Research

Author for correspondence:

Pierre Pinson

e-mail: ppin@dtu.dk

Electric demand response and bounded rationality – Mean-field control for large populations of heterogeneous bounded-rational agents

Andrea Marín Radoszynski[1], Pierre Pinson[1,2]

[1] Technical University of Denmark, Department of Electrical Engineering, 2800 Kgs. Lyngby, Denmark
[2] Technical University of Denmark, Department of Technology, Management and Economics, 2800 Kgs. Lyngby, Denmark

The increased penetration of renewable energy sources into existing power systems induces challenges in supply-demand balancing. Demand-side flexibility is seen as an option to accommodate variability and limited predictability from renewable energy generation. Heat pumps at residential level, if well coordinated, can be one of those flexibility sources. The complexity involved is high though, since their coordinated operation combines control, population effects, and the fact agents may actually not behave as rational decision-makers. We describe here a coordinated control framework that accounts for those aspects altogether. Decentralised model predictive control for large populations of heterogeneous agents is employed. As the cost to be minimised is affected by the population behaviour as a whole through the electricity price, the decentralised control is re-thought as a mean-field game. Existence and uniqueness of a Nash equilibrium are discussed while the Picard-Banach algorithm is used as a solution approach. It is extended to the case of bounded-rational agents. The impact on system dynamics of modelling agents as bounded rational is illustrated through numerical simulations.

1. Introduction

Energy systems, as well as our approach to their design and operations, are evolving rapidly and substantially. This comes with the deployment of renewable energy generation capacities, distributed energy resources more generally, e.g., with electric vehicles, heat pumps and battery storage units, as well as the liberalization of electricity markets. The role of consumers is also changing, at least by becoming more active agents on the demand side of electricity markets, if not by upgrading their assets in order to also produce and store electric energy (so-called "prosumers").

Demand side management is a concept that has been considered for various applications in power system operation and electricity markets – see [1] for instance. Roughly speaking, its objective is to incentivise, steer and reward demand-side flexibility mostly through different forms of financial incentives [2]. It therefore relies on the elasticity of electricity consumers, though other approaches based on e.g. behavioural science and behavioural economics have also been considered more recently [3]. Today, demand side management is seen as a way to support balancing of power systems with high shares of renewable energy sources. It is referred to as demand response when the aim is to affect electricity consumption in the short-term (time scales in the order of minutes to hours), in order to adapt to the availability of renewable energy generation, and to potentially compensate for forecast errors. For a recent review of the benefits and challenges of demand response, the reader is referred to [4].

Demand response relies on the interaction between a central coordination mechanism and the consumers with controllable loads. This leads to decentralised control, where agents make decisions based on both their local information, as well as partial information about the state of the system [5,6]. Decentralised control allows for decision authority and scalability, while in certain cases, it may also allow for privacy preservation. A key issue in systems with large numbers of responsive assets is to assess and accommodate their aggregate effect on system demand and resulting electricity prices. Game theory naturally comes as a relevant approach to study the interaction among agents within a demand response environment. However, it becomes intractable as the number of agents increases [7]. Aggregative games, and more specifically mean-field (often abbreviated MF) games are appealing, by assuming that the number of agents tends towards infinity. In mean-field games, the decisions of each agent depend on the average state of the population instead of on one-to-one interactions. Extensive overviews of mean-field games and their applications can be found in [8,9]. Mean-field theory has been applied to many distributed optimization problems in energy systems, e.g., electric vehicle charging [10], demand response for populations of thermostatically controlled loads [11,12], control of storage devices [13], and congestion control for networks of shared resources [14].

A common and important assumption is that agents are rational decision-makers. This is most likely no longer valid when agents are human though. Human decision-makers are bounded rational and deviate from the optimal response described in normative economic theories [15]. The field of behavioural economics aims at unveiling and better appraising human biases in decision-making. Considering energy applications, in addition to cognitive limitations, decision-making normally occurs under time constraints and incomplete information, leading to systematic biases and sub-optimal decision-making [16]. As of today, even if home energy management systems automatize the interaction with home appliances, the consumer is able to change various settings, e.g., temperature setpoints and discomfort bounds, repeatedly – hence deviating from previously computed optimal scheduling. The way energy consumers interact with and utilize "smart appliances" is still fairly unclear [17]. For instance, a study in the USA monitored the routines of 35.471 programmable thermostats in four main utilities, showing that only 47% actually utilize the programmable options. The other half of the programmable thermostats were used as manual thermostats, since the users did not use the "scheduling" option [18]. Similarly, another more recent study on individual occupant comfort using heating setpoints [19] showed that some occupants were not fully familiarized with the control of the floor heating

system. Thus, they repeatedly changed floor heating set points in order to achieve the desired comfort, eventually deviating optimal usage trajectories. In view of the evolution of energy systems with a more proactive role of consumers, it is of utmost importance to better account for behaviour in the design of mechanisms and operations [20]. An unaddressed problem in the demand response literature is whether bounded rationality impacts equilibria and corresponding agent decisions.

