Delft University of Technology

Delft Center for Systems and Control

Technical report 12-022

Reverse Stackelberg games, Part II: Results and open issues*

N. Groot, B. De Schutter, and H. Hellendoorn

If you want to cite this report, please use the following reference instead:
N. Groot, B. De Schutter, and H. Hellendoorn, "Reverse Stackelberg games, Part II: Results and open issues," *Proceedings of the 2012 IEEE International Conference on Control Applications*, Dubrovnik, Croatia, pp. 427–432, Oct. 2012.

Delft Center for Systems and Control
Delft University of Technology
Mekelweg 2, 2628 CD Delft
The Netherlands
phone: +31-15-278.24.73 (secretary)
URL: `https://www.dcsc.tudelft.nl`

Reverse Stackelberg Games, Part II: Results and Open Issues

Noortje Groot, Bart De Schutter, and Hans Hellendoorn

Abstract— **A reverse Stackelberg game formulation can be adopted as a means to structure hierarchical control problems. Here, a leader player announces a mapping of the follower's decision space into the leader's decision space, after which a follower player determines his optimal decision variable. In the companion paper of this survey entitled 'Reverse Stackelberg Games, Part I: Basic Framework', an introduction to the game has been provided with a clarification of the description of this game as it is studied in different research areas. In the current paper, an overview is provided of several main developments in the field. These contributions are categorized according to several aspects that are inherent to the formulation of the game, and they are briefly analyzed. Finally, several open issues are brought forward that are relevant for further research.**

I. Introduction

Hierarchical control approaches can be applied to large-scale problems that are too complex to solve in a centralized manner, or to problems in which a natural hierarchy exists [1]. In order to deal with such problems, the Stackelberg game [2] can be used as a framework for optimization. In this game, a leader and follower player act sequentially in determining their decision variables. An example of such hierarchical control problems can be found in road tolling where vehicles make their route choices based on the tolls set by a road authority.

The current survey is particularly aimed at the reverse Stackelberg game [3], which is also known as incentives [4], [5] in a control-theoretic framework and more recently as an inverse Stackelberg game [6], [7]. Compared to the original Stackelberg game, in the reverse game, the type of leader action is generalized from making a direct decision to determining a *function* that maps the follower's decision space into the leader's decision space[1]. Thus, although the leader remains the first to act by proposing a leader function, her actual decision variable will not be determined until the follower acts and proposes his decision or control input [3]. This game is useful for applications like road tolling, where in order to minimize congestion, the road authority (leader) determines a toll function that is dependent on the actual flow of vehicles (followers) on the relevant road sections [8]. In this way, the leader has a larger influence over the followers to reach her objective. Nonetheless, in order to apply the – in general complex – reverse Stackelberg game to real-life control problems, several steps still need to be made.

N. Groot, B. De Schutter, and H. Hellendoorn are with the Delft Center for Systems and Control, Delft University of Technology, Delft, The Netherlands. E-mail: n.b.groot@tudelft.nl

[1]Since the Stackelberg game is a special case of the reverse Stackelberg game, i.e., in case the leader function is a constant, several remarks made in this survey regarding the reverse game also hold for the original game.

In the companion paper entitled 'Reverse Stackelberg Games, Part I: Basic Framework', the appearance of the reverse Stackelberg game in the field of control as well as in e.g., an economic context was pointed out and several areas of application were provided. The aim of the current paper is to give a coherent overview of the main results in this particular game. Moreover, we bring forward several open issues that point at the potential for continued research on the reverse Stackelberg game.

This paper is structured as follows. In Section II an overview is provided of main results in current literature, classified along several characteristics that are inherent to the definition of a reverse Stackelberg game, i.e., contributions are considered that (1) involve either static or dynamic cases; (2) look into continuous-time differential games; (3) deal with stochastic scenarios; (4) consider partial, nonnested information; (5) perform a sensitivity analysis; and that (6) consider multi-level games with multiple players on each layer. Finally, an elaborate listing of open issues is presented in Section III.

II. An Overview of Results

In the current section, an overview of contributions in the area of reverse Stackelberg games is provided, categorized into several aspects as is also depicted in Fig 1. Each reverse Stackelberg game includes with the general description as provided in Section II of the companion paper a specification of (1) time elements, e.g., the duration of the game, (2) leveling, and (3) information and uncertainty [4], [5].

