Stackelberg Equilibrium Seeking

Silvia Cianchi, **Sergio Grammatico**

Delft Center for Systems and Control
TU Delft
KTH Vetenskap och Konst

Game On! Seminar @ KTH

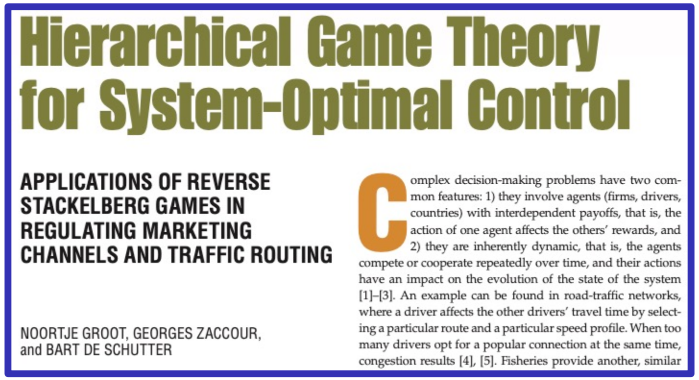
May 12, 2026

Outline

1. Introduction

Hierarchical Games for System-Optimal Control

Groot, Zaccour, De Schutter, IEEE CSM 2017

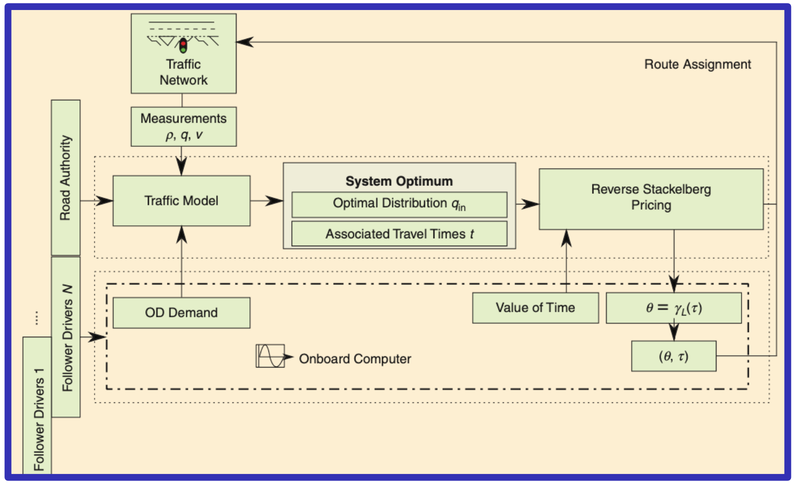


Hierarchical Game Theory for System-Optimal Control

APPLICATIONS OF REVERSE STACKELBERG GAMES IN REGULATING MARKETING CHANNELS AND TRAFFIC ROUTING

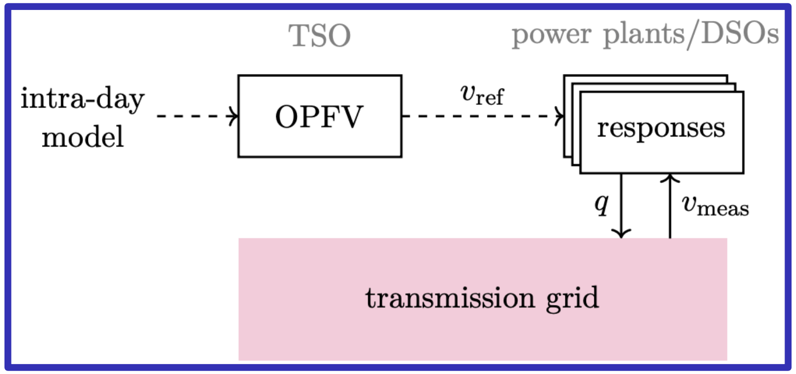
NOORTJE GROOT, GEORGES ZACCOUR, and BART DE SCHUTTER

Complex decision-making problems have two common features: 1) they involve agents (firms, drivers, countries) with interdependent payoffs, that is, the action of one agent affects the others' rewards, and 2) they are inherently dynamic, that is, the agents compete or cooperate repeatedly over time, and their actions have an impact on the evolution of the state of the system [1]–[3]. An example can be found in road-traffic networks, where a driver affects the other drivers' travel time by selecting a particular route and a particular speed profile. When too many drivers opt for a popular connection at the same time, congestion results [4], [5]. Fisheries provide another, similar



Traffic networks

Power grids



Jiang, Bolognani, Belgioioso, Voltage support procurement in transmission grids: Incentive design via online bilevel games," IEEE CDC 2025

This Talk

Cianchi, Sanjab, **SG**, "Induced Stackelberg equilibrium seeking via iterative Tikhonov regularization", arXiv, 2026

Cianchi, Baghbadorani, Sanjab, **SG**, "Learning-based Stackelberg equilibrium seeking with application to demand-side energy management", arXiv, 2026

Silvia Cianchi

Reza Rahimi Baghbadorani

Stackelberg Equilibrium Problem

- N follower agents, 1 leader agent
- Each follower $i \in \mathcal{I}$ has decision variable $\mathbf{x}_i \in \mathbb{R}^n$
- The leader has decision variable $\mathbf{y} \in \mathbb{R}^m$
- The followers play a (lower-level) Nash game: $\forall\, i \in \mathcal{I} : \min_{\mathbf{x}_i \in \mathcal{X}_i(\mathbf{x}_{-i})} J_i(\mathbf{x}_i, \mathbf{x}_{-i}; \mathbf{y})$
- v-GNEs of the followers: $\mathcal{E}(\mathbf{y}) := \mathrm{SOL}(F(\cdot;\mathbf{y}), \Omega)$