Our main contribution here is hence to show how bounded rationality may be accounted for within a decentralised model predictive control setup for demand response seen as a mean-field game, as well as its impact on equilibria. The existence and uniqueness of a Nash equilibrium are preserved when generalizing to this bounded rationality case, while it allows to simulate different forms of bounded rationality. Note, however, that we are not aiming to hedge against bounded rationality as could be done in a robust optimization framework [21]. Similarly, we are not looking at how bounded rationality specifically affects decisions (here for demand response), e.g., in the way the sparse-max model of [22] is employed.

The coordination of a large population of heat pumps through dynamic pricing is used as a representative problem, with interesting complexities coming from constraints on system state (i.e., indoor temperature within dwellings) and the heat pumps themselves (power setpoints and ramping constraints). Network-related aspects are not considered in our problem setup, which is in line with a broad range of use cases related to demand response, e.g., peak shaving and balancing support, for which grid constraints are not accommodated. Electricity consumers are increasingly facing dynamic pricing of electric energy, whether grid-related constraints are considered or not. We adopt a mean-field decentralised control approach in which agents respond optimally to common price sequences issued by a non-profit central coordinator. These depend both on the predicted inelastic demand and total generation capacity, as-well as on the overall population flexible consumption. The objective is to find the price sequence that drives the population to the Nash equilibrium. The mean-field control of thermostatically controlled loads is performed through model predictive control (often abbreviated MPC) at an agent-based level and on a rolling-window setup, hence allowing to adapt to information updates. Agents update their optimal control trajectories every time forecasts are revised. We eventually relax the rational agent assumption and assume that agents can deviate from their optimal (if rational) trajectory. We then characterise the resulting equilibria and compare them with the ideal rational case based on simulations. The existence and uniqueness of these equilibria are proven using tools from variational inequalities. The Nash equilibria are obtained using algorithms proposed in [23] for deterministic mean-field games with heterogeneous convex constraints. Finally, our approach is applied to different case studies where we explore the impact of bounded rationality and forecast accuracy on decisions and system outcomes.

The remaining of the paper is organised as follows: Section 2 describes the decentralised model predictive control setup and related mean-field game. Section 3 covers the relevant tools from variational inequalities to study the Nash Equilibrium, as well as the solution approach. Section 4 presents the concept of bounded rationality in demand response and the way it is considered in the present work. Section 5 gathers results from numerical simulations to illustrate and analyse the impact of bounded rationality on demand response within our decentralised model predictive control framework. Finally, Section 6 gathers a set of conclusions and perspectives for future work.

2. From decentralised model predictive control to the related mean-field game

We first present the coordination problem as a whole. The local optimization problems at agent level are then described, to eventually close with a discussion of the price determination, which makes the problem a mean-field game. All agents are so far considered rational decision-makers.

(a) Coordination problem

Let us consider the problem of coordinating a large population of N heat pumps through dynamic pricing (with N large). Each heat pump is sized according to the heat demand of the corresponding household – the heat pump design problem is not considered here, as emphasis is placed on operational problems only. A heat pump converts electricity into heat according to its coefficient of performance. The heat dynamics of the building allow to partly decouple electricity and heat in time, i.e. the building and air masses have enough thermal inertia to shift the heat peak in time to an hour when the electricity price is cheaper. Heat pumps then provide flexibility to the electricity system side, which may help to accommodate variability and limited predictability of renewable power generation, or different types of power system contingencies.

Being at a given time w within a simulation period of interest $\mathcal{W}$, a forward-looking time window with τ lead times (referred to as the planning horizon) is considered in order to account for the time-binding constraints of the heat pumps, as well as the temporal dynamics of the indoor temperature in dwellings. In practice for instance, if aiming to steer the population of heat pumps over the next 6 hours with a 15-minute temporal resolution, this yields 24 lead times. This also yields τ prices p_t, one for each lead time t, $t \in \mathcal{T} = \{0, \ldots, \tau - 1\}$. When relevant, this price sequence will be summarized by $\boldsymbol{p}$, with $\boldsymbol{p} = [p_0, \ldots, p_{\tau-1}]^\top$. In a model predictive control framework with rolling window, however, only the decisions at lead time $t = 0$ are implemented and binding – decisions being (i) the electricity price for the central coordinator, and (ii) the electricity consumption of the heat pump at the agent level. Then, the window is shifted by one step, and the process continues.

Each household equipped with a heat pump is considered as an agent in the game, with index i. The setup is illustrated in Figure 1. Agents compete to procure electricity to cover their heat demand in the most cost-effective way – we are dealing with a non-cooperative game. This is while the population of electricity consumers is being coordinated by a non-profit central entity through prices. That central coordinator could be an aggregator, the supervisory node of an energy community [24] or a local electricity market. For a given agent i, $i \in \mathcal{N} = \{1, \ldots, N\}$, electricity demand is the sum of an inflexible (and must-serve) component d_i and of a flexible (and elastic) component u_i. The average inflexible demand (over the population of N agents) is an input to the problem of the central coordinator, in the form of a series of load forecasts $\hat{\boldsymbol{d}} = [\hat{d}_0, \ldots, \hat{d}_{\tau-1}]^\top$ for the coming τ lead times. The reason why one works with averages instead of sums is due to the mean-field formulation that will be explained towards the end of this section. The other input to the central coordinator is κ, the total generation capacity scaled by the number of agents N.

