University of Groningen

Dynamic Coordination Games with Activation Costs

Ramirez, Stefanny; Bauso, Dario

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

[Link to publication in University of Groningen/UMCG research database](#)

Dynamic Coordination Games with Activation Costs

Stefanny Ramirez¹ [ID] · **Dario Bauso²,³**

Abstract

Motivated by inventory control problems with set-up costs, we consider a coordination game where each player's dynamics is an inventory model characterized by a controlled input and an uncontrolled output. An activation cost is shared among active players, namely players who control their dynamics at a given time. At each time, each player decides to be active or not depending on its inventory level. The main contribution of this paper is to show that strategies at a Nash equilibrium have a threshold structure on the number of active players. Furthermore, we provide an explicit expression for the lower and upper threshold is given both in the deterministic case, namely when the exogenous signal is known, and in the single-stage game. The relevance of the above results is discussed in the context of inventory control where Nash equilibrium reordering strategies imply that a single retailer reorders only if jointly with a number of other retailers and will reorder to restore a pre-assigned inventory level.

Keywords Dynamic games · Inventory control

1 Introduction

This paper studies a discrete-state discrete-time dynamic game where players have to coordinate actions within a finite horizon window [2,3]. Each player's dynamics is an inventory model characterized by a controlled input and an uncontrolled output. The output flow is an uncontrolled exogenous signal. The input flow is controlled by the player and is subject to an activation cost. The state of the player is the accumulated discrepancy between input flow and output flow. The activation cost is shared among active players, namely those players who control their dynamics at a given time. The possibility of sharing the activation cost

✉ Stefanny Ramirez
    s.g.ramirez.juarez@rug.nl

    Dario Bauso
    d.bauso@rug.nl

¹ ENTEG, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands

² Jan C. Willems Center for Systems and Control, ENTEG, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands

³ Dipartimento di Ingegneria, Università di Palermo, 90128 Palermo, Italy

determines the need for coordination of control strategies on the part of the players. We study the cases under deterministic and stochastic disturbances. All results can be extended to the vector case by using the robust decomposition approach in [4, Section 3]. Applications arise in coordinated replenishment [8], and opportunistic maintenance [7].

1.1 Contribution

This study contributes in different ways to advance the theory on dynamic coordination games with activation costs for the control. An example of two-threshold strategy is the (s, S) strategy used in inventory control, see [6] and [5, Chapter 4]. We recall that (s, S) strategies are strategies where replenishments occur anytime the inventory level goes below a lower threshold s. Replenishments bring back the inventory level up to a higher threshold S. In particular, we highlight the following results.

- Strategies at a Nash equilibrium have a threshold structure. We obtain this result in two steps. First, we prove that Nash equilibria are associated with (s, S) strategies via $\mathscr{K}$-convex analysis. Second, we view the (s, S) strategies as threshold strategies on the number of active retailers.
- Lower and upper thresholds have an explicit expression in the deterministic case, namely when the exogenous signal is known, or in single-stage games.
- We corroborate our results with a numerical analysis of a stylized inventory model.

This paper is organized as follows. In Sect. 2, we introduce the dynamic inventory game. In Sect. 4, we first show that all Nash equilibrium strategies have a two-threshold structure with a reorder level and an order-up-to level. We then provide a dual interpretation of such strategies as threshold strategies on the number of active players. In Sect. 5, we specialize our results to the case of single stage coordination game. In Sect. 6, we provide numerical analysis. Finally, in Sect. 7, we draw conclusions and discuss future works.

2 Dynamic Inventory Coordination Game

Consider a set of n retailers $\Gamma = \{1, \ldots, n\}$. At stage $t = 0, \ldots, N-1$, the ith retailer holds inventory $x_i^t \in \mathbb{Z}$, faces a stochastic demand $\omega_i^t \in \mathbb{Z}_+$ and orders a quantity $u_i^t \in U_i^t \subseteq \mathbb{Z}_+$, where U_i^t denotes the set of admissible decisions, $\mathbb{Z}$ the set of integers, and $\mathbb{Z}_+$ the set of nonnegative integers. Thus, for all retailers $i \in \Gamma$, inventory x_i^t, which we refer to as the *state* of retailer i, evolves according to a linear finite-state, discrete-time model of the form:

$$x_i^{t+1} = x_i^t + u_i^t - \omega_i^t. \tag{1}$$

Here, we assume that there are no delays between orders and deliveries. For all retailers, we also suppose that the inventory at hand plus inventory ordered may not exceed the storage capacity denoted by C_{store}. Hence, we have $x_i^t + u_i^t \leq C_{\text{store}}$. We also assume that $C_{\text{store}} \geq x_0^i$ so to exclude an empty set of feasible orders.

Now, for each time t, let us introduce the vector of the retailers' decisions $u^t = [u_i^t]_{i \in \Gamma}$ and the vector of decisions of all retailers other than i $u_{-i}^t = [u_j^t]_{j \in \Gamma, j \neq i} \in U_{-i}^t$, where U_{-i}^t denotes the Cartesian product of all sets U_j^t $j \neq i$. At each stage, the ith retailer has a cost

$$g_i(x_i^t, u_i^t, u_{-i}^t) = \frac{K}{1+\sum_{j=1, j\neq i}^{n} \delta(u_j^t)} \delta(u_i^t) + c u_i^t \tag{2}$$
$$+ p\mathbb{E}\{\max(0, -x_i^{t+1})\} + h\mathbb{E}\{\max(0, x_i^{t+1})\},$$

where $\mathbb{E}\{.\}$ indicates expectation, $K \geq 0$ represents the transportation cost, $c \geq 0$ is the purchase cost per stock unit, $h \geq 0$ is the penalty on holding, $p \geq 0$ the penalty on shortage. The term $\delta(u_i^t)$ is one if the ith retailer replenishes, i.e., is active, and zero otherwise.

We henceforth denote by a^t the *number of active retailers* at stage t, i.e.:

$$a^t := \sum_{j=1}^{n} \delta(u_j^t).$$

Note that the term $\frac{K}{1+\sum_{j=1, j\neq i}^{n} \delta(u_j^t)} \delta(u_i^t)$, which describes the fixed cost paid by retailer i in (2), is equal to $\frac{K}{a^t}$ if retailer i is active and equal to zero otherwise.

After introducing the N stage decision vectors $\mathbf{u_i}^{0\sim N-1} = [u_i^0, \ldots, u_i^{N-1}]$ and $\mathbf{u_{-i}}^{0\sim N-1} = [u_{-i}^0, \ldots, u_{-i}^{N-1}]$, and denoting by $\Phi_i(x_i^N)$ a penalty term on final state, the cost over the horizon from 0 to N is of the form

$$J_i\left(x_i^0, \mathbf{u_i}^{0\sim N-1}, \mathbf{u_{-i}}^{0\sim N-1}\right) = \Phi_i\left(x_i^N\right) + \sum_{t=0}^{N-1} g_i\left(x_i^t, u_i^t, u_{-i}^t\right). \tag{3}$$

A challenging issue in the definition of the stage cost (2) is its dependence on the number of active retailers through the term $\frac{K}{1+\sum_{j=1, j\neq i}^{n} \delta(u_j^t)} \delta(u_i^t)$. This term establishes that the transportation cost K is equally divided by all active retailers. This in turn implies that the cost of one retailer also depends on the decisions of all other retailers. Conditions (1)–(3) describe the dynamics and the costs of our game.

Other concepts we will make use of in the rest of the paper are Nash equilibrium strategies and $\mathcal{K}$-convexity which we briefly recall next.

