A Distributionally Robust LQR for Systems with Multiple Uncertain Players

I. Tzortzis, C.D. Charalambous, and C.N. Hadjicostis
University of Cyprus
Department of Electrical and Computer Engineering

60th CONFERENCE ON DECISION AND CONTROL, DECEMBER 13-17, 2021

OUTLINE

- Introduction and Motivation
- Problem Formulation
- MinMax Stochastic Control
- Drop-Shipping Problem
- Concluding Remarks

Introduction - Motivation

Introduction - Motivation

High-Level Objective

Develop a distributionally robust LQR approach, applicable to discrete-time dynamical systems composed of several uncertain players with ambiguous distribution information

A Short Review

- LQR provides state feedback control policies for systems stability and optimal performance
  - Tracking control systems, networked control systems, fault-tolerant control systems, power systems, etc
- Multiple sources of uncertainty with ambiguous distribution information require:
  - Properly understand the effect of these sources of uncertainty
  - Study their impact on the performance of the LQR
- In specific applications, estimating system's response probability distribution, under ambiguity, might be insufficient

Introduction - Motivation

An Illustration

- *m etailers*: sell a homogeneous product to consumers
- *Distribution center*: fulfilment and delivery of the product
  - Places orders u_k based on a stochastic demand $\{w_k^i, \mu_{w_k^i}\}$ provided by each of the etailers
  - Equal amount of confidence placed between etailers
- *Objective*: maximize profit with minimum inventory investment, without impacting customer satisfaction levels due to shortage of inventory

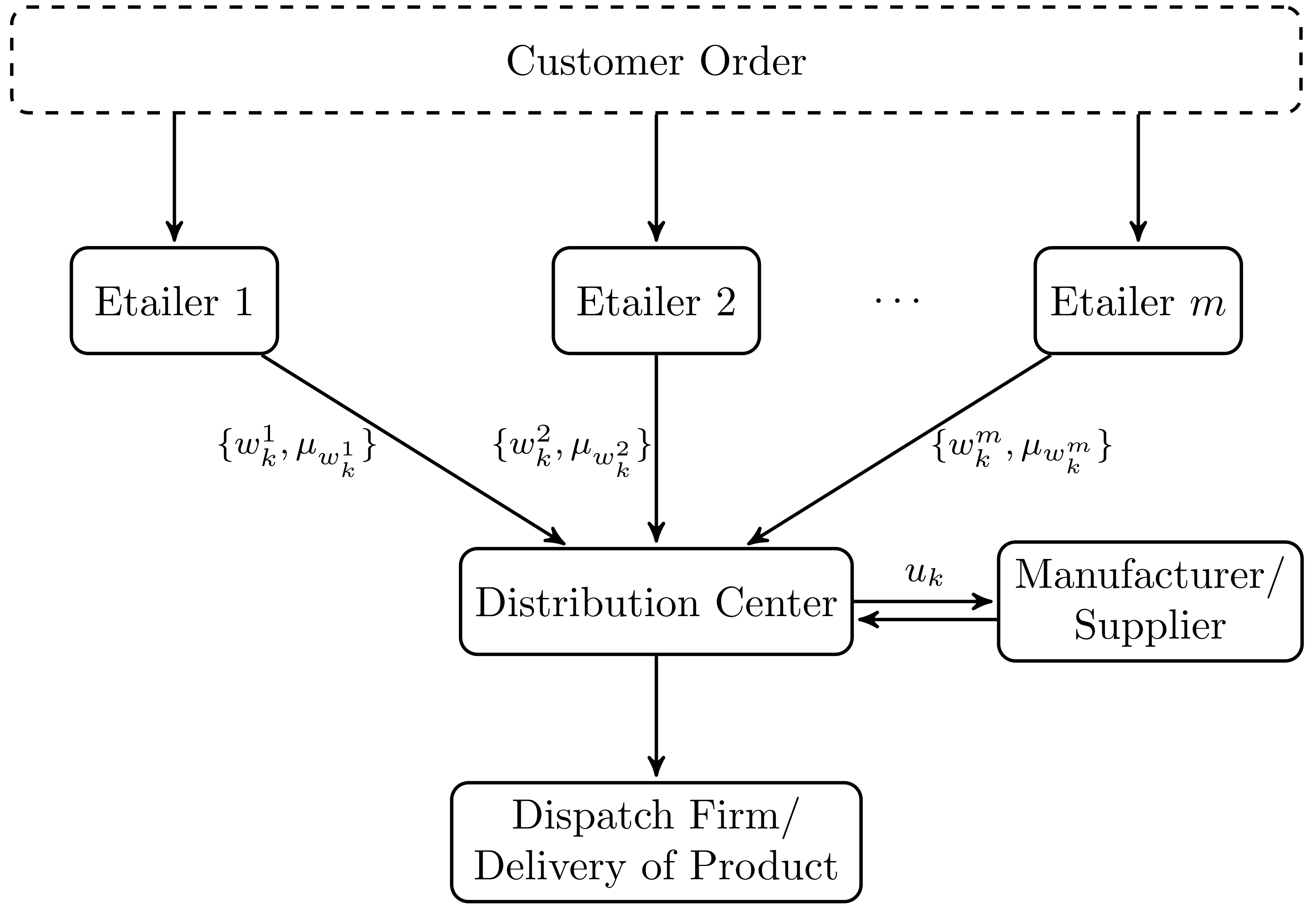


Fig.1: Drop-shipping model

Question

How to deal with the case in which the distribution center has varying confidence for etailers stochastic demand?