For simplicity, the flexible component only relates to heat pumps. This may be readily extended to more advanced setups with additional flexible appliances, though at the cost of modelling complexity. In addition, having appliances that require integer variables in the modelling of their operations would affect convexity of the problems involved, and hence the results that we describe. This could potentially be accommodated through some form of convex relaxation, though not considered here and left for future work.

At the agent level, decisions are affected by the local system dynamics, but also by the prices p_t, $t \in \mathcal{T}$, communicated by the central coordinator. By facing dynamic prices, each agent has the incentive to adjust her consumption profile according to her inherent flexibility. Following a decentralised model predictive control approach, each agent i computes her optimal heat pump consumption $\boldsymbol{u}_i = [u_i^0, \ldots, u_i^{\tau-1}]^\top$ throughout a common planning horizon and communicates it to the central coordinator. Based on the aggregated response of the population, the central coordinator updates the price sequence $\boldsymbol{p}$ and communicates them to all agents. This process is repeated until a Nash equilibrium is reached, i.e., where no agent has an incentive to deviate unilaterally from her strategy.
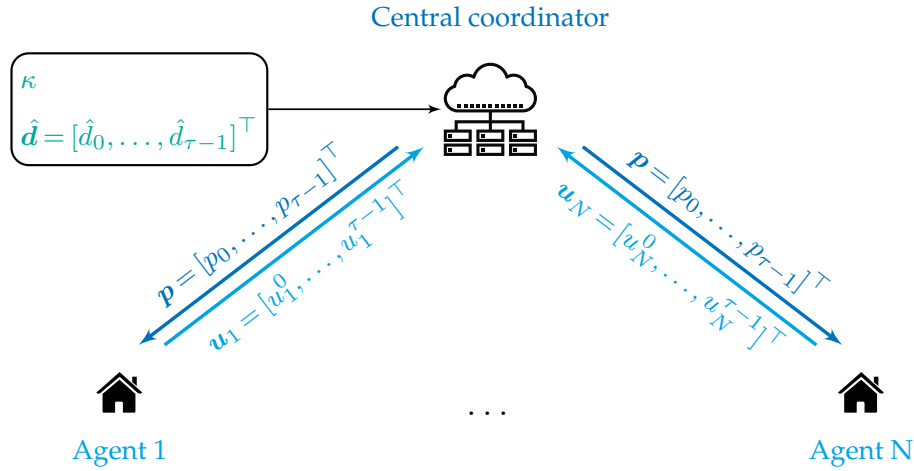
Figure 1. Setup and interfaces between the central coordinator and the population of agents.

(b) Local optimization problem

Assume that the heat pump dynamics can be described by a deterministic discrete-time linear model, so that

$$s_i^{t+1} = A_i^t s_i^t + B_i^t u_i^t + C_i^t, \quad \forall t \in \mathcal{T}, \ i \in \mathcal{N}, \tag{2.1}$$

where $s_i^t \in \mathcal{S}_i^t$ and $u_i^t \in \mathcal{U}_i^t$ are the state and input (or, decision) variables, respectively, of agent i at time t. From now on, we refer to $\boldsymbol{s}_i = [s_i^1, \ldots, s_i^\tau]^\top$ and $\boldsymbol{u}_i = [u_i^0, \ldots, u_i^{\tau-1}]^\top$ as the sequences of states and input variables along the planning horizon. Note that in the dynamical system model (2.1), the constant C_i^t summarizes the impact of exogenous ambient conditions at time t, which also drive that dynamical system. In the simplest case, C_i^t is time-invariant and representing the long-term indoor temperature when heat pumps are off. In the more general case, C_i^t is a linear function of meteorological drivers, e.g., outdoor temperature, solar irradiance and wind, parameterized by the building characteristics. The detailed formulation of heat dynamics model for heat pump control that we eventually used can be found in [25] and references therein. Finally, for an extensive coverage of more general dynamical models for the heat dynamics of buildings, the reader is referred to [26] among others.

In the above, the state variable represents the indoor temperature, which is bounded by predefined agent comfort setpoints $[\underline{s}_i, \overline{s}_i]$, also referred to as comfort bounds. For instance, the indoor temperature must be within $\underline{s}_i = 19°C$ and $\overline{s}_i = 22°C$. A_i^t, B_i^t and C_i^t are the coefficients that describe the heat dynamics of the dwelling, considering thermal resistances, heat capacities, coefficient of performance, temperature of the building envelope, ambient temperature, window area and solar irradiance. The initial indoor temperature s_i^0 is known and the objective of each agent is to keep it close to a tracking temperature $\tilde{s}_i$ set by the agent throughout the planning horizon. The tracking temperature is considered time-invariant, but the time-varying case could be similarly handled. At each lead time, the control has to decide the power u_i^t it should consume in order to keep the indoor temperature within the comfort bounds $[\underline{s}_i, \overline{s}_i]$ and as close as possible to the tracking temperature. The sequence of input variables $\boldsymbol{u}_i$ is also constrained by the heat pump capacity c_i, and ramping limits $\underline{\delta u}_i, \overline{\delta u}_i$, defining the feasibility set $\mathcal{U}_i$, i.e.,

$$\mathcal{U}_i = \left\{ \boldsymbol{u}_i \in \mathbb{R}^\tau \mid u_i^t \in [0, c_i], \ \underline{\delta u}_i \leq u_i^{t+1} - u_i^t \leq \overline{\delta u}_i \right\}. \tag{2.2}$$

