Distributed Control of Micro-storage Devices with Mean Field Games

Antonio De Paola, *Student Member, IEEE,* David Angeli, *Fellow, IEEE,*
and Goran Strbac, *Member, IEEE*

Abstract—This paper proposes a fully distributed control strategy for the management of micro-storage devices that perform energy arbitrage. For large storage populations the problem can be approximated as a differential game with infinite players (Mean Field Game). Through the resolution of coupled partial differential equations (PDEs), it is possible to determine, as a fixed point, the optimal feedback strategy for each player and the resulting price of energy if that strategy is applied. Once this price is calculated, it can be communicated to the devices which are able to independently determine their optimal charge profile. Simulation results are provided, calculating the fixed point through numerical integration of the PDEs. The original model is then extended in order to consider additional elements such as multiple population of devices and demand uncertainty.

Index Terms—Energy arbitrage, micro-storage, mean field games.

I. INTRODUCTION

ONE of the key features of the Smart Grid paradigm is the increasing participation of end customers in system operation. In this context, future households may be equipped with storage devices that would be able to exchange energy with the network. The benefits of this scenario must be considered at two different levels: every single participant (agent) will be able to maximise their profit by optimally charging/discharging energy storage on the basis of energy prices, while the power system will benefit from a reduction in the peak demand and in the reliance on fossil fuel peaking plants. In this context the main challenge is to design a fully distributed control strategy which optimizes the profit of the single device and, at the same time, takes into account the changes in prices introduced by the whole storage population, in order to avoid synchronicity phenomena (i.e. all devices charge when price is low and the peak demand is simply shifted). This problem has been the subject of significant research: in [1] the energy suppliers adopt an adaptive pricing scheme, broadcasting the price of energy for the devices in advance of each daily period, in order to better predict the global storage behaviour. A game theory approach is used in [2] to calculate the performance bounds of the devices, which are then used as benchmark for the proposed adaptive strategy. Game theory has also been applied to the similar case of electric vehicles, calculating a cost-minimizing scheduling which fills the valleys in electric load profiles. Some elements, such as uncertainty in the utilization

A. De Paola, D. Angeli and G. Strbac are with the Department of Electrical and Electronic Engineering, Imperial College London, London, SW7 2AZ UK (e-mail: ad5709@imperial.ac.uk; d.angeli@imperial.ac.uk; g.strbac@imperial.ac.uk).

of vehicles and unidirectionality of the charging, do not apply to the case of storage, but the chosen approaches remain of interest. In [3] the mean behaviour of the vehicle population is used to determine the charging profile of each agent, achieving a Nash Equilibrium that coincides with the globally optimal strategy if the vehicles are identical. Similarly, [4] proposes an iterative strategy where each agent minimizes its own cost function (which takes into account the result of the previous iteration), reaching equilibrium with certain valley-filling properties. A different approach is adopted in [5] to integrate flexible demand in the electricity market, using a two-level iterative process with Lagrange relaxation.

This paper designs a distributed control for the devices, modelling them as competitive players that interact through the changes in the energy prices driven by their operation strategies of charging/discharging energy storage. If a high number of devices are considered, the contribution of the single element can be assumed negligible and only the global effect of the population must be taken into account. The energy arbitrage can therefore be approximated by a differential game with infinite players (Mean Field Game). This kind of games, independently developed by Huan-Caines-Malhamè [6] and Lasry-Lions [7], are described in their basic formulation by two coupled partial differential equations (PDEs): one Hamilton-Jacobi-Bellman (HJB) equation, which returns the optimal control of the agents, and one Fokker-Planck (FP) equation which describes their distribution. In the energy arbitrage case, by solving these coupled PDEs, it is possible to determine a Nash equilibrium for the population of devices. By sending the corresponding price to the agents, these are able to independently calculate their optimal strategy. The choice to model storage with PDEs is not new and has been already adopted, for example in [8], to estimate the value of stochastic storage in a centralized framework. The use of Mean Field Games (MFGs) to describe large populations of competing agents in the power system is present in [9] and [10], which consider respectively electric vehicles and suppliers/consumers dynamics. The novelty of the presented work is the application of the MFG equations to the design of a distributed control of storage, calculating a feedback optimal charge profile which explicitly consider state constraints. Furthermore, the original model is extended by considering additional factors such as multiple populations of devices and uncertainty in demand.

The paper is structured as follows: Section II introduces the model of the storage devices and the energy market, approximates the energy arbitrage as a non-cooperative game for infinite players and determines the corresponding PDEs.

The resolution strategy, the adopted numerical methods and the corresponding simulation results are in Section III, while Section IV contains some extensions to the original model.

II. MEAN FIELD GAME

The energy arbitrage problem is considered within a competitive game framework: each device (agent) exchanges energy with the network aiming at maximizing its own pay-off. The main challenge is to perform such maximization with a distributed control strategy which takes into account the change in energy prices introduced by the agents and achieves a Nash equilibrium. Our analysis is focused on micro-storage devices, which are expected to be present in the network in the order of millions. We originally assume homogeneity of storage: all devices have the same energy and power rating. Instead of separately considering the individual agent, the number of devices is approximated as infinite: the influence of the single device on the energy price becomes negligible and only the effect of the whole storage population (mean field) must be taken into account. The problem is then approached as a differential game with infinite players (Mean Field Game), described by coupled partial differential equations.

A. Modelling of the arbitrage problem

The electricity market has been abstracted with the monotonic increasing function Π which associates, to a given value of aggregate demand $D(t)$, the corresponding energy price $p(t) = \Pi(D(t))$. The power demand D will be given by two different components: the inflexible demand D_i and the contribution D_s of the storage population. The inflexible profile is modelled with the function $D_i(t)$, initially assumed to be known without uncertainties, while the expression for D_s will be introduced later. Each of the infinite devices in the storage population is described by the following equations:

$$
\begin{aligned}
\dot{S}(t) &= r(t) \\
y(t) &= r(t) + \gamma r^2(t)
\end{aligned}
\tag{1}
$$

The variables S and r denote respectively the state of charge (between 0 and 1) and the charging rate of the device. In order to take into account the efficiency of storage, quadratic losses parametrized by γ are introduced in the expression of y, which represents the normalized power (in p.u. over seconds) exchanged by the device with the network. If we denote respectively by E_r and P_r the energy and power rating of the devices, the constraints on the charging rate r are:

$$
\begin{aligned}
0 &\le S(t) \le 1 \\
r^{MIN} = -\frac{P_r}{E_r} &\le r(t) \le \frac{P_r}{E_r} = r^{MAX}
\end{aligned}
\quad \forall\, t \in [0, T]
\tag{2}
$$