Definition 1 Decisions $\mathbf{u_i}^\star$ are at a *Nash equilibrium*, if it holds for all $i \in \Gamma$

$$J_i\left(x_i^0, \mathbf{u_i}^\star, \mathbf{u_{-i}}^\star\right) \leq J_i\left(x_i^0, \mathbf{u_i}, \mathbf{u_{-i}}^\star\right) \ \forall \mathbf{u_i} \in U_i^0 \times \cdots \times U_i^{N-1}.$$

For the inventory problem, once at a Nash equilibrium, no retailer benefits from changing its replenishment decisions. The following definition of $\mathcal{K}$-convexity is borrowed from [6].

Definition 2 A function $f : \mathbb{R} \to \mathbb{R}$ is $\mathcal{K}$-*convex*, where $\mathcal{K} \geq 0$, if

$$\mathcal{K} + f(z+y) \geq f(y) + z\left[\frac{f(y) - f(y-b)}{b}\right] \ \forall z \geq 0, \ b > 0, \ y.$$

$\mathcal{K}$-convexity is used in [6] and reiterated in [5] to prove optimality of (s, S) strategies. We will make use of $\mathcal{K}$-convexity to prove the main result of this work.

In the following section, we consider threshold strategies according to which, given an inventory level x_i^t, there exists a threshold $l_i^t \in \{1, 2, \ldots, n\}$, such that retailer i reorders only if the number of active players a^t is greater than or equal to such a threshold. Such strategies are given by

$$\mu_i(x_i^t, a^t) = \begin{cases} \text{reorder} & \text{if } a^t \geq l_i^t, \\ \text{do not reorder} & \text{if } a^t < l_i^t. \end{cases} \tag{4}$$

As main result, we will show that all Nash equilibrium strategies have the threshold structure (4). To emphasize that l_i^t depends on x_i^t, we sometimes write $l_i^t(x_i^t)$.

Note that orders depend on the history of the game as they are function of the state variable x_i^t which in turn depends on the past orders of the retailer as in (1). Orders of a single retailer also depend on her competitors' orders through variable a^t.

An additional concept that is important to explain is the one of subgame perfect equilibrium. We have borrowed and adapted from [9] the following definition:

Definition 3 A subgame perfect equilibrium is a n-tuple of strategies $\mathbf{u}^\star = [\mathbf{u_i}^\star]_{i \in \Gamma}$, such that for every $i \in \Gamma$ and every history x_i^t, we have that:

$$J_i\left(x_i^t, \mathbf{u_i}^\star, \mathbf{u_{-i}}^\star\right) \le J_i\left(x_i^t, \mathbf{u_i}, \mathbf{u_{-i}}^\star\right) \ \forall \ \mathbf{u_i} \in U_i^0 \times \cdots \times U_i^{N-1}.$$

Therefore, note that, if strategy (4) returns a Nash equilibrium, then such equilibrium is also subgame perfect in the sense that the strategy returns the optimal order depending on the current state x_i^t and irrespective of the fact that past orders might not be optimal.

To simplify the proofs and the graphs plotted in the following figures, in the rest of the paper, we assume that the penalty term on the final state $\Phi(x_i^n)$ is null. However, the results that we prove still hold if $\Phi(x_i^n)$ is a generic convex function with a minimum in $x_i^n = 0$.

3 On the Generality of the Model

Consider an n-dimensional inventory model characterized by discrete states $x^t \in \mathbb{Z}^n$, integer controls $u^t \in \mathbb{Z}_+^n$, and binary controls $y^t \in \{0, 1\}^n$, and discrete stochastic disturbances $w^t \in \mathbb{Z}_+^n$, where $t = 0, 1, \ldots$ is the time index. The evolution of the state is described by a linear discrete-time (difference) equation in the general form (5) below, where A and E are matrices of compatible dimensions and $x(0) = \xi_0 \ge 0$ is a given initial state. Integer and binary controls are linked through the general *capacity constraints* (6), where the (scalar) parameter c is an upper bound on control, with the inequalities in (5) and (6) to be interpreted component-wise.

$$x^{t+1} = Ax^t + Ew^t + u^t, \tag{5}$$

$$0 \le u^t \le cy^t, \quad y^t \in \{0, 1\}^n. \tag{6}$$

The above dynamics are characterized by two discrete valued control variables per each state. Starting from nonnegative initial states, we wish to control the state to remain confined to the positive orthant, which may describe a safety region in engineering applications or reflect the desire to prevent shortfalls in inventory applications.

A common situation is where the disturbance seeks to push the state out of the desired region. Its value is given at the beginning and fixed that way. Each column of matrix E establishes how each disturbance component influences the evolution of the state vector. Then, it is reasonable to assume $Ew(k) < 0$, where the inequality is to be interpreted component-wise.

With regard to (5), we can isolate the dependence of one component state on the other ones and rewrite (5) in a way that establishes similarity with standard lot sizing models [10]:

$$x^{t+1} = x^t + Bx^t + Ew^t + u^t. \tag{7}$$

Equation (7) is a straightforward representation of (5) where

$$B := A - I =: \{b_{ij}\}, \ b_{ij} = a_{ij} - \delta_{ij},$$

$$\delta_{ij} := \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{otherwise.} \end{cases}$$

To preserve the nature of the problem, which has stabilizing control actions playing against unstabilizing disturbances, we assume that the influence of other states on state i is relatively

"weak." In other words, we assume that the influence of Bx^t is small if compared with the unstabilizing effects of disturbances captured by the term Ew^t. This is captured by assuming that the sum $\Delta x^t + Ew^t$ has same (negative) sign of Ew^t, namely

$$Bx^t + Ew^t < 0,$$

where inequality is again component-wise and it holds almost everywhere. Essentially, the states' mutual dependence expressed by Bx^t only emphasizes or reduces "weakly" the destabilizing effects of the disturbances. In the following, we present a robust decomposition approach that translates dynamics (7) into n scalar dynamics in "lot sizing" form [10].

With the term "robust decomposition" we mean a transformation through which dynamics (7) are replaced by n independent uncertain lot sizing models of the form (8) where x_i^t is the inventory, d_i^t the demand, u_i^t the reordered quantity and $\mathscr{D}_i^t \subset \mathbb{R}$ denotes the uncertainty set:

$$x_i^{t+1} = x_i^t - d_i^t + u_i^t, \quad d_i^t \in \mathscr{D}_i^t. \tag{8}$$

Recall that in (7) the disturbance is given at the beginning and fixed that way. We use those values of the disturbance to determine set $\mathscr{D}_i^t$ in (8), as explained in the following. Replacing (7) with (8) is possible once we relate the demand d_i^t to the current values of all other state components and disturbances as expressed below:

$$\begin{aligned} d_i^t &= -\left[\sum_{j=1}^n b_{ij} x_j^t + \sum_{j=1}^n E_{ij} w_j^t\right] \\ &= -\left[\langle B_{i\bullet} x^t\rangle + \langle E_{i\bullet} w^t\rangle\right], \end{aligned} \tag{9}$$

where we denote by $B_{i\bullet}$ the ith row of the matrix B, with the same convention applying to $E_{i\bullet}$.

In other words, we assume that the influence that all other states have on state i enters into Eq. (8) through demand d_i^t defined in (9).

Following the decomposition, each lot sizing model is controlled by an agent i (whose state is x_i) who plays against a virtual opponent which selects a worst-case demand, which can be viewed as a *two-player game*.

Our next step is to make the n dynamics in the form (8) mutually independent. Toward that end, we introduce X^t as the set of x^t and observe that this set is bounded for bounded d_i^t. The set X^t can be defined in two steps. First, we assume that the states never leave a given region, then we compute the worst-case vector x^t in the region, namely the vector x^t that, once substituted in (9), has the effect of pushing the ith state out of the safe region. Then, we check whether the trajectory still lies within the region.