Introduction - Motivation

Methodology

- To study the effect of ambiguous sources of uncertainty, a sequential hierarchical game with multiple uncertain players is adopted
- Each player is prescribed by a nominal probability distribution (p.d.), and categorised according to an uncertainty level of confidence
- The sequence of leader - followers is sequentially decided during the different time intervals
- LQR control policy is derived by following a dynamic programming approach

Main Contributions

- A maximizing, time-varying, probability distribution for each player
  - Leads to an optimal equilibrium solution
- A closed-form expression of the robust LQR control policy
  - Preserves its linearity
  - Its evaluation is performed based on multiple players maximizing probability distribution

Problem Formulation

Problem Formulation

Model Description

- Discrete-time control system:

$$x_{k+1} = A_k x_k + B_k u_k + \sum_{i=1}^{m} C_k^i w_k^i, \qquad x_0 = x \tag{1}$$

  - $x_k \in \mathcal{X}_k$: state process
  - $u_k \in \mathcal{U}_k$: control process
  - A_k, B_k, C_k^i : matrices of compatible dimensions
  - $w_k^i \in \mathcal{W}_k^i$: independent sequence of random vectors with unknown p.d. $\nu_{w_k^i}(dw)$

- Basic random vectors $\{x_0, w_0^1, \ldots, w_{N-1}^1, \ldots, w_0^m, \ldots, w_{N-1}^m\}$ are all assumed to be mutually independent
- Markov feedback control policies $g_k : \mathcal{X}_k \longmapsto \mathcal{U}_k$ such that $u_k^g = g_k(x_k^g)$

Problem Formulation

Ambiguity Class

Given a collection of nominal p.d. $\mu_{w_k^i}(dw)$, the corresponding collection of true p.d. $\nu_{w_k^i}(dw)$ is modelled by

$$\mathbb{B}_{R_k^i}(\mu_{w_k^i}) \triangleq \{\nu_{w_k^i} \in \mathcal{M}_1(\mathcal{W}_k^i) : ||\nu_{w_k^i}(\cdot) - \mu_{w_k^i}(\cdot)||_{TV} \leq R_k^i\} \tag{2}$$

- $\mathcal{M}_1(\mathcal{W})$: set of p.d. on $\mathcal{W}$
- $||\alpha - \beta||_{TV} \triangleq \sup_{P \in \mathcal{P}(\mathcal{W})} \sum_{F_i \in P} |\alpha(F_i) - \beta(F_i)|$, where $\alpha, \beta \in \mathcal{M}_1(\mathcal{W})$ and $\mathcal{P}(\mathcal{W})$ finite partitions on $\mathcal{W}$
- $R_k^i \in [0,2]$: It can be interpreted as the level of confidence on player's i nominal p.d.

Stochastic Control Problem

Define the N-stage expected cost by

$$J_N(g, \nu) \triangleq \mathbb{E}_\nu^{g,x}[\sum_{k=0}^{N-1} (x_k^T Q_k x_k + u_k^T R_k u_k) + x_N^T Q_N x_N] \tag{3}$$

- $\mathbb{E}_\nu^{g,x}[\cdot]$: indicates dependence on policy $g \in G$, for a given $x_0 = x$, and induced by the p.d. $\nu \triangleq \{\nu_{w_k^i}\}$

Problem Formulation

MinMax Stochastic Control Problem

Find an optimal control policy $g^* \in G$, and a maximizing p.d. $\nu^* \in \mathbb{B}_{R_k^i}$ within the total variation distance ambiguity set, which solve the optimization problem

$$J^* = J_N(g^*, \nu^*) = \min_{g \in G} \max_{\nu_{w_k^i} \in \mathbb{B}_{R_k^i}} J_N(g, \nu)$$

Remark

By appropriately adjusting the total variation distance parameter $R_k^i \in [0,2]$ we can control:

- The size of the ambiguity set
- The degree of conservatism of the optimization problem

MinMax Stochastic Control

MinMax Stochastic Control

Dynamic Programming

- Apply dynamic programming to obtain:

$$V_N(x_N) = x_N^T Q_N x_N$$

$$V_k(x) = \min_{u_k \in \mathcal{U}_k(x)} \max_{\nu_{w_k^i} \in \mathbb{B}_{R_k^i}, i=1,\ldots,m} \mathbb{E}_\nu^{g,x}\left[x^T Q_k x + u_k^T R_k u_k + V_{k+1}(x_{k+1})\right]$$

$$= \min_{u_k \in \mathcal{U}_k(x)} \left\{x^T Q_k x + u_k^T R_k u_k + \max_{\nu_{w_k^i} \in \mathbb{B}_{R_k^i}, i=1,\ldots,m} \mathbb{E}_\nu^{g,x}\left[V_{k+1}(x_{k+1})\right]\right\}$$

$$= \min_{u_k \in \mathcal{U}_k(x)} \left\{x^T Q_k x + u_k^T R_k u_k + \max_{\nu_{w_k^i} \in \mathbb{B}_{R_k^i}, i=1,\ldots,m} \mathbb{E}_\nu^{g,x}\left[\ell_k(x_k, u_k, w_k^1, \ldots, w_k^m))\right]\right\}$$

