Project Deliverable

Project Number:	Project Acronym:	Project Title:
611281	DYMASOS	Dynamic Management of Physically Coupled Systems of Systems

Instrument:	Thematic Priority
Collaborative Project	ICT

Title

Preliminary report on modeling methods

Contractual Delivery Date:	Actual Delivery Date:
Month 9	Month 9

Start date of project:	Duration:
October, 1st 2013	36 months

Organization name of lead contractor for this deliverable:	Document version:
ETH	V4.0

Dissemination level (Project co-funded by the European Commission within the Seventh Framework Programme)		
PU	Public	X
PP	Restricted to other programme participants (including the Commission	
RE	Restricted to a group defined by the consortium (including the Commission)	
CO	Confidential, only for members of the consortium (including the Commission)	

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement No 611281

Authors (organizations) :
Basilio Gentile (ETH), Sergio Grammatico (ETH), John Lygeros (ETH).

Reviewers (organizations) :
Sebastian Engell (TUDO).

Abstract :

The Work Package 1 of the DYMASOS project is dedicated to the “population dynamics based approach to the management of systems of systems”, that is, modeling, system identification, state estimation, and control of large-scale systems of systems (SOS).
A centralized control scheme cannot be used to achieve the goals set by the DYMASOS project, mainly because the number of systems under consideration is too high for centralized computations and decisions, and some users may be unwilling to provide private information; consequently, a central coordinator should keep track of an anonymous state for the whole population of systems and steer such state via macroscopic incentives, such as pricing, penalties, bulk commands. The theoretical framework of mean field control, and of mean field games, satisfies these requirements: each individual system (or agent) seeks a dynamical evolution that minimizes a cost-function depending not only on its own behavior, but also on the average behavior of the overall population.
We present and extend the modeling framework of mean field control in order to make it applicable to the use cases of the DYMASOS project, such as the demand-side management for a population of electric loads and the charging control problem for a population of plug-in electric vehicles, in the presence of time-varying, user-defined, charging constraints. The modeling framework envisions dynamic pricing schemes that allow the population of systems to dynamically evolve towards safe and socially optimal operation conditions.

Keywords :

Large-scale populations of dynamical systems, mean field control, mean field games, constrained linear quadratic optimal control, dynamic pricing, demand-side management of flexible loads, thermostatically controlled loads, plug-in electric vehicles, smart power grid.

This project has received funding from the European Union’s Seventh Framework Programme for research, technological development and demonstration under grant agreement No 611281

Revision History

The following table describes the main changes done in the document since it was created.

Revision	Date	Description	Author (Organisation)
V1.0	May 2014	Creation	Basilio Gentile (ETH), Sergio Grammatico (ETH)
V2.0	June 2014	Contributions	Basilio Gentile (ETH), Sergio Grammatico (ETH), John Lygeros (ETH)
V3.0	June 2014	Review	Radoslav Paulen (TUDO), Vahab Rostampour (TUDO)
V4.0	June 2014	Final edit	Basilio Gentile (ETH), Sergio Grammatico (ETH), John Lygeros (ETH)

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement No 611281

The DYMASOS Project

The well-being of the citizens in Europe depends on the reliable and efficient functioning of large interconnected systems, such as electric power systems, air traffic control, railway systems, large industrial production plants, etc. Such large systems consist of many interacting components. The sub-systems are usually managed locally and independently, according to different policies and priorities. The dynamic interaction of the locally managed components gives rise to complex behaviour and can lead to large-scale disruptions as e.g. black-outs in the electric grid.

Large interconnected systems with autonomously acting sub-units are called systems of systems. DYMASOS addresses systems of systems where the elements of the overall system are coupled by flows of physical quantities, e.g. electric power, steam or hot water, etc.

Within the project, new methods for the distributed management of large physically connected systems with local management and global coordination will be developed.

The DYMASOS Consortium consists of:

Participant no.	Participant organisation name	Participant short name	Country
1	Technische Universität Dortmund	TUDO	Germany
2	BASF SE	BASF	Germany
3	HEP-Operator distribucijskog sustava d.o.o	HEP	Croatia
4	INEOS Köln GmbH	INEOS	Germany
5	University of Seville	USE	Spain
6	University of Zagreb - Faculty of Electrical Engineering and Computing	UNIZG-FER	Croatia
7	ETH Zürich	ETH	Switzerland
8	RWTH Aachen University	RWTH	Germany
9	inno TSD	inno	France
10	Optimizacion Orientada a la Sostenibilidad SL	IDENER	Spain
11	euTeXoo GmbH	TEX	Germany
12	Ayesa Advanced Technologies SA	Ayesa AT	Spain

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement No 611281

Acronyms and Definitions

Acronym	Defined as
[SOS]	Systems of Systems
[MFG]	Mean Field Games
[MFC]	Mean Field Control
[ODE]	Ordinary Differential Equation
[PEVs]	Plug-in Electric Vehicles
[TCLs]	Thermostatically Controlled Loads

This project has received funding from the European Union’s Seventh Framework Programme for research, technological development and demonstration under grant agreement No 611281

Preliminary report on modeling methods

Contents

1 Introduction

The Work Package 1 of the DYMASOS project [1] is dedicated to the "*population dynamics based approach to the management of systems of systems*", that is, modeling, system identification, state estimation, and control of large-scale systems of systems (SOS). In particular, Task 1.1 in WP1 aims to look into modeling techniques that describe SOS as large interconnected systems, hence consisting of interacting "agents"; some of the techniques investigated so far are described in this deliverable D1.1, whose objective is *to provide a preliminary report on modeling methods.*

A centralized control scheme cannot be used to achieve the goals set by DYMASOS, mainly because in general the number of systems under consideration is too high for centralized computations and decisions; in addition to that, in some applications (such as systems of thermostatically controlled loads) the users may not be willing to provide information about their daily habits to a central coordinator unit. The WP1 suggests to

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement No 611281

overcome these difficulties by *allowing the central coordinator (for example, the distribution system operator in power grids scenarios) to keep track of an anonymous state for the whole population of systems (for instance, in form of a probability distribution) and to steer this state by providing macroscopic incentives, such as pricing, penalties, bulk commands.*

A modeling framework that inherently satisfies these requirements is that of mean field control (MFC) [2, 3], also called mean field games (MFGs) [4, 5]. In this framework, each system (or agent) seeks a dynamical evolution that minimizes a cost-function depending not only on its own behavior, but also on the average behavior of the overall population. The number of agents is assumed to be very large and indeed the fundamental mathematical results on MFC are derived for an infinite number of agents. As envisioned by the description of Task 1.1, these results describe the evolution of the state distribution, given the systems parameters and cost functions.