Consequently, the strategy of agent $i \in \mathcal{N}$ gathers both state and input variables within a single vector $\boldsymbol{x}_i = [\boldsymbol{s}_i; \boldsymbol{u}_i]$. Those are necessarily linked to the dynamics of the system. The strategy of

each agent $i \in \mathcal{N}$ belongs to the convex and compact strategy set

$$K_i = \left\{ [s_i; u_i] \in \mathbb{R}^{2\tau} \mid s_i^{t+1} = A_i^t s_i^t + B_i^t u_i^t + C_i^t, \ s_i^t \in [\underline{s}_i, \overline{s}_i], \ u_i \in \mathcal{U}_i \right\}, \tag{2.3}$$

described by the model for the dynamical system, and the natural boundaries for the state and input variables.

Each agent i is commonly assumed rational, hence willing to find the strategy x_i that minimises the cost function J_i given that all other agents are consuming according to $x_{-i} = \{x_j\}_{j \in \mathcal{N} \setminus \{i\}}$. This is the aspect we will aim at relaxing in a later part of this work. In practice x_{-i} is obviously not known to agent i, but its effect is summarized through the price sequence broadcast by the central coordinator. The cost function of agent $i \in \mathcal{N}$ is defined as

$$J_i(x_i, x_{-i}) = \sum_{t=0}^{\tau-1} \left(p_t u_i^t + \alpha_i (s_i^{t+1} - \tilde{s}_i)^2 \right). \tag{2.4}$$

We use a game-theoretical notation here since, as will be further described in the next paragraph, the decisions of agent i are necessarily influenced by those of all other agents j, $j \neq i$, through the sequence of electricity prices. That cost function is the sum of two terms: the electricity procurement cost and the discomfort term. This discomfort term penalizes the (squared) deviations of the indoor temperature from the tracking temperature by a factor $\alpha_i > 0$. This part of the loss function is based on local information only, and does not relate to the other agents. In contrast, for the calculation of the electricity procurement cost, one needs the sequence of electricity prices p_t, $t \in \mathcal{T}$. This sequence is communicated by the central coordinator, and is a function of the decisions of the population as a whole.

Eventually, the local optimization problem to be solved by agent i, every time a price sequence is communicated, is

$$\min_{x_i \in K_i} \quad J_i(x_i, x_{-i}). \tag{2.5}$$

The solution of this problem is denoted $x_i^\star(p)$, as it readily depends on the price sequence communicated by the central coordinator.

(c) Price determination and mean-field game

The electricity prices p_t, $t \in \mathcal{T}$, are calculated based on the predicted inelastic demand $\sum_i \hat{d}_i^t$, on the total generation capacity of the system κ_N, and on the aggregated consumption of the heat pumps ($\sum_i u_i^t$). We follow here an approach similar to that used in [10]. The electricity price p_t is assumed to be a linear function of the ratio between total demand and total generation capacity κ_N at time t. The total demand is the sum of the inflexible demand of the system and the aggregated demand from the population of heat pumps. The total generation capacity is a function of the population size N (hence the use of the index N for the related variable). This assumption of a linear relationship may be deemed acceptable if considering that the energy prices received by the consumers are directly proportional to wholesale energy prices, and overlooking temporal constraints. And, even if it was not linear, that assumption could be justified by the idea of a local linearization around an operating point in the market.

Consequently, the electricity price at time $t \in \mathcal{T}$ can be expressed as

$$p_t = \frac{\beta}{\kappa_N} \sum_{i=1}^{N} \left(\hat{d}_i^t + u_i^t \right), \tag{2.6}$$

where $\beta > 0$ is the coefficient for the linear function, $\hat{d}_i^t$ is the predicted inflexible demand for agent i at time t, u_i^t is the flexible demand from the heat pump of agent i at time t and κ_N is the total generation capacity.

The basic idea when rethinking this problem setup as a mean-field game is to consider the case where N tends towards infinity. There, one has

$$\lim_{N\to\infty} \frac{\kappa_N}{N} = \kappa\,, \tag{2.7}$$

i.e., asymptotically the total generation capacity is the scaled value κ, while

$$\lim_{N\to\infty} \frac{\sum_{i=1}^{N} \hat{d}_i^t}{N} = \hat{d}_t\,, \quad \forall t. \tag{2.8}$$

This goes along common practice, which is that, instead of predicting consumption for each household individually, one aims at predicting the total consumption of these households readily (here, in a scaled version). Finally one can also define $\bar{u}_t$ as

$$\bar{u}_t = \lim_{N\to\infty} \frac{\sum_{i=1}^{N} u_i^t}{N}\,, \quad \forall t, \tag{2.9}$$

so that the sequence $\bar{u} = [\bar{u}_0, \ldots, \bar{u}_{\tau-1}]$ is the average electricity consumption profile for heat pumps across the whole population. Then, the individual profiles are to be seen as local deviations due to variations in initial states, constraints, etc.