Stackelberg Equilibrium Problem (+)

- **Stackelberg game**:

$$\min_{\mathbf{y}\in\mathcal{Y}} J_0(\mathbf{y},\mathbf{x}) \quad \text{s.t. } \mathbf{x}\in\mathcal{E}(\mathbf{y})$$

- If the followers' reaction is unique, i.e. $\mathbf{x}^*(\mathbf{y}) := \{\mathcal{E}(\mathbf{y})\}$

  then Stackelberg game = bilevel optimization problem: $\min_{\mathbf{y}\in\mathcal{Y}} J_0(\mathbf{y},\mathbf{x}^*(\mathbf{y}))$

- Global minima are Stackelberg equilibria

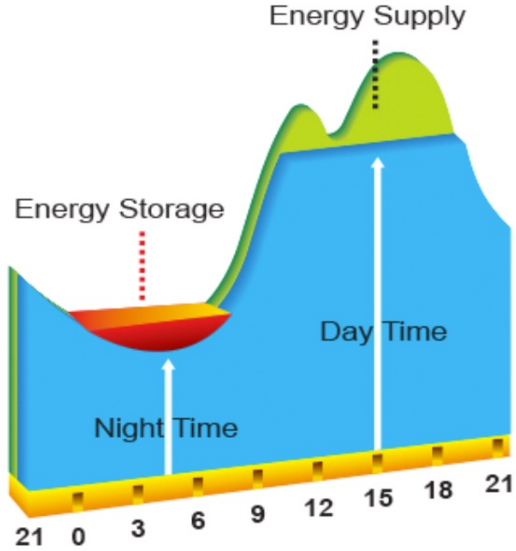
  due to nonconvexity, stationary points are "good enough"

Illustrative application in smart energy systems

Hierarchical demand side management

- DSO (leader): $\begin{cases} \max_{\mathbf{y}} & \text{revenue}(\mathbf{y}, \sum_{j \in \mathcal{J}} \mathbf{x}_j) \\ \text{s.t.} & \mathbf{x} \in \mathcal{E}(\mathbf{y}) \end{cases}$

- Prosumers (followers): $\begin{cases} \min_{\mathbf{x}_i} & \text{discomfort}_i(\mathbf{x}_i) + \text{cost}_i(\mathbf{x}_i, \mathbf{y}) \\ \text{s.t.} & \mathbf{x}_i \in \{\text{limits}\left(\sum_{j \in \mathcal{J}} \mathbf{x}_j\right)\} \end{cases}$



Cianchi, Sanjab, **SG**, A two-part pricing mechanism for demand side management, CoDIT, 2025

2. Induced Stackelberg Equilibrium

Induced Stackelberg equilibrium: Setting & Definition

- Followers reaction set-valued: Stackelberg equilibrium well defined?

- uniqueness via strongly convex selection: $\mathbf{x}_\phi^*(\mathbf{y}) := \arg\min_{\mathbf{x}\in\mathcal{E}(\mathbf{y})} \phi(\mathbf{x})$

- **Induced Stackelberg equilibrium** = global minima of $\min_{\mathbf{y}\in\mathcal{Y}} J_0(\mathbf{y}, \mathbf{x}_\phi^*(\mathbf{y}))$

  (due to nonconvexity, we seek a stationary point)

Cianchi, Sanjab, **SG**, "Induced Stackelberg equilibrium seeking via Iterative Tikhonov regularization," arXiv, 2026

Induced Stackelberg Equilibrium: Toy Example

- Low-level game:

$$\begin{cases} \min_{x_1 \in \mathcal{X}_1} & J_1(x_1, x_2; y) := \frac{1}{2}(y+\epsilon)x_1^2 + x_1 x_2 \\ \min_{x_2 \in \mathcal{X}_2} & J_2(x_2, x_1; y) := \frac{1}{2}x_2^2 + (y+\epsilon)x_1 x_2 \end{cases}$$

- Low-level v-GNEs:

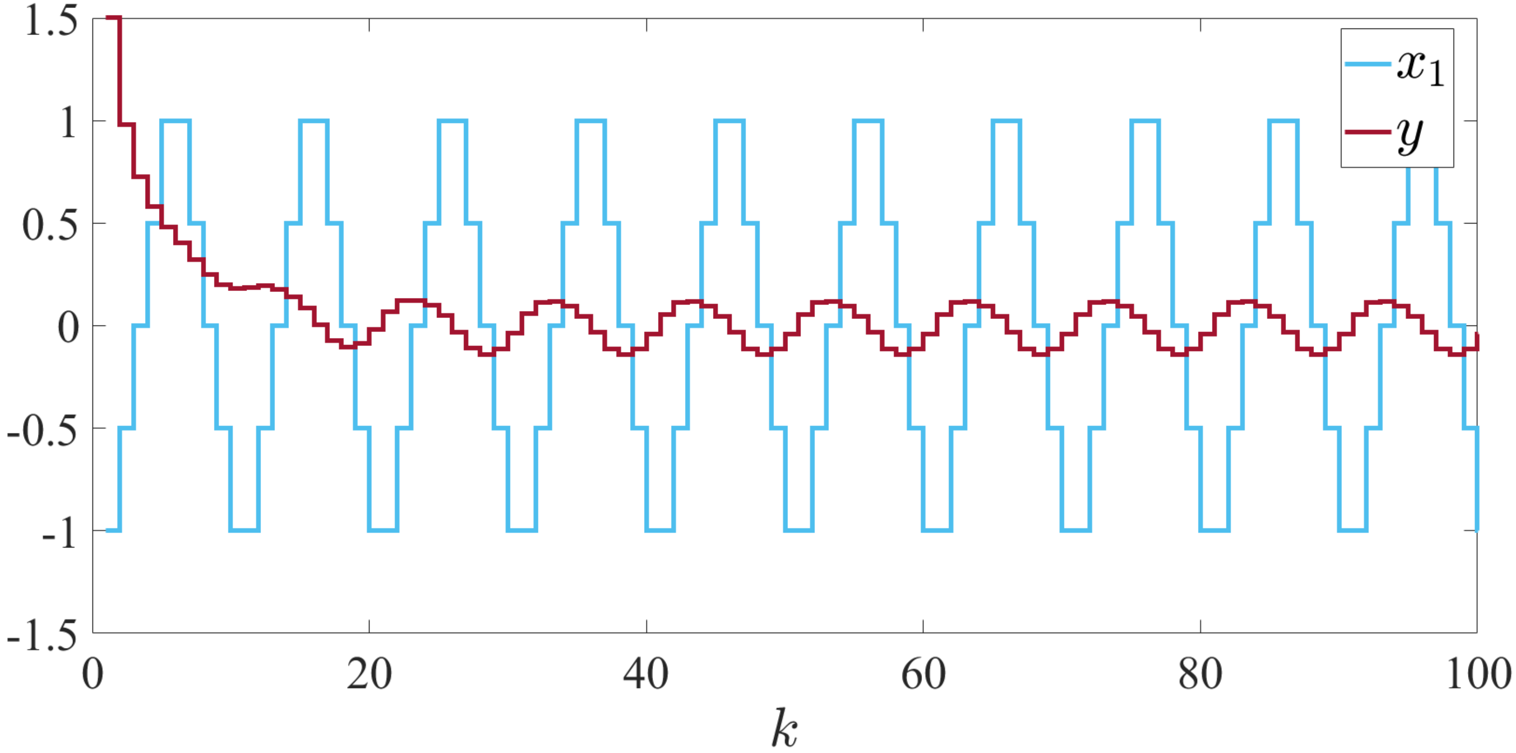
$$\mathcal{E}(y) = \{\mathbf{x} \in \mathrm{int}(\mathcal{X}_1) \times \mathrm{int}(\mathcal{X}_2) \mid x_2 = -(y+\epsilon)x_1\}$$

- High-level optimization:

$$\begin{cases} \min_{y} & J_0(y, \mathbf{x}) := y^2 + y(x_1 + x_2) \\ \text{s.t.} & \mathbf{x} \in \mathcal{E}(y) \end{cases}$$

Induced Stackelberg Equilibrium: Toy Example (+)

- Non-uniqueness of followers v-GNE may generate persistent oscillations!



Induced Stackelberg Equilibrium

- Unique selection of followers v-GNE via **iterative Tikhonov regularization**

- Followers: $\forall\, i \in \mathcal{I}: \min_{\mathbf{x}_i \in \mathcal{X}_i(\mathbf{x}_{-i})} J_i(\mathbf{x}_i, \mathbf{x}_{-i}; \mathbf{y}) + \beta\phi(\mathbf{x})$

- vanishing regularization: $(\mathbf{x}_\beta(\mathbf{y}))_{\beta \to 0}$

Cianchi, Sanjab, **SG**, "Induced Stackelberg equilibrium seeking via iterative Tikhonov regularization," arXiv, 2026

Benenati, Ananduta, **SG**, "Optimal selection and tracking of generalized Nash equilibria in monotone games," TAC, 2023

Induced Stackelberg Equilibrium Seeking

Algorithm 1 Zeroth-order induced SE seeking

Data: Time horizon K, Initial conditions $\mathbf{y}_0 \in \mathbb{R}^m$, Step sizes (η_k), Perturbation radius (δ_k), Regularization parameters (β_k)

```
for k = 1 to K − 1 do
    Sample v_k ∼ Unif(S(R^T));
    Assign ŷ_k = y_k + v_k δ_k;
    Collect x_k = x_{β_k}(y_k);
    Collect x̂_k = x_{β_k}(ŷ_k);
    Compute ĝ_k = (m/δ_k)(J_0(y_k, x_k) − J_0(ŷ_k, x̂_k)) v_k;
    Update y_{k+1} = y_k − η_k ĝ_k;
end for
```

Leader agent sets parameter of incentive and queries followers

Cianchi, Sanjab, **SG**, "Induced Stackelberg equilibrium seeking via iterative Tikhonov regularization," arXiv, 2026

Maheshwari, Cheng, Sastry, Ratliff, Mazumdar, "Follower agnostic learning in Stackelberg games," CDC 2024

Convergence Theorem

- Expected distance from stationary point decreases (Theorem 1)