- *Induction hypothesis:* Suppose that $V_t(x) = x^T P_t x + x^T F_t + r_t$ holds for $t = k+1, \ldots, N$. By characterizing the solution of the inner and outer optimization of the minimax dynamic equation, show that

$$V_k(x) = x^T P_k x + x^T F_k + r_k, \quad P_k = P_k^T \geq 0,\ F_k = F_k^T \geq 0,\ r_k \in \mathbb{R} \tag{4}$$

MinMax Stochastic Control

Inner Optimization

- Consider a sequential game with m players, where each player $i = 1,2,\ldots,m$ is identified by:
  - Its nominal probability distribution $\mu_{w_k^i}$
  - Its total variation distance parameter $R_k^i \in [0,2]$
- In such a game, one player acts as the leader (L) and the remaining $m-1$ players act as the followers (F)

Definition

Let $\mathcal{M} = \{1,2,\ldots,m\}$ and $\mathcal{N} = \{L, F_1, F_2, \ldots, F_{m-1}\}$ be two finite sets with $|\mathcal{M}| = |\mathcal{N}|$. A classification function $\phi : \mathcal{M} \longmapsto \mathcal{N}$ is a bijective function from $\mathcal{M}$ into $\mathcal{N}$

MinMax Stochastic Control

m-Player Problem

Find an optimal classification function $\phi_k : \mathscr{M} \longmapsto \mathscr{N}$ for each k, and a maximizing p.d. $\nu_{w_k^t}$ for all players $t \in \mathscr{N}$, to solve

$$\max_{\substack{\phi_k : \mathscr{M} \longmapsto \mathscr{N} \\ \nu_{w_k^t} \in \mathbb{B}_{R_k^t}, \forall t \in \mathscr{N}}} \mathbb{E}_{\nu}^{g,x}\left[\ell_k(x_k, u_k, w_k^L, w_k^{F_1} \ldots, w_k^{F_{m-1}})\right]$$

Remark

- The classification of the m players is sequentially decided during the different time intervals
- For a fixed classification, the hierarchical model can be defined as an m-stage game model
- Each player finds its maximizing p.d. for fixed p.d. of its predecessors, and given the optimal p.d. of its successors
- The solution forms an equilibrium distribution $(\nu^{*,\phi}_{w_k^L}, \nu^{*,\phi}_{w_k^{F_1}}, \ldots, \nu^{*,\phi}_{w_k^{F_{m-1}}})$

MinMax Stochastic Control

Theorem

Let ϕ_k be a fixed classification function defined on $\mathscr{M}$ with range $\mathscr{N}$. The maximizing, time-varying, p.d. for each player $t \in \mathscr{N}$ is given by

$$\nu^*_{w_k^t}(\Sigma^0(k,t)) = \mu_{w_k^t}(\Sigma^0(k,t)) + \frac{\alpha_k^t}{2},$$

$$\nu^*_{w_k^t}(\Sigma_0(k,t)) = \left(\mu_{w_k^t}(\Sigma_0(k,t)) - \frac{\alpha_k^t}{2}\right)^+,$$

$$\nu^*_{w_k^t}(\Sigma_j(k,t)) = \left(\mu_{w_k^t}(\Sigma_j(k,t)) - \left(\frac{\alpha_k^t}{2} - \sum_{z=1}^{j} \sum_{i \in \Sigma_{z-1}(k,t)} \mu_{w_k^t}(\Sigma_i(k,t))\right)^+\right)^+, \; j = 1,\ldots,r,$$

$$\alpha_k^t = \min\left(R_k^t,\; 2\left(1 - \mu_{w_k^t}(\Sigma^0(k,t))\right)\right)$$

where $\Sigma^0(k,t)$, $\Sigma_j(k,t), j = 0,1,\ldots,r$, denote the identified partition with respect to $w_k^t \in \mathscr{W}_k^t$, $t \in \mathscr{N}$, and $(x)^+ \triangleq \max\{0,x\}$.

MinMax Stochastic Control

Outer Optimization

- Dynamic Programming equation becomes

$$V_k(x) = \min_{u_k \in \mathcal{U}_k(x)} \Big\{ x^T Q_k x + u_k^T R_k u_k + \max_{\nu_{w_k^i} \in \mathbb{B}_{R_k^i}, i=1,\ldots,m} \mathbb{E}_{\nu}^{g,x}\big[\ell_k(x_k, u_k, w_k^1, \ldots, w_k^m))\big] \Big\}$$
$$= \min_{u_k \in \mathcal{U}_k(x)} \Big\{ x^T Q_k x + u_k^T R_k u_k + \mathbb{E}_{\nu^*}^{g,x}\big[\ell_k(x_k, u_k, w_k^1, \ldots, w_k^m)\big] \Big\} \quad (5)$$

  - $\mathbb{E}_{\nu^*}^{g,x}[\,\cdot\,]$ denotes expectation with respect to the maximizing p.d. $\nu^*_{w_k^{\phi_k^{-1}(t)}}$ for all $t \in \mathcal{N}$
  - $\phi_k^{-1}(t)$ denotes the i-th player in $\mathcal{M}$

- Differentiating with respect to u_k, we obtain $u_k^* = -L_k x - S_k$ for $x_k = x$
  - Preserves its linearity similar to the classical case
  - Its evaluation is performed based on the multiple players maximizing p.d.