MFC has been applied to different research areas, such as, among others, production planning in large-scale dynamic markets [6], syncronization of populations of dynamic oscillators in power networks [7], demand-side management for a population of electric loads [8, 9]. This last application is particularly relevant for the HEP case study of WP5 in DYMASOS, as we will explain in Section 3.2. Moreover, MFC has been recently used to model charging policies for a large population of plug-in electric vehicles [10]; Sections 2.1 and 3.1 will point out how such models relate to the AYESA case study within the DYMASOS project.

The theory of MFC has been developed recently, mainly from a pure mathematical point of view, leaving open the chance for engineering scientific research and technological transfer to industrial applications. For instance, the relationship between the way in which the central unit steers the average (population) state by intervening in the agents' cost functions, acting on pricing signals, and the resulting agents' dynamic evolution has not been fully investigated yet. Moreover, the dynamical systems of main interest in DYMASOS are coupled by "*flows of physical quantities (e.g., electric power, steam or hot water, materials in a chemical plant, gas, potable water)*", whereas MFC theory does not explicitly include yet neither local and population-level constraints, nor any information about the network structure.

The following preliminary report on modeling methods is organized as follows: Section 2 introduces the basic linear-quadratic mean field control modeling framework and discusses some of its extensions; Section 3 tailors such modeling techniques to two applications relevant for the DYMASOS use cases, namely systems of plug-in electric vehicles and of thermostatically controlled loads; motivated by the goal of filling part of the gap between the proposed mean field game model and a real system of plug-in electric vehicles, we propose in Section 4 some further generalizations of the mean field control model by

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement No 611281

introducing time-dependent and vehicle-specific constraints on the variables of the model.

2 Preliminary modeling of large populations of dynamical systems

Mean field games were independently proposed by two groups of researchers; the general nonlinear theory has been introduced by Prof. Jean-Michel Lasry (Université Paris Dauphine) and Prof. Pierre-Louis Lions (Collège de France, Paris); the linear-quadratic framework has been instead proposed by Prof. Minyi Huang (Carleton University, Ottawa), Prof. Peter Caines (McGill University, Montreal) and Prof. Roland Malhamé (McGill University, Montreal). Here we concentrate more on the latter approach, which, as discussed below, seems more relevant for the case study that involves plug-in electric vehicles. We touch upon the former approach in Section 3.2, where we discuss mean field control of populations of thermostatically controlled loads, a topic of potential interest for the DYMASOS case study on electrical energy distribution.

2.1 Linear quadratic mean field control

In the seminal works [2, 3], the following dynamical systems, in which the agent dynamics are linear (and scalar for ease of presentation), are considered:

$$\dot{x}_i(t) = a_i x_i(t) + b\, u_i(t)\,, \qquad t \geq 0\,, \tag{1}$$

where $i \in \{1, \ldots, N\}$ indexes the ith agent, x_i and u_i are state and input of agent i, and N is the number of agents. Each agent seeks a control law $u_i(\cdot)$ that minimizes its cost function

$$J_i(u_i(\cdot), v(\cdot)) := \int_0^\infty e^{-\rho t} \left((x_i(t) - v(t))^2 + r u_i^2(t) \right) dt\,, \tag{2}$$

where $v(t) := \gamma(\bar{x}(t) + \eta)$ is an affine function of the average state evolution $\bar{x}(t) := \frac{1}{N}\sum_{i=1}^N x_i(t)$, and the discount factor ρ and the parameters r, γ, η are positive real numbers.

Minimizing the cost (2) consists in finding a trade-off between tracking an affine function of the average state and reducing the magnitude of the control input. For example, the dynamics in (1) and the cost (2) can be used as a simple model for the production planning in a firm: the state x_i represents the production level, while u_i is the action of increasing or decreasing the production level [3, Section II.A]; if we assume the price of the product to be affine in the average production level $\bar{x}$ and the desired production level v to be proportional to the price, then v is also an affine function of $\bar{x}$. This reasoning, along with the penalization of the inherently costly actions u_i, can be indeed formalized into the cost (2).

We will now sketch the derivations that lead to the main results on the system (1)-(2). The problem of minimizing the cost (2) under the dynamics specified by equation (1) is difficult to solve, because the cost function of agent i also depends on the state evolution of all the other agents. If instead $v(\cdot)$ is replaced by a given function of time $x^*(\cdot)$, then each agent has to solve the following standard optimal tracking control problem:

$$\min_{u_i(\cdot)} J_i(u_i(\cdot), x^*(\cdot)) = \min_{u_i(\cdot)} \int_0^\infty e^{-\rho t} \left((x_i(t) - x^*(t))^2 + r u_i^2(t) \right) dt. \tag{3}$$

It is possible to derive explicitly the control law that minimizes the cost (3) under the dynamics (1); its expression is [11, Chapter 11], [12, Chapter 4]

$$\hat{u}_i(t) = -\frac{b}{r}(\Pi_i x_i(t) + s_i(t)) \tag{4}$$

where $\Pi_i \succ 0$ is the solution to the algebraic Riccati equation

$$\rho \Pi_i = 2 a_i \Pi_i - \frac{b^2}{r} \Pi_i^2 + 1\,,$$

and the costate $s_i(\cdot)$ is the solution to the ODE

$$\dot{s}_i(t) = \left(-a_i + \frac{b^2}{r} \Pi_i + \rho \right) s_i(t) + x^*(t)\,.$$

The quantity $-a_i + \frac{b^2}{r}$ is assumed to be positive. If now we require the signal to be tracked x^* to be exactly the affine function of the average $\gamma(\bar{x} + \eta)$ introduced before, we obtain the system of equations [3, Equations (4.6)–(4.9)]

$$\dot{s}_i(t) = \left(-a_i + \frac{b^2}{r} \Pi_i + \rho \right) s_i(t) + x^*(t) \tag{5a}$$

$$\dot{x}_i(t) = \left(a_i - \frac{b^2}{r} \Pi_i \right) x_i(t) - \frac{b^2}{r} s_i(t) \tag{5b}$$

$$\bar{x}(t) = \frac{1}{N} \sum_{i=1}^{N} x_i(t) \tag{5c}$$

$$x^*(t) = \gamma(\bar{x}(t) + \eta)\,, \tag{5d}$$

where Equation (5b) results from the substitution of the optimal tracking control (4) into the system dynamics (1).