A. Static Versus Dynamic

There are not so many results from a control theoretic perspective that consider the static reverse Stackelberg game. As is also mentioned in [7], a legitimate reason for studying

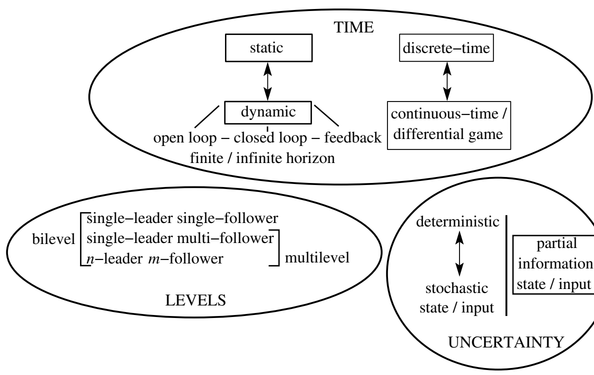

Fig. 1. Overview of characteristics within the reverse Stackelberg Game

the dynamic case is that it more often occurs in the real-life settings that represent the intended application areas of the game. On the contrary, results within the theory of incentives and especially multilevel programming are often based on single-stage problems [9], [10].

Nonetheless, an introduction to the static reverse Stackelberg game can be found in [4], as well as in [11] where a static affine leader function is presented, motivated by a situation with the regulating government as a leader that strives to achieve Pareto optimality[2] while multiple followers (firms) play according to their Nash strategies[3]. Since objective functions in [11] are assumed to be quadratic hence strictly convex, an optimal affine strategy can be computed.

Although the static leader functions in [4] are of the form $\gamma_{\mathrm{L}} : \Omega_{\mathrm{F}} \to \Omega_{\mathrm{L}}$, in the multi-stage case the follower decision variable is replaced by a state-dependent leader function, which is only indirectly dependent on the follower input. Also in [14] derivations of such state-dependent leader strategies that are nonetheless called incentives can be found. As explained in the companion paper, such strategies could be seen as closed-loop or feedback strategies of the original Stackelberg game, i.e., they are different from the leader function as defined in the reverse Stackelberg game.

For the static and dynamic, open-loop case, a sufficient condition was derived in [15] for the existence of an optimal affine leader

$$u_{\mathrm{L}} = \gamma_{\mathrm{L}}(u_{\mathrm{F}}) = u_{\mathrm{L}}^{\mathrm{d}} + B(u_{\mathrm{F}} - u_{\mathrm{F}}^{\mathrm{d}}), \tag{1}$$

with $B : \Omega_{\mathrm{F}} \to \Omega_{\mathrm{L}}$ a linear operator, i.e.: $B^* \nabla_{u_{\mathrm{L}}} \mathcal{J}_{\mathrm{F}}(u_{\mathrm{L}}^{\mathrm{d}}, u_{\mathrm{F}}^{\mathrm{d}}) = -\nabla_{u_{\mathrm{F}}} \mathcal{J}_{\mathrm{F}}(u_{\mathrm{L}}^{\mathrm{d}}, u_{\mathrm{F}}^{\mathrm{d}})$, with $B^* : \Omega_{\mathrm{L}} \to \Omega_{\mathrm{F}}$ the adjoint of B. An operator B now exists if $\nabla_{u_{\mathrm{L}}} \mathcal{J}_{\mathrm{F}}(u_{\mathrm{L}}^{\mathrm{d}}, u_{\mathrm{F}}^{\mathrm{d}}) \neq 0$. However, these results are restricted to games where $\mathcal{J}_{\mathrm{F}}$ is convex and locally strictly convex as well as twice continuously differentiable.

Further, in [16] the optimal affine leader function is proven to be unique for LQ dynamic games with u_{F} scalar; for $n_{\mathrm{F}} > 1$, a unique strategy can be found under some conditions regarding the system matrices additional to those in [15]. Algebraic expressions for this unique optimal affine function are derived for both the static and dynamic case.

More recently, in [17] the analysis of the affine incentive structure for a linear-quadratic discrete-time system is continued (both for the finite and infinite horizon case) but instead of using a leader function that is in fact dependent on u_{F}, also there, state feedback is applied for both players. We therefore do not consider this case as a truly reverse Stackelberg game. Such dynamic state-feedback strategies in linear-quadratic (LQ) settings have been studied before in continuous time in e.g., [18].

It should be noted that the dynamic game with a linear state equation and quadratic cost functions is widely used as an illustrative example, e.g., amongst several other references stated in this survey, in [11], [14], [17]. Moreover, existence results for an optimal (affine) leader function also rely on this specific LQ case, as in [15].

B. Continuous-Time Differential Games

While so far the discrete-time reverse Stackelberg game has been considered in this survey, the results can also be extended to the continuous-time differential game. Also in (Stackelberg) differential games, the LQ problem structure is a popular one [19], [20], [21]; it can be written as follows:

$$\dot{x}(t) = A(t)x(t) + B_{\mathrm{L}}(t)u_{\mathrm{L}}(t) + B_{\mathrm{F}}(t)u_{\mathrm{F}}(t), x(t_0) = x_0$$
$$\mathcal{J}_i(u_{\mathrm{L}}, u_{\mathrm{F}}) = \frac{1}{2} x^{t_{\mathrm{f}}}(t_{\mathrm{f}}) Q_{i,t_{\mathrm{f}}} x(t_{\mathrm{f}}) + \frac{1}{2}\int_{t_0}^{t_{\mathrm{f}}} (x(t)^{t_{\mathrm{f}}} Q_i x(t) + u_i(t)^{\mathrm{T}} R_{ii} u_i(t) + u_j(t)^{\mathrm{T}} R_{ij} u_j(t))\, \mathrm{d}t,$$

$i, j \in \{\mathrm{L}, \mathrm{F}\}, i \neq j,\ t \in [t_0, t_{\mathrm{f}}]$, where the matrices are of appropriate dimension, and $Q_{i,t_{\mathrm{f}}} \geq 0, Q_i \geq 0, R_{ij} \geq 0, R_{ii} > 0$.