Subsequently, considering that N is large enough, the electricity price at time $t \in \mathcal{T}$ from (2.6) is approximated as

$$p_t = \frac{\beta}{\kappa} \left(\hat{d}_t + \bar{u}_t \right). \tag{2.10}$$

The optimal decisions of all agents involved are coupled by the common electricity price resulting in a deterministic mean-field game: the decision of each agent depends on the mean input variable across the whole population [23]. We refer to this mean-field game as $\mathcal{G}_A$. At equilibrium, we write $\boldsymbol{p}^\star = [p_0^\star, \ldots, p_{\tau-1}^\star]^\top$ the optimal price sequence, and $\boldsymbol{x}_i^\star(\boldsymbol{p}^\star)$ the optimal strategy of agent i, $i \in \mathcal{N}$. $\boldsymbol{x}^\star(\boldsymbol{p}^\star)$ then concatenates the optimal strategies of all agents.

3. Theoretical results and solution approach

Since the setup considered here is different from what can be found in the literature for MF games for, e.g., fleets of electric vehicles [27] and populations of thermostatically-controlled loads [12], we first aim here at showing existence and uniqueness of Nash equilibrium for the problem at hand, based on known techniques to connect mean field games and variational inequalities. We then describe an adequate solution approach to obtain it.

(a) Existence and uniqueness of a Nash equilibrium

We show existence and uniqueness of a Nash equilibrium by connecting the solution of the mean field game $\mathcal{G}_A$ to the solution of variational inequalities. This means that, independently of the agent heterogeneity, there exists a price sequence $\boldsymbol{p}$ that drives the population to a unique mean consumption profile, based on the individual optimal strategies $\boldsymbol{x}_i^\star$ of all agents. This consumption profile is a Nash equilibrium, implying that each agent does not have an incentive to deviate from her electricity consumption decisions, given the electricity consumption decisions of the other agents. If this price sequence can be computed, it can be used to perform forward and real-time control of heat pump populations. Necessarily, this solution is also a social welfare maximiser, in the sense that it minimises the overall procurement costs for the population of electricity consumers.

Let us first define the variational inequality problem. In the following we write $K = \prod_i K_i$ the cartesian product of the strategy sets of all agents. And, to lighten notations, we use $m = 2\tau N$ as the dimension of K.

Definition 1. *Consider a mapping $F : K \to \mathbb{R}^m$, given a set $K \subseteq \mathbb{R}^m$. A solution set $SOL(K, F)$ to the variational inequality problem $VI(K, F)$ is a vector $\boldsymbol{x}^\star \in K$ such that $\langle F(\boldsymbol{x}^\star), \boldsymbol{x} - \boldsymbol{x}^\star \rangle \geq 0, \, \forall \boldsymbol{x} \in K$.*

A key result that we are to exploit, for instance found in [7], is that non-cooperative games may be readily connected to the solution of a variational inequality problem. Let us write K_{-i} the cartesian product of the strategy sets of all agents except for agent i, i.e., $K_{-i} = \prod_{j \in \mathcal{N} \setminus \{i\}} K_j$.

Theorem 1. *(Nash equilibria and variational inequalities, Theorem 2 in [7]). Suppose that, for the game $\mathcal{G}_A$, the non-empty strategy sets K_i are closed and convex, and that for every fixed $\boldsymbol{x}_{-i} \in K_{-i}$ the payoff functions $J_i(\boldsymbol{x}_i, \boldsymbol{x}_{-i})$ are convex and continuously differentiable in $\boldsymbol{x}_i \in K_i$. Then a point $\boldsymbol{x}^\star \in K$ is a Nash equilibrium for $\mathcal{G}_A$ if and only if it is a solution of the variational inequality $VI(K, F)$, with $K = \prod_i K_i$ and $F(\boldsymbol{x}) = \left(\nabla_{\boldsymbol{x}_i} J_i(\boldsymbol{x}_i, \boldsymbol{x}_{-i})_{i=1}^N \right)$.*

In view of the problem setup described previously, our mean field game $\mathcal{G}_A$ can be written as a variational inequality problem. We consequently use two results from [28] to show existence and uniqueness of the solution of the $VI(K, F)$ problem and, by Theorem 1, to the mean field game $\mathcal{G}_A$.

Theorem 2. *(Corollary 2.2.5 in [28]). Suppose that K is a compact and convex set and that the mapping F is continuous. Then, the set $SOL(K, F)$ is non-empty and compact.*

Theorem 3. *(Theorem 1.3.1 in [28]). Let $F : U \to \mathbb{R}$ be continuously differentiable on the open convex set $U \subseteq \mathbb{R}$. The following three statements are equivalent: (a) there exists a real-valued function θ such that $F(\boldsymbol{x}) = \nabla \theta(\boldsymbol{x}), \, \forall \boldsymbol{x} \in U$; (b) the Jacobian matrix of $F(\boldsymbol{x})$ is symmetric $\forall \boldsymbol{x} \in U$; (c) F is integrable on U.*

For our mean-field game $\mathcal{G}_A$, readily using Theorem 2, we first deduce that the solution to our $VI(K, F)$ problem exists. That solution is a Nash equilibrium for the mean-field game $\mathcal{G}_A$.