Given an arbitrary continuous and bounded function $x^{*(0)}(\cdot)$, one can compute the corresponding $s_i^{(1)}(\cdot)$ from (5a), $x_i^{(1)}(\cdot)$ from (5b), $\bar{x}^{(1)}(\cdot)$ from (5c) and $x^{*(1)}(\cdot) = \mathcal{T} x^{*(0)}(\cdot)$ from (5d). Solving the system of equations (5) means finding an $\hat{x}$ such that $\mathcal{T}\hat{x} = \hat{x}$. We

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement No 611281

denote such $\hat{x}$ as *mean field equilibrium*. It is worth noticing that such computation is almost decentralized, meaning that each $x_i^{(k)}(\cdot)$ is computed under the only knowledge of local variables and just one global variable $x^{*(k)}(\cdot)$, which represents (an affine function of) the mean population state.

Since there are no constraints and the equations (5) are linear, it is possible to compute a closed form expression for $\mathcal{T}x^{*(0)}(\cdot)$ as an expression of $x^{*(0)}(\cdot)$, and show that $\mathcal{T}x^{*(0)}(\cdot)$ is continuous and bounded. Under technical assumptions on the parameters a_i, b, ρ, r, γ, it can be proven that the operator $\mathcal{T}$ is contractive. It follows that there exists a unique continuous and bounded solution $\hat{x}$ to the system (5), i.e., a mean field equilibrium. Moreover, this solution can be computed by means of an algorithmic procedure: starting from an arbitrary continuous and bounded $x^{*(0)(\cdot)}$, define $x^{*(k+1)}(\cdot) = \mathcal{T}x^{*(k)}(\cdot)$; it indeed holds that $\lim_{k\to\infty} x^{*(k)}(\cdot) = \hat{x}(\cdot)$.

We point out that the algorithmic procedure just introduced is carried out offline, before the time instant $t = 0$. Once the agents agree on the function $\hat{x}(\cdot)$, then in the interval $[0, \infty)$ each of them implements the optimal tracking of such $\hat{x}(\cdot)$.

It is interesting to relate the solution $\hat{x}(\cdot)$ of the system (5) to the initial problem represented by equations (1) and (2). To this end, let us define a set of controls $\{u_1(\cdot), \ldots, u_N(\cdot)\}$ as an ε-*Nash equilibrium*, $\varepsilon > 0$, if for any fixed $1 \le i \le N$ we have

$$J_i(u_i, v(u_1, \ldots, u_i, \ldots, u_N)) \le J_i(u_i', v(u_1, \ldots, u_i', \ldots, u_N)) + \varepsilon \tag{6}$$

for any alternative set of control strategies $\{u_1(\cdot)', \ldots, u_N(\cdot)'\}$. In words, in the configuration given by $\{u_1(\cdot), \ldots, u_N(\cdot)\}$, any agent can improve its cost function by at most ε when unilaterally deviating from its control strategy. When $\varepsilon = 0$ the definition of ε-Nash equilibrium reduces to that of Nash equilibrium. It is now possible to state the second important result of the work in [3]: when the signal to track is $\hat{x}(\cdot)$, the corresponding set of optimal inputs $\{\hat{u}_1(\cdot), \ldots, \hat{u}_N(\cdot)\}$ is an ε-Nash equilibrium, with ε being a decreasing function of the population size N.

2.2 Extensions of linear quadratic mean field control

In Section 2.1 we introduced the simplest formulation of linear-quadratic mean field games. We will now list and explain some of its extensions present in the literature.

The work [3] proposes a stochastic formulation of equations (1) and (2)

$$dx_i = (a_i x_i + b u_i)dt + \sigma_i dw_i\,, \qquad t \ge 0\,,$$
$$J_i(u_i(\cdot), v(\cdot)) := \mathbb{E}\int_0^\infty e^{-\rho t}\left[(x_i(t) - v(t))^2 + r u_i^2(t)\right] dt\,,$$

where $\{w_i\}_{i=1}^N$ are independent standard scalar Wiener processes and $\mathbb{E}$ denotes the expectation operator, with respect to w_i. Due to the presence of the stochastic term w_i, the

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement No 611281

state x_i is a random variable itself. Nonetheless, the results on existence and uniqueness of an ε-Nash equilibrium that have been presented in Section 2.1 still hold if in equation (5b) we replace x_i with its expected value $\mathbb{E}\, x_i$.

Moreover, work [3] assumes that the sequence a_i is drawn from an empirical distribution function $F(a)$, with $a \in \mathcal{A}$, the range space. This implies that we can rewrite equations (1)-(5b) with a, x_a, u_a, Π_a, s_a instead of a_i, x_i, u_i, Π_i, s_i. The computation of the average state evolution in equation (5c) now becomes

$$\bar{x}(t) := \int_{\mathcal{A}} x_a(t) dF(a)\,. \tag{7}$$

Considering the system (5) with equation (5c) replaced by (7), then under some regularity assumptions on the empirical distribution $F(a)$, the results on existence and uniqueness of an ε-Nash equilibrium are still valid.