Boundedness of X^t means that there exists a scalar $\phi > 0$ such that $\|x\|_\infty \le \phi$ for all $x \in X^t$. In view of this, it is possible to decompose the system by replacing the current demand d_i^t by the maximal or minimal demand as computed below:

$$\overline{d}_i^t = \max_{\xi \in X^t}\left\{-\langle B_{i\bullet}\xi\rangle - \langle E_{i\bullet}w^t\rangle\right\} = \sum_j [B_{ij}]_- \phi - \langle E_{i\bullet}w^t\rangle$$

$$\underline{d}_i^t = \min_{\xi \in X^t}\left\{-\langle B_{i\bullet}\xi\rangle - \langle E_{i\bullet}w^t\rangle\right\} = \sum_j [B_{ij}]_+ \phi - \langle E_{i\bullet}w^t\rangle,$$

where $[B_{ij}]_+$ denotes the positive part of B_{ij}, i.e., $\max\{B_{ij}, 0\}$ and $[B_{ij}]_-$ the negative part.

From the above preamble, we derive the uncertainty set as

$$\mathscr{D}_i^t = \{\eta \in \mathbb{R} : \underline{d}_i^t \le \eta \le \overline{d}_i^t\}.$$

Likewise, (10) describes the demand that would push the state out of the positive orthant in the longest time.

4 Nash Equilibrium Strategies

In this section, we show that all Nash equilibrium strategies are threshold strategies of type (4): *retailer i reorders only if the number of active retailers is greater than or equal to a given threshold.* For the general model explained in Sect. 3, proving that strategies at a Nash equilibrium have a threshold structure is not straight forward, for that reason in this section the results are given for a single retailer i. To show this, in the next subsection we prove the optimality of the (s, S)-like strategies via $\mathcal{K}$-convex analysis (see the definition in [5], chapter 4). We recall from [5] that (s, S) strategies are strategies where replenishments occur anytime the inventory level goes below a lower threshold s. Replenishments bring back the inventory level up to a higher threshold S [6]. This is formally stated below where $\mu(.)$ is the strategy, x the inventory, and s and S lower and upper thresholds, respectively:

$$\mu(x) = \begin{cases} S - x & \text{if } x < s, \\ 0 & \text{if } x \geq s. \end{cases} \tag{10}$$

We refer to (s, S)-like strategies as (s, S) strategies whose thresholds depend on the players and on time, i.e., we will have $s := s_i^t$ and $S := S_i^t$ for fixed i and t.

In Theorem 1, we prove the optimality of (s, S)-like strategies. Before doing this, we need some preliminary analysis which is inspired by [5, Chapter 4].

Let $K^t(u_{-i}^t) = \frac{K}{1 + \sum_{i \in \Gamma, j \neq i} \delta(u_j^t)}$ be the transportation cost charged to each retailer i that replenishes at stage t. Fix decisions $u_{-i}^0, \ldots, u_{-i}^{N-1}$ of all retailers other than i over the horizon and denote such decisions $\bar{u}_{-i}^0, \ldots, \bar{u}_{-i}^{N-1}$. Similarly, denote the resulting transportation costs by $K^0, \ldots, K^{N-1}$. Note that K^t is a function of u_{-i}^t but for ease of notation sometimes we omit the dependence. Then, let us rewrite the stage cost (2) for retailer i as

$$\begin{aligned} g_i(x_i^t, u_i^t, \bar{u}_{-i}^t) &= K^t \delta(u_i^t) + c u_i^t \\ &\quad + p\mathbb{E}\{\max(0, -x_i^{t+1})\} + h\mathbb{E}\{\max(0, x_i^{t+1})\}. \end{aligned}$$

Now, we can write the cost-to-go from stage t to the final stage recursively using dynamic programming and the Bellman equation. Let us use the superscript t to indicate the iteration. Then, we have

$$\begin{aligned} v_i^t(x_i^t, \bar{u}_{-i}^{t \sim N-1}) &= \min_{u_i^t \in U}[g_i(x_i^t, u_i^t, \bar{u}_{-i}^t) \\ &\quad + \mathbb{E}\{v_i^{t+1}(x_i^{t+1}, \bar{u}_{-i}^{t+1 \sim N-1})\}], \quad t = 0, \ldots N-1, \end{aligned} \tag{11}$$

$$J_i^N(x_i^N) = 0, \tag{12}$$

where $J_i^0(x_i^0, \bar{u}_{-i}^0)$ is equal to the cost $J_i(x_i^0, u_{-i}^0)$ introduced in (3). Being $y_i^t = x_i^t + u_i^t$, the instantaneous inventory position, i.e., the inventory level just after the order has been issued, let us define the new function

$$\begin{aligned} G_i^t(y_i^t, \bar{u}_{-i}^{t+1 \sim N-1}) &= c y_i^t + p\mathbb{E}\{\max(0, -(y_i^t - \omega_i^t))\} \\ &\quad + h\mathbb{E}\{\max(0, y_i^t - \omega_i^t)\} + \mathbb{E}\{v_i^{t+1}(x_i^{t+1}, \bar{u}_{-i}^{t+1 \sim N-1})\}, \end{aligned}$$

and rewrite the Bellman Eq. (11) as follows

$$\begin{aligned} v_i^t(x_i^t, \bar{u}_{-i}^{t \sim N-1}) &= -c_i x_i^t + \min_{y_i^t \geq x_i^t}[K^t(u_{-i}^t) \\ &\quad + G_i^t(y_i^t, \bar{u}_{-i}^{t+1 \sim N-1}), \; G_i^t(x_i^t, \bar{u}_{-i}^{t+1 \sim N-1})]. \end{aligned} \tag{13}$$

Note that if we can show that v_i^{t+1} is $\mathcal{K}$-convex with $\mathcal{K} = K^t$ then G_i^t is also $\mathcal{K}$-convex for $\mathcal{K} = K^t$ and the Bellman Eq. (13) has a unique minimizer. Indeed, it has been proved in [5], chapter 4.2, that $\mathcal{K}$-convexity of $G_i^t(y_i^t, \bar{u}_{-i}^{t+1})$ implies $\mathcal{K}$-convexity of $v_i^t(x_i^t, \bar{u}_{-i}^t)$. This represents a *sufficient optimality* condition for the (s, S)-like strategies with thresholds depending on time t, that is, $s := s_i^t$ and $S := S_i^t$, where s_i^t and S_i^t satisfy:

$$S_i^t = \arg\min_\gamma G_i^t(\gamma, \bar{u}_{-i}^{t+1 \sim N-1}),$$

$$G_i^t(s_i^t, \bar{u}_{-i}^{t+1 \sim N-1}) = G_i^t(S_i^t, \bar{u}_{-i}^{t+1 \sim N-1}) + K^t(u_{-i}^t).$$

The meaning of s_i^t and S_i^t is exactly the same as in the (s, S) strategies (cfg. [5]), that is, s_i^t represents the minimum threshold on inventory level below which retailers replenish to restore the inventory up to level S_i^t. Now, let us call $\underline{s}_i^t$, the threshold which corresponds to the assumption that the ith retailer is charged the whole transportation cost, i.e.,

$$G_i^t\left(\underline{s}_i^t, \bar{u}_{-i}^{t+1 \sim N-1}\right) = G_i^t\left(S_i^t, \bar{u}_{-i}^{t+1 \sim N-1}\right) + K.$$

In the above condition, we have set $K^t = K$.

Analogously, let us denote by $\bar{s}_i^t$ the threshold computed as if all retailers would share equally the transportation cost, i.e.,

$$G_i^t(\bar{s}_i^t, \bar{u}_{-i}^{t+1 \sim N-1}) = G_i^t(S_i^t, \bar{u}_{-i}^{t+1 \sim N-1}) + \frac{K}{n}.$$

In essence, in the condition above, each retailer is charged a transportation cost $K^t = \frac{K}{n}$, namely one nth of the full cost K. Hence, we have $\underline{s}_i \leq s_i^t \leq \bar{s}_i$.