Alternatively, the efficiency of the devices can be expressed through the parameter k, defined such that the linear loss kr is equivalent to the chosen quadratic one for $r = r^{MIN}$:

$$
-\gamma (r^{MIN})^2 = kr^{MIN} \rightarrow k = -\gamma r^{MIN}
\tag{3}
$$

Remark 1: All the results provided in this section can be rescaled and used to approximate the scenario with a finite

number of players N. In this case the state and rate of charge remain unaltered and it is possible to define also the stored energy E and the charging power P. Assuming E_r and P_r equal for all devices, it yields:

$$
E(t) = E_r \cdot S(t) \qquad P(t) = E_r \cdot r(t)
\tag{4}
$$

The optimal power profile and the profit of the single device in the finite case will be defined in Section IV when simulation results are presented.

The function m is now used to describe the energy distribution within the storage population: given two arbitrary values S_1 and S_2 for the state of charge, $\int_{S_1}^{S_2} m(t, S)\, dS$ will denote the fraction of devices for which $S_1 \le S(t) \le S_2$. Since the optimal control is going to be calculated in feedback form, it is assumed that the charging rate of the devices $r(t, S)$ is a function of time and current state of charge. By considering the aggregated energy capacity $E^{TOT} = N \cdot E_r$ as a rescaling factor, the variation to total demand introduced by the storage population can now be defined:

$$
\begin{aligned}
D_s(t) &= E^{TOT} \int_0^1 m(t, S) \left[r(t, S) + \gamma r^2(t, S) \right]\, dS \\
&= E^{TOT} \int_0^1 m(t, S) y(t, S)\, dS
\end{aligned}
\tag{5}
$$

We finally introduce the cost C minimized by each agent:

$$
\begin{aligned}
C(S(0), r(\cdot)) &= \int_0^T \Pi \left(D_i(t) + D_s(t) \right) \left[r(t) + \gamma r^2(t) \right] dt \\
&+ \Psi(S(T)) = \int_0^T p(t) y(t)\, dt + \Psi(S(T))
\end{aligned}
\tag{6}
$$

The integral in C represents the cost sustained by the device while charging/discharging energy in the considered time interval $[0, T]$ while the final cost function Ψ takes into account the final state of charge and avoids, for example, the total discharge of the device. Notice that the different agents interact between each other by varying D_s and therefore the price of energy. With the proposed formulation which considers infinite agents, such variations do not depend on the single device but only on the charge distribution m of the whole population.

B. Derivation of the coupled PDEs

It is now possible to introduce the partial differential equations that describe the Mean Field Game and are used to determine the decentralized control. In order to calculate the optimal rate of charge r^* which minimizes C, an Hamilton-Jacobi-Bellman equation is considered:

$$
\begin{aligned}
-\partial_t V(t, S) &= \min_{r \in \mathcal{R}(S)} \left[p_m^*(t)(r + \gamma r^2) + \partial_S V(t, S) r \right] \\
V(T, S) &= \Psi(S)
\end{aligned}
\tag{7}
$$

In particular the function p_m^* represents the price of energy resulting from all devices (with distribution m) applying the optimal charging profile r^*:

$$
\begin{aligned}
p_m^*(t) &= \Pi \left(D_i(t) + E^{TOT} \int m(t, S) y^*(t, S)\, dS \right) \\
y^*(t, S) &= r^*(t, S) + \gamma r^{*2}(t, S)
\end{aligned}
\tag{8}
$$

while for the interval of admissible controls $\mathcal{R}(S)$ it yields:

$$\mathcal{R}(S) = \begin{cases} [0, r^{MAX}] & \text{if} \quad S = 0 \\ [r^{MIN}, r^{MAX}] & \text{if} \quad 0 < S < 1 \\ [r^{MIN}, 0] & \text{if} \quad S = 1 \end{cases} \quad (9)$$

Notice that (7) is a PDE in the unknown function V which, according to the dynamic programming principle [11], can be considered as the value function of the optimization problem and is equal to C for $t = 0$:

$$V(t, x) = \min_{r(\cdot)} \int_t^T p_m^*(\tau)[r(\tau) + \gamma r^2(\tau)] \, d\tau + \Psi(S(T))$$
$$S(t) = x$$
$$(10)$$

The optimal control r^*, at each time instant, is the argument of the minimum in (7) and therefore must satisfy the following:

$$r^*(t, S) = \operatorname*{sat}_{\mathcal{R}(S)} \left[-\frac{p_m^*(t) + \partial_S V(t, S)}{2\gamma p_m^*(t)} \right] \quad (11)$$

Since $p_m^*(t)$ depends on $r^*(t, \cdot)$, there is no closed expression for the optimal control. On the other hand it is possible to prove, under mild assumptions, that a fixed point for equations (8) and (11) always exists and is unique. Furthermore, the numerical calculation of such fixed point is straightforward, for example adopting the procedure described in the next section.

It has been shown how the optimal control r^* depends on the charge distribution m through the energy price p_m^* but it is important to point out how an analogue relationship holds in the opposite sense. In particular, m depends on the charge profile of the devices through the following conservation law:

$$\partial_t m(t, S) = -\partial_S [r^*(t, S)m(t, S)]$$
$$m(0, S) = m_0(S)$$
$$(12)$$

The solution to the arbitrage problem as a Mean Field Game consists then in the couple (V, m) (or equivalently (r^*, m)), which represents a fixed point for equations (7) and (12). In other words, it is necessary to determine an optimal charge profile r^* which minimizes the cost function C for a given distribution m and, at the same time, induces such distribution. In practical implementations of the proposed control strategy, the Mean Field Game solution can be calculated in a centralized manner following the procedure detailed in the next section. Once this has been determined, the distributed control can be implemented with a one-way communication channel: the resulting energy price p_m^* is broadcast to the devices, which will be able to independently calculate their optimal charge profile by solving only equation (7).

