A Generalized Two-Agent Location Problem: Asymmetric Dynamics and Coordination

Journal of Optimization Theory and Applications

ISSN 0022-3239
Volume 148
Number 2

J Optim Theory Appl (2010)
148:336-363
DOI 10.1007/s10957-010-9750-x

J Optim Theory Appl (2011) 148: 336–363
DOI 10.1007/s10957-010-9750-x

A Generalized Two-Agent Location Problem: Asymmetric Dynamics and Coordination

Boaz Golany · Konstantin Kogan · Uriel G. Rothblum

Published online: 14 October 2010

Abstract We generalize a static two-agent location problem into dynamic, asymmetric settings. The dynamics is due to the ability of the agents to move at limited speeds. Since each agent has its own objective (demand) function and these functions are interdependent, decisions made by each agent may affect the performance of the other agent and thus affect the overall performance of the system. We show that, under a broad range of system's parameters, centralized (system-wide optimal) and non-cooperative (Nash) behavior of the agents are characterized by a similar structure. The timing of these trajectories and the intermediate speeds are however different. Moreover, non-cooperative agents travel more and may never rest and thus the system performance deteriorates under decentralized decision-making. We show that a static linear reward approach, recently developed in Golany and Rothblum (Nav. Res. Logist. 53(1):1–15, 2006), can be generalized to provide coordination of the moving agents and suggest its dynamic modification. When the reward scheme is applied, the agents are induced to choose the system-wide optimal solution, even though they operate in a decentralized decision-making mode.

Keywords Nash equilibrium · Dynamic control · Coordination · Homeland security

Communicated by Q. Zhao.

B. Golany (✉) · U.G. Rothblum
Faculty of Industrial Engineering and Management, Technion, Israel Institute of Technology, Haifa 32000, Israel
e-mail: golany@ie.technion.ac.il

K. Kogan
Department of Management, Faculty of Social Sciences, Bar Ilan University, Ramat Gan 52900, Israel

1 Introduction

Coordination among agents, engaged in similar activities and sharing some common objectives, is an issue of major importance in many different environments. For example, modern supply chains are composed of many agents that are located at different nodes in their respective supply networks. While each agent may have its individual objective function (e.g., maximize its profits), they share the common objective of optimizing the performance of the chain as a whole. Homeland security is another environment, where many agents may need to perform certain individual tasks that, together, serve a single overall objective. Clearly, in such settings it is necessary to achieve some coordination in order to accomplish solutions that are optimal, from the point of view of the system as a whole.

The ultimate way of achieving coordination is through centralized control, which takes into account the capabilities of all the agents, their possible courses of action, and the impact that their actions may have on the entire system. In contrast, when decentralized control is practiced, each agent makes its own local decisions, which often lead to outcomes that are unfavorable from the perspective of the system as a whole. Indeed, the sub-optimality of decentralized decision-making is a well-known phenomenon. But, decentralized decision-making is often inevitable, due to the number of agents, their geographical dispersion and various constraints on the communications among them. Furthermore, letting agents pursue what they perceive to be their best course of action may increase their motivation and identification with the mission of the organization. These behavioral benefits may have an important impact on the organization and its long-term goals.

The issue of coordination among agents in decentralized decision-making environments has attracted research attention for years and many mechanisms were proposed in order to achieve coordination, that would yield a global optimal (or near optimal) solution. Examples of such mechanisms can be found in Cachon [1, 2], and Taylor [3] in the context of *supply chain management*; Naor [4], Dolan [5], Mendelson and Whang [6] and Avi-Itzhak et al. [7] in *queuing systems*; Beckmann et al. [8] and Cole et al. [9] in *transportation systems*. Recently, Golany and Rothblum [10] suggested a general mechanism (referred to as "linear reward scheme"), that can be employed to obtain the desired coordination under certain conditions and that have broad applicability. In particular, they demonstrated implementations of their approach in five examples taken from the supply chain literature.

The studies described above explore static situations. In the current paper, we explore the virtue of coordination among agents acting in *dynamic* environments. This is a relatively new area with only a few recorded results. Most of the work reported, thus far, was focused on developing heuristics that would enable dynamic coordination in the routing of moving objects. For example, see Frentinos et al. [11] in coordinating *autonomous agricultural vehicles*; and Shima et al. [12] in coordinating *pilotless planes*.

In this paper, we generalize a particular 2-agent dynamic problem, whose static origin can be traced to Hotelling's "law of competition" (Hotelling [13]) and later referred to as the ice-cream sellers problem (see, e.g. Munson et al. [14])—a situation in which two ice-cream sellers need to select their locations along a one dimensional axis. The dynamics in the present paper have an interesting edge, caused by the fact that the agents can move at finite and not necessarily equal speeds. If the speeds were allowed to be infinite, the problem could easily be reduced to the static case (i.e., the agents would have immediately "jumped" to the optimal points of the static case and there would have been no dynamics). Also, the dynamics would have been rather straightforward, if the maximal speeds of the two agents were identical, since in this case, each of the agents would travel at its maximal speed to the optimal locations of the static problem and then drift there, if the demand is time-dependent. The asymmetry in the agents' maximal speeds causes the non-trivial dynamics, where the faster agent has to turn back at some point in time.

We first analyze the difference between the centralized (global optimum) solution and the decentralized solution. In terms of the decentralized solution, we confine our interest to the search for an open-loop Nash equilibrium (OLNE), which implies that *all* the players pre-commit to their control actions throughout the game, i.e., never depart from their initial plan along the equilibrium play. Accordingly, this solution concept is time-consistent and can serve as a benchmark for the assessment of more complex strategies.

In particular, we find the optimal trajectories, that lead the two ice cream sellers from their respective starting points (at the two ends of the route they seek to cover) to the static locations (resting points) they occupy for some time during the planning period and then back to their respective starting points. We further show that, under some conditions (in particular, when the customers' demand is not a function of time), these resting points can be identical to those found in the static problem (see, e.g., Golany and Rothblum [10]). If demand is a function of time, the sellers may never rest and will have to continue moving throughout the planning period. Finally, we generalize the static linear reward approach, developed by Golany and Rothblum [10] to provide coordination of the dynamic agents. We further suggest an approach which imposes linear penalties/rewards for speeding. The approach will make the global optimal solution a Nash equilibrium trajectory—thus inducing the sellers to choose it, even when they operate in a decentralized decision-making environment.

The tools we employ in this paper belong to the field of *differential games*. For earlier references on this subject, see Başar and Olsder [15] and Feichtinger and Jørgensen [16]. More recently, differential games were applied, in the context of supply chain management, by Kogan and Tapiero [17]. For a recent survey, on the use of differential games in various supply chain settings, the reader is referred to He et al. [18].

The models developed here may be employed in various environments. While the two scenarios described below are quite different in nature and context, they share the common theme of coordination among moving agents.

Coordination among fishing boats—there is extensive literature on the need for authorities to coordinate among fishing boats operated by independent operators (fishermen). Typically, fishery coordination is applied through quotas, that dictate the quantities of fish (of various species), which each ship is allowed to catch at specified locations (see, e.g., Bess [19]). These quotas are computed so as to prevent severe depletion of certain fish stock, as well as to encourage the fisherman to explore areas strategically important for the authorities. Thus, a fishing boat that has used up its quota of a certain kind of fish is forced to move to another location where it might find other species and fish. Similarly, coordination can be applied with respect to the location of the various boats along a given timeline. Suppose the boats are restricted to fish along the shoreline due to weather conditions or the existence of certain types of fish. Without coordination, the authorities may find the boats congregating near the harbor and avoiding other areas due to convenience considerations. It is easy to use Global Positioning System (GPS) technology, that will report the exact coordinates, speed and direction of each boat and control them, either through a central controller or through some incentive mechanism (e.g., paying larger prices for fish caught in remote areas and penalizing boats, that stay near the harbor by paying smaller prices for the fish they catch).

Coordination among border patrol units—the US-Mexico border is long and difficult to patrol. In spite of tremendous efforts made by various US government agencies, it is estimated that hundreds of thousands of illegal immigrants, from Mexico and other Central and South American countries, cross this border every year. The inability of these agencies to stop the flood of immigrants has prompted suggestions to outsource the task of patrolling the border and stopping the intruders to private security firms (see, e.g., Carafano et al. [20] and Wheatley and Doty [21]). One way to incentivize the firms to excel in performing the task is to pay them for each intruder they are able to apprehend. Without coordination among the patrol units, one may find them congregating at certain locations along the borderline, where they expect to encounter relatively large numbers of intruders, leaving large stretches of the border completely exposed. To ensure complete coverage of the border, one can rely on GPS technology to report on the exact location and speed of each unit at any time and impose a coordination scheme, that is based on the location (or speed & direction) of the units at different times. Thus, for every intruder, that is apprehended along the optimal travel plan computed by some central command, the unit will get the full price, while intruders that will be apprehended in locations and times, that do not agree with the optimal travel plan, will yield only smaller returns.

The rest of the paper is organized as follows. In Sect. 2, we formulate the two-agent problem. Our main results, about globally optimal solutions and Nash equilibrium solutions, are stated in Sect. 3. Section 4 contains results about the use of rewards and penalties for converting optimal solutions into Nash equilibria; in particular, we use time-discretization to interpret this approach. The proofs of the results of Sects. 3–4 are given in Appendix A. Finally, Sect. 5 contains some concluding remarks.

2 Problem Statement

Let us have two agents indexed by i, $i = 1, 2$. Both agents can move forward and backward along a one-dimensional route[1] which is normalized to be $[0, 1]$. The coordinate of each agent at time t is given by $X_i(t)$, $i = 1, 2$ and the speed at which it moves is denoted by $u_i(t)$. The two agents start at their bases in coordinates 0 and 1,

[1] This route doesn't necessarily have to be a line segment. Whatever shape the route has, we can mathematically "stretch" it along a line (under proper regularity conditions).

respectively, at time $t = 0$ and must return to their bases by time T. The motion equations of the agents within the time interval $[0, T]$ are then given by:

$$\dot{X}_1(t) = u_1(t), \qquad X_1(0) = X_1(T) = 0, \tag{1}$$

$$\dot{X}_2(t) = -u_2(t), \qquad X_2(0) = X_2(T) = 1. \tag{2}$$

Each agent can move in both directions and is constrained by a maximum speed U_i

$$-U_i \le u_i(t) \le U_i, \quad i = 1, 2. \tag{3}$$

In what follows, we assume that the agents never cross each other's path[2]:

$$X_2(t) \ge X_1(t). \tag{4}$$

We refer to $u_1(t)$'s and $u_2(t)$'s as the *control variables* and to $X_1(t)$'s and $X_2(t)$'s as the location (state) variables.

Note that (1)–(3) can be viewed as a modification of the Transportation-Location-Allocation over time problems presented by Cooper [22], Tapiero [23], Tapiero and Soliman [24] and later by Cavalier and Sherali [25]. However, the location-allocation problems are based on meeting destination, in-storage goods, demand $X_2(t)$, with the source in-storage goods $X_1(t)$ and therefore, $u_1 = u_2$, while in our settings, the agents are independent and compete with each other.

We assume that the potential customers are uniformly spread along the route [0, 1] at rate $\beta(t)$, that is, $\beta(t)$ measures the number of customers per unit of distance at time t. A customer considers only the agent that is closer to him and realization of the potential demand decreases linearly, as the distance to that agent increases. This setting corresponds to Hotelling's law of competition (Hotelling [13]). Namely, at distance d from the closer agent, $\alpha \cdot d$ of the $\beta(t)$, potential customers at time t lose their interest in the service offered by that agent. So, given that the coordinates of the agents at time t are $X_1(t)$ and $X_2(t)$, the instantaneous demands per time unit they face at that time are given by the cumulative number of customers that the first (second) agent serves, starting from his disposition, $X_1(t)(X_2(t))$, to his base, coordinate 0 (1), and to the point of equal distance between the two agents $(X_1(t) + X_2(t))/2$:

$$D_1[X_1(t), X_2(t)] = \int_0^{X_1(t)} (\beta(t) - \alpha(X_1(t) - s))ds + \int_{X_1(t)}^{\frac{X_1(t)+X_2(t)}{2}} (\beta(t) - \alpha(s - X_1(t)))ds, \tag{5}$$

$$D_2[X_1(t), X_2(t)] = \int_{X_2(t)}^{1} (\beta(t) - \alpha(s - X_2(t)))ds + \int_{\frac{X_1(t)+X_2(t)}{2}}^{X_2(t)} (\beta(t) - \alpha(X_2(t) - s))ds. \tag{6}$$

[2] In fact, we do not impose inequality (4) explicitly. Rather, each time an optimal solution is found, we verify that (4) holds and impose restrictions on system's parameters if needed.