- The optimal cost for the minimax stochastic control problem is given by
$J^* = V_0(x_0) = x_0^T P_0 x_0 + x_0^T F_0 + r_0$

The Drop-Shipping Problem

THE DROP-SHIPPING PROBLEM

- PARAMETERS OF INTEREST:
  - x_k: stock available at the beginning of k-th period
  - u_k: stock ordered at the beginning of k-th period
  - w_k^i: demand of etailer $i = 1,\ldots,m$ during k-th period

    with given nominal p.d. $\mu_{w_k^i}$
  - h_k, c_k, p_k: holding, ordering, and shortage cost per unit item, respectively

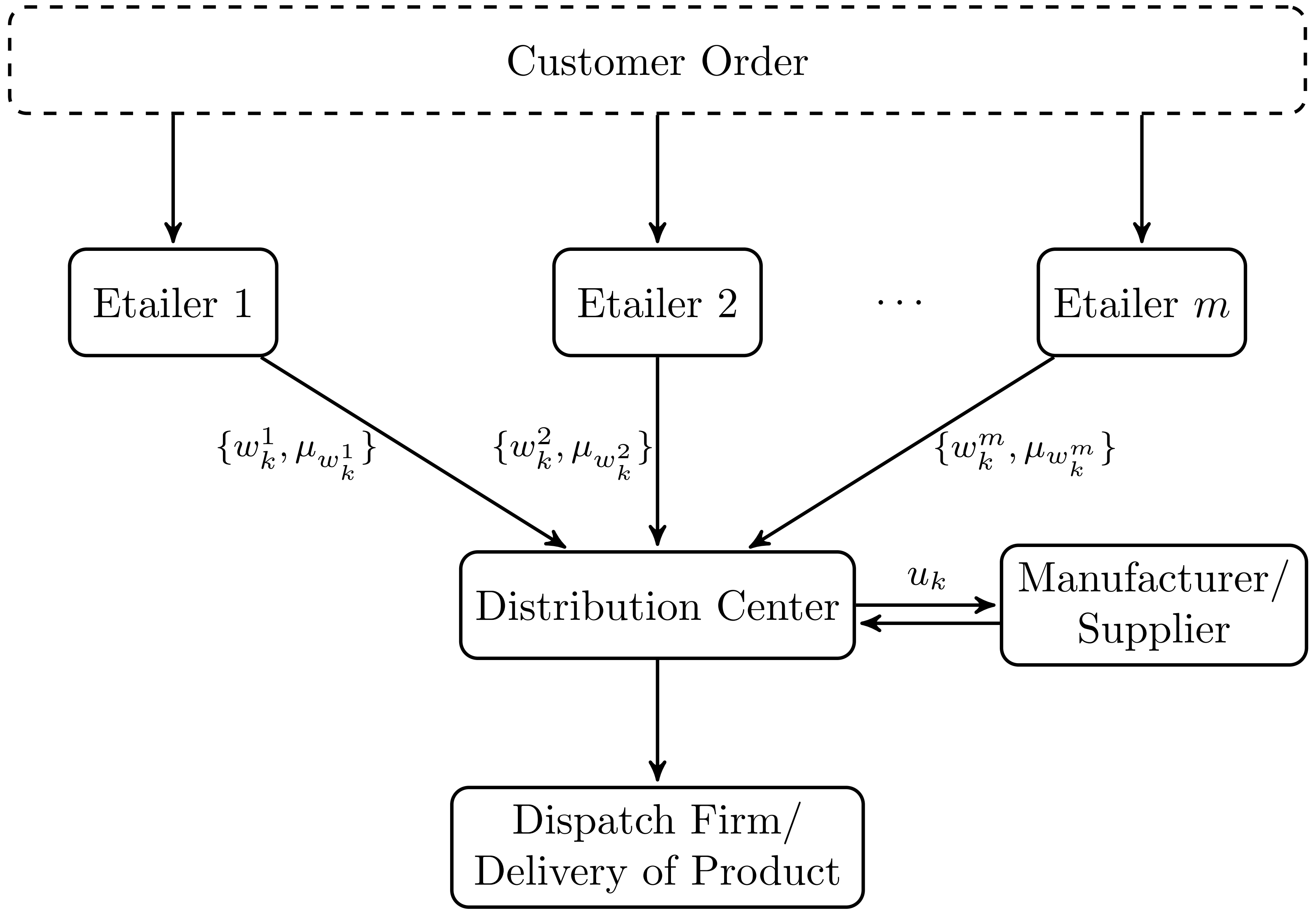


Fig.1: Drop-shipping model

- STATE DYNAMICS: $x_{k+1} = \max\left(0, x_k + u_k - \sum_{i=1}^{m} w_k^i\right) \quad (6)$

- PAYOFF: $\sum_{k=0}^{N-1} \left(c_k u_k + p_k\left(\min\left(0, x_k + u_k - \sum_{i=1}^{m} w_k^i\right)\right)^2 + h_k\left(\min\left(0, \sum_{i=1}^{m} w_k^i - x_k - u_k\right)\right)^2\right) \quad (7)$