It is difficult to theoretically prove the existence and uniqueness of the centralized solution of the Mean Field Game. Current mathematical literature on the topic provides results in this sense only for much simpler classes of systems. In the simulations described in the next sections the case of multiple equilibria has never been experienced. Nevertheless, if this were the case, it would be always possible to choose the equilibrium which maximizes some global index on the resulting energy price function or aggregate demand profile.

III. SIMULATION RESULTS

The fixed point for the Mean Field Game described by (7) and (12) and the resulting distributed control have been calculated by numerical integration of the PDEs. In this respect, it is important to point out that not only the two equations are interdependent, but they are also integrated in two different directions. In fact, for the dynamic programming principle, the HJB equation must be integrated backward in time while the conservation law is integrated forward. For this reason an iterative strategy has been chosen, solving one equation at a time and using the result as a starting point for the next integration. All steps of the numerical resolution are detailed below:

1) The initial guess for the distribution function and the demand variation introduced by storage are denoted respectively as $\tilde{m}$ and $\tilde{D}_s$ and are defined as follows:

$$\tilde{m}(t, S) = m_0(S)$$
$$\tilde{D}_s(t) = 0 \qquad \forall \ t \in [0, T] \quad (13)$$

2) The HJB equation (7) is integrated backward in time starting from $V(T, S) = \Psi(S)$ and assuming $m = \tilde{m}$. At each time step t the fixed point for equations (8) and (11) which determines the optimal control r^* is calculated according to steps a) to d):

   a) The energy price is initially assumed equal to $\tilde{p}(t) = \Pi(D_i(t))$.
   b) Optimal control estimate $\bar{r}(t, \cdot)$ is calculated with equation (11), where $p_m^*(t)$ is replaced with $\tilde{p}(t)$.
   c) The new price estimate $\bar{p}$ is defined as follows:

$$\bar{p}(t) = \Pi \left(D_i(t) + E^{TOT} \int \tilde{m}(t, S) \bar{y}(t, S) \, dS \right)$$
$$(14)$$

   with $\bar{y}(t, S) = \bar{r}(t, S) + \gamma \bar{r}^2(t, S)$.
   d) The norm $|\bar{p}(t) - \tilde{p}(t)|$ is evaluated. Given a certain price error tolerance ϵ_p, if $|\bar{p}(t) - \tilde{p}(t)| < \epsilon_p$ the iterations are stopped and step 3 is executed. If the condition is not satisfied, the initial price estimation is reset with $\tilde{p}(t) = \bar{p}(t)$ and steps 2.b-c are repeated.

3) A new estimate $\bar{m}$ is obtained integrating forward equation (12) and assuming $r^* = \bar{r}$. The resulting demand variation $\bar{D}_s(t)$ introduced by storage is calculated by replacing $\bar{m}$ and $\bar{r}$ in (5).

4) The following function norm is evaluated:

$$\|\bar{D}_s - \tilde{D}_s\|_1 = \int_0^T \left| \bar{D}_s(t) - \tilde{D}_s(t) \right| dt \quad (15)$$

   Given a certain demand error tolerance ϵ_D, if $\|\bar{D}_s - \tilde{D}_s\|_1 < \epsilon_D$ step 5 is executed. Otherwise, the initial estimates are updated with $\tilde{D}_s = \bar{D}_s$ and $\tilde{m} = \bar{m}$ and steps 2-3 are repeated.

5) The solution of the MFG corresponds to the results of the last iteration:

$$r^* = \bar{r} \qquad m = \bar{m} \quad (16)$$

The numerical integration has been performed using finite difference schemes. The time and state of charge steps Δt and ΔS are chosen and the corresponding vectors are partitioned:

$$t_i = i \cdot \Delta t \quad S_j = j \cdot \Delta S \tag{17}$$

The value of a function $f(t,S)$ in $(i\Delta t, j\Delta S)$ will be hereby denoted as f_j^i. In the integration of the HJB an upwind method is used [12]. The control $\bar{r}$ in step 2.b is calculated using central derivatives:

$$\bar{r}_j^i = \operatorname*{sat}_{\mathcal{R}(S_j)} \left[-\frac{\bar{p}^i + \frac{V_{j+1}^i - V_{j-1}^i}{2\Delta S}}{2\gamma\bar{p}^i} \right] \tag{18}$$

The integration of the HJB equation (7) at the time step $i-1$ is performed with the following numerical scheme:

$$\begin{aligned}
V_j^{i-1} =\ & \Delta t \cdot \bar{r}_j^i \left[\frac{1 + sign(\bar{r}_j^i)}{2} \frac{(V_{j+1}^i - V_j^i)}{\Delta S} \right] \\
& + \Delta t \cdot \bar{r}_j^i \left[\frac{1 - sign(\bar{r}_j^i)}{2} \frac{(V_j^i - V_{j-1}^i)}{\Delta S} \right] \\
& + \Delta t \cdot \bar{p}^i \left[\bar{r}_j^i + \gamma \left(\bar{r}_j^i \right)^2 \right] + V_j^i
\end{aligned} \tag{19}$$

Notice that, with this approach, the calculation of the partial derivative $\partial_S V$ at the state boundaries is straightforward: when $j = 1$ the forward derivative, calculated using V_1^i and V_2^i, is always adopted and the same holds for the backward derivative when $j = 1/\Delta S$. For the integration of the conservation law (12) the Lax-Friedrich method is used [13]. An artificial viscosity with coefficient ϵ is introduced and the following numerical scheme is adopted:

$$\begin{aligned}
\bar{m}_j^{i+1} =\ & \bar{m}_j^i - \frac{\Delta t}{2\Delta S} \left[\bar{r}_{j+1}^i \bar{m}_{j+1}^i - \bar{r}_{j-1}^i \bar{m}_{j-1}^i \right] \\
& + \epsilon \left[\bar{m}_{j+1}^i - 2\bar{m}_j^i + \bar{m}_{j-1}^i \right]
\end{aligned} \tag{20}$$

Notice that (20) can be easily extended to the energy boundaries ($j = 1$ and $j = 1/\Delta S$) by taking into account that the sum of the viscosity term over j must be equal to 0.