The following theorem establishes the optimality of (s, S)-like strategies, where each pair of thresholds is valid on different intervals of inventory levels.

Theorem 1 *Let K^t be nondecreasing. Solutions of the Bellman Eq. (11) are at most N different (s, S)-like strategies (s_i^t, S_i^t), $t = 0, \dots, N-1$, where $S_i^t \in \{\sum_{\hat{t}=t}^{t+j} \omega_i(\hat{t}), j = t, \dots, N-1-t\}$ and threshold s_i^t verifies $G_i^t(s_i^t, \bar{u}_{-i}^{t+1}) = G_i^t(S_i^t, \bar{u}_{-i}^{t+1}) + K^t$.*

Proof The proof is by induction. Assume $J_i^N(x_i^N) = 0$, and consider the convex function

$$G_i^{N-1}\left(y_i^{N-1}, \bar{u}_{-i}^N\right) = cy_i^{N-1} + p\mathbb{E}\{\max\left(0, -\left(y_i^{N-1} - \omega_i^{N-1}\right)\right)\}$$
$$+ h\mathbb{E}\{\max\left(0, y_i^{N-1} - \omega_i^{N-1}\right)\}. \quad (14)$$

Then, we say that $G_i^{N-1}(\cdot)$ is convex and hence, it is also $\mathcal{K}$-convex where $\mathcal{K} = K^{N-1}$ as shown in Fig. 1. Here, we also use the notation

$$H_i^{N-1}(x_i) := p\mathbb{E}\{\max(0, -(x_i^{N-1} - \omega_i^{N-1})\}$$
$$+ h\mathbb{E}\{\max(0, x_i^{N-1} - \omega_i^{N-1})\}.$$

The above reasoning on $\mathcal{K}$-convexity implies that the piecewise linear function

$$v_i^{N-1}\left(x_i^{N-1}, \bar{u}_{-i}^{N-1}\right) = -c_i x_i^{N-1} + \min_{y_i^{N-1} \geq x_i^{N-1}}\left[K^{N-1}\left(u_{-i}^{N-1}\right)\right.$$
$$\left. + G_i^{N-1}\left(y_i^{N-1}, \bar{u}_{-i}^N\right), G_i^{N-1}\left(x_i^{N-1}, \bar{u}_{-i}^N\right)\right] \quad (15)$$

is K^{N-1}-convex, with a global minimum at $S_i^{N-1} := \arg\min_\gamma G_i^{N-1}(\gamma, \bar{u}_{-i}^N)$ (in the deterministic case if the cost of purchase is relatively small then $S_i^{N-1} = \omega_i^{N-1}$) (see, e.g., Fig. 1).

To obtain S_i^{N-1}, let a probability distribution function $\phi^{N-1} : \mathbb{Z}_+ \rightarrow [0, 1]$ be given, namely $\omega \mapsto \phi^{N-1}(\omega)$ where $\phi^{N-1}(\omega)$ is the probability that $\omega_i^{N-1} = \omega$ for all $\omega \in \mathbb{Z}_+$. Then, the cost of reordering is given by

$$K^{N-1}(u_{-i}^{N-1}) - c_i x_i^{N-1} + G_i^{N-1}(\gamma, \bar{u}_{-i}^N)$$
$$= K^{N-1}(u_{-i}^{N-1}) + c_i u_i^{N-1}$$
$$+ p\mathbb{E}\{\max(0, -(\gamma - \omega_i^{N-1}))\} + h\mathbb{E}\{\max(0, \gamma - \omega_i^{N-1})\}$$
$$= K^{N-1}(u_{-i}^{N-1}) + c_i(\gamma - x_i^{N-1}) + h\mathbf{E}_h^{N-1}(\gamma) + p\mathbf{E}_s^{N-1}(\gamma),$$

where $\mathbf{E}_h^t(\gamma)$ and $\mathbf{E}_s^t(\gamma)$ are the expected holding and shortage, respectively, defined as:

$$\mathbf{E}_h^t(\gamma) := \mathbb{E}\left\{\max\left(0, \gamma - \omega_i^t\right)\right\},$$
$$\mathbf{E}_s^t(\gamma) := \mathbb{E}\left\{\max(0, -\left(\gamma - \omega_i^t\right))\right\}.$$

Let the discrete difference operator be given, $\frac{d}{dS}$ and let us apply such an operator to function $G_i^{N-1}(\gamma, \bar{u}_{-i}^N) = c_i(\gamma - x_i^{N-1}) + h\mathbf{E}_h(\gamma) + p\mathbf{E}_s(\gamma)$. Then, we have

$$\frac{d}{d\gamma} G_i^{N-1}(\gamma, \bar{u}_{-i}^N) := G_i^{N-1}(\gamma + 1, \bar{u}_{-i}^N) - G_i^{N-1}(\gamma, \bar{u}_{-i}^N)$$
$$= c_i + h\Phi_\omega^{N-1}[\gamma] - p\left(1 - \Phi_\omega^{N-1}[\gamma]\right).$$

where

$$\Phi_\omega^t[\gamma] := \sum_{\omega=0}^\gamma \phi_\omega^t, \quad 1 - \Phi_\omega^t[\gamma] := \sum_{\omega=\gamma+1}^\infty \phi_\omega^t.$$

In the above equations, we make use of the following conditions

$$\sum_{\omega=0}^{\gamma+1}(\gamma + 1 - \omega)\phi_\omega^{N-1} = \sum_{\omega=0}^\gamma(\gamma + 1 - \omega)\phi_\omega^{N-1}$$
$$= \sum_{\omega=0}^\gamma(\gamma - \omega)\phi_\omega^{N-1} + \sum_{\omega=0}^\gamma \phi_\omega^{N-1},$$
$$\sum_{\omega=\gamma+2}^\infty(\omega - \gamma - 1)\phi_\omega^{N-1} = \sum_{\omega=\gamma+1}^\infty(\omega - \gamma - 1)\phi_\omega^{N-1}$$
$$= \sum_{\omega=\gamma+1}^\infty(\omega - \gamma)\phi_\omega^{N-1} - \sum_{\omega=\gamma+1}^\infty \phi_\omega^{N-1}. \tag{16}$$

The order-up-to level S_i^{N-1} is the optimal γ, which is obtained from solving

$$\min_\gamma \left\{\gamma \,|\, \frac{d}{d\gamma} G_i^{N-1}(\gamma, \bar{u}_{-i}^N) \geq 0\right\}$$
$$= \min_\gamma \left\{\gamma \,|\, c_i + h\Phi_\omega^{N-1}[\gamma] - p\left(1 - \Phi_\omega^{N-1}[\gamma]\right) \geq 0\right\}.$$

From the above, we then obtain

$$S_i^{N-1} = \arg\min_\gamma \left\{\gamma \,|\, \Phi_\omega^{N-1}[\gamma] \geq \frac{-c_i + p}{h + p}\right\}.$$

To obtain s_i^{N-1}, let us consider the cost of not reordering, which is given by

$$-c_i x_i^{N-1} + G_i^{N-1}(x_i^{N-1}, \bar{u}_{-i}^N) = p\mathbb{E}\{\max(0, -(x_i^{N-1} - \omega_i^{N-1}))\}$$
$$+ h\mathbb{E}\{\max(0, x_i^{N-1} - \omega_i^{N-1})\}$$
$$= h\mathbf{E}_h(x_i^{N-1}) + p\mathbf{E}_s(x_i^{N-1})$$

Also, we have

$$s_i^{N-1} := \arg\min_{x_i^{N-1}} \left\{ x_i^{N-1} \mid h\mathbf{E}_h\left(x_i^{N-1}\right) + p\mathbf{E}_s\left(x_i^{N-1}\right) \right.$$
$$\leq K^{N-1}\left(u_{-i}^{N-1}\right) - c_i \left(S_i^{N-1} - x_i^{N-1}\right)$$
$$+ h\mathbf{E}_h\left(S_i^{N-1}\right) + p\mathbf{E}_s\left(S_i^{N-1}\right) \bigg\}.$$

Now, we are going to assume that the statement is true for some $t = m$, and we are going to proof that it is also valid for $t = m - 1$.