In [22] conditions are developed under which the reverse Stackelberg game with memory from stage $\tau \in [t_0, t]$ at stage $t \in [t_0, t_{\mathrm{f}}]$, has the following optimal solution:

$$u_{\mathrm{L}}(t) = \gamma_{\mathrm{L}}(u_{\mathrm{F}})(t) = u_{\mathrm{L}}^{\mathrm{d}}(t) + \int_{t_0}^{t} R(t, \tau)(u_{\mathrm{F}} - u_{\mathrm{F}}^{\mathrm{d}})(\tau)\mathrm{d}\tau, \tag{2}$$

which is the continuous-time equivalent to the affine function (1). However, in order to find a matrix $||R|| < \infty$ as required for a leader function, a necessary condition is that $||\nabla_{u_{\mathrm{F}}} \mathcal{J}_{\mathrm{F}}(u_{\mathrm{L}}^{\mathrm{d}}, u_{\mathrm{F}}^{\mathrm{d}})(t)|| < c(t_{\mathrm{f}} - t), t \to t_{\mathrm{f}}$, with c a constant [22].

In order to relax this requirement, in [20] the leader strategy with memory is instead defined by using the Lebesgue-Stieltjes integral:

$$u_{\mathrm{L}}(t) = \gamma_{\mathrm{L}}(u_{\mathrm{F}})(t) = u_{\mathrm{L}}^{\mathrm{d}}(t) + \int_{t_0}^{t} [d_\theta \eta(t, \theta)](u_{\mathrm{F}} - u_{\mathrm{F}}^{\mathrm{d}})(\theta), \tag{3}$$

with $\theta \in [t_0, t], t \in [t_0, t_{\mathrm{f}}]$ and where $\eta : [t_0, t_{\mathrm{f}}] \times \mathbb{R} \to \mathbb{R}^{(n_x + n_{\mathrm{F}})}$ represents the vector of available information. Analytically solvable necessary and sufficient conditions are obtained for the optimality of this strategy in case of convex cost functions. Different representations of $\eta(t, \theta)$ are considered in [20]; however they are dependent on the initial state x_0 rather than explicitly on $u_{\mathrm{F}}(t)$, for which the authors state that stringent conditions on the game parameters would be needed. Likewise, in [23] strategies dependent on $u_{\mathrm{F}}(t)$ and past values of x are suggested. However, the dependence on $u_{\mathrm{F}}(t)$ is not explicitly adopted in the derivation of (sufficient) conditions of [23].

In [20] also time-delay strategies are considered in which $\gamma_{\mathrm{L}}(t)$ depends linearly on $u_{\mathrm{F}}(\theta), t - \delta \leq \theta \leq t - \sigma, \delta \leq \sigma$, with $\delta > 0, \sigma \geq 0$ constants, where the integral in γ_{L} according (3) is evaluated from $t - \delta$ to $t - \sigma$. The necessary and sufficient existence conditions are shown to extend to this time-delay case, where the strategy $\gamma_{\mathrm{L}}(t)$ for $t \leq t_0 + \delta$, and $t > t + \sigma$ is distinguished between, hence γ_{L} is discontinuous.

Also in [21] conditions are developed for the existence of an optimal affine strategy in a continuous-time LQ game.

[2]No player in the Pareto equilibrium is able to unilaterally deviate from the Pareto optimal decisions without making another player worse off [12].

[3]When players act simultaneously, in the Nash equilibrium no player can improve his/her situation by unilaterally deviating from the decisions associated with the Nash equilibrium [13].

There, not only the leader function is taken to be dependent on $u_{\rm F}$; the same strategy applies to the second player in the game:

$$\gamma_i(u_j) = u_j^{\rm d} + D_j(u_i - u_i^{\rm d}), i, j = 1, 2, i \neq j,$$

with $D_j : \Omega_i \to \Omega_j$, $D_j \neq 0, j = 1, 2$ for scalar variables u_1, u_2. Although the derivation of this strategy is useful for the leader in the reverse Stackelberg game, it does not follow a true Stackelberg setting but rather a cooperative game with equivalent players.

C. Deterministic Versus Stochastic

Under stochastic reverse Stackelberg games we usually consider the case in which the state variable of the game includes random components; in general, the state variable is assumed to have a known distribution, often Gaussian with zero mean [3], [24].