We further assume that each of the two agents is capable of covering at least half the route [0, 1] in terms of demand, i.e., $\beta(t)-\frac{\alpha}{2}\geq 0$ at any $t\in[0,T]$. This model is a direct extension of the corresponding static formulation in Golany and Rothblum [10].

Integrating the right-hand side of (5), we have for agent 1,

$$D_1[X_1(t),X_2(t)]=\frac{1}{2}\beta(t)X_1(t)+\frac{1}{4}\alpha X_1(t)X_2(t)-\frac{5}{8}\alpha X_1(t)^2 + \frac{1}{2}\beta(t)X_2(t)-\frac{1}{8}\alpha X_2(t)^2. \tag{7}$$

Similarly, by integrating (6), we have for agent 2,

$$\begin{aligned}D_2[X_1(t),X_2(t)]&=\beta(t)(1-X_2(t))+\alpha X_2(t)-\frac{1}{2}\alpha-\frac{1}{2}\alpha X_2(t)^2\\&\quad+\beta(t)\frac{X_2(t)-X_1(t)}{2}-\alpha X_2(t)\frac{X_2(t)-X_1(t)}{2}\\&\quad+\frac{1}{2}\alpha X_2(t)^2-\alpha\frac{[X_1(t)+X_2(t)]^2}{8}.\end{aligned}$$

Ignoring constants that do not affect the optimization (but keeping the same notation $D_2[X_1(t),X_2(t)]$ for convenience) and exercising simple algebraic operations, we get:

$$\begin{aligned}D_2[X_1(t),X_2(t)]&=\left[\alpha-\frac{1}{2}\beta(t)\right]X_2(t)+\frac{1}{4}\alpha X_2(t)X_1(t)-\frac{5}{8}\alpha X_2(t)^2\\&\quad-\frac{1}{2}\beta(t)X_1(t)-\frac{1}{8}\alpha X_1(t)^2.\end{aligned}\tag{8}$$

We assume the existence of a unit profit per customer. Hence, instantaneous/total demand represents instantaneous/total profit. Thus, the objective of each individual agent i of maximizing sales/profit during the given time interval is given by

$$\max_{u_i}\int_0^T D_i[X_1(t),X_2(t)]\,dt. \tag{9}$$

In a centralized system, there is a single decision-maker (DM), whose decisions bind both agents. The objective function of the centralized DM is then to maximize the overall sales/profit given by

$$\max_{u}\int_0^T\sum_{i=1}^{2}D_i[X_1(t),X_2(t)]\,dt \tag{10}$$

subject to (1)–(4). Taking into account expressions (7) and (8), the joint instantaneous demand in the centralized problem takes the following form

$$\begin{aligned}&D_1[X_1(t),X_2(t)]+D_2[X_1(t),X_2(t)]\\&\quad=\alpha\left[X_2(t)+\frac{1}{2}X_2(t)X_1(t)-\frac{3}{4}X_2(t)^2-\frac{3}{4}X_1(t)^2\right].\end{aligned}\tag{11}$$

Without loss of generality, we assume that the first agent is faster than the second, i.e., $U_1 > U_2$. We also assume that the planning horizon T is long enough (a condition, that will be specifically stated in the context of our formal results and the agents' maximum speed is not extremely low (equivalently, $1/U_2$ is not extremely large); this allows the agents to attain the most general optimal behavior and return back to their bases within the time frame of negligible discounting effect.

3 System-Wide Optimal Solution and Nash Equilibrium Solutions

In a centralized system, one addresses the problem of maximizing (10), subject to (1)–(4). The next result presents a unique optimal solution to this problem. The proof of this result, given in the Appendix, derives characterizing conditions for optimality (those of the Maximum Principle), which are then uniquely solved.

Theorem 3.1 *Let* $T \geq \frac{1}{2U_2}$ *and* $\tau_1 = \frac{1}{3U_1+U_2}$, $\tau_2 = \frac{1}{4U_2}$, $\tau_{-2} = T - \tau_2$, $\tau_{-1} = T - \tau_1$. *Then there exists a unique system-wide optimal solution of* (10), *subject to* (1)–(4), *which is piecewise constant and is given by*:

$$u_1^*(t) = \begin{cases} U_1, & 0 \leq t < \tau_1, \\ -\frac{U_2}{3}, & \tau_1 \leq t < \tau_2, \\ 0, & \tau_2 < t \leq \tau_{-2}, \\ \frac{U_2}{3}, & \tau_{-2} \leq t < \tau_{-1}, \\ -U_1, & \tau_{-1} \leq t \leq T, \end{cases} \qquad u_2^*(t) = \begin{cases} U_2, & 0 \leq t < \tau_2, \\ 0, & \tau_2 \leq t < \tau_{-2}, \\ -U_2, & \tau_{-2} \leq t \leq T. \end{cases}$$

The optimal agents' speeds and locations in the various regions, defined by Theorem 3.1, are depicted in Fig. 1. Specifically, the optimal behavior of the two agents includes two trajectories. The first one is to move towards each other as fast as possible, until a certain relationship between their coordinates is reached. Then, the faster agent turns back and both agents move in the same direction, until they reach the coordinates $X_1^*(t) = 1/4$ and $X_2^*(t) = 3/4$ (which are the optimal locations in the static problem). At this point in time, both agents stop and stay at their respective coordinates, as long as there is sufficient time for the slow agent to get back to its base by time T at its maximum speed. The other part of the optimal trajectory is symmetric to the one described above. Namely, the faster agent first moves forward at an intermediate speed and turns back, only when there is enough time to get to its base at its maximum speed, while the slow agent moves all the time to its base at its maximum speed.

Theorem 3.1 presents the optimal solution for the control problem of the two mobile agents, assuming that the planning horizon, T, is larger than a minimal value, $T \geq 1/(2U_2)$. The solution we obtain has the property that the agents never cross each other. If the condition, $T \geq 1/(2U_2)$, does not hold, then we have a special case of the general optimal solution, described by Theorem 3.1. In such a case, the agents will not have enough time to attain the optimal locations $X_1^*(t) = 1/4$ and $X_2^*(t) = 3/4$. Accordingly, the agents will move as far as they can and turn short of the optimal locations to get back to their basis by time T.

In decentralized settings, we consider the Nash equilibrium solution, where each agent maximizes its objective function (9), subject to constraints (1)–(4), assuming

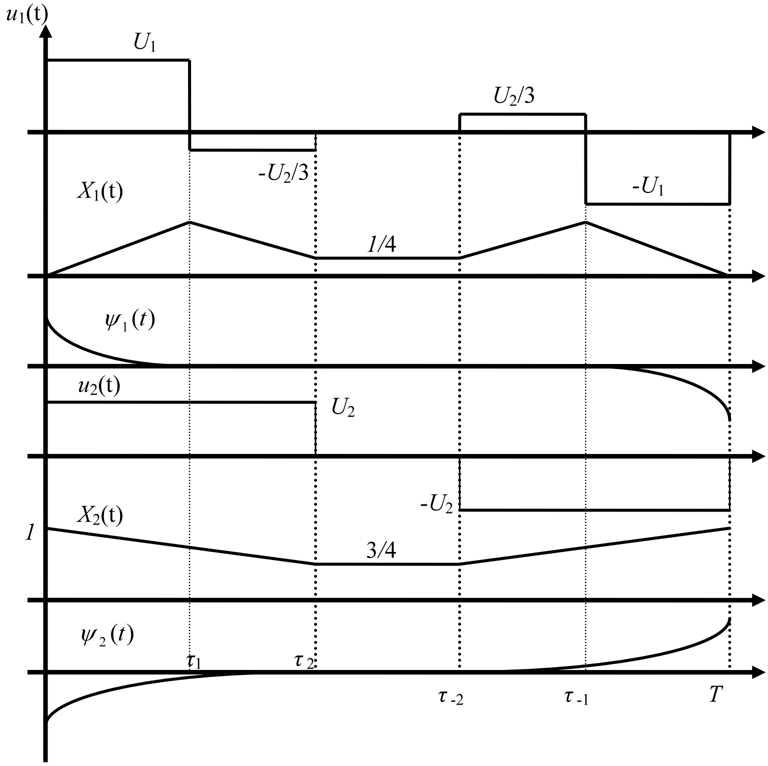

Fig. 1 System-wide optimal motion of the agents over time.

that the values of the decision variables of the other agents are given; of course, (7)–(8) can be substituted in (9) to obtain a more explicit optimization problem.

The next result presents a unique Nash equilibrium solution. The proof of this result, presented in Appendix A, considers the problem of each agent with the actions of the other agent considered as data. Here too, we apply characterizing conditions (of the Maximum Principle) which are then uniquely solved.

Theorem 3.2 *Let $\beta(t)$ be differentiable for $t \in [0,T]$, $|\frac{\dot\beta(t)}{3\alpha}| < U_2$. $T \geq 1/U_2$ and t_1, t_2, t_{-2}, t_{-1}, uniquely determined by*: $t_1(U_2+5U_1) = \frac{2\beta(t_1)}{\alpha}+1$, $1-U_2t_2 = \frac{5}{6}-\frac{\beta(t_2)}{3\alpha}$, $(T-t_{-1})(U_2+5U_1) = \frac{2\beta(t_{-1})}{\alpha}+1$, $1-U_2(T-t_{-2}) = \frac{5}{6}-\frac{\beta(t_{-2})}{3\alpha}$. *Then, there exists a unique Nash equilibrium, which is given by*:

$$u_1^N = \begin{cases} U_1, & 0 \leq t < t_1, \\ \frac{2\dot\beta(t)}{5\alpha} - \frac{U_2}{5}, & t_1 \leq t < t_2, \\ \frac{\dot\beta(t)}{3\alpha}, & t_2 \leq t < t_{-2}, \\ \frac{2\dot\beta(t)}{5\alpha} + \frac{U_2}{5}, & t_{-2} \leq t < t_{-1}, \\ -U_1, & t_{-1} \leq t \leq T, \end{cases} \qquad u_2^N = \begin{cases} U_2, & 0 \leq t < t_2, \\ \frac{\dot\beta(t)}{3\alpha}, & t_2 \leq t < t_{-2}, \\ -U_2, & t_{-2} \leq t \leq T. \end{cases}$$

Note that graphically, the Nash solution looks very similar to the system-wide solution depicted in Fig. 1. Both solutions are piecewise constant, if $\beta(t)$ is time invariant and both have the same number of break points, as discussed in the next section. Unlike the system-wide optimal solution, the Nash solution, of Theorem 3.2, involves an assumption that $|\dot{\beta}(t)/3\alpha| < U_2$, i.e., the ratio between the rate of change at time t of the number of available customers, to the rate of loss of the customers per unit of distance, should be lower than three times the maximum speed of the slow agent. This condition ensures that the solution is feasible, i.e., $u_2^N = |\dot{\beta}(t)/3\alpha| < U_2$, and thereby, the slow agent is able to trace optimally any change in time of the potential customers $\beta(t)$ (the system-wide optimal solution is independent of $\beta(t)$ and the agents do not compete by moving when they are at the optimal location). (If $u_2^N = |\dot{\beta}(t)/3\alpha| < U_2$, then, of course, the same holds for the fast agent, $u_1^N = |\dot{\beta}(t)/3\alpha| < U_1$.) If this condition is not satisfied, we have a special case, characterized by no intermediate speed of the slow agent. That is, the slow agent will utilize the so-called "bang-bang" strategy and move back and forth around the optimal location at maximum speed before getting back to its basis.

4 Coordination

From Theorems 3.1 and 3.2, we have found that, under some assumptions, the Nash (non-cooperative) solution has a similar structure to the system-wide optimal solution. In particular, both solutions contain the same number of break points in which the control functions change. The main difference is in the timing (i.e., in the values of the break points) and the intermediate speed of the faster agent at time intervals $[t_1, t_2]$ and $[t_{-2}, t_{-1}]$. The following proposition shows that the solutions of the centralized and decentralized problems do not converge to the same point, that is, the distances which the agents travel, their coordinates over time and their speeds are different.