Proposition 1. *For the non-cooperative mean field game $\mathcal{G}_A$ related to the coordination of a population of heat pumps within a demand-response framework, a Nash equilibrium exists.*

Proof. The strategy set of each agent K_i is formed by the upper and lower temperature of the dwelling, as well as the power bounds and ramping constraints of the heat pump. This satisfies the compactness and convexity of K_i, and so it does for the strategy set of the game $K = \prod_i K_i$. The vector $\boldsymbol{x}$ stands for the decision variables of all agents stacked together $\boldsymbol{x} = [\boldsymbol{x}_1, \dots, \boldsymbol{x}_N]^\top, \boldsymbol{x} \in \mathbb{R}^m$. Furthermore the map $F(\boldsymbol{x}) \in \mathbb{R}^m$ is continuous, since the cost functions J_i are quadratic and thus, continuous over $\boldsymbol{x}$ and convex over $\boldsymbol{x}_i$ for fixed values of $\boldsymbol{x}_{-i} \in K_{-i}$. Therefore, the solution of the game $\mathcal{G}_A$ exists by Theorem 2. $\square$

Following Theorem 3, we rewrite the variational inequality problem as an optimization problem and study its properties to eventually deduce the uniqueness of the Nash equilibrium.

Proposition 2. *For the non-cooperative mean field game $\mathcal{G}_A$ related to the coordination of a population of heat pumps within a demand-response framework, the Nash equilibrium is unique.*

Proof. If the conditions of Theorem 2 are satisfied, the game $\mathcal{G}_A$ can be expressed as a $VI(K, F)$ problem. Following Theorem 3 we can write an equivalent optimization problem that solves $VI(K, F)$. To show that the solution to the $VI(K, F)$ is unique (and by equivalence, uniqueness of

Nash equilibria for $\mathcal{G}_A$), we compute the Jacobian $\nabla F(\boldsymbol{x}) \in \mathbb{R}^{m \times m}$ as

$$
\begin{bmatrix}
\operatorname{diag}(2\alpha_1) & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots \\
\mathbf{0} & \operatorname{diag}\left(\frac{2\beta}{\kappa N}\right) & \mathbf{0} & \operatorname{diag}\left(\frac{\beta}{\kappa N}\right) & \mathbf{0} & \operatorname{diag}\left(\frac{\beta}{\kappa N}\right) & \dots \\
\mathbf{0} & \mathbf{0} & \operatorname{diag}(2\alpha_2) & \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots \\
\mathbf{0} & \operatorname{diag}\left(\frac{\beta}{\kappa N}\right) & \mathbf{0} & \operatorname{diag}\left(\frac{2\beta}{\kappa N}\right) & \mathbf{0} & \operatorname{diag}\left(\frac{\beta}{\kappa N}\right) & \dots \\
\mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \operatorname{diag}(2\alpha_3) & \mathbf{0} & \dots \\
\mathbf{0} & \operatorname{diag}\left(\frac{\beta}{\kappa N}\right) & \mathbf{0} & \operatorname{diag}\left(\frac{\beta}{\kappa N}\right) & \mathbf{0} & \operatorname{diag}\left(\frac{2\beta}{\kappa N}\right) & \dots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{bmatrix}.
$$

Since the Jacobian of F is symmetric and condition (b,c) of Theorem 3 holds, the complete version of Theorem 1.3.1 in [28, p. 14] states that there exists a scalar function $\theta(\boldsymbol{x})$ for an arbitrary vector $\boldsymbol{x}^0$ in the open convex set $U \in \mathbb{R}^m$ given by

$$
\begin{aligned}
\theta(\boldsymbol{x}) &= \int_0^1 F(\boldsymbol{x}^0 + v(\boldsymbol{x} - \boldsymbol{x}^0))^\top (\boldsymbol{x} - \boldsymbol{x}^0) dv \\
&= \int_0^1 \begin{pmatrix} 2\alpha_i(v\boldsymbol{s}_i - \tilde{\boldsymbol{s}}_i) \\ \frac{\beta}{\kappa N} v \boldsymbol{u}_i + \frac{\beta}{\kappa N} \sum_{j=1}^N v \boldsymbol{u}_j + \frac{\beta}{\kappa} \hat{\boldsymbol{d}} \end{pmatrix}_{i=1,N}^\top \begin{pmatrix} \boldsymbol{s}_i \\ \boldsymbol{u}_i \end{pmatrix}_{i=1,N} dv \\
&= \int_0^1 \sum_{i=1}^N \Bigg(-2\alpha_i \sum_{t=1}^\tau \tilde{s}_i s_i^t + \frac{\beta}{\kappa} \sum_{t=0}^{\tau-1} \hat{d}_t u_i^t \\
&\qquad\qquad + \Bigg[2\alpha_i \sum_{t=1}^\tau (s_i^t)^2 + \sum_{t=0}^{\tau-1} \frac{\beta}{\kappa N} \Big((u_i^t)^2 + u_i^t \sum_{j=1}^N u_j^t \Big) \Bigg] v \Bigg) dv \\
&= -2 \sum_{t=1}^\tau \sum_{i=1}^N \alpha_i \tilde{s}_i s_i^t + \frac{\beta}{\kappa} \sum_{t=0}^{\tau-1} \sum_{i=1}^N \hat{d}_t u_i^t + \sum_{t=1}^\tau \sum_{i=1}^N \alpha_i (s_i^t)^2 + \sum_{t=0}^{\tau-1} \frac{\beta}{2\kappa N} \Big(\sum_{i=1}^N u_i^t \Big)^2,
\end{aligned}
$$