In [3] the two-player, static reverse Stackelberg case has been analyzed with a randomly (Gaussian) distributed variable ξ. The leader's cost function is now defined as $\mathcal{J}_{\rm L} = {\rm E}[L_{\rm L}(\gamma_{\rm L}, u_{\rm F}, \xi)]$, with $L_{\rm L} : \Omega_{\rm L} \times \Omega_{\rm F} \times \mathcal{X} \to \mathbb{R}$, i.e., now, the expected value ${\rm E}[L_{\rm L}]$ is optimized; $\mathcal{J}_{\rm F}$ is defined accordingly.[4]

The static LQG case is examined in detail in [3] and sufficient conditions are obtained under several additional assumptions on the game parameters. In the description of how the results translate to a multistage version of the game, however, a state feedback strategy is adopted, i.e., the leader's strategy is no longer formulated as a function of the follower's decision as is applicable in a reverse Stackelberg game.

In [24] a stochastic closed-loop reverse Stackelberg game is considered with a leader function directly dependent on the current or previous value of $u_{\rm F}$, or on the partial information that the follower signals to the leader. It is shown that the three problems can be solved in a similar way, and they lead to an optimal solution in case a LQG problem is studied.

Several results on the stochastic case consider also incomplete information, as will be shown in Section II-D below.

D. Incomplete Information

Another variant of uncertainty is in the lack of complete information; in particular within the theory of incentives, different types of information asymmetry are studied.

In [3], [4] both a nested and nonnested stochastic reverse Stackelberg game is considered, where the random variable ξ is assumed to follow a Gaussian distribution. While in the nested case, the follower's information is a subset of the information that the leader possesses, in the nonnested case, the leader does not have access to all follower knowledge. In the latter case, the leader is generally unable to compute her globally optimal solution. In order to arrive at a feasible desired equilibrium, the restrictive assumption is made that $L_{\rm L}$ is in fact independent of $u_{\rm L}$ [4]. Assuming that $\Omega_{\rm F}$ is known, the desired leader solution is now defined as a $\gamma_{\rm L} : \Omega_{\rm F} \to \Omega_{\rm L}$ such that

$$\arg\min_{u_{\rm F} \in \Omega_{\rm F}} E[L_{\rm F}(\gamma_{\rm L}(u_{\rm F}), u_{\rm F}, \xi)] = \arg\min_{u_{\rm F} \in \Omega_{\rm F}} E[L_{\rm L}(u_{\rm F}, \xi)].$$

Also in [5] an overview is provided of possible incomplete information structures in a stochastic setting, both for a static and a dynamic game. There, the leader strategy is taken to be a function of the available information, which does not in all cases include $u_{\rm F}$. In other words, the case is analyzed in which the leader cannot observe the follower's decision. The focus of [5] is therefore, as in the theory of incentives, on the follower not acting truthfully in the case of incomplete information.

In a setting with multiple followers, a stochastic, random state reverse Stackelberg game is considered in [25] where the leader has access to a linear combination of the followers' actions. Here, the followers' cost functions are again strictly convex and continuously differentiable. An affine leader function is computed that is based on this random linear combination; it is shown that the performance yielded by the leader is equivalent to the performance that applies in case she would be able to observe the followers' individual actions. In [26] this result is expanded to deal with more than two levels of hierarchy, according to the technique from [4] described in Section II-F below.

In case no knowledge is available of $\mathcal{J}_{\rm F}$ or of the follower's reaction curve, an iterative learning procedure may have to be adopted to arrive at a close-to optimal leader decision [27] or leader function in the reverse case [28]. For this purpose the use of a genetic algorithm is proposed and compared with a standard gradient approach for off-line computation of an incentive strategy [28].

It should be noted that, when adopting this iterative procedure, the game would have to be repeated until convergence is reached and the resulting strategy yields a sufficient performance for the leader. This requires a setting in which a start-up period with suboptimal policies for the leader would be possible. Further, it may be more realistic in real-life control settings to enforce the follower to communicate $\mathcal{J}_{\rm F}$ to the leader rather than to perform several off-line iterations of the game in order to make up for the missing information. Additionally, it may take many iterations to reach the true optimum that is verified according known solutions to rather simple problem instances. Finally, when using the genetic approach, in general no analytical suboptimality bounds can be obtained for the leader's neither for the the follower's performance.