Another variant proposed in [2] is the introduction of an extra term in the dynamics of each system i. Namely, the average population state $\bar{x}(t) = \frac{1}{N}\sum_{i=1}^{N} x_i(t)$ enters linearly in the state evolution of each agent:

$$\dot{x}_i(t) = a_i x_i(t) + b\, u_i(t) + \alpha \bar{x}(t)\,, \qquad t \geq 0\,.$$

This generalization only results in the term $\bar{x}(t)$ appearing in the costate equation (5a) as follows [2, Proposition 9.1]:

$$\dot{s}_i(t) = \left(-a_i + \frac{b^2}{r} + \rho\right) s_i(t) + x^*(t) - \alpha \Pi_i \bar{x}(t)\,.$$

The results on existence and uniqueness of the ε-Nash equilibrium still hold. Likewise, the book chapter [2] shows that these results still hold if the affine function $\gamma(\cdot + \eta)$ appearing in equation (5d) is substituted with a general Lipschitz function.

The work in [13] starts from the model in [2], but the population of agents is divided into two classes, leaders and followers. By minimizing its cost function, each leader tries to track a convex combination of the leaders' average and of a fixed reference function $h(\cdot)$. The followers have no information about the reference $h(\cdot)$, which is only known by the leaders. By minimizing its cost function, each follower seeks to track a convex combination of the followers' average and of the leaders' average. Each agent (being it a leader or a follower) evolves dynamically according to (1); the cost function of a leader differs from the cost function of a follower and their expressions are respectively

$$J_i^L(u_i^L(\cdot), \bar{x}^L(\cdot)) = \limsup_{T\to\infty} \frac{1}{T} \int_0^T \Big(x_i^L(t) - \big(\lambda h(t) + (1-\lambda)\bar{x}^L(t)\big)\Big)^2 + r\Big(u_i^L(t)\Big)^2 dt \tag{8a}$$

$$J_i^F(u_i^F(\cdot), \bar{x}^L(\cdot)) = \limsup_{T\to\infty} \frac{1}{T} \int_0^T \Big(x_i^F(t) - \big(\mu \bar{x}^F(t) + (1-\mu)\bar{x}^L(t)\big)\Big)^2 + r\Big(u_i^F(t)\Big)^2 dt, \tag{8b}$$

where $\bar{x}^L$ and $\bar{x}^F$ are the leaders' and followers' average state evolutions. By neglecting the discount factor $e^{-\rho t}$ and presenting instead "$\limsup_{T\to\infty} \frac{1}{T}(\cdots)$", the costs (8) weight only the steady-state behavior of the systems.

Generalizing the results presented in Section 2.1, in the work [13] it is proven that, if the convex combination coefficients λ and μ are respectively close enough to 0 and 1, then there exists an ε_1-Nash equilibrium for the leaders and an ε_2-Nash equilibrium for the followers. The Nash-equilibria can be found as the limit point of an iterative algorithm based on two systems (one for the leaders and one for the followers) which are both similar to the system (5).

The leaders-followers formulation can be used as a basic model for the stock-market [14]: each investor would like to guess other investors' guesses and the investors are divided into well-informed ones (as institutional investors) and followers (as retail investors). Then the costs (8a) and (8b) represent the goal of the two classes: given $\lambda \simeq 1$, then for the institutional investors it is important to track $h(\cdot)$, which represents the piece of information they possess about the stock-market trend; given $\mu \simeq 0$, then for the followers it is important to track the institutional investors' average behavior $\bar{x}^L$.

Another possible application of the leaders-followers model is in multi-vehicle coordination control [15], in a situation in which the goal of all the vehicles is to track a reference trajectory, which is available only to some of them (the leaders). The remaining vehicles (the followers) can then track the leaders' trajectories in order to reach their goal. Other similar applications of the leaders-followers model are in flocking [16], formation control [17], dynamic industry models [18] and social opinion models with a very large number of leaders [19].

Linear-quadratic MFC methodologies have been recently adapted and applied to the context of *coalition games* [20], therefore we foresee some potential impact in the WP2 of the DYMASOS project.

3 Mean field control for DYMASOS use cases

We now outline how MFC approaches can be applied to the use cases of the DYMASOS project, that are described in deliverable D5.1 [21]. In particular, we focus on the constrained charging problem for a population of plug-in electric vehicles [21, Chapter 2] and on the demand-side management problem for a population of thermostatically controlled loads [21, Chapter 5].

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement No 611281

3.1 Mean field control for plug-in electric vehicles

In [10], Ma, Callaway and Hiskens use MFC techniques to model the problem of charging a large population of cost-coupled plug-in electric vehicles (PEVs). The population is composed by N electric vehicles; each of them has a battery whose state of charge at time t is described by the variable $x_i(t)$ with values in $[0,1]$: 0 corresponds to an empty battery, and 1 to a fully charged battery. The time evolution of battery i is specified by the discrete-time system

$$x_i(t+1) = x_i(t) + b_i u_i(t)\,, \qquad t = 0, \dots, T-1\,, \tag{9}$$

where $u_i(t)$ is the charging control for vehicle i at time t and b_i is the charging efficiency. Unlike [3, 22, 23], the charging control cannot assume negative values and its magnitudes over the time interval $[0, T-1]$ have to guarantee a full charge at the end of the interval. Denoting $\mathbf{u}_i = \{u_i(0), \dots, u_i(T-1)\}$, this can be expressed as

$$\mathbf{u}_i \in \mathcal{U}_i := \left\{ \mathbf{u}_i \in \mathbb{R}^T \mid \; u_i(t) \geq 0,\; \sum_{t=0}^{T-1} u_i(t) = b_i^{-1}(1 - x_i(0)) \right\}, \tag{10}$$

where $x_i(0)$ is the initial state of charge of vehicle i.
Each vehicle seeks a control sequence $\mathbf{u}_i$ that minimizes its cost function

$$J_i(\mathbf{u}_i) := \sum_{t=0}^{T-1} \left\{ p\left(\frac{d_t + \bar{u}(t)}{c}\right) u_i(t) + \delta\left(u_i(t) - \bar{u}(t)\right)^2 \right\}. \tag{11}$$

Let us explain the terms appearing in the cost (11): the quantity $\bar{u}(t)$ is the average PEVs electricity demand at time instant t, i.e., the total PEVs electricity demand divided by the number N of PEVs; in the same way, the quantity d_t is the total non-PEVs electricity demand at time instant t, divided by the number N of PEVs. The quantity $p(\cdot)$ is the price of the electricity, which is a function of the ratio between the average electricity demand $d_t + \bar{u}(t)$ and the average network capacity c. Hence the first term of the sum in equation (11) can be thought of as the electricity bill of the vehicle i over the entire time interval. The second term is instead a penalty cost due to the deviation of charging control i from the average charging control, weighted by the positive real parameter δ. This second term does not present a physical interpretation, but it guarantees the validity of the theoretical results derived by [10]. Indeed these results hold for small values of δ, hence the contribution of the second term in the cost (11) is small. The cost function in (11) is similar to the one considered in PEVs case study [21, Section 2.10.3], where the main difference is represented by the term $\delta(u_i(t) - \bar{u}(t))^2$, which however can be made arbitrarily small by selecting $\delta > 0$ small enough.