Regarding the parameters of the system, the inflexible demand D_i has been chosen equal to a 24-h UK demand profile [14] shown in Fig. 1. The price function $\Pi(D)$ has been defined on the basis of the values used in [15] as shown in Fig. 2. In this case study a population of $N = 10^6$ devices has been considered, with the following parameters:

$$P_r = 2.5KW \quad E_r = 25KWh \quad k = 0.25 \tag{21}$$

The approximation with infinite players is then introduced, defining the boundary conditions of the PDEs: m_0 has been chosen as a gaussian distribution and Ψ as a quadratic function which penalizes final values of S different from $1/2$:

$$m_0(S) = \frac{1}{2\pi\sigma^2} e^{-\frac{\left(S - \frac{1}{2}\right)^2}{2\sigma^2}} \quad \Psi(S) = c\left(S - \frac{1}{2}\right)^2 \tag{22}$$

The time step, state of charge step and the other parameters in the numerical integration are:

$$\begin{aligned}
\Delta t = 0.02h \quad \Delta S = 0.004 \\
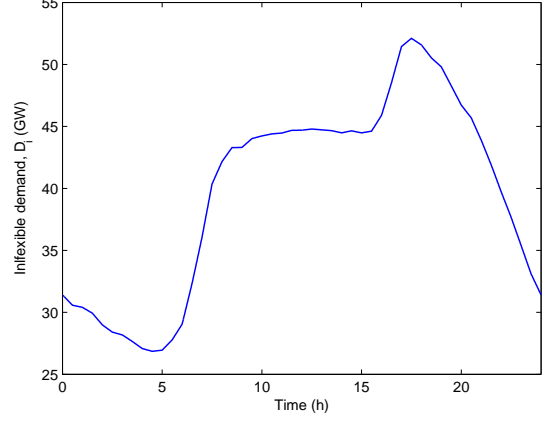\sigma = 1.2 \quad c = 1000
\end{aligned} \tag{23}$$

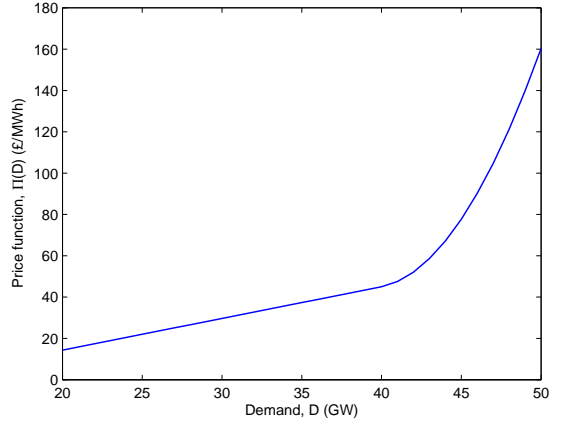Fig. 1. Chosen values of D_i over a time interval of 24h.

Fig. 2. Price of energy $\Pi(D)$ with respect to aggregate demand D.

Notice that, once the Mean Field Game has been solved, the results for a finite number of devices can be obtained following Remark 1: the energy level E and the charging power P for the single device can be calculated by multiplying S and r by E_r. In a similar way, it is possible to obtain their energy distribution m_s, optimal power profile P^* and cost function C_s. A numerical solution to the Mean Field Game described by (7) and (12) is now obtained by following the procedure detailed at the beginning of this section. Simulations have been run in a MATLAB environment on a HP Z600 machine equipped with an Intel Xeon CPU (frequency of $2.4GHz$) and 12 GB of RAM. The whole resolution procedure is completed, for the specified parameters, in about 7 seconds. The inflexible demand profile D_i and the aggregate demand $D_i + D_s$ at each iteration of the backward/forward integration are shown in Fig. 3. Notice that the proposed resolution strategy converges in 3 iterations when an error tolerance $\epsilon_D = 1000$ is chosen. Furthermore, the presence of storage devices performing energy arbitrage introduces in the demand profile a considerable peak shaving/valley filling. The optimal power profile $P^*(t, E) = r^*(t, \frac{E}{E_r})E_r$ is shown in Fig. 4. The resolution of the HJB equation returns an optimal control in feedback form: since the charge profile is not only a function

of time but it also depends on the current energy of the device, greater robustness is guaranteed. The energy stored in the devices, for different values of $E(0)$, is shown in Fig. 5: notice that, in general, the trajectories have similar trends across time and charging occurs in the first hours of the day when energy prices are low. The only exception are devices with high values of $E(0)$: in this case the initial charging is limited by the constraint on the maximum energy E_r. Furthermore, it is important to point out that the whole population, given the particular choice of Ψ (equal for all devices), converges at time T towards the energy value $\frac{E_r}{2}$. The possibility to introduce different final constraints, for example imposing that devices must have the same initial and final energy, is considered in the next section. In Fig. 6 the original price of energy $p(t) = \Pi(D_i(t))$ is compared to $p_m^*(t)$, obtained when the storage population applies the optimal charge profile r^*. Following the variations introduced in the aggregate demand profile (shown in Fig. 3) and the chosen price function Π, the proposed operation strategy achieves a considerable price reduction during peak times. If one considers definition (10) for the value function V, the profit $G_s = -C_s$ of the single device as a function of its initial energy $E(0)$ can be defined as follows:

$$G_s(E(0)) = -V\left(0, \frac{E(0)}{E_r}\right) \cdot E_r \qquad (24)$$

and it is shown, for the last iteration, in Fig. 7. As previously mentioned, in the current formulation all devices will tend towards the same final value $\frac{E_r}{2}$ of stored energy. This means that G_s will be bigger for devices with higher values of $E(0)$ that have more energy available to exchange.
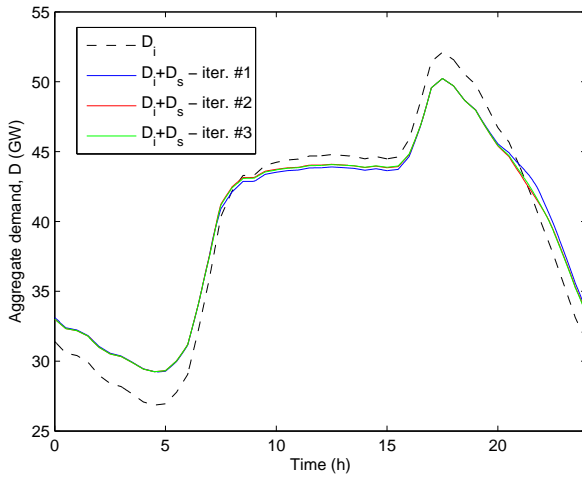

Fig. 4. Optimal power profile P^* as a function of time and current stored energy E.

Fig. 5. Stored energy $E(t)$ of the single device across time for different initial values $E(0)$.