which then results in the following optimization problem:

$$
\min_{\boldsymbol{x} \in K} \quad \left(\sum_{t=1}^\tau \sum_{i=1}^N \alpha_i \Big(-2\tilde{s}_i s_i^t + (s_i^t)^2 \Big) + \frac{\beta}{\kappa} \sum_{t=0}^{\tau-1} \Big(\sum_{i=1}^N \hat{d}_t u_i^t + \frac{1}{2N} \Big(\sum_{i=1}^N u_i^t \Big)^2 \Big) \right). \tag{3.1}
$$

Problem (3.1) has a unique solution due to its convex and compact constraint set, and strict convexity of the objective function, being the sum of strictly convex functions [29]. Since the solution of (3.1) defines the solution set $\mathrm{SOL}(K, F)$, the solution of the original mean field game $\mathcal{G}_A$ is also unique. $\qquad \square$

(b) Solution approach

At the agent level, a requirement for computing the optimal strategy is that each agent i has access to the electricity price, which is broadcast by the central coordinator. We use this information structure because it allows agents to compute their optimal strategy without having access to other agent states or aggregation parameters. They all respond to a macroscopic incentive $\boldsymbol{p}$, the price sequence, that accounts for their contribution $\boldsymbol{u}_i$. Therefore, the objective of the mean-field control is to find the price sequence $\boldsymbol{p}^\star$, such that the set of resulting strategies $\{\boldsymbol{x}_i^\star\}_{i \in \mathcal{N}}$ of the individual agents yield an almost mean-field Nash equilibrium of the game $\mathcal{G}_A$. The definition of an almost mean-field Nash equilibrium (also referred to as ε-Nash equilibrium) can be found in [23]. Basically, it slightly relaxes the definition of a Nash equilibrium for the case of mean-field games. For a mean-field ε-Nash equilibrium, agent i has a benefit of maximum ε (with ε small) if altering her own strategy, given the strategies of all other agents. For $\varepsilon = 0$ one obtains a classical Nash equilibrium.

The solution approach that is used here is inspired from the developments in [23], and references therein. As the population size N tends towards infinity (or at least, with N large enough), the mean-field Nash equilibrium becomes a mean-field ε-Nash Equilibrium with $\varepsilon = \mathcal{O}(1/N)$. In order to compute the price sequence that yields the ε-Nash equilibrium both Picard-Banach and Mann algorithms could be seen as relevant. The Picard-Banach algorithm was already used in [10] to compute optimal price signals for the decentralised charging of fleets of electric vehicles. Even-though our problem does not strictly fulfil the conditions of the Picard-Banach algorithm for finding a fixed point, we have observed from numerical simulations that it actually converged, and even faster than the Mann algorithm. Therefore, we only describe and focus on the Picard-Banach algorithm here. Further theoretical analysis should be performed to understand why it converges and performs so well with our problem structure.

Remember that when describing the problem setup and related game, emphasis was placed on a planning horizon with τ time steps. However, the start of this planning horizon is at a given time $w \in \mathcal{W}$, with $\mathcal{W}$ the simulation period. The problem is not solved a single time only. It is solved successively on sliding windows. For a given time $w \in \mathcal{W}$ even if the problem is solved for the whole period $[w, \ldots, w + \tau - 1]$, only the decisions for time w is implemented (or "binding") – i.e., both price at the central coordinator level, and electricity consumption for heat pumps at the agent level. Then, the window is moved by one step with new information available, and the problem is solved for that new window. The process iterates this way until the end of the simulation period $\mathcal{W}$.

Because we place ourselves at a given time w that conditions the information available, state of the system, etc., this is here reflected in our notations with "$\cdot_{|w}$" indicating that the related variable is conditioned by the information available at time w. The solution approach for the decentralised mean-field model predictive control (hence, the MPC described) consists of the following steps:

(MPC1) The central coordinator issues (or receives from a vendor) a revised forecast of the inflexible demand $\hat{\boldsymbol{d}}_{|w}$ over the planning horizon $\mathcal{T}$. Based on the forecast and initial estimates of the heat pump consumption $\boldsymbol{u}_{i|w}$, the central coordinator calculates the price sequence $\boldsymbol{p}_{|w}$ following (2.10) and broadcast it to the agents. These initial estimates $\boldsymbol{u}_{i|w}$ are available from the solution at previous time step $w - 1$;

(MPC2) Each agent computes her own optimal strategy $\boldsymbol{x}_{i|w}^{\star}(\boldsymbol{p})$ over the planning horizon by minimizing her cost function (2.4), and then communicates the optimal heat pump consumption $\boldsymbol{u}_{i|w}^{\star}(\boldsymbol{p})$ to the central coordinator.

(MPC3) The central coordinator collects the individual strategies and updates the price signals $\boldsymbol{p}_{|w}$. The updated prices are broadcast to all agents.