As it was the case in the previous section, minimizing the cost (11), subject to the dynamics (9) and the constraints (10), is a difficult problem, because the cost function (11)

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement No 611281

of vehicle i depends on the control strategies of all the other vehicles. The procedure to overcome this difficulty assumes that the population is composed by an infinite number of vehicles and closely resembles the scheme proposed by in [3]; this procedure can be interpreted as a "*dynamic pricing scheme for selling energy*", as also proposed in the DYMASOS case study [21, Section 2.3.2].

First, the average PEVs electricity demand $\bar{\mathbf{u}}$ is replaced by a fixed function $\mathbf{z}$, so each agent can solve the minimization problem separately and determine its optimal charging strategy in response to $\mathbf{z}$; then, it is imposed that the average of the optimal charging strategies in response to $\mathbf{z}$ must be $\mathbf{z}$ itself, that is:

$$\mathbf{u}_i^\star(\mathbf{z}) := \arg\min_{\mathbf{u}_i \in \mathcal{U}_i} \sum_{t=0}^{T-1} \left(p\left(\frac{1}{c}(d_t + z_t)\right) u_i(t) + \delta(u_i(t) - z_t)^2 \right) \tag{12a}$$

$$\mathbf{z} = \bar{\mathbf{u}}^\star \, . \tag{12b}$$

Given a function $\mathbf{z}$, [10, Lemma 4.1] characterizes the unique solution of the optimization problem (12a), namely

$$u_i^\star(t) = \frac{1}{2\delta} \max\left\{0, A(\mathbf{z}) - p\left(\frac{1}{c}(d_t + z_t)\right) + 2\delta z_t\right\} , \tag{13}$$

where $A(\mathbf{z})$ is a real number dependent on the function $\mathbf{z}$, which is not known in closed form.

The system (12) can be interpreted in the following way: given a function $\mathbf{z}^{(0)}$, each agent computes its optimal charging strategy $\mathbf{u}_i^{\star(1)}$ by means of (13); then $\bar{\mathbf{u}}^{\star(1)} = \Gamma \mathbf{z}^{(0)}$ is computed by averaging over all the $\mathbf{u}_i^{\star(1)}$. A solution to the system (12) is a function $\hat{\bar{\mathbf{u}}}$ such that $\Gamma\hat{\bar{\mathbf{u}}} = \hat{\bar{\mathbf{u}}}$. If the price function $p(\cdot)$ is continuously differentiable and strictly increasing, then under technical assumptions on the value of the parameter δ, [10, Theorem 4.2] proves that the system (12) possesses a unique solution $\hat{\bar{\mathbf{u}}}$. Moreover, this solution can be computed by means of an algorithmic procedure: starting from an arbitrary function $\bar{\mathbf{u}}^{(0)}$, define $\bar{\mathbf{u}}^{(k+1)} = \Gamma\bar{\mathbf{u}}^{(k)}$; it holds that $\lim_{k\to\infty} \bar{\mathbf{u}}^{(k)} = \hat{\bar{\mathbf{u}}}$. We point out that the algorithmic procedure just introduced is carried out offline, before the time instant $t = 0$. Once the agents agree on the function $\hat{\bar{\mathbf{u}}}$, then in the interval $[0, T]$ each of them implements the optimal charging strategy in response to the PEV demand $\hat{\bar{\mathbf{u}}}$.

It is now interesting to relate the original problem given by equations (9), (10) and (11) to the uniquely defined function $\hat{\bar{\mathbf{u}}}$. Let us define a Nash equilibrium exactly as we did in (6) (but with $\varepsilon = 0$); if the price function $p(\cdot)$ is continuously differentiable and strictly increasing, then under some technical assumption on the value of the parameter δ, [10, Theorem 4.2] also guarantees that the set of optimal charging control strategies in response to the PEV demand $\hat{\bar{\mathbf{u}}}$ is a Nash equilibrium.

We finally remark that in the PEVs case study of the DYMASOS project it is of interest to model the spatial evolution of the agents behavior and in turn of the associated price function [21, Section 2.9]. The MFC framework actually allows for local interactions between systems according to the approach proposed in [22], which can be further adapted to the mean field control of large populations of PEVs.

3.2 Mean field games for thermostatically controlled loads

So far we investigated different mean field game models, but all of them share a common framework of linear dynamics and quadratic cost. A theory that generalizes this to nonlinear dynamics and cost was developed by Lasry and Lions [4], [24]. Here we will concentrate on the work [9] by Bagagiolo and Bauso, which introduces a model with linear switched dynamics and nonlinear cost for a system of thermostatically controlled loads. This modeling approach is of particular interest for the HEP case study in the DYMASOS project, because it concerns the *demand-side management* of *flexible loads* [21, Sections 3.1.1, 3.3.1].