Let $\eta_k = \bar{\eta}(k+1)^{-1/2}m^{-1}, \delta_k = \bar{\delta}(k+1)^{-1/4}m^{-1/2}$, such that $\bar{\eta} \leq m/2\tilde{\ell}$, and $\beta_k = \bar{\beta}(k+1)^{-\alpha}$ with $\alpha > \dfrac{1}{2}$.
Then, the iterations $(\mathbf{y}_k)_{k\in[K]}$ of Algorithm 1 are such that

$$\frac{1}{\sum_{k=1}^{K}\eta_k}\sum_{k=1}^{K}\eta_k\mathbb{E}\left[\left\|\nabla J_0(\mathbf{y}_k, \mathbf{x}_\phi^*(\mathbf{y}_k))\right\|^2\right] \leq \mathcal{U}(K)$$

Cianchi, Sanjab, **SG**, "Induced Stackelberg equilibrium seeking via iterative Tikhonov regularization," arXiv, 2026

Induced Stackelberg Equilibrium: Application example

- Followers = users in daily P2P energy trading $\quad \mathbf{x}_{i,h} = \left[p^{\mathrm{g}}_{i,h}, p^{\mathrm{mg}}_{i,j}, p^{\mathrm{st}}_{i,j}, \{p^{\mathrm{tr}}_{(i,j),h}\}_{j\in\mathcal{N}_i}, \theta_{i,h}\right]$

$$J_i(\mathbf{x};\mathbf{y}) = \sum_{h=1}^{24} f^{\mathrm{g}}_{i,h}(p^{\mathrm{g}}_{i,h}) + f^{\mathrm{tr}}_{i,h}(\{p^{\mathrm{tr}}_{(i,j),h}\}_{j\in\mathcal{N}_i};\mathbf{y}) + f^{\mathrm{mg}}_{i,h}(p^{\mathrm{mg}}_{i,h}, p^{\mathrm{mg}}_{-i,h})$$

- Leader = community manager setting trading prices $\quad J_0(\mathbf{y},\mathbf{x}) = \sum_{i\in\mathcal{I}} J_i(\mathbf{x},\mathbf{y}) + \lambda\|\mathbf{y}-\bar{\mathbf{y}}\|^2$

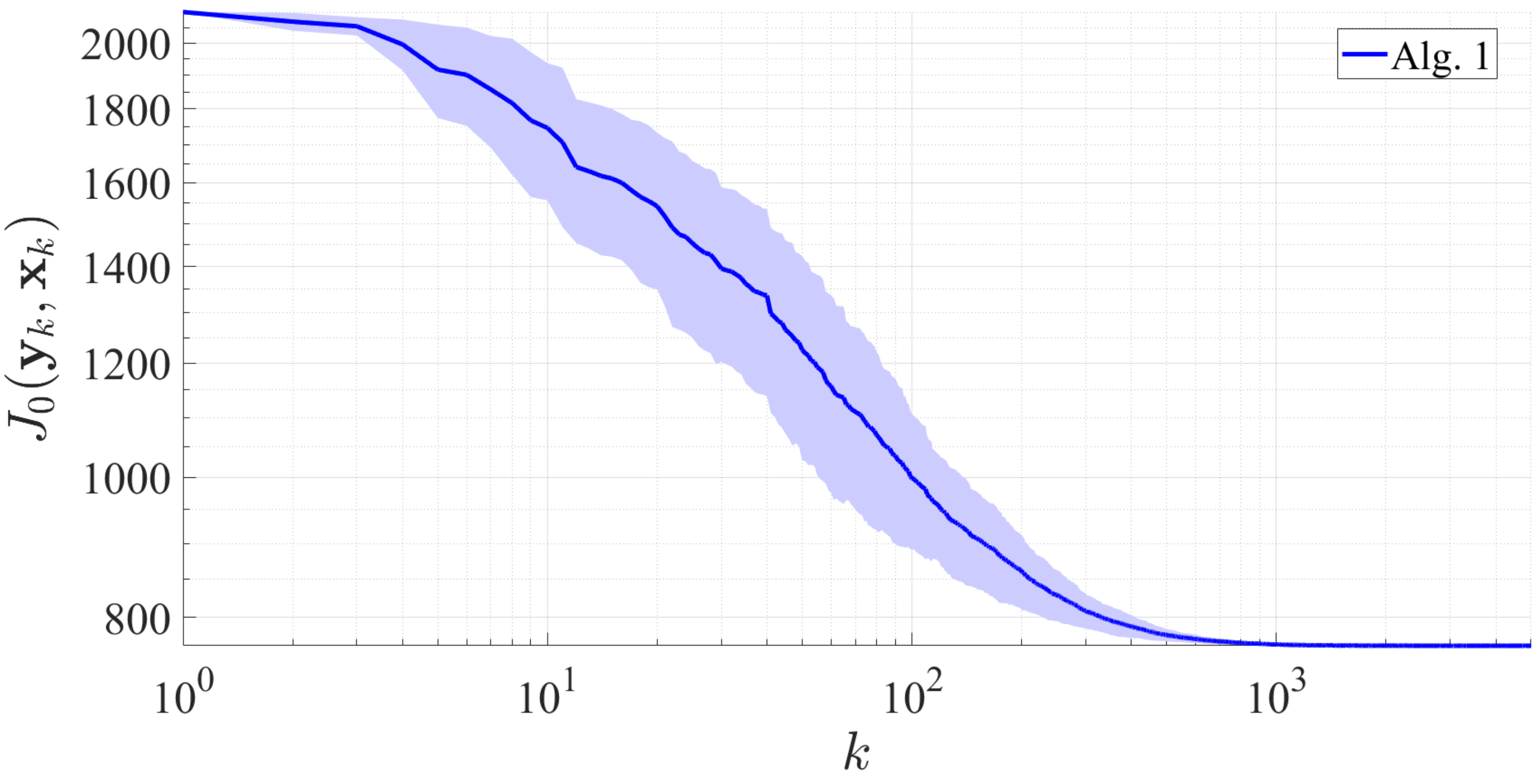
- Leader incentivizes users to reduce grid load/imbalances & enhance renewable energy