Fig. 3. Aggregate demand profile at each iteration of the MFG-solving procedure.

Fig. 6. Comparison between the original energy price $\Pi(D_i(t))$ (blue) and p_m^* (red), resulting from the solution of the Mean Field Game.

IV. MODEL EXTENSIONS

The distributed control for the energy arbitrage problem, described by the Mean Field Game with equations (7) and (12), is extended in this section in order to consider additional elements or different constraints.
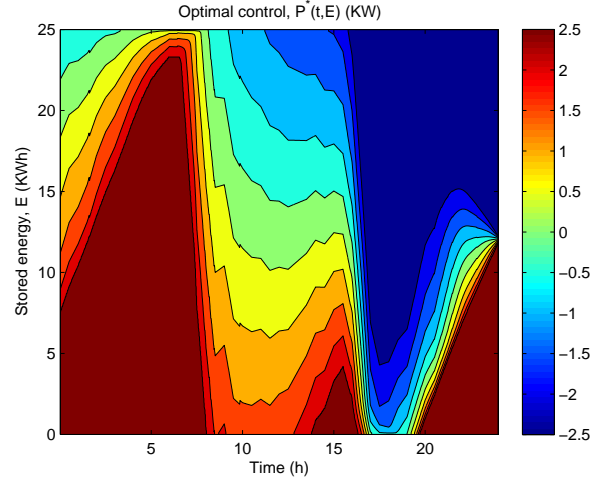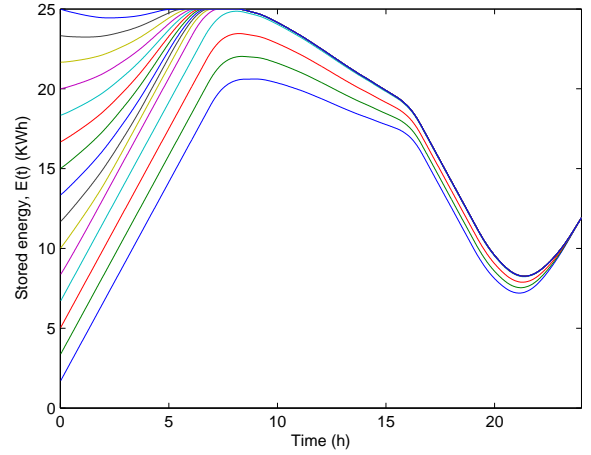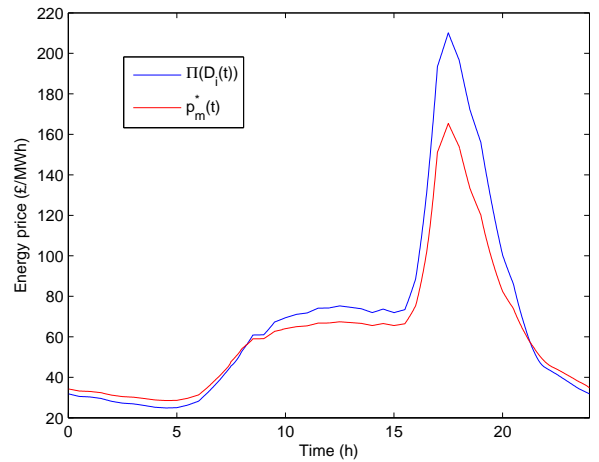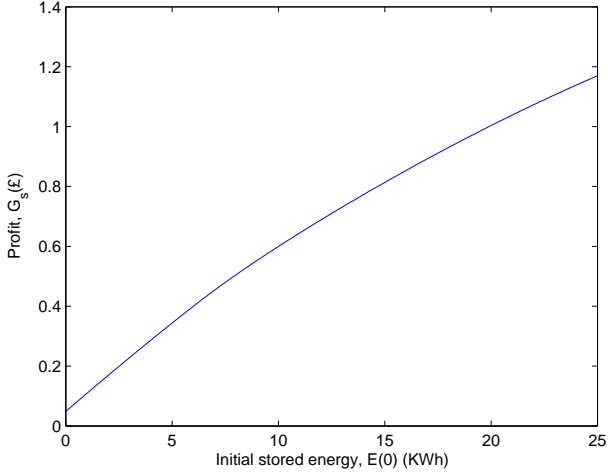
Fig. 7. Profit G_s of the single device as a function of the initial stored energy.

A. Cyclic Constraints

In the original formulation the final penalty function $\Psi(S)$ is the same for all devices which in turn induces convergence towards similar values of E at $t = T$. It might be desirable to impose, for each device, a cyclic constraint $E(0) = E(T)$. In order to do so, it is necessary to introduce an additional state variable I_s, defined as follows:

$$I_s(t) = \int_0^t P(\tau)\,d\tau \qquad -E_r \leq I_s(t) \leq E_r \qquad (25)$$

Notice that $I_s(t)$ represents the total variation of stored energy in the interval $[0, t]$ and, to impose a cyclic constraint, it must be required $I_s(T) = 0$. For the case of infinite players, the equivalent state variable $I(t) = \int_0^t r(\tau)\,d\tau$ (in p.u.) is introduced in the Mean Field Game equations. For the conservation law, taking into account that the optimal control will depend from time and the two state variables, it yields:

$$\partial_t m(t, S, I) = -\partial_S \left[r^*(t, S, I) m(t, S, I) \right] \\ -\partial_I \left[r^*(t, S, I) m(t, S, I) \right] \qquad (26)$$

Similarly, for the HJB equation:

$$-\partial_t V(t, S, I) = \min_{r \in \mathcal{R}(S)} \left[p_m^*(t)(r + \gamma r^2) \\ + \left(\partial_S V(t, S, I) + \partial_I V(t, S, I) \right) r \right] \qquad (27)$$

In this case the demand variation D_s introduced by the storage population and considered in the calculation of p_m^* is defined as follows:

$$D_s(t) = E^{TOT} \int \int m(t, S, I)[r^*(t, S, I) + \gamma r^{*^2}(t, S, I)]\,dI\,dS \qquad (28)$$

Regarding the boundary conditions of the PDEs, for the conservation law (26) it is sufficient to impose the same initial distribution in S, with $I = 0$. Denoting by δ the Dirac delta, it yields:

$$m_0(S, I) = \frac{1}{2\pi\sigma^2} e^{-\frac{\left(S - \frac{1}{2}\right)^2}{2\sigma^2}} \cdot \delta(I) \qquad (29)$$