In most cases the capability for electric heating or cooling appliances of storing thermal energy is greater than the capability of a battery of storing chemical energy; hence the interest in modeling of systems composed by a large number of such thermostatically controlled loads. If we define $x(t)$ to be the temperature of a single cooling appliance at time t, then the linear switched dynamic model proposed by [9] reads as

$$\dot{x}(t) = \begin{cases} -\alpha\,(x(t) - x_{\text{on}}) & \text{if } \; u(t) = 1 \\ -\alpha\,(x(t) - x_{\text{off}}) & \text{if } \; u(t) = 0\,, \end{cases} \quad 0 \le t \le T\,, \tag{14}$$

where $x(0) = x_0$ and the rate $\alpha > 0$ is a given real number; when the control signal is $u(t) = 1$ (i.e. when the appliance is ON), then the temperature decreases exponentially towards the steady-state x_{on}; when instead the control signal is $u(t) = 0$ (i.e. when the appliance is OFF), then the temperature increases exponentially towards the steady-state x_{off}. Let us define $\bar{m}(t)$ as the mean temperature over all the appliances at time t and the instantaneous cost g as

$$\begin{aligned} g(x, u, \bar{m}) := u\,(W_{\text{on}} + h[\bar{m} - m_{\text{ref}}]_+) + \\ (1-u)\,(W_{\text{off}} + k[\bar{m} - m_{\text{ref}}]_-) + q(x - x_{\text{ref}})^2\,. \end{aligned} \tag{15}$$

Then the cost function for each appliance is

$$J(u(\cdot)) = \int_0^T g\,(x(t), u(t), \bar{m}(t))\,dt\,. \tag{16}$$

Let us explain the terms appearing in (15). The quantity W_{on} is the power consumed when the appliance is ON, while W_{off} is the power consumed when the appliance is OFF. The scalar $[y]_+$ is equal to zero when y is negative and is equal to y itself when y is positive; in the cost, $u \cdot h[m - m_{\text{ref}}]_+$ penalizes an appliance which is ON when the mean temperature is above the reference temperature m_{ref}. This has a precise meaning which is related to the frequency of the system: if we assume that, given a reference frequency ω_{ref}, we have $(\omega(t) - \omega_{\text{ref}}) \sim (m_{\text{ref}} - m(t))$, then the term $h[m - m_{\text{ref}}]_+$ penalizes the appliances that contribute to the deviation of the system frequency $\omega(t)$ from the reference frequency ω_{ref}. The same idea underlies the term $u \cdot k[m - m_{\text{ref}}]_-$. Finally, the last term in the cost (15) penalizes the deviation of the temperature from the reference x_{ref}.

In order to state its main results, in [9] the domain of u is convexified in the dynamics (14) and in (15); namely, the input, instead of being either 0 or 1, can assume all the values in the interval $[0, 1]$ and it is now interpreted as the probability of each appliance to be ON. Moreover, it is introduced the important concept of temperature distribution by making use of the function $m : [x_{\text{on}}, x_{\text{off}}] \times [0, +\infty[\to [0, +\infty[, (x, t) \mapsto m(x, t)$. If we have a continuum of appliances, for each time instant t, the function $m(\cdot, t)$ is the probability density function that describes the distribution of the temperatures of all the appliances.

In Sections 2.1 and 3.1, we introduced the concept of mean field equilibrium for the systems (1)-(2) and (9)-(11). We also presented an algorithmic procedure that guarantees the convergence to such mean field equilibrium under some technical assumptions on the problem data. The same ideas apply to the system of thermostatically controlled loads (14)-(16), in a slightly different way. Given a fixed temperature distribution $m(x, t)$, then every appliance computes its optimal control strategy in response to it, by means of a partial differential equation named Hamilton-Jacobi-Bellman equation, which is an analogous of (5a)-(5b) or (13). Then, given the optimal controllers just computed, a new expression for the temperature distribution $m(x, t)$ is derived, by means of a second partial differential equation, named Fokker-Planck-Kolmogorov equation. For the system of coupled partial differential equations, it is actually possible to introduce a concept of mean field equilibrium analogous to the one presented in Section 2.1.

Having introduced the concept of mean field equilibrium for the system (14)-(16), we can state the two main contributions given by [9]; first, it is proven that the system (14)-(16) possesses a mean field equilibrium. Second, it is proven that the mean field equilibrium is stable, in the sense that the states of all agents, initially at different values, will eventually converge to the equilibrium.

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement No 611281

4 Generalized modeling framework for large populations of dynamical systems in presence of constraints

In this section we present a generalized framework to model the behavior of a large population of coupled dynamical systems, which takes into account the presence of heterogeneous state and input constraints for each agent of the population. We then apply such modeling framework to the DYMASOS case study in WP5, whose preliminary description is in [21, Chapter 2], concerning decentralized charging algorithms of PEVs in presence of heterogeneous user-defined charging constraints.

4.1 Large populations of constrained linear dynamical systems

In order to guarantee numerical tractability of the optimal control problem, we consider here discrete-time linear dynamics over a finite-time horizon, subject to convex constraints as follows:

$$\begin{aligned} &x_i(t+1) = A_i x_i(t) + B_i u_i(t)\,, && t = 0,1,\dots,T-1\,, \\ &\text{s.t.} \quad x_i(t) \in \mathcal{X}_t, \quad u_i(t) \in \mathcal{U}_t, && \forall t \in \{0,1,\dots,T-1\} \end{aligned}$$

where $i \in \{1,\dots,N\}$ indexes the ith agent, $x_i \in \mathbb{R}^n$ and $u_i \in \mathbb{R}^m$ are state and input of agent i, and N is the number of agents; $\mathcal{X}_t \subseteq \mathbb{R}^n$ and $\mathcal{U}_t \subseteq \mathbb{R}^m$ are time-varying state and input convex constraints sets. Let us consider a reference sequence $\{z_t\}_{t=0}^{T-1}$, which we interpret as a given function $z_t = \Phi(\bar{x}(t))$ of the average state $\bar{x}(t) = \sum_{i=1}^{N} x_i(t)$. Each agent seeks a control sequence $\mathbf{u}_i = \{u_i(t)\}_{t=0}^{T-1}$ and hence an admissible state sequence $\mathbf{x}_i = \{x_i(t)\}_{t=1}^{T}$ solving the *constrained* optimal control problem:

$$\min_{\mathbf{x}_i,\mathbf{u}_i} \sum_{t=0}^{T-1} \|x_i(t) - z_t\|_{Q_t}^2 + \|u_i(t)\|_{R_t}^2 + 2\left(P_t z_t + p_t\right)^\top u_i(t) \tag{18a}$$

$$\text{s.t.} \quad x_i(t) \in \mathcal{X}_t, \quad u_i(t) \in \mathcal{U}_t, \qquad \forall t \in \{0,1,\dots,T-1\} \tag{18b}$$

$$x_i(t+1) = A_i x_i(t) + B_i u_i(t) \quad \forall t \in \{0,1,\dots,T-1\}. \tag{18c}$$