The Drop-Shipping Problem

- MiniMax Stochastic Control:

$$\min_{u_k \in \mathcal{U}_k(x_k)} \max_{\substack{\nu_{w_k^i} \in \mathbb{B}_{R_k^i} \\ k=0,1,\ldots,N-1 \\ i=1,\ldots,m}} \mathbb{E}_{\nu}^{g,x}\Big[\sum_{k=0}^{N-1}\Big(c_k u_k + p_k\Big(\min\Big(0, x_k + u_k - \sum_{i=1}^{m} w_k^i\Big)\Big)^2 + h_k\Big(\min\Big(0, \sum_{i=1}^{m} w_k^i - x_k - u_k\Big)\Big)^2\Big)\Big] \qquad (8)$$

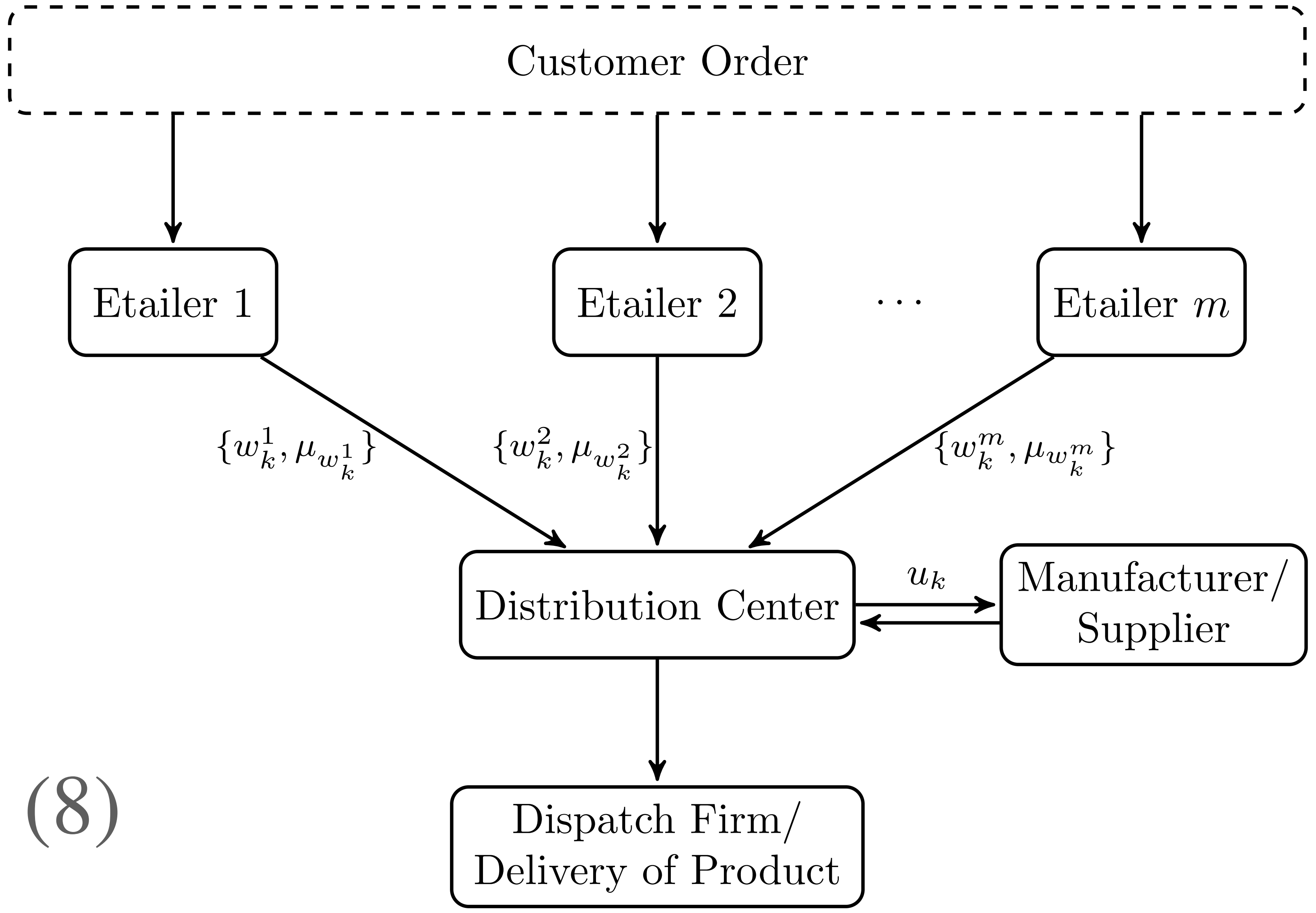
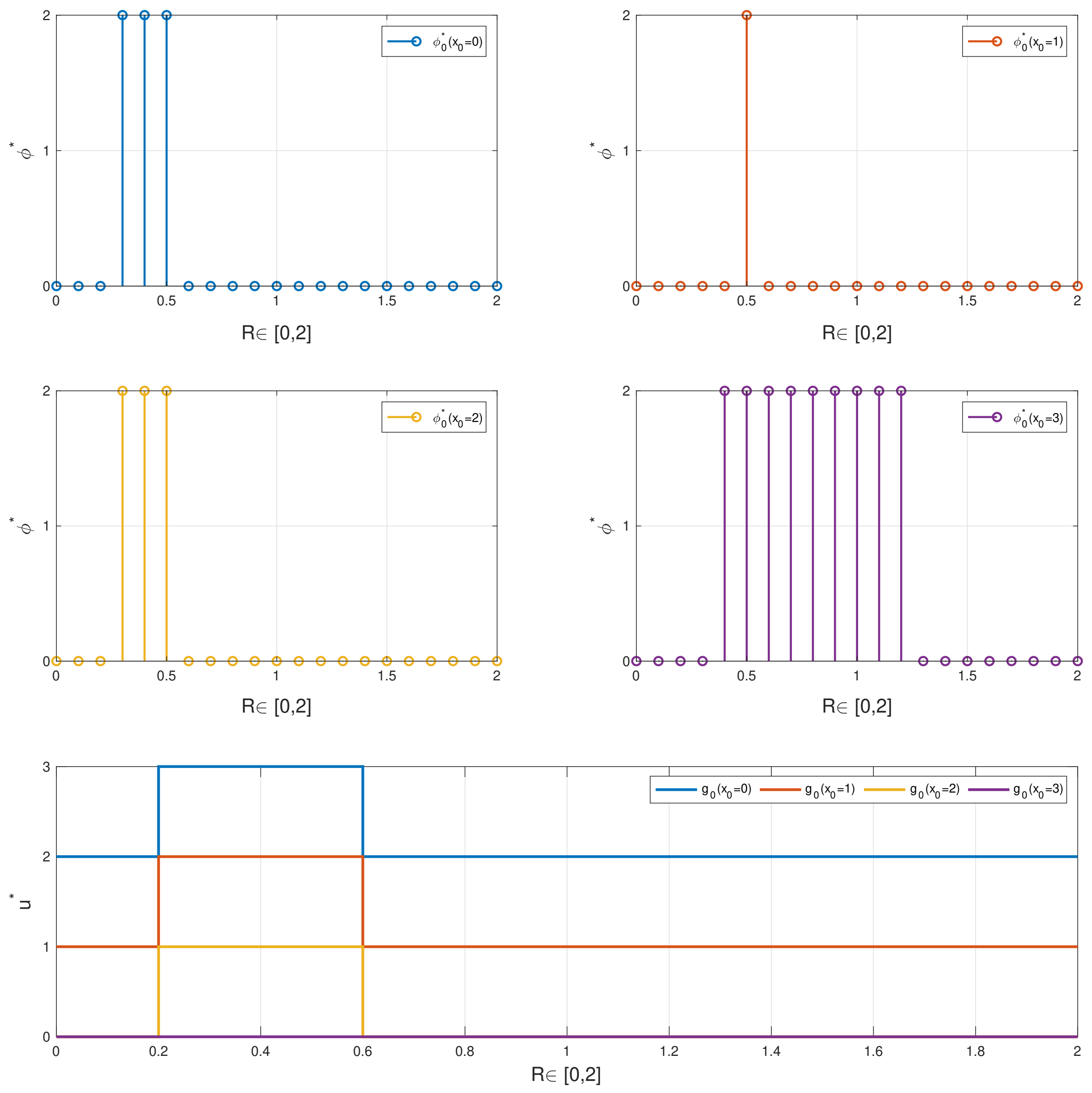