Proposition 4.1 *Assume the conditions of Theorems* 3.1 *and* 3.2 *hold*. *Then*, *compared with the system-wide optimal solution*, *in the Nash equilibrium solution*, *the agents travel more time*, *come closer to each other at* $[t_2, t_{-2}]$ (*as compared to the minimal distance between them during* $[\tau_2, \tau_{-2}]$) *even if* $\beta(t)$ *is time-invariant*, *and the faster agent always moves slower during the intervals* $[t_1, t_2]$ *and* $[t_{-2}, t_{-1}]$, *if* $\dot{\beta}(t) = 0$, *i.e.*, *during the intervals of intermediate speed levels.*

In what follows, we introduce instantaneous linear rewards/penalties $B_i(t)$ (dollars for unit of speed) on the controls $u_i(t)$ variables of the agents. Specifically, when agents 1 and 2 select, respectively, controls $u_1(t)$ and $u_2(t)$ at time t, the instantaneous penalty that agent 1 pays is $B_1(t)u_1(t)$ and the instantaneous reward that agent 2 gets is $B_2(t)u_2(t)$ (we set penalties on agent 1 and rewards on agent 2 to indicate that the underlying aim is to slow them down). These adjustments influence the agents' instantaneous returns in the following way:

$$\hat{D}_1[X_1(t), X_2(t)] = D_1[X_1(t), X_2(t)] - B_1(t)u_1(t), \tag{12}$$

$$\hat{D}_2[X_1(t), X_2(t)] = D_2[X_1(t), X_2(t)] + B_2(t)u_2(t). \tag{13}$$

The next result shows that with a specific class of reward/penalty functions, the optimal solution becomes a Nash equilibrium solution.

Theorem 4.1 *Let* $T \geq \frac{1}{2U_2}$, τ_1, τ_2, τ_{-2}, τ_{-1} *be determined as in Theorem* 3.1, $(X_1^*(\cdot), X_2^*(\cdot))$ *be the optimal solution established in Theorem* 3.1 *and* A_1 *and* A_2 *arbitrary constants. Suppose we modify* (9), *by replacing the instantaneous demand* $D_i[X_1(t), X_2(t)]$ *of each agent* $i = 1, 2$ *with* $\hat{D}_i[X_1(t), X_2(t)]$, *as defined by* (12)–(13). *If* $B_1(\tau_1) = A_1$, $B_2(\tau_2) = A_2$,

$$\dot{B}_1(t) = \frac{X_2^*(t) - X_1^*(t)}{4}\alpha - \frac{\beta(t)}{2}, \quad \tau_1 \leq t < \tau_{-1},$$

$$\dot{B}_2(t) = \frac{\beta(t)}{2} - \alpha\frac{X_2^*(t) - X_1^*(t)}{4}, \quad \tau_2 \leq t < \tau_{-2},$$

$$B_1(t) < A_1 - \int_t^{t_1}\left[-\frac{\beta(\tau)}{2} - \frac{1}{4}\alpha X_2^*(\tau) + \frac{5}{4}\alpha X_1^*(\tau)\right]d\tau, \quad 0 \leq t < \tau_1,$$

$$B_1(t) > A_1 + \int_{t_1}^{t_{-1}}\left[\frac{X_2^*(\tau) - X_1^*(\tau)}{4}\alpha - \frac{\beta(\tau)}{2}\right]d\tau$$
$$+ \int_{t_{-1}}^{t}\left[-\frac{\beta(\tau)}{2} - \frac{1}{4}\alpha X_2^*(\tau) + \frac{5}{4}\alpha X_1^*(\tau)\right]d\tau, \quad \tau_{-1} < t \leq T,$$

$$B_2(t) > A_2 - \int_t^{t_2}\left[\frac{\beta(\tau)}{2} - \alpha - \frac{1}{4}\alpha X_1^*(\tau) + \frac{5}{4}\alpha X_2^*(\tau)\right]d\tau, \quad 0 \leq t < \tau_2,$$

$$B_2(t) < A_2 + \int_{t_2}^{t_{-2}}\left[\frac{\beta(\tau)}{2} - \alpha\frac{X_2^*(\tau) - X_1^*(\tau)}{4}\right]d\tau$$
$$+ \int_{t_{-2}}^{t}\left[\frac{\beta(\tau)}{2} - \alpha - \frac{1}{4}\alpha X_1^*(\tau) + \frac{5}{4}\alpha X_2^*(\tau)\right]d\tau, \quad \tau_{-2} < t \leq T,$$

then the system-wide optimal solution of the original problem $(X_1^*(\cdot), X_2^*(\cdot))$ *is a unique Nash equilibrium solution of the modified problem.*

Theorem 4.1 defines a class of functions that perfectly coordinate the two mobile agents. An important special case of this class has the two marginal penalty functions satisfying

$$\dot{B}_1(t) = \frac{\partial D_2}{\partial X_1} = \frac{X_2^*(t) - X_1^*(t)}{4}\alpha - \frac{\beta(t)}{2} < 0, \tag{14}$$

$$\dot{B}_2(t) = \frac{\partial D_1}{\partial X_2} = \frac{\beta(t)}{2} - \alpha\frac{X_2^*(t) - X_1^*(t)}{4} > 0. \tag{15}$$

It is easy to observe that these expressions satisfy Theorem 4.1 for the intervals $\tau_1 \leq t < \tau_{-1}$ and $\tau_2 \leq t < \tau_{-2}$, respectively. Specifically, (14)–(15) imply that the marginal penalty (reward), $\dot{B}_1(t)$ and $\dot{B}_2(t)$, that one agent incurs for increasing (decreasing) its speed by one unit, is equal to the loss (gain) that this change causes to the other agent per unit of distance.

Corollary 4.1 *Let $T \geq \frac{1}{2U_2}$ and suppose we modify* (9), *by replacing the instantaneous demand $D_i[X_1(t), X_2(t)]$ of each agent $i = 1, 2$ with $\hat{D}_i[X_1(t), X_2(t)]$, as defined by* (12)–(13), *with $B_1(t)$ and $B_2(t)$ satisfying* (14)–(15) *for $0 \leq t \leq T$. Then, the modified problem has a unique Nash equilibrium solution, which coincides with the unique system-wide optimal solution of the original problem.*

Based on Corollary 4.1, we can choose, for example, constants A_1 and A_2 so that $B_1(T) = B_2(T) = 0$. Then,

$$B_1(0) = -\int_0^T \left[\frac{X_2^*(t) - X_1^*(t)}{4}\alpha - \frac{\beta(t)}{2}\right] dt = K_1 > 0,$$

$$B_2(0) = -\int_0^T \left[\frac{\beta(t)}{2} - \alpha\frac{X_2^*(t) - X_1^*(t)}{4}\right] dt = K_2 < 0,$$

$$A_1 = -\int_{t_1}^T \left[\frac{X_2^*(t) - X_1^*(t)}{4}\alpha - \frac{\beta(t)}{2}\right] dt,$$

$$A_2 = -\int_{t_1}^T \left[\frac{\beta(t)}{2} - \alpha\frac{X_2^*(t) - X_1^*(t)}{4}\right] dt$$

and, more importantly, for $0 \leq t < T$

$$B_1(t) = K_1 + \int_0^t \left[\frac{X_2^*(\tau) - X_1^*(\tau)}{4}\alpha - \frac{\beta(\tau)}{2}\right] d\tau > 0,$$

$$B_2(t) = K_2 + \int_0^t \left[\frac{\beta(\tau)}{2} - \alpha\frac{X_2^*(\tau) - X_1^*(\tau)}{4}\right] d\tau < 0.$$

This case (when both agents are penalized over the entire planning period, as $B_1(t) \cdot u_1(t)$ is subtracted from the demand of the first agent, and $B_2(t) \cdot u_2(t)$ is added to the demand of the second agent) is of special interest.

We next provide an interpretation for the conditions that were shown in Corollary 4.1 to be sufficient for converting the unique optimal solution into a Nash equilibrium. The interpretation we derive builds on conditions, developed in Rothblum [26] and Golany and Rothblum [10], for finite dimensional optimization problems, yielding a continuous variant of those conditions.

The method we employ is time discretization, while maintaining continuous location variables. Specifically, assume that the time interval $[0, T]$ is split into N equal intervals of size $\Delta = T/N$. The speeds of agent $i = 1, 2$ along these intervals are assumed to be constant and given by $u_{i,h}$, $h = 0, \ldots, N-1$—these are the decision variables of agent i. The marginal increments of the locations of agent i are then given by $\Delta u_{i,h}$. The resulting location of agent i at time $\tau\Delta$, $\tau = 0, \ldots, N$ is denoted

$X_i(\tau\Delta)$ and is given by

$$X_1(\tau\Delta) = \Delta\left[\sum_{h=0}^{\tau-1} u_{1,h}\right], \tag{16}$$

$$X_2(\tau\Delta) = 1 - \Delta\left[\sum_{h=0}^{\tau-1} u_{2,h}\right]. \tag{17}$$

The constraints on the instantaneous speeds, available from (3), are

$$-U_i \le u_{i,h} \le U_i, \quad h = 0, \dots, N-1, i = 1, 2. \tag{18}$$

The instantaneous demand, that the agents face at time $\tau\Delta$, are functions of their location and are given by (7)–(8) (developed in the continuous-time framework). Using the discretization framework, these expressions can be rewritten by:

$$D_1(X_1(\tau\Delta), X_2(\tau\Delta)) = \frac{\beta(\tau\Delta)[X_1(\tau\Delta) + X_2(\tau\Delta)]}{2} - \frac{\alpha X_1(\tau\Delta)^2}{2} - \frac{[\alpha(X_2(\tau\Delta) - X_1(\tau\Delta)]^2}{8}, \tag{19}$$

$$D_2(X_1(\tau\Delta), X_2(\tau\Delta)) = \frac{\beta(\tau\Delta)[1 - X_1(\tau\Delta) - X_2(\tau\Delta)]}{2} - \frac{\alpha[1 - X_2(\tau\Delta)]^2}{2} - \frac{[\alpha(X_2(\tau\Delta) - X_1(\tau\Delta)]^2}{8}. \tag{20}$$

For $i = 1, 2$, let u_i be the N-vector $(u_{i,0}, \dots, u_{i,N-1})$ and let $V_i(u_1, u_2)$ be the utility function of agent i when the selected actions of agents are given by the N-vectors u_1 and u_2, respectively. The utility functions are expressed through total demands and are given by

$$V_1(u_1, u_2) = \sum_{h=1}^{N} \Delta D_1(X_1(h\Delta), X_2(h\Delta)), \tag{21}$$

$$V_2(u_1, u_2) = \sum_{h=1}^{N} \Delta D_2(X_1(h\Delta), X_2(h\Delta)) \tag{22}$$

where $X_1(\cdot)$ and $X_2(\cdot)$ are given by (16)–(17). Evidently, the functions $V_1(u_1, u_2)$ and $V_2(u_1, u_2)$ are (jointly) strictly concave in u_1 and u_2 (they are trivially strictly concave in $X_1(\cdot)$ and $X_2(\cdot)$, and these depend linearly on the u_i's). Consequently, the optimization problem, where the sum of the V_i's is maximized subject to (18), has a unique optimal solution, which we denote (u_1^*, u_2^*); further, the results of Rothblum [26] apply and show that this optimal solution becomes a Nash equilibrium, if per-unit rewards are imposed on the instantaneous speeds. Imposing these rewards here is done, by adding to the utility function of agent i, the expression $\sum_\tau C_{i,\tau} \cdot u_{i,\tau}$, where

$$C^*_{1,\tau} = \frac{\partial V_2(u_1, u_2)}{\partial u_{1,\tau}}\bigg|_{(u_1,u_2)=(u_1^*,u_2^*)}, \tag{23}$$

$$C^*_{2,\tau} = \frac{\partial V_1(u_1, u_2)}{\partial u_{2,\tau}}\bigg|_{(u_1,u_2)=(u_1^*,u_2^*)}. \tag{24}$$