The cost term $\|x_i(t) - z_t\|_{Q_t}^2$, with $Q_t \succcurlyeq 0$ for all t, represents the penalty on the tracking error of the reference signal $\{z_t\}$; $\|u_i(t)\|_{R_t}^2$, with $R_t \succcurlyeq 0$ for all t, penalizes the input magnitude; the term $2\left(P_t z_t + p_t\right)^\top$ (with $P_t \in \mathbb{R}^{m\times n}$, $p_t \in \mathbb{R}^m$ for all t), can be interpreted as time-varying price. This is a generalization of the electricity bill term $p\left(\frac{d_t+\bar{u}(t)}{c}\right) u_i(t)$ appearing in (11). Indeed, the entire problem formulation (18) is a threefold generalization of (9)-(11): first, the cost appearing in (18a) comprises the cost (11) as a special case (to see this, set $Q_t = 0$ in (18a)); second, the constraints (18b) generalize the ones in (10); third, (18) presents a multi-variate formulation.

Moreover, the formulation (18) can be seen as a generalization of the linear-quadratic model proposed in (1)-(2); indeed, once we set $P_t = p_t = 0$ in (18a), and we discard the constraints (18b), then the two problems coincide (a part from the fact that one is formulated in continuous time, infinite horizon, whereas the other is formulated in discrete time, finite horizon).

In order to cast the optimization problem (18) in the mean field control theory, we need to impose that the reference sequence to be tracked equals a given function of the average optimal solution of (18); in symbols, $z_t = \Phi(\bar{x}^{\star}(t))$, where the star subscript refers to optimality. The way in which a solution is achieved is exactly the same proposed in Section 2.1: a central operator collects the optimal state evolutions $\{\mathbf{x}_i^{\star}\}_{i=1}^{N}$ of (18) and broadcasts the reference sequence $\{z_t\}_{t=0}^{T-1}$, which is a function Φ of the optimal average state evolution. The new optimal state evolution is computed by every agent, then the new optimal average state is collected by the central operator which finally broadcasts the new reference sequence.

While the agents are interested to the optimal response to the given reference signal, the central authority may be in general interested in the following goals: convergence of the population to a socially advantageous (mean field) equilibrium, for instance a Nash equilibrium; convergence of the population state to a desired state and/or target set; constraints satisfaction in the "population state space". We remark that these goals must be achieved in a decentralized fashion, or eventually distributed.

Addressing a mean field control problem for a large population of dynamical systems, each one wishing to optimize its dynamical evolution according to (18), generates open mathematical questions because it is not possible any more (unlike in the unconstrained linear quadratic case) to compute the closed-form state evolution analytically. Therefore it is of both academic and practical interest to investigate the properties of such generalized model for large populations of systems.

4.2 Application to plug-in electric vehicles: decentralized constrained charging management

We show now how the charging control problem for plug-in electric vehicles [10] can be generalized in presence of charging constraints, for instance set by the vehicles' owners themselves [25]. Such charging constraints, not considered in [10], are of particular interest in the PEVs case study of the DYMASOS project [21, Section 2.10.3].

Given the vehicles dynamics in (9), let us consider generalized, time dependent and

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement No 611281

vehicle dependent, charging constraints as follows (to be compared with (10)):

$$\mathbf{u}_i \in \mathcal{U}_i := \left\{ \mathbf{u}_i \in \mathbb{R}^T \mid 0 \leq u_i(t) \leq U_i(t), \sum_{t=0}^{T-1} u_i(t) = b_i^{-1}(1 - x_i(0)) \right\}, \tag{19}$$

for some $U_i(t) \geq 0$, $i = 1, \ldots, N$, $t = 0, \ldots, T-1$. In particular, with $U_i(t) = 0$ the agent i asks not to be charged at time interval t. As in (11), agent i seeks to minimize its own bill, that is $\sum_{t=0}^{T-1} p\left(\frac{1}{c}(d_t + \bar{\mathbf{u}}_t)\right) u_i(t)$, plus a cost due to the deviation from the average charging control, which helps to guarantee the validity of the technical results in [10, 25]. We remark that each vehicle owner is indeed modeled as an individual agent (system), along the lines of the AYESA test case [21, Section 2.7].

The main technical issue here is to find algorithms that guarantee convergence to a Nash equilibrium, and this goal becomes even more challenging in presence of the charging constraints in (19).

In particular, an important goal is the convergence to a "valley-filling" equilibrium, meaning that the total power demand remains constant over the time horizon as specified in the DYMASOS case study [21, Section 2.9.1], which is shown to be "socially optimal" in the unconstrained case [10, Section III]. In a more-realistic setting with model uncertainties, external disturbances such as forecasts mismatch and charging constraints, MFC has been successfully implemented in receding horizon [26], at least to relieve the demand peaks in peak day hours, as outlined in the DYMASOS case study [21, Section 2.3.1].

In [25] we show that convergence can be guaranteed even in generalized constrained case under the same technical assumptions in [10], via more general price-updating algorithms. These are shown to provide substantial benefits to the convergence properties for the charging control problem [25, Section V]. We believe that such results can be extended to the generalized (mean field) population control problems modeled in this section [27], in particular in the DYMASOS PEVs setting [21, Section 2.7.1] where the price function is modeled as an affine function of the power demand.

5 Conclusion

This deliverable provides a preliminary report on some modeling techniques for population-based approaches to the dynamic management of systems of systems; such modeling techniques are indeed the object of investigation of Task 1.1 in WP1 of the DYMASOS project. We mainly focused on mean field control, a new emerging research field, whose main purpose is to describe, analyze and control systems with a large number of agents interacting via their cost functions. We first presented the linear quadratic framework, its extensions and generalizations; we showed how such modeling scheme is particularly suited for describing systems of plug-in electric vehicles, whose goal is to recharge their

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement No 611281

batteries optimally in terms of both individual and social objectives. We also introduced a non-linear formulation of mean field control and applied it to the modeling of thermostatically controlled loads, discussing about the state evolution of such system. Since the dynamic management of populations of plug-in electric vehicles and flexible eletric loads are of particular interest in two DYMASOS use cases, we believe that mean field control theory has high potential to impact the DYMASOS project.