$$\phi(\mathbf{x}) = \sum_{h=1}^{24} \left\|\mathbf{p}^{\mathrm{g}}_h - \bar{\mathbf{p}}^{\mathrm{g}}\right\|^2 + \left\|\mathbf{p}^{\mathrm{mg}}_h\right\|^2 + \left\|\theta_h - \bar{\theta}\right\|^2 + \left\|\mathbf{p}^{\mathrm{tr}}_h\right\|^2 + \left\|\mathbf{p}^{\mathrm{st}}_h\right\|^2$$

Cianchi, Sanjab, **SG**, "Induced Stackelberg equilibrium seeking via iterative Tikhonov regularization," arXiv, 2026

Belgioioso, Ananduta, **SG**, Ocampo-Martinez, "Operationally-safe peer-to-peer energy trading in distribution grids: A game-theoretic market-clearing mechanism," IEEE TSG, 2022

Induced Stackelberg Equilibrium Seeking converges



3. Stackelberg Equilibrium Learning & Seeking

Linear-Quadratic Stackelberg Equilibrium Problem

- Followers: $$\forall i : \begin{cases} \min\limits_{\mathbf{x}_i \in \mathcal{X}_i} & J_i(\mathbf{x}_i; \mathbf{y}) := ||\mathbf{x}_i - \mathbf{r}_i||^2 + (\mathbf{p} + \mathbf{y})^\top \Pi_i A_i \mathbf{x}_i, \\ \text{s.t.} & \sum_{i=1}^N A_i \mathbf{x}_i \leq \mathbf{b} \end{cases}$$

- v-GNE set: $$\mathcal{E}(\mathbf{y}) = \{\mathbf{x}^*(\mathbf{y})\} \text{ with } \mathbf{x}^*(\mathbf{y}) = \mathrm{proj}_C(\Theta_A \mathbf{y} - \Theta_b)$$

- Leader: $$\begin{cases} \min\limits_{\mathbf{y} \in \mathcal{Y}} & J_0(\mathbf{y}, \mathbf{x}) := -\mathbf{y}^\top \Lambda A \mathbf{x} + \mu(\mathbf{1}^\top \mathbf{y} - \bar{y})^2 + \lambda \|\mathbf{y}\|^2 \\ \text{s.t.} & \mathbf{x} \in \mathcal{E}(\mathbf{y}) \end{cases}$$

Cianchi, Baghbadorani, Sanjab, **SG**, "Learning-based Stackelberg equilibrium seeking with application to demand-side energy management," arXiv, 2026

Stackelberg Equilibrium Learning & Seeking

- Stackelberg game: $\min_{\mathbf{y} \in \mathcal{Y}} J_0(\mathbf{y}, \mathbf{x}^*(\mathbf{y}))$
- Leader interacts with followers via queries (zeroth order)
- Leader knows parametric structure of followers' reaction: $\mathbf{x}^*(\mathbf{y}; \Theta_A, \Theta_b) := \mathrm{proj}_C(\Theta_A \mathbf{y} - \Theta_b)$

Learning-Based Stackelberg Equilibrium Seeking:

1. Zeroth-order algorithm with datase collection
2. Parameter estimation
3. Solution of the estimated Stackelberg game

Zeroth order Algorithm with Data Collection

1. Zeroth-order algorithm with data collection

Algorithm 2 Zeroth-order algorithm

Data: Number of iterations K, Initial condition $\mathbf{y}_0 \in \mathbb{R}^T$, Step sizes $(\eta_k)_{k\in[K]}$, Perturbation radius $(\delta_k)_{k\in[K]}$

for $k = 1$ to $K-1$ **do**
- Sample $\mathbf{v}_k \sim \text{Unif}(\mathcal{S}(\mathbb{R}^T))$;
- Assign $\tilde{\mathbf{y}}_k = \mathbf{y}_k + \mathbf{v}_k\delta_k$;
- Collect $\mathbf{x}_k = \mathbf{x}^*(\mathbf{y}_k)$;
- Collect $\tilde{\mathbf{x}}_k = \mathbf{x}^*(\tilde{\mathbf{y}}_k)$;
- Compute $\hat{\mathbf{g}}_k = \frac{T}{\delta_k}(J_0(\mathbf{y}_k, \mathbf{x}_k) - J_0(\tilde{\mathbf{y}}_k, \tilde{\mathbf{x}}_k))\mathbf{v}_k$;
- Update $\mathbf{y}_{k+1} = \mathbf{y}_k - \eta_k\hat{\mathbf{g}}_k$;

end for

Dataset collection: $\mathcal{D} := \{(\mathbf{y}_k, \mathbf{x}_k)\}_k$

Convergence Theorem

- Expected distance depends on initial function value (Theorem 1)