Consider now the convex function (see Fig. 2 which illustrate the example of $t = N - 2$)

$$G_i^{m-1}(y_i^{m-1}, \bar{u}_{-i}^m) = c_i y_i^{m-1} + p\mathbb{E}\left\{\max\left(0, -\left(y_i^{m-1} - \omega_i^{m-1}\right)\right)\right\}$$
$$+ h\mathbb{E}\left\{\max\left(0, y_i^{m-1} - \omega_i^{m-1}\right)\right\} + \mathbb{E}\left\{v_i^m\left(x_i^m, \bar{u}_{-i}^m\right)\right\}$$
$$= c_i y_i^{m-1} + h\mathbf{E}_h\left(y_i^{m-1}\right) + p\mathbf{E}_s\left(y_i^{m-1}\right)$$
$$+ \sum_{\omega=0}^{\infty} v_i^m\left(y_i^{m-1} - \omega, \bar{u}_{-i}^m\right) \phi_\omega^{m-1}. \tag{17}$$

We know that G_i^{m-1} is $\mathscr{K}$-convex, with $\mathscr{K} = K^{m-1}$. This property implies that the function

$$v_i^{m-1}\left(x_i^{m-1}, \bar{u}_{-i}^{m-1}\right) = -c_i x_i^{m-1} + \min_{y_i^{m-1} \geq x_i^{m-1}} \left[K^{m-1}\left(u_{-i}^{m-1}\right) \right.$$
$$\left. + G_i^{m-1}\left(y_i^{m-1}, \bar{u}_{-i}^m\right), G_i^{m-1}\left(x_i^{m-1}, \bar{u}_{-i}^m\right) \right], \tag{18}$$

is K^{m-1}-convex, with a global minimum at $S_i^{m-1} := argmin_\gamma G_i^{m-1}(\gamma, \bar{u}_{-i}^m)$. It is important to notice that we can ensure the existence of a unique minimum value in (18) thanks to the nondecreasing property of K^{m-1}.
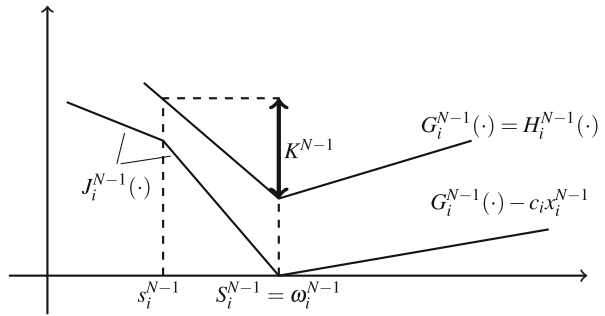
The cost of reordering for $t = m - 1$ is given by

$$K^{m-1}\left(u_{-i}^{m-1}\right) - c_i x_i^{m-1} + G_i^{m-1}\left(\gamma, \bar{u}_{-i}^m\right)$$
$$= K^{m-1}\left(u_{-i}^{m-1}\right) + c_i u_i^{m-1}$$
$$+ p\mathbb{E}\left\{\max\left(0, -\left(\gamma - \omega_i^{m-1}\right)\right)\right\} + h\mathbb{E}\left\{\max\left(0, \gamma - \omega_i^{m-1}\right)\right\}$$
$$= K^{m-1}(u_{-i}^{m-1}) + c_i\left(\gamma - x_i^{m-1}\right) + h\mathbf{E}_h^{m-1}(\gamma) + p\mathbf{E}_s^{m-1}(\gamma).$$

Applying operator $\frac{\mathrm{d}}{\mathrm{d}\gamma}$ to function $G_i^{m-1}(\gamma, \bar{u}_{-i}^m)$, we have

$$\frac{\mathrm{d}}{\mathrm{d}\gamma} G_i^{m-1}(\gamma, \bar{u}_{-i}^m) := G_i^{m-1}\left(\gamma + 1, \bar{u}_{-i}^m\right) - G_i^{m-1}\left(\gamma, \bar{u}_{-i}^m\right)$$
$$= c_i + h\Phi_\omega^{m-1}[\gamma] - p\left(1 - \Phi_\omega^{m-1}[\gamma]\right) +$$
$$\sum_{\omega=0}^{\infty} \left[v_i^m(\gamma + 1 - \omega, \cdot) - v_i^m(\gamma - \omega, \cdot)\right] \phi_\omega^{m-1}.$$

Fig. 1 A qualitative plot of functions $G_i^{N-1}(\cdot)$ and $v_i^{N-1}(\cdot)$ obtained from (14) and (15), respectively

Hence, the order-up-to level S_i^{m-1} is the optimal γ, which is obtained from solving

$$S_i^{m-1} = \arg\min_{\gamma} \left\{ \gamma \mid c_i + h\Phi_\omega^{m-1}[\gamma] - p\left(1 - \Phi_\omega^{m-1}[\gamma]\right) \right.$$
$$\left. + \sum_{\omega=0}^{\infty} \left[v_i^m(\gamma + 1 - \omega, \cdot) - v_i^m(\gamma - \omega, \cdot)\right] \phi_\omega^{m-1} \geq 0 \right\}.$$

To obtain s_i^{m-1}, let us consider the cost of not reordering, which is given by

$$-c_i x_i^{m-1} + G_i^{m-1}\left(x_i^{m-1}, \bar{u}_{-i}^m\right)$$
$$= h\mathbf{E}_h\left(x_i^{m-1}\right) + p\mathbf{E}_s(x_i^{m-1}) + \sum_{\omega=0}^{\infty} v_i^m\left(x_i^{m-1} - \omega, \bar{u}_{-i}^m\right) \phi_\omega^{m-1}.$$

Then, we have

$$s_i^{m-1} := \arg\min_{x_i^{m-1}} \left\{ x_i^{m-1} \mid -c_i x_i^{m-1} + G_i^{m-1}(x_i^{m-1}, \bar{u}_{-i}^m) \right.$$
$$+ \sum_{\omega=0}^{\infty} v_i^m(x_i^{m-1} - \omega, \bar{u}_{-i}^m)\phi_\omega^{m-1}$$
$$\leq K^{m-1}(u_{-i}^{m-1}) - c_i S_i^{m-1} + G_i^{m-1}(S_i^{m-1}, \bar{u}_{-i}^m)$$
$$\left. + \sum_{\omega=0}^{\infty} v_i^m(S_i^{m-1} - \omega, \bar{u}_{-i}^m)\phi_\omega^{m-1} \right\}.$$

The above can be rewritten as

$$s_i^{m-1} := \arg\min_{x_i^{m-1}} \left\{ x_i^{m-1} \mid h\mathbf{E}_h(x_i^{m-1}) + p\mathbf{E}_s(x_i^{m-1}) \right.$$
$$+ \sum_{\omega=0}^{\infty} v_i^m(x_i^{m-1} - \omega, \bar{u}_{-i}^m)\phi_\omega^{m-1}$$
$$\leq K^{m-1}(u_{-i}^{m-1}) - c_i(S_i^{m-1} - x_i^{m-1})$$
$$+ h\mathbf{E}_h(S_i^{m-1}) + p\mathbf{E}_s(S_i^{m-1})$$
$$\left. + \sum_{\omega=0}^{\infty} v_i^m(S_i^{m-1} - \omega, \bar{u}_{-i}^m)\phi_\omega^{m-1} \right\}.$$

Thus by induction backwards in time, we have proved Theorem 1. $\quad\square$

We can reinterpret the (s, S)-like strategies as threshold strategies on the number of active retailers. The result is that all Nash equilibrium strategies have the threshold structure (4).