Let $(X_1^*(\cdot), X_2^*(\cdot))$ be the outcome of applying (16)–(17) to (u_1^*, u_2^*). Observing that

$$\frac{\partial X_i(h\Delta)}{\partial u_{j,\tau}} = \begin{cases} \Delta, & j=1,\ h>\tau, \\ -\Delta, & j=2,\ h>\tau, \\ 0, & \text{otherwise} \end{cases} \tag{25}$$

we have from (23)–(24) that

$$\begin{aligned} C^*_{1,\tau} &= \frac{\partial V_2(u_1, u_2)}{\partial u_{1,\tau}}\bigg|_{(u_1,u_2)=(u_1^*,u_2^*)} \\ &= \sum_{h=\tau+1}^{N} \Delta \left\{ \frac{-\beta(h\Delta)}{2} + \frac{\alpha[X_2^*(h\Delta) - X_1^*(h\Delta)]}{4} \right\} \Delta, \end{aligned} \tag{26}$$

$$\begin{aligned} C^*_{2,\tau} &= \frac{\partial V_1(u_1, u_2)}{\partial u_{2,\tau}}\bigg|_{(u_1,u_2)=(u_1^*,u_2^*)} \\ &= \sum_{h=\tau+1}^{N} \Delta \left\{ \frac{-\beta(h\Delta)}{2} + \frac{\alpha[X_2^*(h\Delta) - X_1^*(h\Delta)]}{4} \right\} \Delta. \end{aligned} \tag{27}$$

Using the representation of integrals as limits of Rieman-sums, we conclude that the reward mechanism, that converts an optimal solution to a Nash equilibrium, is expressed by $\int_0^T C_i^*(t) u_i(t) dt$, where

$$C_1^*(t) = C_2^*(t) = \int_t^T \left(-\frac{\beta(s)}{2} + \frac{\alpha[X_2^*(s) - X_1^*(s)]}{4} \right) ds \tag{28}$$

here $X_i^*(S)$ is the limit of $X_i^*(h\Delta)$, when $\Delta \to 0$ and $h\Delta = s$. Continuity arguments show that $X_i^*(\cdot)$ is the optimal solution to the original control problem and therefore (28) is consistent with the optimal $B_i(\cdot)$'s of Corollary 4.1 (except for a sign change due to the minus in (12)).

Finally, we make two observations about the expression of the penalties that are given in (28) (or equivalently in (12)–(13)). As $\int_0^T u_1(t)dt = X_1(T) = 0$ and $\int_0^T u_2(t)dt = 1 - X_2(T) = 0$, we have that with $h(s)$ as the integrand in (28),

$$\begin{aligned} \int_0^T C_i^*(t) u_i(t) dt &= \int_0^T \left[\int_t^T h(s) ds \right] u_i(t) dt \\ &= \left[\int_0^T h(s) ds \right] \left[\int_0^T u_i(t) dt - \int_0^T \left[\int_0^t h(s) ds \right] u_i(t) dt \right. \end{aligned}$$

$$= 0 - \int_0^T \left[\int_0^t h(s)ds \right] u_i(t)dt.$$

Thus, an alternative expression for the rewards/penalties is given by $C_i^{\#}(t) = -\int_0^t h(s)ds$. The second observation uses integration by parts to convert expressions of rewards or penalties on the agents' speeds to expressions that refer to their locations. Such a transformation is of special importance in terms of the potential applications discussed in the Introduction. Specifically, based on $C_i^*(T) = 0$ and $X_1(0) = X_1(T) = 0$, $X_2(0) = X_2(T) = 0$, we have that

$$\int_0^T C_1^*(t)u_1(t)dt = C_1^*(t)X_1(t)\Big|_0^T - \int_0^T \dot{C}_1^*(t)X_1(t)dt = 0 - \int_0^T \dot{C}_1^*(t)X_1(t)dt$$

and

$$\begin{aligned}\int_0^T C_2^*(t)u_2(t)dt &= C_2^*(t)[-X_2(t)]\Big|_0^T - \int_0^T \dot{C}_2^*(t)[-X_2(t)]dt \\ &= C_2^*(0) + \int_0^T \dot{C}_2^*(t)X_2(t)dt.\end{aligned}$$

So, instantaneous reward/penalty $C_i^*(t)$ on the speed $u_i(t)$ of agent i at time t is converted to instantaneous reward/penalty $\dot{C}_i^*(t)$ on the location $X_i(t)$ of agent i at time t. The new expressions are useful when agents' locations are easier to observe than their speeds. These two approaches are well known in real-life, as they correspond to penalties for speeding, when the speed limits depend on rush hours as well as location charges, which are time-dependent in some parking lots or municipal zones.

5 Conclusions

We consider a dynamic system, where two asymmetric mobile agents compete for the same customers to maximize their profits. The competition is modeled as a differential game and conditions are determined when Nash equilibrium exists. This non-cooperative behavior is contrasted to an integrated system, with a single DM maximizing the overall profit. We derive that both system-wide optimal and Nash solutions involve four breaking points that define regions in which the agents exercise different speeds. We show that, despite these structural similarities, the timing and the speeds of the agents may differ, depending on whether the decisions are centralized or not. Specifically, we prove that non-cooperative agents travel more time, come closer to each other, even if the demand is time invariant, and the fast agent always moves slower during the intervals of intermediate speed levels. Moreover these differences deepen when the demand is time dependent.

To improve the system's performance in a non-cooperative environment, we employ a class of linear rewards/penalties on the speeds of the agents. As in the static case, these penalties are derived for each agent, according to the marginal loss it causes the other agent, by increasing its speed by one unit. We prove that these penalties ensure that the non-cooperative agents will choose a *unique* Nash equilibrium solution, which coincides with the global optimal solution, which would have been imposed on them, if there was a central decision-maker. We further show that it is easy to convert the speed-dependent rewards (penalties) to location-dependent rewards (penalties). Such conversion might be useful, if the location of each agent is readily observable while its speed is not.

Our work is one of only a few attempts in literature to coordinate dynamic non-cooperative systems. It shows that though such coordination is more complex, compared to that in the corresponding static systems, it is still possible to develop efficient schemes. Future research directions would include such challenges as coordination under multiple agents and multiple coordinate dimensions, non-identical agents starting from multiple bases and completing their tasks at given destinations (analogous to the multi-commodity work of Tapiero and Soliman [24]), agents' acceleration limits and stochastic fluctuations in the coordinates of moving agents. Also, while we allowed asymmetry only on the "supply-side" (i.e., the agents), in future work we intend to explore asymmetry on the "demand-side" as well (i.e., customers that are not necessarily spread uniformly along the route or customers' distribution that changes over time).

Acknowledgements The research of the first two authors was partially supported by the Gordon Center for Systems Engineering and the Center for Science & Technology of Security, both at the Technion, Israel Institute of Technology, and by a grant for Promotion of Research at the Technion.

Appendix A: Proofs

A.1 Proof of Theorem 3.1

In this subsection, we prove Theorem 3.1, deriving the solution for the problem of maximizing (10), subject to (1)–(4). For brevity, we refer to controls that solve this problem simply as optimal controls.

Lemma A.1 *$u_1(\cdot)$ and $u_2(\cdot)$ are optimal controls, if and only if they satisfy (3) and there exist variables $X_1(\cdot)$, $X_2(\cdot)$ and $\psi_1(\cdot)$, $\psi_2(\cdot)$, such that for every time t, (1)–(2) and the following equations are satisfied.*

$$\dot{\psi}_1(t) = -\frac{1}{2}\alpha X_2(t) + \frac{3}{2}\alpha X_1(t), \tag{29}$$

$$\dot{\psi}_2(t) = -\alpha - \frac{1}{2}\alpha X_1(t) + \frac{3}{2}\alpha X_2(t), \tag{30}$$

$$u_1(t) = \begin{cases} U_1, & \textit{if } \psi_1(t) > 0, \\ -U_1, & \textit{if } \psi_1(t) < 0, \end{cases} \tag{31}$$

$$u_2(t) = \begin{cases} -U_2, & \textit{if } \psi_2(t) > 0, \\ U_2, & \textit{if } \psi_2(t) < 0. \end{cases} \tag{32}$$

Proof We first establish (joint) concavity of the objective function in (10), as a function of the control variables, $u_1(\cdot)$ and $u_2(\cdot)$. As (1)–(2) establish linear dependence of each $X_1(t)$ and $X_2(t)$ on the corresponding control variables, and the dependence of the objective function on $X_1(\cdot)$, $X_2(\cdot)$ is through the integral of the total instantaneous demand over the interval $[0, T]$, it suffices to show for each t that the total instantaneous demand at that time is a concave function of $X_1(t)$, $X_2(t)$. Indeed, using (11), for each t the corresponding Hessian is negative definite as it equals

$$\begin{bmatrix} -\frac{3}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{3}{2} \end{bmatrix}.$$

The conditions stated in the lemma are the standard conditions that are derived from the Maximum Principle; under concavity, these are known to characterize optimal controls (see p. 61 of Sethi and Thompson [27] for necessity and p. 64 for sufficiency, the latter referring to Sydsaeter ([28], p. 187)). □

The common use of the "Maximum Principle" refers just to the necessity of the conditions for optimal controls (stated in Lemma A.1 for our problem).

We refer to the variables $\psi_1(t)$ and $\psi_2(t)$ that appear in the conditions of Lemma A.1 as the *co-state variables*. The next lemma provides implications of the conditions that are shown in Lemma A.1 to characterize optimal controls.

Lemma A.2 *Let $u_1(\cdot)$, $u_2(\cdot)$, $X_1(\cdot)$, $X_2(\cdot)$ and $\psi_1(\cdot)$, $\psi_2(\cdot)$ be control variables, state variables and co-state variables, respectively, that satisfy the characterizing conditions of Lemma A.1; then for every nonempty interval $I \subseteq [0, T]$*

(i) *if $-U_1 < u_1(t) < U_1$ for $t \in I$, then $\psi_1(t) = 0$, implying that $-1X_2(t) + 3X_1(t) = 0$ and $u_2(t) = -3u_1(t)$ for $t \in I$,*
(ii) *if $-U_2 < u_2(t) < U_2$ for $t \in I$, then $\psi_2(t) = 0$, implying that $3X_2(t) - X_1(t) = 2$ and $3u_2(t) = -u_1(t)$ for $t \in I$,*
(iii) *if both $-U_1 < u_1(t) < U_1$ and $-U_2 < u_2(t) < U_2$ for $t \in I$, then $\psi_1(t) = \psi_2(t) = 0$, implying that $X_1(t) = 1/4$, $X_2(t) = 3/4$ and $u_2(t) = u_1(t) = 0$ for $t \in I$.*

Proof If $u_1(t)$ takes intermediate values everywhere on $I \subseteq [0, T]$, then (31) implies that $\psi_1(t) = 0$ for $t \in I$; differentiating with respect to t and using (29), we conclude that $0 = \dot{\psi}_1(t) = -\frac{1}{2}\alpha X_2(t) + \frac{3}{2}\alpha X_1(t)$ for $t \in I$. Next, on differentiation, the equations $-\frac{1}{2}\alpha X_2(t) + \frac{3}{2}\alpha X_1(t) = 0$ and accounting for (1)–(2), it follows that $u_2(t) = -3u_1(t)$ for $t \in I$. This establishes (i). Also, similar arguments, using (30) and (32) instead of (29) and (31), establishes (ii).

Finally, (iii) follows from combining (i) and (ii), and observing that $-1X_2(t) + 3X_1(t) = 0$ and $3X_2(t) - X_1(t) = 2$, imply that $X_1(t) = 1/4$ and $X_2(t) = 3/4$. Also, if $u_2(t) = -3u_1(t)$ and $3u_2(t) = -u_1(t)$, then necessarily $u_2(t) = u_1(t) = 0$. □

Lemma A.3 *Let $k \geq 1$, $X_1(\cdot)$, $X_2(\cdot)$, $u_1(\cdot)$, $u_2(\cdot)$ satisfy (1)–(3) and $u_1(t) = U_1$, $u_2(t) = U_2$ for t in the time-interval $[0, \tau]$.*

If $\frac{X_2(\tau)}{X_1(\tau)} = k$, then $X_1(\tau) > \frac{1}{k+1}$, $X_2(\tau) > \frac{k}{k+1}$ and $\tau = \frac{1}{kU_1+U_2}$.