Let $\eta_k = \bar{\eta}(k+1)^{-1/2}T^{-1}, \delta_k = \bar{\delta}(k+1)^{-1/4}T^{-1/2}$, such that $\bar{\eta} \leq T/2\tilde{\ell}$. Then, the iterations $(\mathbf{y}_k)_{k\in[K]}$ of Algorithm 2 are such that

$$\frac{1}{\sum_{k=1}^{K}\eta_k}\sum_{k=1}^{K}\eta_k\mathbb{E}\left[\|\nabla J_0(\mathbf{y}_k)\|^2\right] \leq \mathcal{U}(J_0(\mathbf{y}_0), K, \bar{\eta}, \bar{\delta}).$$

Cianchi, Baghbadorani, Sanjab, **SG**, "Learning-based Stackelberg equilibrium seeking with application to demand-side energy management," arXiv, 2026

Inverse Problem for Parameter Estimation

2. Parameter estimation (to improve next initial condition, hence convergence)

$$\begin{cases} \min\limits_{\hat{\Theta}_A, \hat{\Theta}_b} & \frac{1}{2K} \sum_{k=1}^{2K} \left\| J_0(\mathbf{y}_k, \mathbf{x}_k) - J_0(\mathbf{y}_k, \mathbf{x}^*(\mathbf{y}_k; \hat{\Theta}_A, \hat{\Theta}_b)) \right\|^2 \\ \text{s.t.} & \mathbf{x}^*(\mathbf{y}_k; \hat{\Theta}_A, \hat{\Theta}_b) = \text{proj}_{\mathcal{C}}(\hat{\Theta}_A \mathbf{y}_k - \hat{\Theta}_b) \end{cases} \longrightarrow \mathbf{x}^*\left(\mathbf{y}; \hat{\Theta}_A, \hat{\Theta}_b\right)$$

(Theorem 2, informal):

> For each data point $(\mathbf{y}_k, \mathbf{x}_k) \in \mathcal{D}$, there exists a region, $\mathcal{B}(\mathbf{y}_k, \rho)$, where the estimation $\mathbf{x}^*(\mathbf{y}; \hat{\Theta}_A, \hat{\Theta}_b)$ is sufficiently accurate.

Cianchi, Baghbadorani, Sanjab, **SG**, "Learning-based Stackelberg equilibrium seeking with application to demand-side energy management," arXiv, 2026

Estimated Stackelberg Game solved by Leader agent

3. Solution of the estimated Stackelberg game

Leader minimizes $J_0\left(\mathbf{y}; \mathbf{x}^*\left(\mathbf{y}; \hat{\Theta}_A, \hat{\Theta}_b\right)\right)$

Algorithm 3 Difference of Convex Programs

Data: Initial condition $\mathbf{z}_0$

1: **while** $t \leq t_{\max}$ **do**

2: Solve $\mathbf{z}_{t+1} \in \arg\min_{\mathbf{z} \in \mathcal{Z}} \quad g(\mathbf{z}) - \langle \nabla p(\mathbf{z}_t), \mathbf{z} \rangle$

3: **end while**

Return: $\hat{\mathbf{y}} = \mathbf{y}_{t_{\max}}$

keep iterates in trust region!

Tao, An, "Convex analysis approach to DC programming: theory, algorithms and applications," Acta Mathematica, 1997

Stackelberg Equilibrium Learning & Seeking (+)

Tradeoff between learning and seeking:

Long data collection means slow convergence (ZOM) but also accurate learning

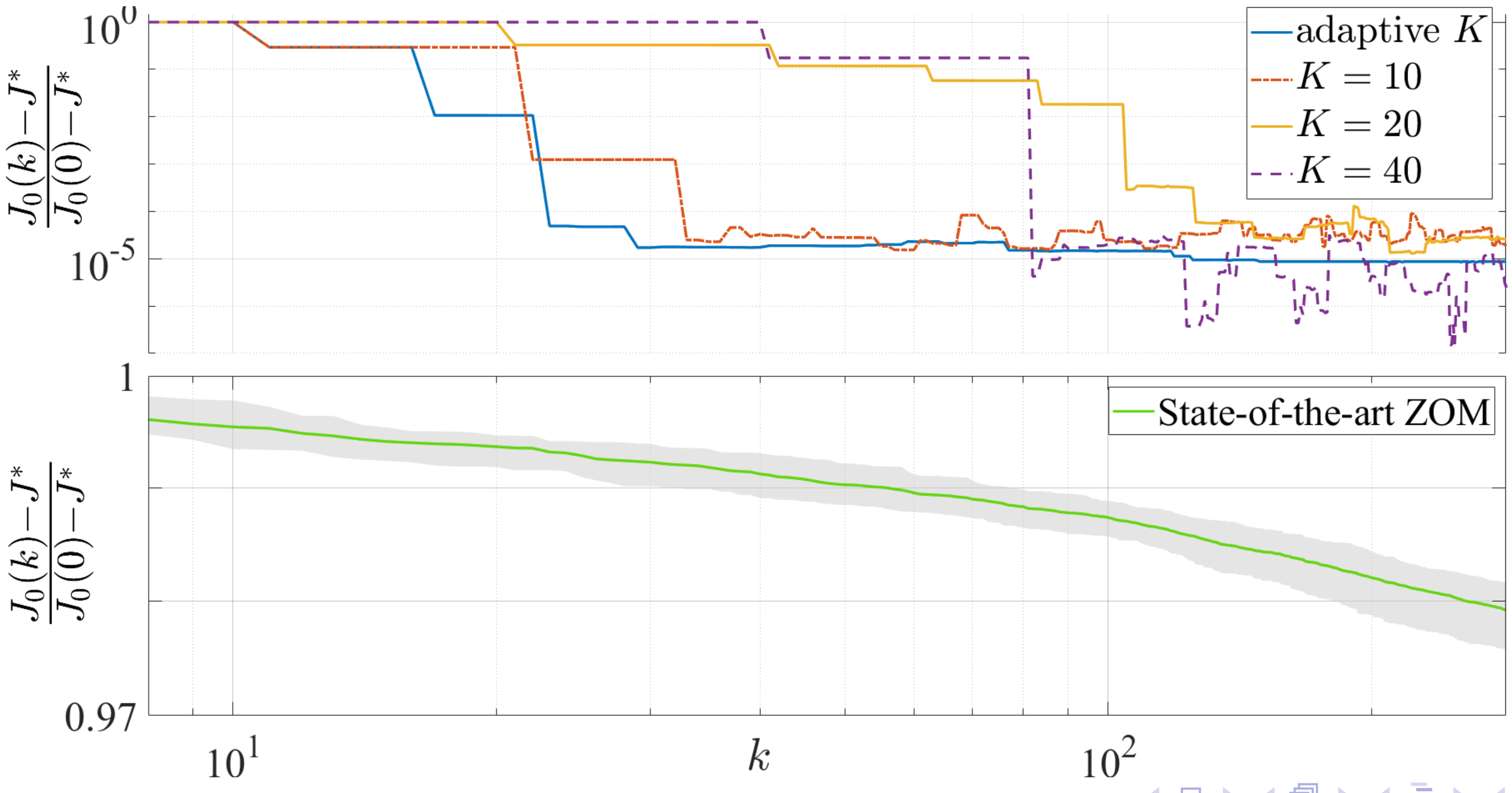
Adaptive K (Remark 1):

- Zeroth-order algorithm runs for K_1 iterations
- Next round of zeroth order algorithm: at least K_2 iterations, based on convergence rate

Cianchi, Baghbadorani, Sanjab, **SG**, "Learning-based Stackelberg equilibrium seeking with application to demand-side energy management," arXiv, 2026