E. Sensitivity Analysis

Since the set of possible optimal (affine) leader functions is often nonsingular, a minimum sensitivity approach to incentive strategies is developed in [29]. In case the leader does not exactly know one or several of the parameters of $\mathcal{J}_{\rm F}$, she assumes some nominal values of the unknown parameters and based on these, a robust leader function is computed, i.e., the deviation from the nominal values is minimized. Based

[4]It should be noted that ξ is presented as a state vector in [3], defined to represent some unknown elements of the game in both the static and the dynamic case. This state vector ξ should therefore not be confused with the system state variable we use solely in the dynamic, multi-stage game.

on the results from [15], it is known that for each possible value of the unknown parameters, there exists an optimal affine leader function under the assumption of strictly convex cost functions and rational follower behavior. Next to the proposed affine leader function, additional degrees of freedom in the reduction of sensitivity can be introduced by considering also nonlinear terms in the affine strategy [29]. The least-sensitive optimal strategy is taken to belong to the twice continuously differentiable functions with bounded first and second derivatives with respect to u_{F}. It should however be noted that in [29] no explicit analysis is provided of the performance of the leader in case the assumed nominal values characterizing $\mathcal{J}_{\mathrm{F}}$ are incorrect.

In [30] the work of [29] is extended to include stochastic incentive schemes, again where some parameters characterizing the unknown part of $\mathcal{J}_{\mathrm{F}}$ vary around some nominal value and where the state is now a random variable. A smooth strategy is found that results in the desired leader solution, which solution is again based on the assumed nominal values under the assumption of strictly convex and twice continuously differentiable cost functions. Compared to the deterministic case, in the stochastic setting the follower's optimal response is proven to be minimally or even completely insensitive to variations in the unknown parameter values with respect to the nominal values. It should be noted though that this result is only possible in case the leader is assumed to have full access to the follower's information, including u_{F}.

F. Multi-Player, Multi-Level

The static reverse Stackelberg game with multiple leaders or followers is considered in [6], where it is mentioned that not much theory is available with respect to multi-player extensions to the reverse Stackelberg game. Indeed, most cases with multiple followers assume that these play a noncooperative simultaneous Nash game amongst themselves and act as one follower group in response to the single leader [11], where the leader strategy is of the form $\gamma_{\mathrm{L}} : \Omega_{\mathrm{F},1} \times \cdots \times \Omega_{\mathrm{F},n} \rightarrow \Omega_{\mathrm{L}}$ for n followers [5], [31]. This setting is also considered in [6], where in case of multiple leaders, also these are assumed to announce their leader function simultaneously according the Nash equilibrium.

However, in [31] an alternative to multiple followers playing a simultaneous Nash game is presented. There, conditions are developed for the existence of a leader function $\gamma_{\mathrm{L}}(u_{\mathrm{F},1}, ..., u_{\mathrm{F},n})$ under which the n followers' objective functions become identical except for a constant; the problem then reduces to a single-leader single-follower game. The results are however based on the assumption of strict convexity and continuous differentiability of the cost functions. The same idea of identical follower cost functions is also discussed by means of a specific numerical example in earlier work [4]. Here, the leader has one decision variable for each of the n followers, and proposes for each follower $i = 1, ..., n$ a different leader function $u_{\mathrm{L},i} = \gamma_{\mathrm{L},i}(u_{\mathrm{F},1}, ..., u_{\mathrm{F},n})$.

True multi-hierarchy settings in which players perform as a leader and follower simultaneously with respect to the upper respectively lower levels have also been briefly studied in [4]. Sufficient conditions are derived such that the lower-level players are induced to perform as desired for the higher level players, by successively substituting the leader functions in the order of announcement. For a three-level system with player 1 being the upper level leader and with $\gamma_i : \times_{i=1}^{3} \Omega_i \rightarrow \mathbb{R}, i = 1, 2, 3$:

$$\begin{aligned}
(u_1^{\mathrm{d},1}, u_2^{\mathrm{d},1}, u_3^{\mathrm{d},1}) &= \arg \min_{u_1,u_2,u_3} \mathcal{J}_1(u_1, u_2, u_3),\\
(u_2^{\mathrm{d},2}(\gamma_1), u_3^{\mathrm{d},2}(\gamma_1)) &= \arg \min_{u_2,u_3} \mathcal{J}_2(u_2, u_3; \gamma_1),\\
\text{and } u_3^{\mathrm{d},3}(\gamma_2, \gamma_1) &= \arg \min_{u_3} \mathcal{J}_3(u_3; \gamma_2, \gamma_1).
\end{aligned}$$

However, to the authors' best knowledge, cases with multiple leader and multiple followers simultaneously where the Nash concept is not adopted, have not been looked into.

III. Discussion and Open Problems

An overview has been presented to integrate results on reverse Stackelberg games from its origin in the 1970s with more recent contributions to the field. While the reverse Stackelberg game can be adapted w.r.t. several aspects depicted in Fig. 1, which allows it to be flexible in various settings, there are still unresolved problems. In general, these problems stem from the fact that the game is difficult to solve analytically (especially if asymmetric and imperfect information applies) as well as numerically, as shown in the companion paper. In the following, open issues are enumerated to emphasize the potential of the reverse Stackelberg for further research and application in the field of control.