Proof First, note that since both agents move at maximum speed for $t \in [0, \tau]$, then $X_1(\tau) = U_1\tau$, $X_2(\tau) = 1 - U_2\tau$ and thereby $k = \frac{X_2(\tau)}{X_1(\tau)} = \frac{1-U_2\tau}{U_1\tau}$, i.e., $\tau = \frac{1}{kU_1+U_2}$, as stated in this lemma. Next, we obtain the coordinates at $t = \tau$ by substituting the found value for τ into expressions for $X_1(\tau)$ and $X_2(\tau)$, and taking into account that $U_1 > U_2$,

$$X_1(\tau) = \frac{U_1}{kU_1 + U_2} > \frac{U_1}{kU_1 + U_1} = \frac{1}{k+1};$$
$$X_2(\tau) = 1 - \frac{U_2}{kU_1 + U_2} > 1 - \frac{U_2}{kU_2 + U_2} = \frac{k}{k+1}. \qquad \square$$

Lemma A.3 provides simple results with respect to the state variables, when agents that move at their maximum speeds desire to attain a given ratio between their coordinates. The next lemma analyzes the simultaneous behavior of the state and co-state variables.

Lemma A.4 *Let there exist points of time s and l, $s < l$, such that $\psi_1(t) = 0$ for $t \in [s, l]$, $\psi_2(t) = 0$ for $t \geq l$, $X_1(t)$, $X_2(t)$, $\psi_1(t)$, $\psi_2(t)$ satisfy (1), (2), (29) and (30) and $u(t)$ satisfies (3).*

(i) *If $u_1(t) = U_1$ for $t < s$, $u_1(t) = -\frac{U_2}{3}$ for $t \in [s, l]$ and $u_2(t) = U_2$ for $t < l$, then $\psi_1(t) > 0$ for $t < s$ and $\psi_2(t) < 0$ for $t < l$.*
(ii) *If $u_1(t) = U_1$ and $u_2(t) = -U_2$ for $t > l$, then $\psi_1(t) < 0$ and $\psi_2(t) > 0$ for $t > l$.*

Proof Given $\psi_1(t) = 0$ for $t \in [s, l]$, then according to Lemma A.2, $u_2(t) = -3u_1(t)$, $t \in [s, l]$. Since $\dot{\psi}_2(l) = 0$, $u_1(t) = -\frac{U_2}{3}$ and $u_2(t) = U_2$ for $t \in [s, l]$, differentiating (14) we find

$$\ddot{\psi}_2(t) = -\frac{1}{2}\alpha\dot{X}_1(t) + \frac{3}{2}\alpha\dot{X}_2(t) = -\frac{1}{2}\alpha u_1(t) - \frac{3}{2}\alpha u_2(t) < 0.$$

That is, $\dot{\psi}_2(t) > 0$, which, with respect to the fact that $\psi_2(l) = 0$, implies $\psi_2(t) < 0$ for $t \in [s, l]$. Moreover, if $\dot{X}_1(t) = u_1(t) = U_1$ and $\dot{X}_2(t) = -u_2(t) = -U_2$ for $t < s$, then $\ddot{\psi}_2(t) = -\frac{U_1}{2}\alpha - \frac{3U_2}{2}\alpha < 0$, thereby $\dot{\psi}_2(t) > 0$ for $t < s$ and, since we have proved that $\psi_2(t) < 0$ for $t \in [s, l]$, we now have $\psi_2(t) < 0$ for $t < s$, as stated in (i). Similarly, one can verify that

$$\ddot{\psi}_1(t) = -\frac{1}{2}\alpha\dot{X}_2(t) + \frac{3}{2}\alpha\dot{X}_1(t) = \frac{1}{2}\alpha U_2(t) + \frac{3}{2}\alpha U_1(t) > 0$$

and since $\dot{\psi}_1(s) = 0$, we obtain $\psi_1(t) > 0$ for $t < s$, as also stated in (i). Finally, if $u_1(t) = -U_1$ and $u_2(t) = -U_2$ for $t > l$, then

$$\ddot{\psi}_2(t) = \frac{1}{2}\alpha U_1 + \frac{3}{2}\alpha U_2 > 0, \qquad \ddot{\psi}_1(t) = -\frac{1}{2}\alpha U_2(t) - \frac{3}{2}\alpha U_1(t) < 0$$

for $t > l$, while $\dot{\psi}_1(l) = \dot{\psi}_2(l) = 0$. Therefore, $\dot{\psi}_1(t) < 0$ and $\dot{\psi}_2(t) > 0$ for $t > l$ and, taking into account that $\psi_1(l) = 0$ and $\psi_2(l) = 0$, we find that $\psi_1(t) < 0$ and $\psi_2(t) < 0$ for $t > l$, as stated in (ii). $\square$

Proof of Theorem 3.1 Our approach is to construct a solution that satisfies the optimality conditions of Lemma A.1. Consider the following candidate solution $u_1(t) = U_1$ for $0 \le t < t_1$; $u_1(t) = -\frac{U_2}{3}$ for $t_1 \le t < t_2$, $u_1(t) = 0$ for $t_2 \le t < t_{-2}$, $u_1(t) = \frac{U_2}{3}$ for $t_{-2} \le t < t_{-1}$ and $u_1(t) = -U_1$ for $t_{-1} \le t \le T$; $u_2(t) = U_2$ for $0 \le t < t_2$, $u_2(t) = 0$ for $t_2 \le t < t_{-2}$ and $u_2(t) = -U_2$ for $t_{-2} \le t \le T$.

It is easy to observe that this solution consists of two symmetrical parts (see Fig. 1). As a result, the proofs for the two parts are very similar. We, therefore, focus here only on the proof for the first half of the planning horizon. The proof is a straightforward application of Lemmas A.2–A.3.

Specifically, using Lemma A.3 and choosing $k = 3$, we find that the first half of the candidate solution: $u_1(t) = U_1$ for $0 \le t < t_1$, $u_1(t) = -\frac{U_2}{3}$ for $t_1 \le t < t_2$, and $u_2(t) = U_2$ for $0 \le t < t_2$ is always feasible if we set $t_1 = \tau = \frac{1}{3U_1+U_2}$ and $t_2 = \frac{1}{4U_2}$ so that $\frac{X_2(t)}{X_1(t)} = 3$ will be kept for $t_1 \le t < t_2$ (and thereby assumption (4) is met), $X_2(t_2) = 1 - U_2 t_2 = 3/4$ and $X_1(t_2) = 1/4$. Note, that according to the candidate solution, at t_1 the agents have the closest location after moving towards each other. After t_1, the fast agent retreats. Clearly, this general solution can fully realize, over the planning horizon, if the idling period $(u_1(t) = u_2(t) = 0)$ is non-negative, i.e., $t_{-2} \ge t_2$, from where we obtain the last feasibility requirement, $T - 1/2U_2 \ge 0$. This defines the minimum length of the planning horizon, which is ensured under the conditions of the theorem.

Since the found solution is feasible, we next verify the optimality conditions for the first half of the trajectory. Consider the following solution for the co-state variables

$$\psi_1(t) = 0 \quad \text{for } t \in [t_1, t_{-1}] \quad \text{and} \quad \psi_2(t) = 0 \quad \text{for } t \in [t_2, t_{-2}].$$

According to Lemma A.2, with such a co-state solution, statement (i) of this lemma holds for $t \in [t_1, t_2]$ and statement (iii) holds for $t \in [t_2, t_{-2}]$. That is, the suggested candidate solution satisfies the co-state equations (29)–(30) and the optimality conditions (31)–(32) are met for the corresponding periods. Furthermore, we observe that the conditions of Lemma A.4 are met as well. In particular, according to statement (i) (Lemma A.4), $\psi_1(t) > 0$ for $t \in [0, t_1]$ and $\psi_2(t) < 0$ for $t \in [0, t_2]$ and thereby the optimality conditions for the maximum speed of the agents are met. We thus verified optimality conditions (31)–(32) for $t \in [0, t_{-2}]$ for the suggested solution. The proof is complete by denoting $\tau_1 = t_1$, $\tau_2 = t_2$, $\tau_{-2} = T - \tau_2 = t_{-2}$ and $\tau_{-1} = T - \tau_1 = t_{-1}$. □

A.2 Proof of Theorem 3.2

In this subsection, we prove Theorem 3.2, deriving the joint solution for the problems of maximizing (9) for $i = 1, 2$, subject to (1)–(4). For brevity, we refer to controls that jointly solve the two problems as *Nash equilibrium controls*.

Lemma A.5 *Assume that $\beta(t)$ be differentiable for $t \in [0, T]$. Then $u_1(\cdot)$ and $u_2(\cdot)$ are Nash equilibrium controls, if and only if they satisfy* (3) *and there exist variables $X_1(\cdot)$, $X_2(\cdot)$ and $\lambda_1(\cdot)$, $\lambda_2(\cdot)$ such that for every time t,* (1)–(2) *and the following* *equations are satisfied.*

$$\dot{\lambda}_1(t) = -\frac{1}{2}\beta(t) - \frac{1}{4}\alpha X_2(t) + \frac{5}{4}\alpha X_1(t), \tag{33}$$

$$\dot{\lambda}_2(t) = \frac{1}{2}\beta(t) - \alpha - \frac{1}{4}\alpha X_1(t) + \frac{5}{4}\alpha X_2(t), \tag{34}$$

$$u_1(t) = \begin{cases} U_1, & \text{if } \lambda_1(t) > 0; \\ -U_1, & \text{if } \lambda_1(t) < 0; \end{cases} \tag{35}$$

$$u_2(t) = \begin{cases} -U_2, & \text{if } \lambda_2(t) > 0; \\ U_2, & \text{if } \lambda_2(t) < 0. \end{cases} \tag{36}$$

Proof For $i = 1, 2$, the objective function in (9) is a concave function of the corresponding location variables, with the location variables of the other agent given. By the linear dependence on the location variables on the control variables, it follows that the objective function is concave as a function of the control variables. The fact that the conditions stated of our lemma characterize Nash equilibrium solutions now follows from the corresponding arguments of Lemma A.1. □

We shall refer to the variables $\lambda_1(t)$ and $\lambda_2(t)$ that appear in the conditions of Lemma A.5 as the *co-state variables*. The next lemma provides implications of the conditions that are shown in Lemma A.5 to characterize optimal controls.

Lemma A.6 *Assume that $\beta(t)$ be differentiable for $t \in [0, T]$. Let $u_1(\cdot)$, $u_2(\cdot)$, $X_1(\cdot)$, $X_2(\cdot)$ and $\lambda_1(\cdot)$, $\lambda_2(\cdot)$ be control variables, state variables and co-state variables that satisfy the characterizing conditions of Lemma A.5; then for every nonempty interval $I \subseteq [0, T]$,*

(i) *if $-U_1 < u_1(t) < U_1$ for $t \in I$, then $\lambda_1(t) = 0$, implying that $-X_2(t) + 5X_1(t) = \frac{2\beta(t)}{\alpha}$ and $u_2(t) + 5u_1(t) = \frac{2\dot{\beta}(t)}{\alpha}$ for $t \in I$,*
(ii) *if $-U_2 < u_2(t) < U_2$ for $t \in I$, then $\lambda_2(t) = 0$, implying that $-X_1(t) + 5X_2(t) = 4 - \frac{2\beta(t)}{\alpha}$ and $u_1(t) + 5u_2(t) = \frac{2\dot{\beta}(t)}{\alpha}$ for $t \in I$,*
(iii) *if $-U_1 < u_1(t) < U_1$ and $-U_2 < u_2(t) < U_2$ for $t \in I$, then $\lambda_1(t) = \lambda_2(t) = 0$, implying that $X_1(t) = \frac{1}{6} + \frac{\beta(t)}{3\alpha}$, $X_2(t) = \frac{5}{6} - \frac{\beta(t)}{3\alpha}$ and $u_2(t) = u_1(t) = \frac{\dot{\beta}(t)}{3\alpha}$ for $t \in I$.*

Proof The proof of the lemma follows from the arguments proving Lemma A.2. □

An immediate observation from Lemma A.6 is that an equilibrium solution for the two agents is very different from the system-wide optimal solution, even if the potential demand $\beta(t)$ is time invariant. Moreover, unlike the system-wide optimal coordinates, $X_1^*(t_2) = 1/4$ and $X_2^*(t_2) = 3/4$ (where the agents stop for a while), under Nash strategy, the agents will never stop moving, if $\dot{\beta}(t) \neq 0$ for $t \in [0, T]$. More precisely, if condition (iii) of Lemma A.6 is met, the agents will move around according to coordinates, $X_1^N(t) = \frac{1}{6} + \frac{\beta(t)}{3\alpha}$, $X_2^N(t) = \frac{5}{6} - \frac{\beta(t)}{3\alpha}$ (where superscript N designates Nash equilibrium).