The cyclic constraints can be finally imposed by deciding to penalize, for example quadratically with parameter c, the final values of I:

$$\Psi(S, I) = c \cdot I^2 \qquad (30)$$

As in the previous case, the coupled PDEs (26) and (27) have been solved numerically, using the algorithm described at the beginning of Section III. The main difference is that the additional state variable increases considerably the computational complexity of the integration schemes. For this reason, different numerical methods are considered. The step Δr is introduced and the admissible controls are defined as:

$$r_l = [r^{MIN} + l \cdot \Delta r] \qquad l = 0, 1, \ldots, \frac{r^{MAX} - r^{MIN}}{\Delta r} \qquad (31)$$

The time and state of charge vectors are partitioned in such a way that the state variation introduced by r_l is a multiple of ΔS:

$$\Delta S = \Delta r \cdot \Delta t = \Delta I \qquad (32)$$

A notation similar to the one described in Section III is adopted for the numerical schemes, denoting as $f_{j,k}^i$ the value of a function $f(t, S, I)$ in $(i\Delta t, j\Delta S, k\Delta I)$. It is now possible to solve the HJB equation simply as a dynamic programming problem:

$$V_{j,k}^{i-1} = \min_l \left[\bar{p}^i \left[r_l + \gamma r_l^2 \right] + V_{j+\Delta_l, k+\Delta_l}^i \right] \qquad (33)$$

where $\Delta_l = r_l \frac{\Delta t}{\Delta S}$ corresponds to the energy index variation introduced by r_l. Given the control $\bar{r}_{j,k}^i$ and the corresponding energy index variation $\bar{\Delta}_{j,k}^i$, the integration of the conservation law can be easily implemented by initializing $\bar{m}^{i+1}$ at 0 and iterating the following over j and k:

$$\bar{m}_{j+\bar{\Delta}_{j,k}^i, k+\bar{\Delta}_{j,k}^i}^{i+1} = \bar{m}_{j+\bar{\Delta}_{j,k}^i, k+\bar{\Delta}_{j,k}^i}^{i+1} + \bar{m}_{j,k}^i \qquad (34)$$

The scenario described in the previous section has been simulated with the following parameters:

$$P_r = 2.5KW \quad E_r = 25KWh \quad k = 0.25$$
$$\Delta r = 0.0125/h \qquad \Delta t = 0.32h \qquad (35)$$
$$\Delta S = \Delta r \cdot \Delta t = 0.004 \qquad c = 10^5$$

The algorithm for the resolution of the MFG converges, as in the previous case, in about three iterations and, given the additional variable I introduced in the problem, for the parameters in (35) requires about 4 hours to be completed. Notice that, by simply choosing $\Delta r = 0.025/h$ it is possible to complete the same calculations in 1 hour, with a difference in the results which is not significant. The profile of total demand obtained in simulations is very similar to the one in Fig. 3. The corresponding energy trajectories for different values of $E(0)$ are shown in Fig. 8. As expected, each device charges/discharges on the basis of the current price of energy, returning at its initial state of charge for $T = 24h$. In this case the profit G_s of the single device as a function of its initial energy $E(0)$ is equal to $-V\left(0, \frac{E(0)}{E_r}, 0\right) E_r$. A sensitivity analysis has been carried out, comparing the results for different values of E_r (while keeping unaltered the other parameters). The results are shown in Fig. 9. As expected, to

higher energy ratings correspond higher profits. It should also be noticed that, in general, the arbitrage is more profitable for the devices which start with a low state of charge and are therefore able to charge more energy during the first hours, when energy price is lower. A further consideration can be made for the case $E_r = 25KWh$: in this case the highest profit is achieved by devices with initial state of charge within a certain interval. It can be noticed from Fig. 8 that these devices are the ones for which the constraints on minimum and maximum energy are not active in the considered time horizon.
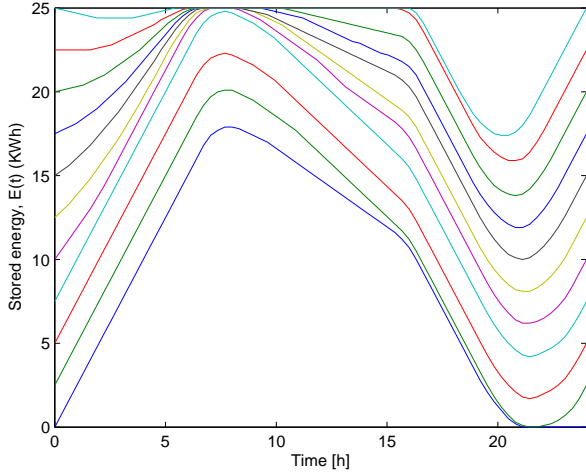

Fig. 8. Stored energy of the devices across time for different values of $E(0)$ ($E_r = 25KWh$).
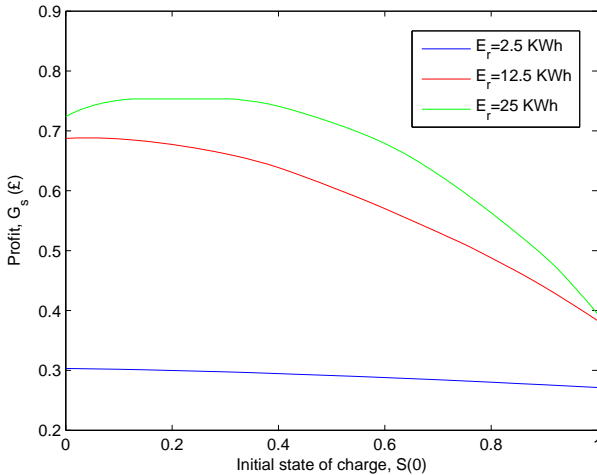

Fig. 9. Profit G_s of the single device as a function of the initial state of charge $S(0)$, for different energy ratings.