- *Convexity assumptions:*
  There is a large body of literature available on dynamic reverse Stackelberg games with linear state equations and quadratic cost functions, for which the affine strategy or leader function has been proven to solve the game to optimality [15], [20], [21], [24]. Although this result is said to be applicable to a 'sufficiently large' number of cases, a (strictly) convex and differentiable cost function and linear constraints will not generally be found in real-life applications. A relaxation of these assumptions should therefore be made.
- *Computational tractability and optimality bounds:*
  Another gap may be found in the lack of focus on numerical tractability of reverse Stackelberg problems, whereas this is important especially in the context of real-life control or optimization applications, given that large-scale control or optimization problems are often mentioned as a reason for studying Stackelberg strategies [32], [33]. Especially in the largely untouched case of nonconvex cost functionals, approximate or suboptimal solutions may be required. However, there is little focus on establishing bounds on the quality of suboptimal solutions for the leader; similarly, the quality of the solution for the follower is not taken into account, whereas bounds on the performance of lower-level controllers would be a relevant addition.
- *Robustness:*

Even in case a simple (affine) leader function suffices for attaining the leader's desired equilibrium, it would be interesting to investigate how sensitive it is to changes in the game parameters. As mentioned in Section II-E, in [29] nonuniqueness of the optimal leader function is used to consider the minimization of the deviation from an estimated nominal parameter value as a secondary objective. Similarly, robustness to unintentional deviations of the follower to the optimal response should be considered, as discussed in [14] for the closed-loop Stackelberg game. Here, the problem is addressed by adopting discontinuous state-dependent closed-loop strategies that are however developed for very particular numerical problem instances; also, it is not clear how much worse the leader is off by punishing the follower from a deviation from his optimal response. Further, the stability of (ϵ-optimal) Stackelberg solutions has been considered [34]; however no similar results can currently be found in the context of reverse Stackelberg games.

- *Desired leader solution*:
  Generally it is assumed that the leader strives after obtaining her – single – global optimum. Irrespective of whether this solution can be obtained by a particular leader function, cases could be investigated in which the leader strives after a broader set of possible solutions. This becomes even more relevant in case the leader solves a multi-objective problem and thus has to find a trade-off between several optima. Similarly, instead of cardinal solutions that follow from optimization of a real-valued objective function, discrete orderings of preferred solutions could be considered. Such ordinal solutions for regular Stackelberg games have also been advocated in [35].
- *Stability:*
  A stability analysis has been made for the continuous-time [18] and later for the discrete-time [17] LQ Stackelberg game with no-memory state feedback. For time-invariant weighting matrices that occur in the cost functions and state equation of the Stackelberg game, sufficient conditions are developed for a leader function that leads to an asymptotically stable system for the infinite-time game. However, it is not guaranteed that such a leader function exists, nor is a direct approach available for the computation of this function due to the complexity of the problem. Nonetheless, similar guarantees on the system stability for the dynamic reverse Stackelberg game have not yet been investigated.
- *Leader-follower role:*
  In most cases the positions of leader and follower are taken to be known in advance and fixed. While recent research in Stackelberg games allocates more flexibility to this role, no similar results can currently be found for the reverse Stackelberg game, to which this flexible role should also be applicable.
  In the original Stackelberg framework, the leader-follower role has been analyzed in [36]. There, it is shown that if two players act sequentially rather than simultaneously as in a Nash game, both players may obtain better, but not worse, results. Note that this provides another benefit of adopting a hierarchical design for decision making as opposed to the alternative solution concept of the Nash equilibrium. Further, the leader role is preferred to the follower role in a closed-loop Stackelberg game, which is however not the case if feedback information applies [37].
  In [38] a discrete-time dynamic Stackelberg game is considered where players in one of two groups take the position of the leader in turns. There, players are allocated to two fixed groups and they take a switching position of leader respectively follower at each stage of the game. Finally, in [39] an open-loop differential Stackelberg game is considered with mixed leadership, meaning that a player can be both leader and follower at the same time, depending on the subset of control variables that are associated with a particular role a-priori. After announcement of the leader decisions, the optimal follower responses are determined simultaneously, i.e., here, the Nash equilibrium concept is adopted.
- *Nonlinear leader functions:*
  While several options are given for possible shapes of leader functions [4], nonlinear function structures are hardly considered in the reverse Stackelberg game, except from those that occur in specific numerical examples [6], [7], [40]. Discontinuous, state-dependent closed-loop Stackelberg strategies are considered in [14]; in case the follower played suboptimally during the previous stage, a leader strategy is adopted that differs from the normal mode. However, it is not analyzed how much the leader's performance is reduced in case of the 'punishment' strategy that follows suboptimal follower behavior. Moreover, although it is considered an incentive strategy, the leader strategies in [14] are not directly dependent on the followers decisions as required for the reverse Stackelberg game.
- *Approximate solutions:*
  Most research is focused on achieving the leader's desired, global, optimum, for which often an optimal affine leader function can be derived. On the other hand, the case in which this is not possible is not much studied. In particular, suboptimal or ϵ-optimal strategies could be investigated, as is done for specific numerical examples in [6], [7].
- *Applications:*
  As has been mentioned in the companion paper, relatively many applications have been considered for the original Stackelberg game. Similar to the flexible leader-follower role, it should however be no problem to apply the reverse game to these problems. Moreover, as also stated in the companion paper, while adopting a reverse Stackelberg approach, the leader player may be able to achieve a better performance then in the original case.
- *Constraints:*
  Most results on incentive Stackelberg strategies do not