There is a similarity, however, with the system-wide solution as well. The total coordinate under condition (iii) remains constant, $X_2^N(t)+X_1^N(t)=X_2^*(t)+X_1^*(t)=1$.

Another similarity with respect to condition (iii) is that the agents do not cross. Indeed, the condition

$$X_2^N(t)-X_1^N(t)=\frac{2}{3}-\frac{2}{3\alpha}\beta(t)\geq 0$$

always holds (recall the sufficient condition, $\beta(t)-\frac{\alpha}{2}\geq 0$, $t\in[0,T]$ for the instantaneous demand to be non-negative). In addition, we observe that the agents can meet under condition (iii), if $\beta(t)-\alpha=0$ (this will happen in the middle of the route, $X_1^N(t)=X_2^N(t)=1/2$). Therefore, to prevent the violation of condition (iii) we henceforth assume (a very natural assumption) that an agent located at one end of the interval is not able to cover more than the demand across the entire interval [0, 1], i.e., $\beta(t)-\alpha\leq 0$. Hence, $\frac{\alpha}{2}\leq\beta(t)\leq\alpha$ for $t\in[0,T]$.

Lemma A.7 *Let* $k\geq 1$, $X_1(\cdot)$, $X_2(\cdot)$, $u_1(\cdot)$, $u_2(\cdot)$ *satisfy* (1)–(3), *and both agents move at maximum speeds during an interval of time* $[0,\tau]$. *If,* $X_1(0)=0$, $X_2(0)=1$, *and* $-X_2(\tau)+kX_1(\tau)=\frac{2\beta(\tau)}{\alpha}$, *then*

$$X_1(\tau)>\frac{1}{k+1}+\frac{2\beta(\tau)}{(k+1)\alpha},\qquad X_2(\tau)>\frac{k}{1+k}-\frac{2\beta(\tau)}{(1+k)\alpha},\qquad \tau=\frac{\frac{2\beta(\tau)}{\alpha}+1}{U_2+kU_1}.$$

Proof First, note that since both agents move at maximum speed for $t\in[0,\tau]$, then $X_1(\tau)=U_1\tau$, $X_2(\tau)=1-U_2\tau$ and thereby $-1+U_2\tau+kU_1\tau=\frac{2\beta(\tau)}{\alpha}$, i.e., $\tau=\frac{\frac{2\beta(\tau)}{\alpha}+1}{U_2+kU_1}$, as stated in this lemma. Next, we obtain the coordinates at $t=\tau$, by substituting the found value for τ into expressions for $X_1(\tau)$ and $X_2(\tau)$, and taking into account that $U_1>U_2$,

$$X_1(\tau)=\frac{\frac{2\beta(\tau)}{\alpha}+1}{U_2+kU_1}U_1>\frac{\frac{2\beta(\tau)}{\alpha}+1}{U_1+kU_1}U_1=\frac{1}{k+1}+\frac{2\beta(\tau)}{(k+1)\alpha};$$

$$X_2(\tau)=1-\frac{\frac{2\beta(\tau)}{\alpha}+1}{U_2+kU_1}U_2>1-\frac{\frac{2\beta(\tau)}{\alpha}+1}{U_2+kU_2}U_2=\frac{k}{1+k}-\frac{2\beta(\tau)}{(1+k)\alpha}.\qquad\square$$

Similar to Lemma A.3, Lemma A.7 shows (by setting $k=5$) that if the agents move at maximum speed within a period of time $[0,\tau]$ to ensure $-X_2(\tau)+5X_1(\tau)=\frac{2\beta(\tau)}{\alpha}$, then the fast agent ($i=1$) will overrun and the slow agent ($i=2$) will not reach the coordinates $X_1^N(t)=\frac{1}{6}+\frac{\beta(t)}{3\alpha}$, $X_2^N(t)=\frac{5}{6}-\frac{\beta(t)}{3\alpha}$, respectively. Note, that due to our assumption, $\frac{\alpha}{2}\leq\beta(t)\leq\alpha$, the agents cannot meet at $t_1=\tau$. Indeed,

$$X_2(\tau)-X_1(\tau)=1-\frac{\frac{2\beta(\tau)}{\alpha}+1}{U_2+5U_1}U_2-\frac{\frac{2\beta(\tau)}{\alpha}+1}{U_2+5U_1}U_1\geq 0,$$

is equivalent to, $1\geq\frac{\beta(\tau)}{\alpha}\frac{U_1+U_2}{2U_1}$, which always holds.

The next lemma describes the simultaneous behavior of the state and co-state variables.

Lemma A.8 *Let there exist points of time s and l, $s < l$, $X_1(t)$, $X_2(t)$, $\lambda_1(t)$, $\lambda_2(t)$ satisfy* (1), (2), (19) *and* (20) *and $u(t)$ satisfies* (3), *and assume that $\lambda_1(t) = 0$, $|\frac{\dot{\beta}(t)}{3\alpha}| < U_2$ for $t \in [s, l]$ and $\lambda_2(t) = 0$ for $t \geq l$.*

(i) *If $u_1(t) = U_1$ for $t < s$, $U_2 + 5u_1(t) = \frac{2\dot{\beta}(t)}{\alpha}$ for $t \in [s, l]$ and $u_2(t) = U_2$ for $t < l$, then $\lambda_1(t) > 0$ for $t < s$ and $\lambda_2(t) < 0$ for $t < l$.*
(ii) *If $u_1(t) = -U_1$ and $u_2(t) = -U_2$ for $t > l$, then $\lambda_1(t) < 0$ and $\lambda_2(t) > 0$ for $t > l$.*

Proof Given $\lambda_1(t) = 0$ for $t \in [s, l]$, then according to Lemma A.6, $u_2(t) + 5u_1(t) = \frac{2\dot{\beta}(t)}{\alpha}$, $t \in [s, l]$. Since $u_1(t) = \frac{2\dot{\beta}(t)}{5\alpha} - \frac{U_2}{5}$ and $u_2(t) = U_2$ for $t \in [s, l]$, differentiating (20) and taking into account that $|\frac{\dot{\beta}(t)}{3\alpha}| < U_2$, we find

$$\ddot{\lambda}_2(t) = \frac{1}{2}\dot{\beta}(t) - \frac{1}{4}\alpha \dot{X}_1(t) + \frac{5}{4}\alpha \dot{X}_2(t) = \frac{2}{5}\dot{\beta}(t) - \frac{6}{5}\alpha U_2(t) < 0.$$

Since $\lambda_2(t) = 0$ for $t \geq l$, then $\dot{\lambda}_2(l) = 0$, that is, from $\ddot{\lambda}_2(t) < 0$, we have $\dot{\lambda}_2(t) < 0$. Accordingly, from $\lambda_2(l) = 0$ and $\dot{\lambda}_2(t) < 0$, we find, $\lambda_2(t) < 0$ for $t \in [s, l)$. Moreover, if $\dot{X}_1(t) = u_1(t) = U_1$ and $\dot{X}_2(t) = -u_2(t) = -U_2$ for $t < s$, then

$$\ddot{\lambda}_2(t) = \frac{1}{2}\dot{\beta}(t) - \frac{1}{4}\alpha U_1(t) - \frac{5}{4}\alpha U_2(t) < 0,$$

thereby $\dot{\lambda}_2(t) > 0$ for $t < s$ and, since we have proved that $\lambda_2(t) < 0$ for $t \in [s, l)$, we now have $\lambda_2(t) < 0$ for $t < s$, as stated in (i). Similarly, one can verify with (19) that

$$\ddot{\lambda}_1(t) = -\frac{1}{2}\dot{\beta}(t) - \frac{1}{4}\alpha \dot{X}_2(t) + \frac{5}{4}\alpha \dot{X}_1(t) = -\frac{1}{2}\dot{\beta}(t) + \frac{1}{4}\alpha U_2(t) + \frac{5}{4}\alpha U_1(t) > 0;$$

and, since $\dot{\lambda}_1(s) = 0$, we obtain $\lambda_1(t) > 0$ for $t < s$, as also stated in (i).

Finally, if $u_1(t) = -U_1$ and $u_2(t) = -U_2$ for $t > l$, then

$$\ddot{\lambda}_1(t) = -\frac{1}{2}\dot{\beta}(t) - \frac{1}{4}\alpha U_2(t) - \frac{5}{4}\alpha U_1(t) < 0,$$

$$\dot{\lambda}_2(t) = \frac{1}{2}\dot{\beta}(t) + \frac{1}{4}\alpha U_1(t) + \frac{5}{4}\alpha U_2(t) > 0 \quad \text{for } t > l,$$

while $\dot{\lambda}_1(l) = \dot{\lambda}_2(l) = 0$. Therefore, $\lambda_1(t) > 0$ and $\lambda_2(t) < 0$ for $t > l$ and, taking into account that $\lambda_1(l) = 0$ and $\lambda_2(l) = 0$, we find that $\lambda_1(t) < 0$ and $\lambda_2(t) > 0$ for $t > l$, as stated in (ii). □

Proof of Theorem 3.2 Similar to the proof of Theorem 3.1, our approach is to construct a solution and verify that it satisfies the optimality conditions of Lemma A.5. Consider the solution, stated in this theorem, as a candidate solution. It is easy to observe that this solution consists of two symmetrical parts. We focus here only on the proof for the first half of the planning horizon.

Specifically, using Lemma A.7 and choosing $k = 5$, we find that the first half of the candidate solution: $u_1(t) = U_1$ for $0 \leq t < t_1$, $u_1(t) = \frac{2\dot{\beta}(t)}{5\alpha} - \frac{U_2}{5} < 0$ for $t_1 \leq t < t_2$, is always feasible if we set $t_1 = \tau = \frac{\frac{2\beta(\tau)}{\alpha}+1}{U_2+5U_1}$ and t_2 so that $-X_2(t) + 5X_1(t) = \frac{2\beta(t)}{\alpha}$ holds, for $t_1 \le t \le t_2$, i.e., $X_2(t_2) = 1 - U_2 t_2 = \frac{5}{6} - \frac{\beta(t_2)}{3\alpha}$ and $X_1(t_2) = \frac{1}{6} + \frac{\beta(t_2)}{3\alpha}$. Clearly, this general solution can be fully realized over the planning horizon, if $t_{-2} \ge t_2$, as ensured by $T \ge 1/U_2$ in the theorem.

Since the found solution is feasible, we next verify the optimality conditions for the first half of the trajectory. Consider the following solution for the co-state variables

$$\lambda_1(t) = 0 \quad \text{for } t \in [t_1, t_{-1}] \quad \text{and} \quad \lambda_2(t) = 0 \quad \text{for } t \in [t_2, t_{-2}].$$

According to Lemma A.6, with such a co-state solution, statement (i) of this lemma holds for $t \in [t_1, t_2]$ and statement (iii) holds for $t \in [t_2, t_{-2}]$. That is, the suggested candidate solution satisfies the co-state equations (19)–(20), and the optimality conditions (21)–(22) are met for the corresponding periods. Furthermore, we observe that the conditions of Lemma A.7 are met as well. In particular, according to statement (i) (Lemma A.7), $\lambda_1(t) > 0$ for $t \in [0, t_1]$ and $\lambda_2(t) < 0$ for $t \in [0, t_2]$ and thereby the optimality conditions for the maximum speed of the agents are met. We thus verified optimality conditions (21)–(22) for $t \in [0, t_{-2}]$ for the suggested solution which completes the proof.