In the following result on a single-stage inventory game (where we have dropped index t), we reinterpret a threshold on the inventory level as a threshold on the number of "active retailers".
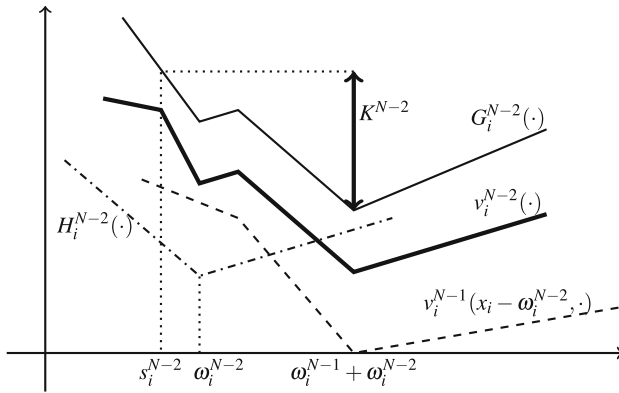
Fig. 2 A qualitative plot of functions $G_i^{N-2}(\cdot)$ and $v_i^{N-2}(\cdot)$ obtained from (17) and (18), respectively

Theorem 2 *For each inventory level x_i, there exists a threshold $l_i \in \{1, 2, \ldots, n\}$, such that the replenishment strategy*

$$\mu_i(x_i, a) = \begin{cases} S_i - x_i, & \text{if } a \geq l_i, \\ 0, & \text{if } a < l_i, \end{cases} \tag{19}$$

is a Nash equilibrium for the single-stage formulation of the inventory game. For the sake of simplicity, we have dropped dependence on time.

Proof From Theorem 1, if $N = 1$, we have a unique multi-period strategy (s_i, S_i). This means that the retailers make decisions according to

$$u_i = \mu_i(x_i) = \begin{cases} S_i - x_i, & \text{if } x_i < s_i, \\ 0, & \text{if } x_i \geq s_i. \end{cases} \tag{20}$$

Note that from $G_i(s_i) = G_i(S_i) + (\frac{K}{a})$ we have that s_i depends on the number of active players a. Now, for given x_i, the idea is to find l_i as the minimum number of active players such that the cost of replenishing does not exceed the cost of not replenishing. This can be expressed by the minimization below (in a single-stage optimization, we can drop the second argument $\bar{u}_{-i}^{t+1}$ from $G_i(., .)$)

$$l_i = \min_{a=1,\ldots,n} \left\{ a \mid G_i(x_i) \geq G_i(s_i), \; G_i(s_i) = G_i(S_i) + (K/a) \right\}. \tag{21}$$

Strategy (20) implies (19) once we compute l_i from (21) for fixed x_i.

In solving (21), we distinguish three cases.

– The inventory level is "low," namely, $x_i < \underline{s}_i$. Then, the optimal decision is "replenish" independently of a. Actually, the minimization (21) returns $l_i = 1$ and as it always holds $a \geq l_i$ we have $\mu_i(x_i, a) = S_i - x_i$.
– The inventory level is "high," namely $x_i \geq \bar{s}_i$. Then, the optimal decision is "do not replenish." Indeed, the minimization (21) is infeasible. With a little abuse of notation, we can take $l_i = n + 1$ so that it always holds $a < l_i$ and therefore also $\mu_i(x_i, a) = 0$.
– The inventory level verifies $\underline{s}_i \leq x_i \leq \bar{s}_i$. To see this, note that the computation of l_i as in (21) leads to $1 \leq l_i \leq n$. Then, if $a \geq l_i$ from (21), we have $x_i < s_i$ which substituted in (20) returns $\mu_i(x_i, a) = S_i - x_i$. Differently if $a < l_i$ from (21), we have $x_i \geq s_i$ which,

again, substituted in (20) returns $\mu_i(x_i, a) = 0$. The obtained strategy for $\mu_i(x_i, a)$ is in accordance with (19), and this concludes the proof.   □

5 Single-Stage Coordination

In this section, we specialize our results to the case of single-stage game. In particular, we provide explicit expressions for the two thresholds, as a function of the probability distribution function which determines the stochastic demand.

Let us start by noting that in the single-stage game function $G_i^t(y_i^t, \bar{u}_{-i}^{t+1 \sim N-1})$ does not depend on $\bar{u}_{-i}^N$ and therefore, we simply write $G_i^t(y_i^t)$:

$$G_i^t\left(y_i^t\right) = cy_i^t + p\mathbb{E}\left\{\max\left(0, -\left(y_i^t - \omega_i^t\right)\right)\right\} + h\mathbb{E}\left\{\max\left(0, y_i^t - \omega_i^t\right)\right\}. \tag{22}$$

Then, we have for the value function

$$v_i^t\left(x_i^t, \bar{u}_{-i}^t\right) = -c_i x_i^t + \min_{y_i^t \geq x_i^t}\left[K^t\left(u_{-i}^t\right) + G_i^t\left(y_i^t\right), G_i^t\left(x_i^t\right)\right]. \tag{23}$$

To obtain S_i^t, consider the cost of reordering, which is given by

$$\begin{aligned}
K^t &\left(u_{-i}^t\right) - c_i x_i^t + G_i^t(\gamma) \\
&= K^t\left(u_{-i}^t\right) + c_i u_i^t + p\mathbb{E}\{\max\left(0, -\left(\gamma - \omega_i^t\right)\right)\} \\
&\quad + h\mathbb{E}\{\max\left(0, \gamma - \omega_i^t\right)\} \\
&= K^t\left(u_{-i}^t\right) + c_i\left(\gamma - x_i^t\right) \\
&\quad + p\mathbb{E}\{\max(0, -(\gamma - \omega_i^t))\} + h\mathbb{E}\{\max\left(0, \gamma - \omega_i^t\right)\} \\
&= K^t(u_{-i}^t) + c_i\left(\gamma - x_i^t\right) + h\mathbf{E}_h(\gamma) + p\mathbf{E}_s(\gamma).
\end{aligned}$$

Let the discrete difference operator be given, $\frac{\mathrm{d}}{\mathrm{d}S}$ and let us apply such an operator to function

$$G_i^t(\gamma) = c_i(\gamma - x_i^t) + h\underbrace{\sum_{\omega=0}^{\gamma}(\gamma - \omega)\phi_\omega^t}_{\mathbf{E}_h(\gamma)} + p\underbrace{\sum_{\omega=\gamma+1}^{\infty}(\omega - \gamma)\phi_\omega^t}_{\mathbf{E}_s(\gamma)}.$$

By applying the difference operator to function $G_i^t(\gamma)$, we then have

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}\gamma}G_i^t(\gamma) &:= G_i^t(\gamma + 1) - G_i^t(\gamma) \\
&= c_i\left(\gamma + 1 - x_i^t\right) + h\sum_{\omega=0}^{\gamma+1}(\gamma + 1 - \omega)\phi_\omega^t \\
&\quad + p\sum_{\omega=\gamma+2}^{\infty}(\omega - \gamma - 1)\phi_\omega^t \\
&\quad - c_i(\gamma - x_i^t) - h\sum_{\omega=0}^{\gamma}(\gamma - \omega)\phi_\omega^t - p\sum_{\omega=\gamma+1}^{\infty}(\omega - \gamma)\phi_\omega^t.
\end{aligned}$$