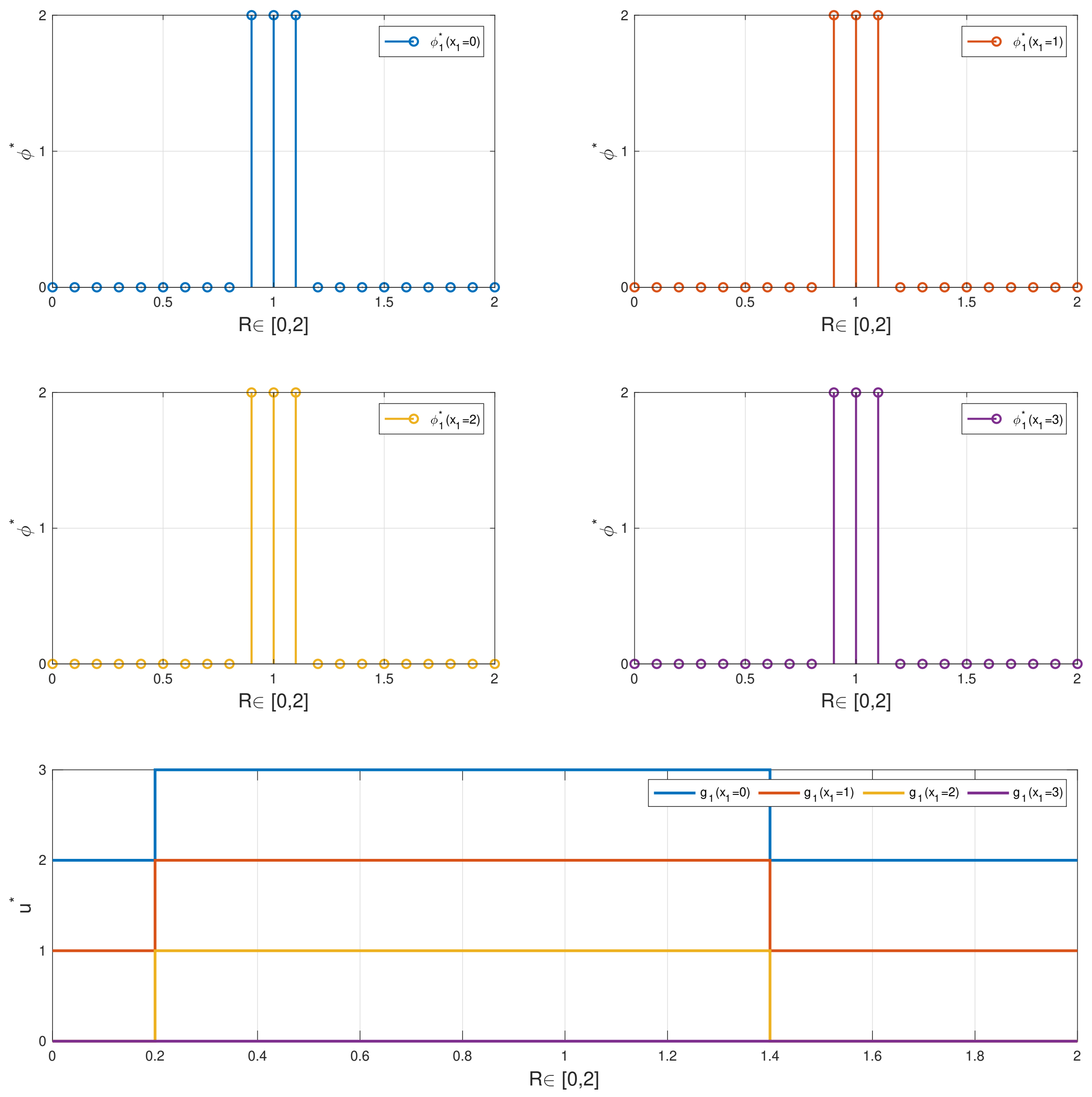
Fig.1: Drop-shipping model

- Input Data:
  - $N = 3$, $m = 2$, $\mathcal{U} = \{0,1,2,3\}$, and $x_k, u_k \in \mathbb{Z}_+$
  - Nominal p.d.: $\mu_{w_k^1} = [0.4\ \ 0.2\ \ 0.4]^T$ and $\mu_{w_k^2} = [0.1\ \ 0.1\ \ 0.8]^T$
  - Maximum capacity: $x_k + u_k \leq 3$, excess demand is lost, and $h_k = c_k = p_k = 1,\ \forall k$
  - Classification function $\phi^{[i]} : \mathcal{M} \longmapsto \mathcal{N}$, $i = 1,2$, $\mathcal{M} = \{1,2\}$, and $\mathcal{N} = \{L, F\}$