Finally, we have found a unique Nash solution for the *selected feasible behavior* of co-state variables $\lambda_1(t)$ and $\lambda_2(t)$. To prove that this solution is always unique, we next show that there is not any other feasible behavior for the co-state variables, which satisfies the equilibrium conditions (21)–(22). Specifically, there are three alternative trajectories for the co-state variables. First alternative occurs when one of the co-state variables (or both) is zero at an initial time interval, rather than for $t \in [t_1, t_{-1}]$ or $t \in [t_2, t_{-2}]$. Then we readily observe that no condition of Lemma A.2 can be met at $t = 0$ and thereby, $X_1(0) = 0$ and $X_2(0) = 1$. The second alternative, is when the trajectories of $\lambda_1(t)$ and $\lambda_2(t)$ are swapped, i.e., $\lambda_2(t) = 0$ for $t \in [t_1, t_{-1}]$ and $\lambda_1(t) = 0$ for $t \in [t_2, t_{-2}]$. Then the slow agent will not be able to reach coordinate $X_2(t) = \frac{5}{6} - \frac{\beta(t)}{3\alpha}$ by time t_1 thus condition (iii) of Lemma A.2 will not be met. The last alternative is determined by a trajectory characterized by the co-state variables, which never attain zero level at an interval of time. Then, according to the equilibrium conditions (21)–(22) both agents must move only at maximum speed forward and backward. If there is a single switch in the direction each agent moves, then, since the agents have to return to their starting points exactly by time T, they move forward during $[0, T/2)$ and backward during $[T/2, T]$. In terms of the optimality conditions (21)–(22), this implies that $\dot{\lambda}_1(T/2) < 0$ and $\dot{\lambda}_2(T/2) > 0$; thereby from (19)–(20) we have

$$5X_1\left(\frac{T}{2}\right) - X_2\left(\frac{T}{2}\right) < \frac{2\beta(T/2)}{\alpha} \quad \text{and} \quad X_1\left(\frac{T}{2}\right) - 5X_2\left(\frac{T}{2}\right) + 4 < \frac{2\beta(T/2)}{\alpha}.$$

We next show that the former of the two inequalities never holds under condition $T \ge 1/U_2$ of this theorem. Indeed, since the agents move only at maximum speeds and $\beta \ge \frac{\alpha}{2}$, this inequality holds only if $5U_1\frac{T}{2} - (1 - U_2\frac{T}{2}) < \frac{2\beta(T/2)}{\alpha} < 1$, which after simple manipulations leads to $5U_1 + U_2 < 4/T$. Taking into account that, $U_1 \ge U_2$, the last inequality holds only if $6U_2 < 4/T$, i.e., $T < 2/(3U_2)$. This, however, cannot hold under condition, $T \geq 1/U_2$. Similarly, one can verify, that switching more than once from moving at maximum speed forward to backward does not satisfy (19)–(22). □

A.3 Proofs of the Results of Sect. 4

Proof of Proposition 4.1 First, note that from the Nash solution provided in Theorem 3.2, we have$1 - U_2 t_2 = \frac{5}{6} - \frac{\beta(t_2)}{3\alpha}$, while the corresponding system-wide optimal break point satisfies $t_2 = \frac{1}{4U_2}$. Recalling that $\frac{\alpha}{2} \leq \beta(t) \leq \alpha$, we readily conclude that the break point, t_2 of the Nash solution, is greater than the corresponding break point of the system-wide optimal solution. Similarly, one can verify that t_{-2} is earlier under the non-cooperative (Nash) strategy. Thus, even if $\dot{\beta}(t) = 0$ for $[t_2, t_{-2}]$ (idle period under any solution), the interval of time over which non-cooperative agents not travel is shorter than that of the system-wide optimal solution, as stated in this proposition. Similarly, from Theorems 3.1 and 3.2, we find that

$$X_1^N(t_2) - X_1^*(t_2) = \frac{1}{6} + \frac{\beta(t)}{3\alpha} - \frac{1}{4} \geq \frac{1}{6} + \frac{\alpha/2}{3\alpha} - \frac{1}{4} = \frac{1}{12} > 0,$$

$$X_2^N(t_2) - X_2^*(t_2) = \frac{5}{6} - \frac{\beta(t)}{3\alpha} - \frac{3}{4} \leq \frac{5}{6} - \frac{\alpha/2}{3\alpha} - \frac{3}{4} = -\frac{1}{12} < 0.$$

That is, if $\dot{\beta}(t) = 0$, the agents don't move for $[t_2, t_{-2}]$ and thus remain closer to each other, compared to the corresponding system-wide optimal coordinates. Of course, if $\dot{\beta}(t) \neq 0$, then the non-cooperative agents don't stop and thereby can become only closer, as stated in the proposition.

Finally, comparing the intermediate speeds of the fast agent, when $\dot{\beta}(t) = 0$

$$u_1^*(t) = -\frac{U_2}{3}, \qquad u_1^N(t) = \frac{2\dot{\beta}(t)}{5\alpha} - \frac{U_2}{5},$$

we immediately find that $u_1^N(t) - u_1^*(t) < 0$, as stated in the proposition. □

We shall use "$\hat{}$" to denote variables, expressions, equation numbers and results which refer to the problem, where for each agent $i = 1, 2$, the original instantaneous demands $D_i(\cdot)$ is replaced with $\hat{D}_i(\cdot)$, e.g., we shall refer to $(\hat{33}) - (\hat{34})$, which uses the corresponding co-state variables. In particular, Lemmas A.5 and A.6 then take the following form.

Lemma A.5′ *Assume that for $i = 1, 2$, the instantaneous demand in objective function* (9) *is replaced with $\hat{D}_i[X_1(t), X_2(t)]$, as defined by* (12)–(13). *Let $\beta(t)$ be differentiable for $t \in [0, T]$. Then $u_1(\cdot)$, $u_2(\cdot)$ constitute a Nash equilibrium solution for the modified problem, if and only if they are satisfy* (3) *and there exist variables $X_1(\cdot)$, $X_2(\cdot)$, $\hat{\lambda}_1(\cdot)$, $\hat{\lambda}_2(\cdot)$ such that for every t* (1), (2), $(\hat{33})$, $(\hat{34})$ *and the following equations are satisfied*:

$$u_1(t) = \begin{cases} U_1, & \text{if } \hat{\lambda}_1(t) > B_1(t); \\ -U_1, & \text{if } \hat{\lambda}_1(t) < B_1(t); \end{cases} \tag{37}$$

$$u_2(t) = \begin{cases} -U_2, & \text{if } \hat{\lambda}_2(t) > B_2(t); \\ U_2, & \text{if } \hat{\lambda}_2(t) < B_2(t). \end{cases} \tag{38}$$

Lemma A.6′ *Assume that for $i = 1, 2$, the instantaneous demand in objective function (9) is replaced with $\hat{D}_i[X_1(t), X_2(t)]$, as defined by (12)–(13). Let $\beta(t)$ be differentiable for $t \in [0, T]$. Let $u_1(\cdot)$, $u_2(\cdot)$, $X_1(\cdot)$, $X_2(\cdot)$ and $\lambda_1(\cdot)$, $\lambda_2(\cdot)$ be control variables, state variables and co-state variables that satisfy the characterizing conditions of Lemma A.5′; then for every nonempty interval $I \subseteq [0, T]$,*

(i) *if $-U_1 < u_1(t) < U_1$ for $t \in I$, then $\hat{\lambda}_1(t) = B_1(t)$, implying that*

$$-X_2(t) + 5X_1(t) = \frac{2\beta(t)}{\alpha} + \frac{4\dot{B}_1(t)}{\alpha} \quad \text{and}$$
$$5u_1(t) + u_2(t) = \frac{2\dot{\beta}(t)}{\alpha} + \frac{4\ddot{B}_1(t)}{\alpha} \quad \text{for } t \in I,$$

(ii) *if $-U_2 < u_2(t) < U_2$ fort $\in I$, then $\hat{\lambda}_2 = B_2(t)$, implying that*

$$-X_1(t) + 5X_2(t) = 4 - \frac{2\beta(t)}{\alpha} + \frac{4\dot{B}_2(t)}{\alpha} \quad \text{and}$$
$$u_1(t) + 5u_2(t) = \frac{2\dot{\beta}(t)}{\alpha} - \frac{4\ddot{B}_2(t)}{\alpha} \quad \text{for } t \in I,$$

(iii) *if $-U_1 < u_1(t) < U_1$ and $-U_2 < u_2(t) < U_2$ for $t \in I$, then $\hat{\lambda}_1(t) = B_1(t)$ and $\hat{\lambda}_2(t) = B_2(t)$, implying that*

$$X_1(t) = \frac{1}{6} + \frac{\beta(t)}{3\alpha} + \frac{5\dot{B}_1(t)}{6\alpha} + \frac{\dot{B}_2(t)}{6\alpha},$$
$$X_2(t) = \frac{5}{6} - \frac{\beta(t)}{3\alpha} + \frac{\dot{B}_1(t)}{6\alpha} + \frac{5\dot{B}_2(t)}{6\alpha};$$
$$u_1(t) = \frac{\dot{\beta}(t)}{3\alpha} + \frac{5\ddot{B}_1(t)}{6\alpha} + \frac{\ddot{B}_2(t)}{6\alpha},$$
$$u_2(t) = \frac{\dot{\beta}(t)}{3\alpha} - \frac{\ddot{B}_1(t)}{6\alpha} - \frac{5\ddot{B}_2(t)}{6\alpha} \quad \text{for } t \in I.$$

As in the proof of Theorem 3.2, there is a control that satisfies the conditions of Lemma A.5′. It has the same structure as that of the original problem, with the co-state variables shifted by $B_i(t)$; so

$$\begin{cases} \hat{\lambda}_1(t) > B_1(t), & 0 \le t < t_1, \\ \hat{\lambda}_1(t) = B_1(t), & t_1 \le t < t_{-1}, \\ \hat{\lambda}_1(t) < B_1(t), & t_{-1} \le t < T, \end{cases} \tag{39}$$

$$\begin{cases} \hat{\lambda}_2(t) < B_2(t), & 0 \le t < t_2, \\ \hat{\lambda}_2(t) = B_2(t), & t_2 \le t < t_{-2}, \\ \hat{\lambda}_2(t) > B_2(t), & t_{-2} \le t < T. \end{cases} \tag{40}$$

Proof of Theorem 4.1 The required conditions for the validity of Theorem 3.2 are invariant of the changes made in defining the modified model (with (12)–(13) replacing the original instantaneous demands); hence Theorem 3.2 establishes uniqueness of the Nash equilibrium for the modified model. Thus, to prove the conclusions of Theorem 4.1, we only need to verify that the suggested penalty/reward functions fully coordinate the agents. We start off from time intervals $t_1 \le t < t_2$ and $t_{-2} \le t < t_{-1}$. According to (39), $\hat{\lambda}_1(t) = B_1(t)$ over a larger interval, $t_1 \le t < t_{-1}$. Recalling the system-wide optimal solution $u_2^*(t) = U_2$ and $U_2 = -3u_1^*(t)$ at interval $t_1 \le t < t_2$, $t_1 = \frac{1}{3U_1+U_2}$ and employing condition (i) of Lemma A.6′ for the interval $t_1 \le t < t_2$ results in

$$-X_2(t) + 5X_1(t) = \frac{2\beta(t)}{\alpha} + \frac{4\dot{B}_1(t)}{\alpha} \quad \text{and} \quad U_2 + 5u_1(t) = \frac{2\dot{\beta}(t)}{\alpha} + \frac{4\ddot{B}_1(t)}{\alpha}.$$

Then, to verify that $t_1 = \frac{1}{3U_1+U_2}$ under the Nash equilibrium, we need

$$-X_2(t_1) + 5X_1(t_1) = \frac{2\beta(t_1)}{\alpha} + \frac{4\dot{B}_1(t_1)}{\alpha}$$

to hold for

$$\dot{B}_1(t_1) = \frac{X_2^*(t_1) - X_1^*(t_1)}{4}\alpha - \frac{\beta(t_1)}{2}.$$

That is,

$$-X_2 * (t_1) + 5X_1 * (t_1) = X_2 * (t_1) - X_1 * (t_1),$$

Therefore, $3X_1^*(t_1) = X_2^*(t_1)$ and substituting $X_1^*(t_1) = U_1 t_1$, $X_2^*(t_1) = 1 - U_2 t_1$ we obtain $3U_1 t_1 = 1 - U_2 t_1$. This equality always holds as $t_1 = \frac{1}{3U_1+U_2}$, that is, t_1 is system-wide optimal.