Further derivations yield

$$
\begin{aligned}
\frac{d}{d\gamma}G_i^t(\gamma) &= c_i\left(\gamma + 1 - x_i^t\right) + h\left[\sum_{\omega=0}^{\gamma}(\gamma - \omega)\phi_\omega^t + \sum_{\omega=0}^{\gamma}\phi_\omega^t\right] \\
&\quad + p\left[\sum_{\omega=\gamma+1}^{\infty}(\omega - \gamma)\phi_\omega^t - \sum_{\omega=\gamma+1}^{\infty}\phi_\omega^t\right] - c_i\left(\gamma - x_i^t\right) \\
&\quad - h\sum_{\omega=0}^{\gamma}(\gamma - \omega)\phi_\omega^t - p\sum_{\omega=\gamma+1}^{\infty}(\omega - \gamma)\phi_\omega^t \\
&= c_i + h\sum_{\omega=0}^{\gamma}\phi_\omega^t - p\sum_{\omega=\gamma+1}^{\infty}\phi_\omega^t \\
&= c_i + h\Phi_\omega^t[\gamma] - p\left(1 - \Phi_\omega^t[\gamma]\right).
\end{aligned}
$$

In the above, we have used the following equalities

$$
\begin{aligned}
\sum_{\omega=0}^{\gamma+1}(\gamma + 1 - \omega)\phi_\omega^t &= \sum_{\omega=0}^{\gamma}(\gamma + 1 - \omega)\phi_\omega^t \\
&= \sum_{\omega=0}^{\gamma}(\gamma - \omega)\phi_\omega^t + \sum_{\omega=0}^{\gamma}\phi_\omega^t, \\
\sum_{\omega=\gamma+2}^{\infty}(\omega - \gamma - 1)\phi_\omega^t &= \sum_{\omega=\gamma+1}^{\infty}(\omega - \gamma - 1)\phi_\omega^t \\
&= \sum_{\omega=\gamma+1}^{\infty}(\omega - \gamma)\phi_\omega^t - \sum_{\omega=\gamma+1}^{\infty}\phi_\omega^t.
\end{aligned}
\tag{24}
$$

The order-up-to level S_i^t is the optimal γ, which is obtained from solving

$$
\begin{aligned}
\min_\gamma &\left\{\gamma\,\Big|\,\frac{d}{d\gamma}G_i^t(\gamma) \geq 0\right\} \\
&= \min_\gamma\left\{\gamma\,\big|\,c_i + h\Phi_\omega^t[\gamma] - p(1 - \Phi_\omega^t[\gamma]) \geq 0\right\}.
\end{aligned}
$$

From the above, we then obtain

$$
S_i^t = \arg\min_\gamma\left\{\gamma\,\Big|\,\Phi_\omega^t[\gamma] \geq \frac{-c_i + p}{h + p}\right\}.
\tag{25}
$$

To obtain s_i^t, let us consider the cost of not reordering, which is given by

$$
\begin{aligned}
-c_i x_i^t + G_i^t\left(x_i^t\right) &= p\mathbb{E}\{\max\left(0, -\left(x_i^t - \omega_i^t\right)\right)\} \\
&\quad + h\mathbb{E}\{\max\left(0, x_i^t - \omega_i^t\right)\} \\
&= h\mathbf{E}_h\left(x_i^t\right) + p\mathbf{E}_s\left(x_i^t\right).
\end{aligned}
\tag{26}
$$

From the above, we then obtain

$$
s_i^t := \arg\min_{x_i^t}\left\{x_i^t\,\big|\,-c_i x_i^t + G_i^t\left(x_i^t\right) \leq K^t\left(u_{-i}^t\right) - c_i S_i^t + G_i^t\left(S_i^t\right)\right\}.
$$

In particular, we have

$$
\begin{aligned}
s_i^t := \arg\min_{x_i^t}\Big\{&x_i^t\,\big|\,h\mathbf{E}_h\left(x_i^t\right) + p\mathbf{E}_s(x_i^t) \\
&\leq K^t\left(u_{-i}^t\right) + c_i\left(S_i^t - x_i^t\right) + h\mathbf{E}_h\left(S_i^t\right) + p\mathbf{E}_s\left(S_i^t\right)\Big\}.
\end{aligned}
\tag{27}
$$

Equations (25) and (27) represent explicit expressions for the two thresholds and fully characterize then the reordering strategy once the probability distribution of the stochastic demand is given.

Once thresholds are obtained, we implement the control u_i^t which is given by

$$u_i^t = \mu(x^t) = \begin{cases} S_i^t - x^t, & \text{if } x^t < s_i^t, \\ 0, & \text{if } x^t \geq s_i^t. \end{cases} \tag{28}$$

The resulting dynamics is then

$$x_i^{t+1} = \begin{cases} S_i^t - \omega_i^t, & \text{if } x_i^t < s_i^t, \\ x_i^t - \omega^t, & \text{if } x_i^t \geq s_i^t. \end{cases} \tag{29}$$

6 Numerical Analysis

We consider an example where the demand $\omega^t \in \Omega := \{0, 1, 2\}$ and is uniformly distributed, namely after introducing the notation ϕ_ω to indicate the probability that $\omega^t = \omega$, we have $\phi_\omega = \frac{1}{3}$ for $\omega = 0, 1, 2$.

Assume that the proportional purchase cost is $c = 1$, the shortage cost is $p = 10$, and the holding cost is $h = 2$. In the case of single stage optimization, we have that the order up to level is given by

$$S = \arg\min_\gamma \left\{ \gamma \mid \Phi_\omega^t[\gamma] \geq \frac{-c + p}{h + p} \right\}.$$

From the above, we obtain $S = 2$. Indeed for $\gamma = 2$, we have

$$\Phi_\omega^t[2] = 1 \geq \frac{-c + p}{h + p} = \frac{3}{4}.$$

Differently, for $\gamma = 1$ it holds

$$\Phi_\omega^t[1] = \frac{2}{3} \ngeq \frac{-c + p}{h + p} = \frac{3}{4},$$

and therefore

$$S = \arg\min_\gamma \left\{ \gamma \mid \Phi_\omega^t[\gamma] \geq \frac{-c + p}{h + p} \right\} = 2.$$

As for the reorder level s, we have

$$s := \arg\min_x \Big\{ x \mid h\mathbf{E}_h(x) + p\mathbf{E}_s(x)$$
$$\leq K^t + c(S - x) + h\mathbf{E}_h(S) + p\mathbf{E}_s(S) \Big\}.$$

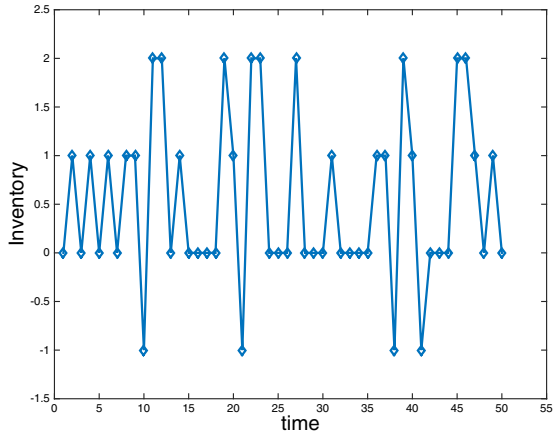We show next that we have $s = 1$. Actually, for $x = 1$ we obtain

$$h\mathbf{E}_h(1) + p\mathbf{E}_s(1) = h\frac{1}{3} + p\frac{1}{3} = 4$$
$$\leq K^t + c + h\mathbf{E}_h(2) + p\mathbf{E}_s(2)$$
$$= K^t + c + h\mathbf{E}_h(2) = K^t + 3,$$

which is satisfied by any $K^t \geq 1$.