This deliverable is an important step towards a thorough and relevant investigation of several models for the description of systems of systems, which is the goal of Task 1.1; moreover, it constitutes a solid basis for Task 2.1, namely for system identification and state estimation for populations of systems of systems.

References

[1] "DYnamic MAnagement of Systems Of Systems (DYMASOS) website," 2013. [Online]. Available: www.dymasos.eu

[2] M. Huang, P. Caines, and R. Malhamé, "Nash equilibria for large-population linear stochastic systems of weakly coupled agents," in *Analysis, Control and Optimization of Complex Dynamic Systems*. Springer, New York, 2005, pp. 215–252.

[3] M. Huang, P. Caines, and R. Malhamé, "Large-Population Cost-Coupled LQG Problems With Nonuniform Agents: Individual-Mass Behavior and Decentralized ε-Nash equilibria," *IEEE Trans. on Automatic Control*, vol. 52, no. 9, pp. 1560–1571, 2007.

[4] J.-M. Lasry and P.-L. Lions, "Mean field games," *Japanese Journal of Mathematics*, vol. 2, no. 1, pp. 229–260, 2007.

[5] A. Bensoussan, J. Frehse, and P. Yam, *Mean field games and mean field type control theory*. Springer, 2013.

[6] V. E. Lambson, "Self-enforcing collusion in large dynamic markets," *Journal of Economic Theory*, vol. 34, pp. 282–291, 1984.

[7] H. Yin, P. Mehta, S. Meyn, and U. Shanbhag, "Synchronization of coupled oscillators is a game," *IEEE Trans. on Automatic Control*, vol. 57, no. 4, pp. 920–935, 2012.

[8] R. Malhamé and C.-Y. Chong, "Electric load model synthesis by diffusion approximation of a high-order hybrid-state stochastic system," *IEEE Trans. on Automatic Control*, pp. 854–860, 1985.

[9] F. Bagagiolo and D. Bauso, "Mean-field games and dynamic demand management in power grids," *Dynamic games and Applications*, vol. 4, no. 2, pp. 155–176, 2014.

[10] Z. Ma, D. Callaway, and I. Hiskens, "Decentralized charging control of large populations of plug-in electric vehicles," *IEEE Trans. on Control Systems Technology*, vol. 21, no. 1, pp. 67–78, 2013.

[11] B. Anderson and J. Moore, *Linear Optimal Control*. Englewood Cliffs, New Yersey: Prentice Hall, 1971.

[12] F. L. Lewis, D. Vrabie, and V. Syrmos, *Optimal control*. Wiley & Sons, 2012.

[13] M. Nourian, P. Caines, R. Malhamé, and M. Huang, "Mean field LQG control in leader-follower stochastic multi-agent systems: Likelihood ratio based adaptation," *IEEE Trans. on Automatic Control*, vol. 57, no. 11, pp. 2801–2816, 2012.

[14] D. Acemoglu and M. K. Jensen, "Robust comparative statics in large static games," in *IEEE Conf. on Decision and Control*, Atlanta, Georgia, USA, 2010, pp. 3133–3139.

[15] W. Ren, "Multi-vehicle consensus with a time-varying reference state," *Systems & Control Letters*, vol. 56, pp. 474–483, 2007.

[16] D. Gu and Z. Wang, "Leader-follower flocking: Algorithms and experiments." *IEEE Trans. on Control Systems Technology*, vol. 17, no. 5, pp. 1211–1219, 2009.

[17] H.-T. Liu, J. Shan, and D. Sun, "Adaptive synchronization control of multiple spacecraft formation flying," *Journal of Dynamic Systems, Measurement, and Control*, vol. 129, no. 3, pp. 337–342, 2006.

[18] G. Y. Weintraub, C. L. Benkard, and B. V. Roy, "Oblivious equilibrium: A mean field approximation for large-scale dynamic games," in *Neural Information Processing Systems*, Vancouver, Canada, 2005.

[19] B. Düring, P. Markowich, J.-F. Pietschmann, and M.-T. Wolfram, "Boltzmann and Fokker-Planck equations modelling opinion formation in the presence of strong leaders," *Proceedings of the Royal Society of London. Series A. Mathematical, Physical and Engineering Sciences*, vol. 465, no. 2112, pp. 3687–3708, 2009.

[20] A. Kizilkale and P. Caines, "Emergence of coalitions in mean field stochatic systems," in *IEEE Conf. on Decision and Control*, Florence, Italy, 2013, pp. 5768–5773.

[21] "Preliminary technical report on the DYMASOS case studies. Deliverable D5.1," 2014.

[22] M. Huang, P. Caines, and R. Malhamé, "The NCE (mean field) principle with locality dependent cost interactions," *IEEE Trans. on Automatic Control*, vol. 55, no. 12, pp. 2799–2805, 2010.

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement No 611281

[23] M. Huang, P. Caines, and R. Malhamé, "Social Optima in Mean Field LQG Control: Centralized and Decentralized Strategies," *IEEE Trans. on Automatic Control*, vol. 57, no. 7, pp. 1736–1751, 2012.

[24] P.-L. Lions and J.-M. Lasry, "Large investor trading impacts on volatility," *Annales de l'Institut Henri Poincare. Non Linear Analysis*, vol. 24, no. 2, pp. 311–323, 2007.

[25] F. Parise, M. Colombino, S. Grammatico, and J. Lygeros, "Mean field constrained charging policy for large populations of plug-in electric vehicles," in *IEEE Conf. on Decision and Control (submitted)*, Los Angeles, California, USA, 2014.

[26] Z. Ma, I. Hiskens, and D. Callaway, "A decentralized MPC strategy for charging large populations of plug-in electric vehicles," in *IFAC World Congress*, Milano, Italy, 2011.

[27] S. Grammatico, F. Parise, M. Colombino, and J. Lygeros, "Decentralized convergence to Nash equilibria in constrained linear quadratic mean field control," *(in preparation for journal submission)*, 2014.

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement No 611281