B. Multiple Populations of Devices

Given the chosen MFG approach, a preliminary hypothesis was the homogeneity of the devices: all agents have the same efficiency, capacity and maximum rate of charge. It

is possible to introduce in the model a parameter α which varies in the population, considering it as an additional state variable and imposing $\dot{\alpha} = 0$. The main problem with this approach is that increases considerably the complexity of the problem. An acceptable trade-off can be obtained if, instead of parameters that vary continuously, we consider a finite number of populations, each with devices of the same kind. Notice in fact that, for the considered arbitrage problem, the different populations interact only through the energy price $\Pi(D_i + D_s)$. The distributed control for this particular case can be obtained by solving in parallel a set of coupled PDEs (26) and (27), one for each population, that share the same price function. The same resolution procedure is followed, with a different expression for the demand variation introduced by storage. Denoting with subscript j the energy distribution, total capacity and optimal control for each of the M populations, it yields:

$$D_s(t) = \sum_{j=1}^{M} E_j^{TOT} \int \int \bar{m}_j(t, S, I) \bar{y}_j(t, S, I) \, dS \, dI \quad (36)$$

where $\bar{y}_j$ is the normalized power exchanged by the devices of the jth population and is defined as:

$$\bar{y}_j(t, S, I) = \bar{r}_j(t, S, I) + \gamma_j \bar{r}_j^2(t, S, I) \quad (37)$$

In the case of a finite number M of populations, each with the same parameters, M coupled PDEs are solved in parallel, with a computational complexity that increases linearly with respect to M. This extension is particularly significant if one considers practical implementations of the proposed control algorithm. In fact, in order to obtain the equilibrium of the Mean Field Game, the centralized entity that performs this calculation must know the initial energy distribution of the devices and the corresponding characteristics (energy/power rating and efficiency). In this respect, it is reasonable to consider that a finite number of device typologies (with known parameters) will be available for commercial purposes. In the proposed formulation to each kind of device will correspond one of the M populations described above.

A scenario with two populations (A and B) has now been simulated, with the following choice of parameters:

$$\begin{aligned} N_A = 5 \cdot 10^5 \quad & E_{r_A} = 20KWh \quad P_{r_A} = 2KW \\ N_B = 5 \cdot 10^5 \quad & E_{r_B} = 30KWh \quad P_{r_B} = 3KW \end{aligned} \quad (38)$$

The equations of the mean field game with cyclic constraints have been integrated using the same boundary conditions and numerical methods presented in the previous subsection. With the hardware/software specifications provided in Section III, the simulation is completed in about 4 hours. As in the previous case, the simulation time is sensibly reduced if one increases the discretization step Δr of the control r. The aggregate demand profile is shown in Fig. 10: notice that, as in previous cases, convergence to the solution is achieved in about 3 steps. Furthermore, the peak shaving/valley filling introduced by storage is comparable to the case when only one population is considered. The profit of the devices as a function of their initial state of charge is compared in Fig. 11 for the

two populations. Notice in particular that, for the devices of population A, with a lower power rating, the corresponding profit G_{s_A} is sensibly lower. Furthermore, the considerations on Fig. 9 about the effect of the initial state of charge can be extended to the present case.
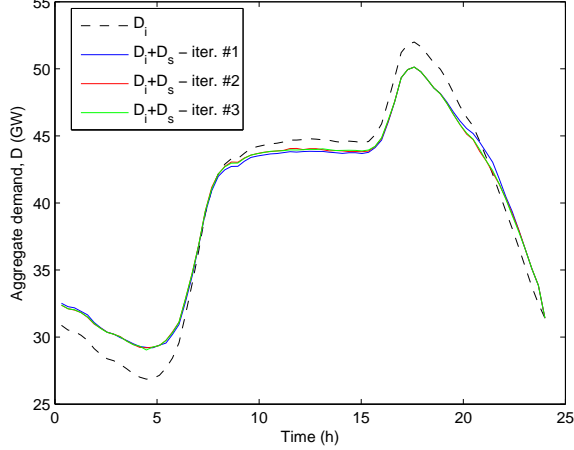

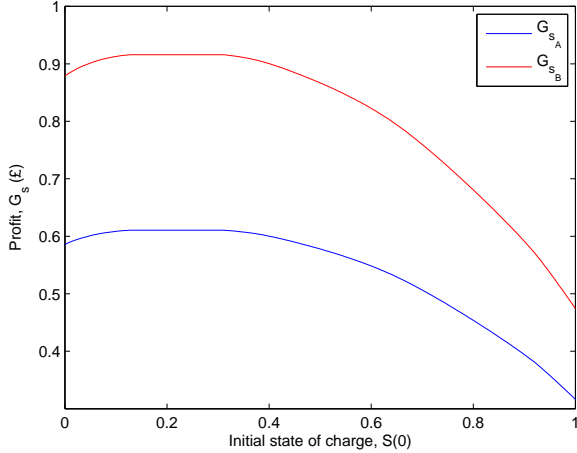Fig. 10. Aggregate demand profile at each iteration of the MFG-solving procedure.

Fig. 11. Profit G_s of the single device as a function of the initial state of charge, for population A (blue) and B (red).

C. Uncertainty in Demand

Another assumption in the initial calculation of the distributed control presented in Section II is that the inflexible demand profile D_i is known without uncertainties. It is possible to consider error in the demand forecast D_f by introducing the following expression:

$$D_f(t) = D_i(t) + \eta(t) \tag{39}$$

where D_i is the actual inflexible demand profile and η is a stochastic process. One possible way to take into account this uncertainty is to use a receding horizon control:

1) At time t the actual demand $D_i(t)$ is measured and the forecast D_f is estimated over the interval $[t, t+T]$.
2) Based on the current distribution $m(t, \cdot)$, the final cost function Ψ is updated:

$$\bar{S}(t) = \int_0^1 m(t, S) S \, dS$$
$$\Psi(S) = c \cdot \left(S - \bar{S}(t)\right)^2 \tag{40}$$

3) The coupled PDEs (7) and (12) are solved considering D_f as profile of inflexible demand.
4) The optimal r^* is applied only at the current time step.
5) Steps 1-4 are repeated for $t + \Delta t$.