Next we verify that the speeds are system-wide optimal under the Nash equilibrium. Substituting into condition (i) of Lemma A.6′, the system-wide optimal solution, $U_2 = -3u_1^*(t)$ for $t_1 \le t < t_2$, we find that

$$U_2 - 5\frac{U_2}{3} = -\frac{2}{3}U_2 = \frac{2\dot{\beta}(t)}{\alpha} + \frac{4\ddot{B}_1(t)}{\alpha}.$$

On the other hand, differentiating the suggested penalty function we have

$$\ddot{B}_1(t) = \frac{-U_2 - (-\frac{U_2}{3})}{4}\alpha - \frac{\dot{\beta}(t)}{2}.$$

Substituting this result into $-\frac{2}{3}U_2 = \frac{2\dot{\beta}(t)}{\alpha} + \frac{4\ddot{B}_1(t)}{\alpha}$, we obtain $-\frac{2}{3}U_2 = -\frac{2}{3}U_2$.

Thus, the system-wide optimal solution at $[t_1, t_2]$ is Nash equilibrium. Similarly one can verify this result for the symmetric interval $[t_{-2}, t_{-1}]$.

Note, that at time intervals $0 \le t < t_1$ and $t_{-1} < t \le T$ we can select any function $B_1(t)$ so that $\hat{\lambda}_1(t) > B_1(t)$ for $0 \le t < t_1$ and $\hat{\lambda}_1(t) < B_1(t)$ for $t_{-1} < t \le T$. Let $B_1(t_1) = A_1$. Then from $\hat{\lambda}_1(t_1) = B_1(t_1)$ and (33) we have

$$\hat{\lambda}_1(t) = A_1 - \int_t^{t_1} \left[-\frac{\beta(\tau)}{2} - \frac{1}{4}\alpha X_2^*(\tau) + \frac{5}{4}\alpha X_1^*(\tau) \right] d\tau \quad \text{for } 0 \le t \le t_1,$$

which along with the requirement $\hat{\lambda}_1(t) > B_1(t)$ readily results in the corresponding condition of Theorem 4.1 for $0 \le t < t_1$. Similarly, since

$$B_1(t_{-1}) = A_1 + \int_{t_1}^{t_{-1}} \left[\frac{X_2^*(\tau) - X_1^*(\tau)}{4}\alpha - \frac{\beta(\tau)}{2} \right] d\tau,$$

we find the symmetric condition for $t_{-1} < t \le T$, which ensures $\hat{\lambda}_1(t) < B_1(t)$.

Next, we verify that $B_1(t)$ and $B_2(t)$ coordinate the non-cooperative agents for period $t_2 \le t \le t_{-2}$. According to (39)–(40), the reward/penalty functions are identical to the corresponding shadow prices over this time interval. Both $\hat{\lambda}_2(t) = B_2(t)$ and $\hat{\lambda}_1(t) = B_1(t)$ for $t_2 \le t \le t_{-2}$ are ensured in condition (iii) of Lemma A.6′.

Specifically, we need,

$$-X_2(t_2) + 5X_1(t_2) = \frac{2\beta(t_2)}{\alpha} + \frac{4\dot{B}_1(t_2)}{\alpha} \quad \text{and}$$
$$-X_1(t_2) + 5X_2(t_2) = 4 - \frac{2\beta(t_2)}{\alpha} + \frac{4\dot{B}_2(t_2)}{\alpha}.$$

Substituting the suggested penalty function at t_2, it is easy to observe that the above equations hold for $X_1^*(t_2) = 1/4$, $X_2^*(t_2) = 3/4$.

We thus satisfied conditions (iii) of Lemma A.6′ at $t_2 = \frac{1}{4U_2}$ with respect to the coordinates. To maintain these conditions along $t_2 \le t \le t_{-2}$, we also need to meet conditions (iii) with respect to the speeds,

$$u_2(t) + 5u_1(t) = \frac{2\dot{\beta}(t)}{\alpha} + \frac{4\ddot{B}_1(t)}{\alpha} \quad \text{and} \quad u_1(t) + 5u_2(t) = \frac{2\dot{\beta}(t)}{\alpha} - \frac{4\ddot{B}_2(t)}{\alpha}.$$

However, the system-wide optimal speeds at $t_2 \le t \le t_{-2}$ are zero. Therefore, we have

$$\frac{2\dot{\beta}(t)}{\alpha} + \frac{4\ddot{B}_1(t)}{\alpha} = 0 \quad \text{and} \quad \frac{2\dot{\beta}(t)}{\alpha} - \frac{4\ddot{B}_2(t)}{\alpha} = 0.$$

Consequently, we observe that these equalities always hold if we employ the fact that

$$\ddot{B}_2(t) = \frac{\dot{\beta}(t)}{2} - \alpha\frac{\dot{X}_2^*(t) - \dot{X}_1^*(t)}{4} = \frac{\dot{\beta}(t)}{2},$$
$$\ddot{B}_1(t) = \frac{\dot{X}_2^*(t) - \dot{X}_1^*(t)}{4}\alpha - \frac{\dot{\beta}(t)}{2} = -\frac{\dot{\beta}(t)}{2}.$$

Finally, at time intervals $0 \le t < t_2$ and $t_{-2} \le t \le T$ we can select any function $B_2(t_2)$ so that $\hat{\lambda}_2(t) < B_2(t)$ for $0 \le t < t_2$ and $\hat{\lambda}_2(t) > B_2(t)$ for $t_{-2} < t \le T$. Therefore, the last condition of this theorem is proven by setting $B_2(t_2) = A_2$, which, with respect to (20), leads to

$$\hat{\lambda}_2(t) = A_2 - \int_t^{t_2} \left[\frac{\beta(\tau)}{2} - \alpha - \frac{1}{4}\alpha X_1^*(\tau) + \frac{5}{4}\alpha X_2^*(\tau) \right] d\tau \quad \text{for } 0 \le t < t_2. \qquad \square$$

Proof of Corollary 4.1 Assume that $B_1(t)$ and $B_2(t)$ satisfy (14)–(15) for $0 \le t \le T$. Comparing $\dot{B}_1(t)$ and $\dot{\hat{\lambda}}_1(t)$, we observe that

$$\begin{aligned} \dot{B}_1(t) &= \frac{X_2^*(t) - X_1^*(t)}{4}\alpha - \frac{\beta(t)}{2} > \dot{\hat{\lambda}}_1(t) \\ &= -\frac{\beta(t)}{2} + \frac{\alpha X_2^*(t)}{4} - \frac{5\alpha X_1^*(t)}{4}, \quad 0 \le t < t_1 \end{aligned} \tag{41}$$

holds if $X_2^*(t) > 3X_1^*(t)$ for $0 \le t < t_1$. Recalling the system-wide optimal motion $X_2^*(t) = 1 - U_2 t$ and $X_1^*(t) = U_1 t$, inequality (34) takes the following form, $t < \frac{1}{3U_1+U_2}$, $0 \le t < t_1$, which always holds as $t_1 = \frac{1}{3U_1+U_2}$. Consequently, taking into account that $\hat{\lambda}_1(t_1)$ is chosen so that $B_1(t_1) = \hat{\lambda}_1(t_1)$, we conclude from $\dot{B}_1(t) > \dot{\hat{\lambda}}_1(t)$, for $0 \le t < t_1$, that $B_1(t) < \hat{\lambda}_1(t)$ for this time interval. As a result, the condition of Theorem 4.1

$$B_1(t) < A_1 - \int_t^{t_1} \left[-\frac{\beta(\tau)}{2} - \frac{1}{4}\alpha X_2^*(\tau) + \frac{5}{4}\alpha X_1^*(\tau) \right] d\tau$$

holds by choosing a constant $A_1 = B_1(t_1)$. Similarly, one can verify the corresponding conditions for $B_1(t)$, $t_{-1} < t \le T$, as well as for function $B_2(t)$ for $0 \le t < t_2$ and $t_{-2} < t \le T$. $\square$

References

1. Cachon, G.P.: Competitive supply chain inventory management. In: Tayur, S., Ganesham, R., Magazine, M. (eds.) Quantitative Models for Supply Chain Management. Kluwer International, Dordrecht (1999)
2. Cachon, G.P.: Supply chain coordination with contracts. In: Graves, S., de Kok, T. (eds.) The Handbook of Operations Research and Management Science: Supply Chain Management. Kluwer Academic, Dordrecht (2003)
3. Taylor, T.A.: Supply chain coordination under channel rebates with sales effort effects. Manag. Sci. **48**(8), 992–1007 (2002)
4. Naor, P.: The regulation of queue size by levying tolls. Econometrica **37**, 15–24 (1969)
5. Dolan, R.J.: Incentives and mechanisms for priority queuing problems. Bell J. Econ. **9**, 421–436 (1978)
6. Mendelson, H., Whang, S.: Optimal incentive-compatible priority pricing for the M/M/1 queue. Oper. Res. **38**(5), 870–883 (1990)
7. Avi-Itzhak, B., Golany, B., Rothblum, U.G.: Strategic equilibrium vs. global optimum for a pair of competing servers. J. Appl. Probab. **43**, 1165–1172 (2006)
8. Beckmann, M., McGuire, C.B., Winsten, C.B.: Studies in the Economics of Transportation. Yale University Press, New Haven (1956)
9. Cole, R., Dodis, Y., Roughgarden, T.: Pricing network edges for heterogeneous selfish users. In: Proceedings of the 35th Annual ACM Symposium on Theory Computing (STOC), pp. 521–530 (2003)

10. Golany, B., Rothblum, U.G.: Inducing coordination in supply chains through linear reward schemes. Nav. Res. Logist. **53**(1), 1–15 (2006)
11. Ferentinos, K.P., Arvanitis, K.G., Sigrimis, N.: Heuristic optimisation methods for motion planning of autonomous agricultural vehicles. J. Glob. Optim. **23**(2), 155–170 (2002)
12. Shima, T., Rasmussen, S.J., Sparks, A.G., Passino, K.M.: Multiple task assignments for cooperating uninhabited aerial vehicles using genetic algorithms. Comput. Oper. Res. **33**, 3252–3269 (2006)
13. Hotelling, H.: Stability in competition. Econ. J. **39**(153), 41–57 (1929)
14. Munson, C.L., Hu, J., Rosenblatt, M.J.: Teaching the costs of uncoordinated supply chains. Interfaces **33**(3), 24–39 (2003)
15. Başar, T., Olsder, G.L.: Dynamic Noncooperative Game Theory. Academic Press, London (1982)
16. Feichtinger, G., Jørgensen, S.: Differential game models in management science. Eur. J. Oper. Res. **14**, 137–155 (1983)
17. Kogan, K., Tapiero, C.S.: Supply Chain Games: Operations Management and Risk Valuation. Springer, Boston (2007)
18. He, X., Prasad, A., Sethi, S.P., Gutierrez, G.J.: A survey of Stackelberg differential game models in supply and marketing channels. J. Syst. Sci. Syst. Eng. **16**(4), 385–413 (2007)
19. Bess, R.: New Zealand seafood firm competitiveness in export markets: the role of the quota management system and aquaculture legislation. Mar. Policy **30**(4), 367–378 (2006)
20. Carafano, J.J., Walsh, B.W., Muhlhausen, D.B., Keith, L.P., Gentilli, D.D.: Better, faster and cheaper border security, a policy paper. The Heritage Foundation (2006)
21. Wheatley, E., Doty, R.: Outsourcing the Dirty Work: The Use of Private Security Firms at the Mexico/US Border. Paper presented at the annual meeting of the ISA's 49th Annual Convention, Bridging Multiple Divides, San Francisco (2008)
22. Cooper, L.: Solutions of generalized locational equilibrium models. J. Reg. Sci. **7**(1), 1–18 (1967)
23. Tapiero, C.S.: Transportation-location-allocation problems over time. J. Reg. Sci. **14**(3), 377–384 (1971)
24. Tapiero, C.S., Soliman, M.A.: Multi-commodity transportation problems over time. Networks **2**, 311–327 (1972)
25. Cavalier, T.M., Sherali, H.D.: Sequential location-allocation problems on chains and trees with probabilistic link demands. Math. Program. **32**(3), 249–277 (1985)
26. Rothblum, U.G.: Optimality vs. equilibrium: inducing coordination by linear rewards and penalties. Unpublished manuscript (2005)
27. Sethi, S.P., Thompson, G.L.: Optimal Control Theory: Applications to Management Science and Economics, 2nd edn. Kluwer Academic, Dordrecht (2000)
28. Sydsaeter, K.: Topics in Mathematical Analysis for Economists. Academic Press, Dordrecht (1981)