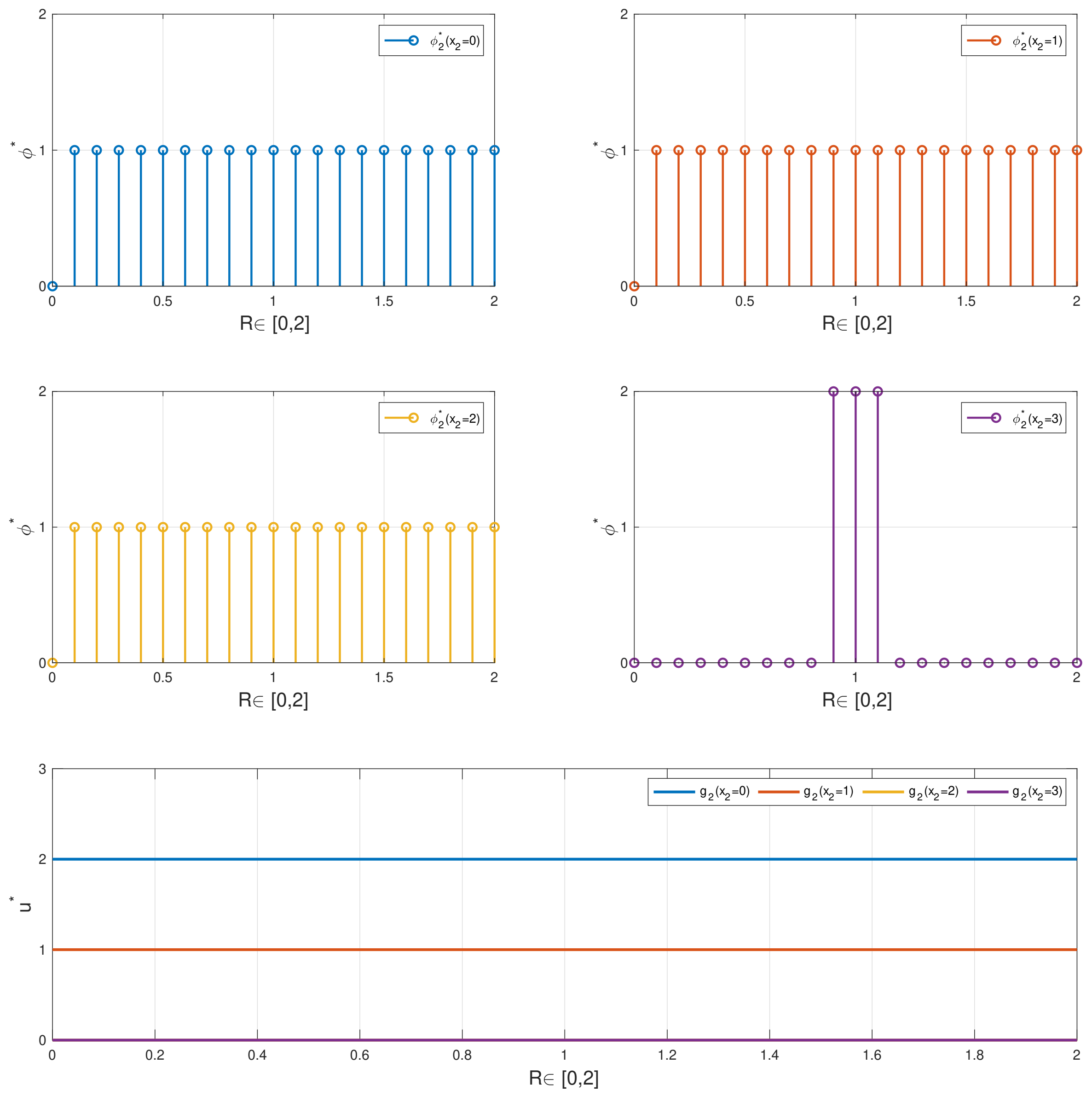
THE DROP-SHIPPING PROBLEM



$k = 0$



$k = 1$

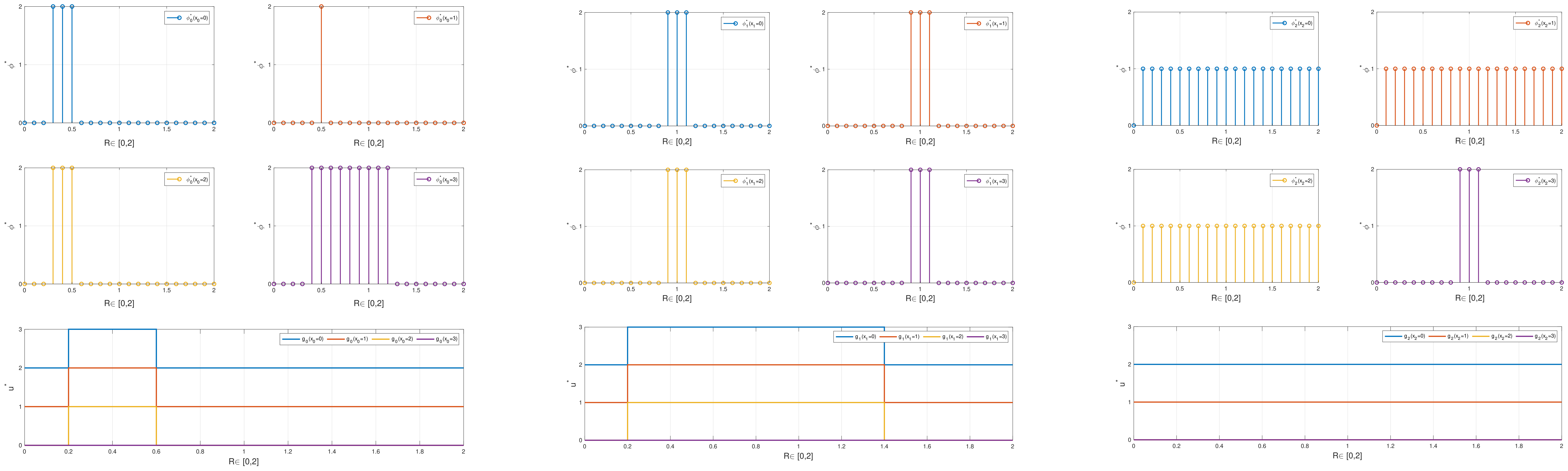


$k = 2$

Two cases are considered:

- For stage $k = 0$: $R_k^1 = R_k^2 > 0$: equal amount of confidence between the two players
- For stage $k = 1{,}2$: $R_k^1 > R_k^2 > 0$: player 1 less reliable compared to player 2

The Drop-Shipping Problem



Optimal Classification

$$\phi_k^*(x_k) = \begin{cases} 2, & \text{player 1 = L, player 2 = F} \\ 1, & \text{player 1 = F, player 2 = L} \\ 0, & \text{both classifications are optimal} \end{cases}$$

Concluding Remarks

Concluding Remarks

- A robust LQR method for discrete-time dynamical systems with multiple sources of uncertainty of ambiguous distribution information is proposed
- To study the effects of ambiguity a sequential hierarchical game with multiple uncertain players is studied
- A maximizing equilibrium policy characterising each player's optimal policy is derived
- A robust LQR feedback control policy is derived
- The proposed solution is illustrated through an application to the drop-shipping retail fulfilment model

Concluding Remarks

- A robust LQR method for discrete-time dynamical systems with multiple sources of uncertainty of ambiguous distribution information is proposed
- To study the effects of ambiguity a sequential hierarchical game with multiple uncertain players is studied
- A maximizing equilibrium policy characterising each player's optimal policy is derived
- A robust LQR feedback control policy is derived
- The proposed solution is illustrated through an application to the drop-shipping retail fulfilment model

Acknowledgement:
This work was partially funded by the European Regional Development Fund and the Republic of Cyprus through the Cyprus Research and Innovation Foundation
(Project: POST-DOC/0916/0139)

RESEARCH
& INNOVATION
FOUNDATION

Structural Funds
of the European Union in Cyprus