take into account constraints, i.e., except for trivial boundary constraints like $\gamma_{\mathrm{L}}(\cdot) \geq 0$ and $\gamma_{\mathrm{L}}(0) = 0$. This prevents from considering general control applications and should therefore be considered as well.

IV. ACKNOWLEDGMENTS

Research supported by the BSIK project "Next Generation Infrastructures (NGI)", the European Union Seventh Framework Programme [FP7/2007-2013] under grant agreement no. 257462 HYCON2 Network of Excellence, and by the European COST Action TU0702 and TU1102.

REFERENCES

[1] R. Scattolini, "Architectures for distributed and hierarchical model predictive control: A review," *Journal of Process Control*, vol. 19, no. 5, pp. 723–731, 2009.

[2] H. von Stackelberg, *The Theory of Market Economy*. Oxford, UK: Oxford University Press, 1952.

[3] Y.-C. Ho, P. B. Luh, and R. Muralidharan, "Information structure, Stackelberg games, and incentive controllability," *IEEE Transactions on Automatic Control*, vol. 26, no. 2, pp. 454–460, 1981.

[4] Y.-C. Ho, P. B. Luh, and G. J. Olsder, "A control-theoretic view on incentives," *Automatica*, vol. 18, no. 2, pp. 167–179, 1982.

[5] Y. C. Ho, "On incentive problems," *Systems & Control Letters*, vol. 3, no. 2, pp. 63–68, 1983.

[6] G. J. Olsder, "Phenomena in inverse Stackelberg games, part 1: Static problems," *Journal of Optimization Theory and Applications*, vol. 143, no. 3, pp. 589–600, 2009.

[7] ——, "Phenomena in inverse Stackelberg games, part 2: Dynamic problems," *Journal of Optimization Theory and Applications*, vol. 143, no. 3, pp. 601–618, 2009.

[8] K. Staňková, G. J. Olsder, and M. C. J. Bliemer, "Comparison of different toll policies in the dynamic second-best optimal toll design problem: Case study on a three-link network," *European Journal of Transport and Infrastructure Research*, vol. 9, no. 4, pp. 331–346, 2009.

[9] I. Macho-Stadler and J. D. Pérez-Castrillo, *An Introduction to the Economics of Information: Incentives and Contracts*. Oxford: Oxford University Press, 1997.

[10] B. Colson, P. Marcotte, and G. Savard, "An overview of bilevel optimization," *Annals of Operations Research*, vol. 153, no. 1, pp. 235–256, 2007.

[11] M. A. Salman and J. B. Cruz, Jr., "An incentive model of duopoly with government coordination," *Automatica*, vol. 17, no. 6, pp. 821–829, 1981.

[12] D. Fudenberg and J. Tirole, *Game Theory*. Cambridge, MA, USA: MIT Press, 1991.

[13] J. F. Nash, "Equilibrium points in n-person games," in *Proceedings of the National Academy of Sciences*, vol. 36, no. 1, 1950, pp. 48–49.

[14] B. Tolwinski, "Closed-loop Stackelberg solution to a multistage linear-quadratic game," *Journal of Optimization Theory and Applications*, vol. 34, no. 4, pp. 485–501, 1981.

[15] Y.-P. Zheng and T. Başar, "Existence and derivations of optimal affine incentive schemes for Stackelberg games with partial information: A geometric approach," *International Journal of Control*, vol. 35, no. 6, pp. 997–1011, 1982.

[16] X. Liu and S. Zhang, "Optimal incentive strategy for leader-follower games," *IEEE Transactions on Automatic Control*, vol. 37, no. 12, pp. 1957–1961, 1992.

[17] M. Li, J. B. Cruz, Jr., and M. A. Simaan, "An approach to discrete-time incentive feedback Stackelberg games," *IEEE Transactions on Systems, Man, and Cybernatics - Part A: Systems and Humans*, vol. 32, no. 4, pp. 472–481, 2002.

[18] M. A. Salman and J. B. Cruz, Jr., "Team-optimal closed-loop Stackelberg strategies for systems with slow and fast modes," *International Journal of Control*, vol. 37, no. 6, pp. 1401–1416, 1983.

[19] J. C. Engwerda, *LQ Dynamic Optimization and Differential Games*. Chichester, England: John Wiley & Sons, 2005.