For $x = 0$, we have

$$h\mathbf{E}_h(0) + p\mathbf{E}_s(0) = p\mathbf{E}_s(0) = 10$$
$$\nleq K^t + 2c + h\mathbf{E}_h(2) + p\mathbf{E}_s(2)$$

Fig. 3 Time plot of the microscopic dynamics of a single player

$$= K^t + 2c + h\mathbf{E}_h(2) = K^t + 4,$$

which is satisfied by any $K^t < 6$. For any $K^t < 6$, we then have

$$s := \arg\min_x \left\{ x \,|\, h\mathbf{E}_h(x) + p\mathbf{E}_s(x) \right.$$
$$\left. \leq K^t + c(S - x) + h\mathbf{E}_h(S) + p\mathbf{E}_s(S) \right\} = 1.$$

We can conclude then that for any K^t such that $1 \leq K^t < 6$ we have the reorder level $s = 1$ and the order-up-to level $S = 2$.

Then, from (29) the microscopic dynamics is defined in the bounded support $\{-1, 0, 1, 2\}$, namely $x^t \in \{-1, 0, 1, 2\}$ for all $t \geq 0$ and is given by

$$x^{t+1} = \begin{cases} 2 - \omega^t, & \text{if } x^t = -1, 0, \\ x^t - \omega^t, & \text{if } x^t = 1, 2. \end{cases} \tag{30}$$

Figure 3 displays the time plot of the microscopic dynamics for a single player. In other words, the plot shows the inventory level (the state) of a player. The player's inventory is for most of the time in state 0 and 1, which is in accordance with the greater values of the distribution in those states.

In the following example, we consider a larger instance involving five agents, where the demand of each agent $w^t \in \Omega := \{0, 1, \ldots, 20\}$ and is uniformly distributed.
Assume the same purchase, shortage, and holding costs as in the previous example and consider a transportation cost $K = 120$, which will be divided among the active agents at each time $t \in [0, 50]$.

Figure 4 shows the relation between the inventory levels and the transportation costs that each player is willing to pay in case of reordering as well as the minimum number of active agents in case of replenishment for any inventory level. It is possible to see that the inventory has an inverse relation with the transportation cost and an increase relation with the number of active agents. This means that if the inventory level of agent i is higher, the agent is willing to pay less in case of reordering and hence, it is expected to require a large number of active agents to coordinate with.

The last two figures (Figs. 5, 6) display the inventory level of the five players over time. In Fig. 5, it is possible to see the moment in time when it is most convenient that the players coordinate for replenishment. On the other hand, Fig. 6 exhibits the relation of the inventory
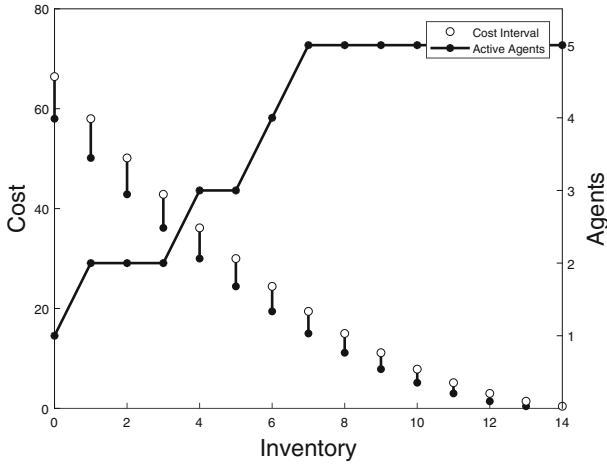
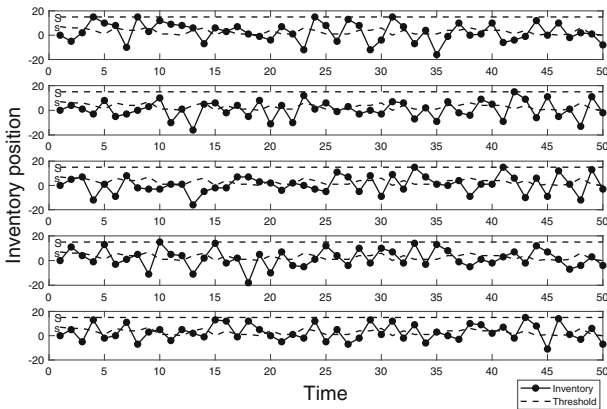**Fig. 4** Transportation cost intervals and active agents at each positive inventory level

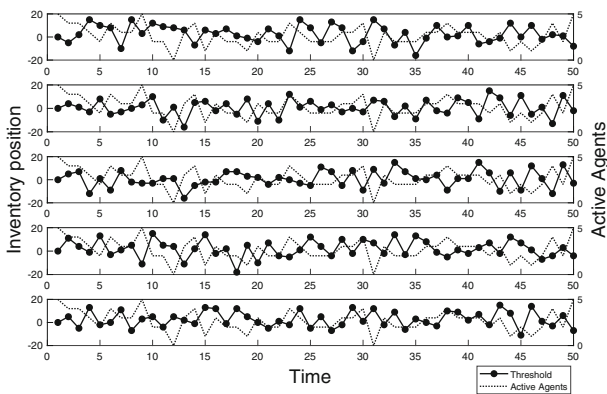**Fig. 5** Time plot of the microscopic dynamics of 5 players

Fig. 6 Time plot of the microscopic dynamics of 5 players and the number of active players

level and the number of active agents at each time. It is clear that the agents reorder when its inventory level is lower or equal to the threshold s, which also depends on the number of active agents, and they reorder up to the upper threshold $S = 15$.

7 Conclusions and Future Works

We first developed an abstraction in the form of a dynamic coordination game model where each player's dynamics is a scalar inventory model characterized by a controlled input and an uncontrolled output. The players have to pay a share of the activation cost to control their dynamics at a given time. First, we showed that if the retailers are rational players, then they benefit from using threshold strategies where the threshold is on the number of active players. We then turned to obtain an explicit expressions for the lower and upper thresholds under specific circumstances. A main key direction for future works is to explore the feasibility of the proposed coordination scheme in multi-vector energy systems (heat, gas, power) with special focus on coalitional bidding in decentralized energy trade. The ultimate goal is to investigate the benefits of aggregating independent wind power producers.

References

1. Başar T, Olsder GJ (1995) Dynamic Noncooperative Game Theory, 2nd edn. Academic Press, London
2. Bauso D, Giarrè L, Pesenti R (2008) Consensus in noncooperative dynamic games: a multi-retailer inventory application. IEEE Trans Autom Control 53(4):998–1003
3. Bauso D, Giarrè L, Pesenti R (2009) Distributed consensus in noncooperative inventory games. Eur J Oper Res 192(3):866–878
4. Bauso D, Zhu Q, Başar T (2016) Decomposition and mean-field approach to mixed integer optimal compensation problems. J Optim Theory Appl 169:606–630
5. Bertsekas DP (1995) Dynamic programming and optimal control, 2nd edn. Athena, Bellmont, MA
6. Clark A, Scarf S (1960) Optimal policies for a multi-echelon inventory problem. Manag Sci 6(4):475–490
7. Dekker R, Wildeman RE, van der Duyn Schouten FA (1997) A review of multi-component maintenance models with economic dependence. Math Methods Oper Res 45:344–357
8. Federgruen A, Groenevelt H, Tijms HC (1984) Coordinated replenishments in a multi-item inventory system with compound poisson demands. Manag Sci 30(3):344–357
9. Osborne Martin J, Rubinstein Ariel (1994) A course in game theory. MIT Press, Cambridge, MA
10. Pochet Y, Wolsey LA (1993) Lot sizing with constant batches: formulations and valid inequalities. Math Oper Res 18(4):767–785