This approach allows to implicitly incorporate uncertainty in the model but, on the other hand, increases the computational complexity by a factor of $T/\Delta t$ since the equations of the Mean Field Game must now be solved at each time step. Denoting by W the Wiener process, the dynamics of η are defined as follows:

$$d\eta_t = \sigma_f \cdot dW_t \tag{41}$$

Notice that the forecast error η is characterized by mean value $\mu_\eta(t) = 0$ and standard deviation $\sigma_\eta(t) = \sigma_f \sqrt{t}$, as similarly shown for example in [16]. The performance of the receding horizon strategy has been evaluated for the case study described in Section III, with a higher time step $\Delta t = 0.2h$. The average daily profit $\bar{G}_s$ of the single device over 200 days has been calculated for different values of σ_f and is shown in Fig. 12. The calculations for each σ_f are completed in about 80 minutes. Notice that, in practical implementations, they would be gradually performed during the whole time interval and not entirely at the beginning. As expected, there is a reduction of $\bar{G}_s$ for increasing values of σ_f. On the other hand such reduction is not very significant: by solving the MFG at each time step, with updated forecasts on demand, the impact of uncertainties on the final result can be considerably reduced.
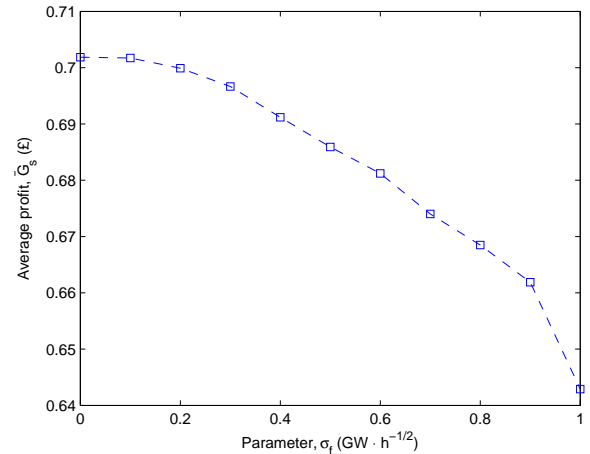

Fig. 12. Average daily profit $\bar{G}_s$ of the single device as a function of the parameter σ_f when the receding horizon control is applied.

V. Conclusion

A new methodology for distributed control of storage that performs energy arbitrage is presented. The changes in the total power demand and energy price introduced by the devices are modelled by describing the storage population as a continuum and the problem is approached as a differential game with infinite players. Once the game is solved by the integration of coupled PDEs, an updated energy price can be communicated to the devices, which are able to calculate their optimal charge profile in a decentralized manner. The methodology is tested through simulations and then expanded to account for additional elements such as cyclic constraints on the state of charge and demand uncertainty.

References

[1] T. D. Voice, P. Vytelingum, S. D. R. A. Rogers, and N. R. Jennings, "Decentralised control of micro-storage in the smart grid," *AAAI-11: Twenty-Fifth Conference on Artificial Intelligence, 2011*, pp. 1421–1426.

[2] P. Vytelingum, T. D. Voice, S. D. Ramchurn, A. Rogers, and N. R. Jennings, "Theoretical and practical foundations of large-scale agent-based micro-storage in the smart grid," *Journal of Artificial Intelligence Research*, vol. 42, no. 1, pp. 765–813, 2011.

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

[4] L. Gan, U. Topcu, and S. H. Low, "Optimal decentralized protocol for electric vehicle charging," *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 940–951, 2013.

[5] D. Papadaskalopoulos and G. Strbac, "Decentralized participation of flexible demand in electricity markets," *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 3658–3666, 2013.

[6] M. Huang, P. E. Caines, and R. P. Malhamé, "Large population stochastic dynamic games: Closed-loop mckean-vlasov systems and the nash certainty equivalence principle," *Communications in Information and Systems*, vol. 6, no. 3, pp. 221–252, 2006.

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

[8] S. D. Howell, P. W. Duck, P. V. Johnson, G. Strbac, A. Hazel, N. Proudlove, and M. Black, "A pde system for modelling stochastic storage in physical and financial systems," *IMA Journal of Management Mathematics*, vol. 22, no. 3, pp. 231–252, 2011.

[9] R. Couillet, S. M. Perlaza, H. Tembine, and M. Debbah, "Electric vehicles in the smart grid: a mean field game analysis," *IEEE Transactions on Control Systems Technology*, vol. 21, no. 1, pp. 67–78, 2013.

[10] A. C. Kizilkale, S. Mannor, and P. E. Caines, "Large scale real-time bidding in the smart grid: A mean field framework," *Proceedings of the 51st IEEE Conference on Decision and Control, 2012*, pp. 3680 – 3687, 2013.

[11] W. H. Fleming and H. M. Soner, *Controlled Markov Processes and Viscosity Solutions*. New York: Springer, 2006, pp. 6–13.

[12] S. Wang, F. Gao, and K. L. Teo, "An upwind finite-difference method for the approximation of viscosity solutions to hamilton-jacobi-bellman equations," *IMA Journal of Mathematical Control and Information*, vol. 17, no. 2, pp. 167–178, 2000.

[13] R. J. LeVeque, *Numerical Methods for Conservation Laws*. Basel: Birkhauser-Verlag, 1990.

[14] National Grid. Metered half-hourly electricity demands. [Online]. Available: http://www.nationalgrid.com/uk/Electricity/Data/Demand+Data/

[15] P. Vytelingum, T. D. Voice, S. D. Ramchurn, A. Rogers, and N. R. Jennings, "Agent-based micro-storage management for the smart grid," *The Ninth International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2010)*, pp. 39–46.

[16] R. Doherty and M. O'Malley, "A new approach to quantify reserve demand in systems with significant installed wind capacity," *IEEE Transactions on Power Systems*, vol. 20, no. 2, pp. 587–595, 2005.

Antonio De Paola (SM12) received the B.Sc. and M.Sc. in Control Engineering from University of Rome Tor Vergata in 2008 and 2011. He is currently a Ph.D. student at Imperial College London, London, U.K. His research interests are optimal control, game theory and their application to smart grids.

David Angeli (F15) is a Reader in Stability of Nonlinear Systems at the Imperial College London. He received his Laurea and Ph.D. degrees from the University of Florence, Italy in 1996 and 2000, respectively. He was later appointed Assistant Professor within the Department of Systems and Computer Science of the same University and Associate Professor since 2005. He served as Associate Editor for the IEEE Transactions on Automatic Control and Automatica. He is an IEEE Fellow since January 2015. His research interests include: stability and dynamics of nonlinear systems and networks, control of constrained systems, biomolecular dynamics and control solutions for the smart grid.

Goran Strbac (M95) is a Professor of Electrical Energy Systems at Imperial College London, London, U.K. His research interests are in modelling and optimization of electricity system operation and investment, economics and pricing, including the integration of new forms of generation and demand side technologies.