[20] H. Ehtamo and R. P. Hämäläinen, "Incentive strategies and equilibria for dynamic games with delayed information," *Journal of Optimization Theory and Applications*, vol. 63, no. 3, pp. 355–369, 1989.

[21] G. Martín-Herrán and G. Zaccour, "Credible linear-incentive equilibrium strategies in linear-quadratic differential games," in *Advances in Dynamic Games and Their Applications*, ser. Annals of the International Society of Dynamic Games, O. Pourtallier, V. Gaitsgory, and P. Bernhard, Eds. Birkhäuser Boston, 2009, vol. 10, no. 4, pp. 1–31.

[22] Y. P. Zheng, T. Başar, and J. B. Cruz, Jr., "Stackelberg strategies and incentives in multiperson deterministic decision problems," *IEEE Transactions on Systems, Man, and Cybernetic*, vol. 14, no. 1, pp. 10–24, 1984.

[23] G. B. Papavassilopoulos and J. B. Cruz, Jr., "Sufficient conditions for Stackelberg and Nash strategies with memory," *Journal of Optimization Theory and Applications*, vol. 31, no. 2, pp. 1233–260, 1980.

[24] T.-S. Chang and Y.-C. Ho, "Incentive problems: A class of stochastic Stackelberg closed-loop dynamic games," *Systems and Control Letters*, vol. 1, no. 1, pp. 16–21, 1981.

[25] D. H. Cansever and T. Başar, "On stochastic incentive control problems with partial dynamic information," *Systems & Control Letters*, vol. 6, no. 1, pp. 69–75, 1985.

[26] T. Başar, "Stochastic incentive problems with partial dynamic information and multiple levels of hierarchy," *European Journal of Political Economy*, vol. 5, no. 2–3, pp. 203–217, 1989.

[27] T. Vallée and T. Başar, "Off-line computation of Stackelberg solutions with the genetic algorithm," *Computational Economics*, vol. 13, no. 1, pp. 201–209, 1999.

[28] ——, "Incentive Stackelberg solutions and the genetic algorithm," *Paper presented at the International Symposium of Dynamic Games and Applications (ISDGA), Maastricht, The Netherlands*, 1998.

[29] D. H. Cansever and T. Başar, "A minimum sensitivity approach to incentive design problems," *Large Scale Systems*, vol. 5, pp. 233–244, 1983.

[30] ——, "Optimum/near optimum incentive policies for stochastic decision problems involving parametric uncertainty," *Automatica*, vol. 21, no. 5, pp. 575–584, 1985.

[31] Y.-w. Jing and S.-y. Zhang, "The solution to a kind of Stackelberg game systems with multi-follower: Coordinative and incentive," in *Analysis and Optimization of Systems*, ser. Lecture Notes in Control and Information Science, A. Bensoussan and J. Lions, Eds. Berlin, Germany: Springer-Verlag, 1988, vol. 111, pp. 593–602.

[32] T. Başar and H. Selbuz, "Closed-loop Stackelberg strategies with applications in the optimal control of multilevel systems," *IEEE Transactions on Automatic Control*, vol. 24, no. 2, pp. 166–179, 1979.

[33] J. B. Cruz, Jr., "Leader-follower strategies for multilevel systems," *IEEE Transactions on Automatic Control*, vol. 23, no. 2, pp. 244–255, 1978.

[34] M. B. Lignola and J. Morgan, "Topological existence and stability for Stackelberg problems," *Journal of Optimization Theory and Applications*, vol. 84, no. 1, pp. 145–169, 1995.

[35] J. B. Cruz, Jr. and M. A. Simaan, "Ordinal games and generalized Nash and Stackelberg solutions," *Journal of Optimization Theory and Applications*, vol. 107, no. 2, pp. 205–222, 2000.

[36] T. Başar, "On the relative leadershop property of Stackelberg strategies," *Journal of Optimization Theory and Applications*, vol. 11, no. 6, pp. 655–661, 1973.

[37] M. Simaan and J. B. Cruz, Jr., "A Stackelberg solution for games with many players," *IEEE Transactions on Automatic Control*, vol. 18, no. 3, pp. 322–324, 1973.

[38] P.-y. Nie, "Dynamic discrete-time multi-leader-follower games with leaders in turn," *Computers and Mathematics with Applications*, vol. 61, no. 8, pp. 2039–2043, 2011.

[39] T. Başar, A. Bensoussan, and S. P. Sethi, "Differential games with mixed leadership: The open-loop solution," *Applied Mathematics and Computation*, vol. 217, no. 3, pp. 972–979, 2010.

[40] H. Shen and T. Başar, "Incentive-based pricing for network games with complete and incomplete information," in *Advances in Dynamic Game Theory*, ser. Annals of the International Society of Dynamic Games, S. Jørgensen, M. Quincampoix, and T. Vincent, Eds. New York, NY, USA: Birkhäuser Boston, 2007, vol. 9, no. V, pp. 431–458.