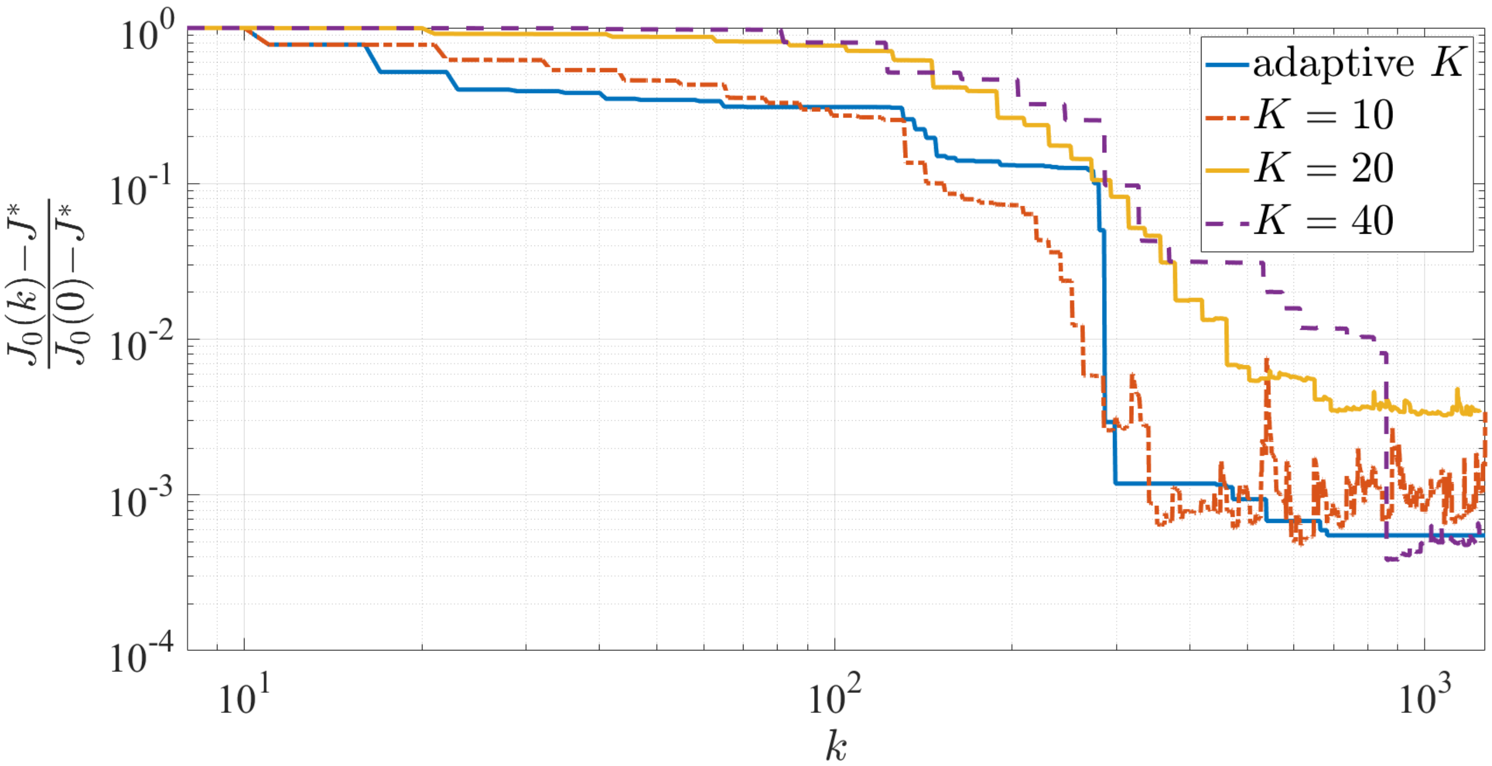
Learning & Seeking converges fast

- Known feasible set

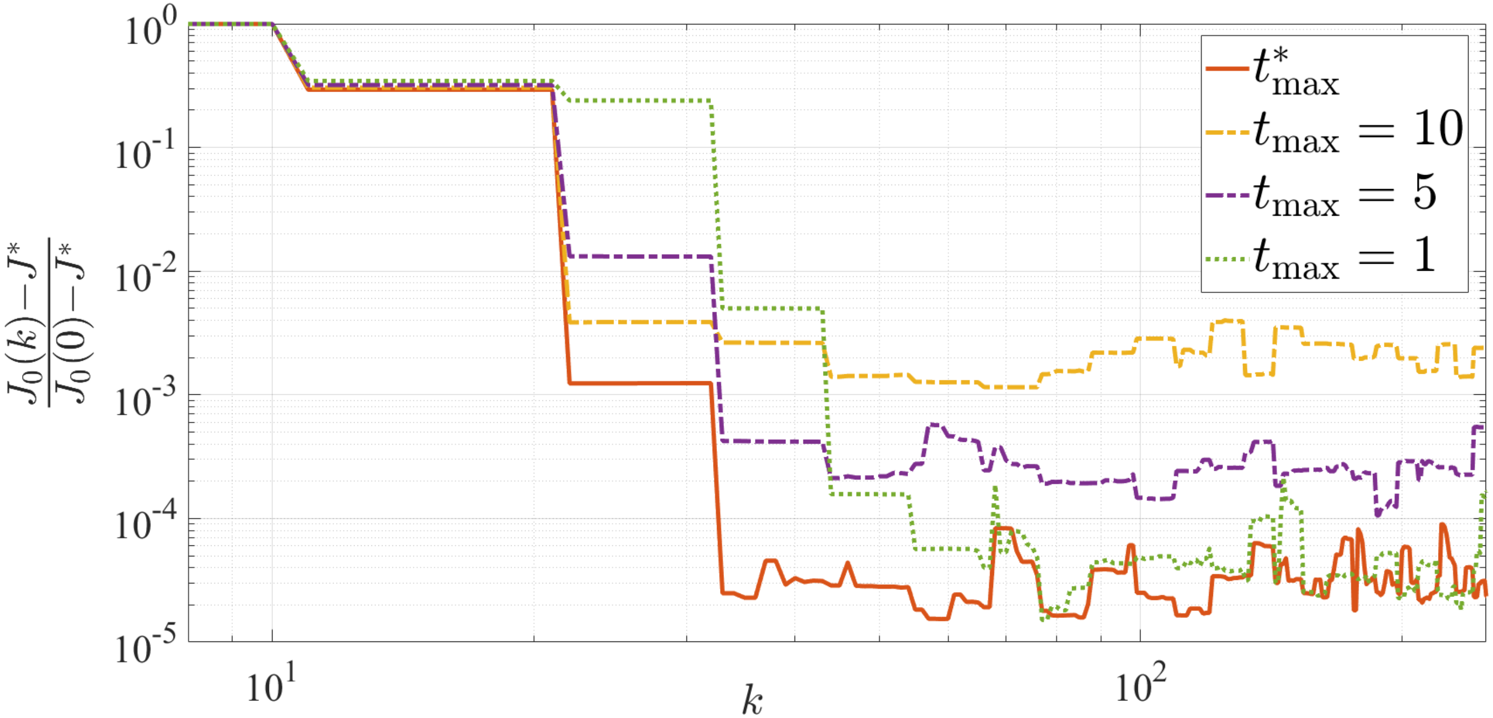


Effect of partially unknown feasible set

- Feasible set learned from collected data



Effect of unknown trust region



4. Conclusion & Outlook

Conclusion and Outlook

- Stackelberg games model hierarchical decision making, leading to bilevel and nonconvex optimization problems
- Equilibrium selection via followers regularization makes Stackelberg equilibrium problems well posed
- Zeroth order Stackelberg equilibrium seeking is slow; Hybrid learning-and-seeking accelerates convergence significantly
- For non-quadratic Stackelberg games, kernel-based learning sounds promising

Acknowledgements

- Collaborators

- Funding agencies

Sergio Grammatico
Associate Professor, Optimal Control & Dynamic Game Theory, TU Delft

s.grammatico@tudelft.nl

sites.google.com/site/**grammaticosergio**

Thank you for your kind attention!