include a URL.
    - The name of the license (e.g., CC-BY 4.0) should be included for each asset.
    - For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
    - If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
    - For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
    - If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New assets**

    Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

    Answer: [Yes]

    Justification: The released code (Item 5) is documented in its README, which describes its contents, installation, tests, and how to run ZOS and certify its iterates. The paper does not release new datasets, pretrained models, or benchmark assets.

    Guidelines:

    - The answer [N/A] means that the paper does not release new assets.
    - Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
    - The paper should discuss whether and how consent was obtained from people whose asset is used.
    - At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and research with human subjects**

    Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

    Answer: [N/A]

    Justification: The paper does not involve crowdsourcing or human-subject experiments.

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
    - Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
    - According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

    Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

    Answer: [N/A]

    Justification: The paper does not involve human subjects, so IRB approval is not applicable.

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
    - Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
    - We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
    - For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage**

    Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does *not* impact the core methodology, scientific rigor, or originality of the research, declaration is not required.

    Answer: [N/A]

    Justification: LLMs are not part of the core methodology, theory, or experiments of this work.

    Guidelines:

    - The answer [N/A] means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
    - Please refer to our LLM policy in the NeurIPS handbook for what should or should not be described.