(MPC4) Repeat (MPC2) and (MPC3) until the optimal consumption of the heat pumps no longer changes (based on a chosen convergence criterion).

(MPC5) Each agent implements only her control decision $u_{i|w}^{\star,0}(\boldsymbol{p}^{\star})$ at timestamp w of the MPC. The state variable evolves following (2.1).

(MPC6) Set $w = w + 1$ and repeat steps (MPC1)-(MPC5) until reaching the end of the simulation window.

This solution approach is presented in algorithmic form in Algorithm 1.

4. Accounting for bounded rationality

After reviewing various aspects of bounded rationality that are of relevance to demand response, we look at a simple, though generic, approach based on the perturbation of the agent optimal strategies. It allows looking at the effect of bounded rationality on outcomes, without having to invest into complex and possibly intractable models.

Algorithm 1: Picard-Banach decentralised model predictive control

---

1 Choose ϵ_{stop}

2 Initialise $s_{i|0}^0, \forall i \in \mathcal{N}$

3 for $w = 1, \ldots, n(\mathcal{W})$ **do**

4 $\quad$ Initialise $\epsilon > \epsilon_{\text{stop}}$

5 $\quad$ Set $k = 1$

6 $\quad$ Calculate $p_{|w}^k$ with $\hat{d}_{|w}$ and $\overline{u}_{|w}^k = \frac{1}{N} \sum_{i=1}^N u_{i|w}^k$

7 $\quad$ **while** $\epsilon > \epsilon_{stop}$ **do**

8 $\quad\quad$ Broadcast $p_{|w}^k$

9 $\quad\quad$ Compute $x_{i|w}^\star(p_{|w}^k) = [s_{i|w}^\star(p_{|w}^k); u_{i|w}^\star(p_{|w}^k)]$ by solving (2.5), $\forall i \in \mathcal{N}$

10 $\quad\quad$ Compute $\overline{u}_{|w}^{k+1} = \frac{1}{N} \sum_{i=1}^N u_{i|w}^\star(p_{|w}^k)$

11 $\quad\quad$ Update $p_{|w}^{k+1}$ using (2.10)

12 $\quad\quad$ Update $\epsilon \leftarrow \left\| \overline{u}_{|w}^{k+1} - \overline{u}_{|w}^k \right\|_2$

13 $\quad\quad$ $k \leftarrow k + 1$

14 $\quad$ Implement $u_{i|w}^{\star,0}$ for time step w, $\forall i \in \mathcal{N}$

15 $\quad$ Compute $s_{i|w}^1$ as in (2.1) and set $s_{i|w+1}^0 = s_{i|w}^1$, $\forall i \in \mathcal{N}$

(a) Bounded rationality in demand response

Different strands in behavioural economics have established a new framework to understand human decision-making, as well as its biases, by fundamentally questioning the rational behaviour of human decision-makers. Based on a large collection of experimental studies, behavioural economists and psychologists have shown that decision-makers are not rational, as normative models of utility maximization would predict [15]. These models assume that decision-makers have full information and can compute optimal decisions. Conversely, experimental evidence shows that decision-makers do not always have known, ordered, and consistent preferences, leading to sub-optimal decisions and systematic biases [16]. Typically decision-making occurs under time constraints and incomplete knowledge. Processing capabilities of decision-makers are restricted by cognitive limitations. Decision-makers are therefore bounded rational, using heuristics and mental shortcuts, specially in situations characterised by high level of complexity and uncertainty [30]. The literature on behavioural economics is now very rich, while easily communicating human biases in decision-making to a broad audience – see [31] for instance.

Thus far, demand response programs that assume consumer rationality do not rely on realistic views of energy consumers. Economists argue that existing theories do not represent energy-related behaviour in the residential sector [32]. In general, consumer information regarding energy use is incomplete and systematically incorrect, e.g., people tend to overestimate the amount of energy they use and the amount that can be saved by adopting energy-efficient technologies [33]. There are also often inconsistencies between observable behaviour and individuals self-reported knowledge, values, attitudes, and intentions [16].

A few examples of such bounded rationality include:

Time inconsistency: decision-makers show a preference towards immediate outcomes. When valuing costs and benefits over time, they apply hyperbolic discounting rather than a constant discount rate as described by normative economic theory. In this way, decisions in the future are farsighted while decisions in the present are shortsighted. This immediacy effect gives rise to time-inconsistencies, e.g. not prioritising time and money to purchase new energy-efficient appliances [16];

Anchoring on defaults: there is a tendency to maintain the status-quo with resistance to change even if alternatives may yield better outcomes. Decision-makers generally choose the default option, specially when the amount of information or complexity increases [16]. An example in the residential energy sector is the selection of an electricity retailer: consumers tend to stick to the default option even if there are better deals in the market;

Social comparisons: decisions are affected to a large extent by social norms and behaviours of others. There is also a tendency to form social comparisons and to evaluate performance or well-being in relative terms [16]. A good example is the result of the campaign carried out by an electricity retailer who managed to influence consumers behaviour by comparing their energy use to their neighbours.

In all cases, the basic idea of bounded rationality when used in decision-making in relation to demand response is that there is going to be a deviation from the optimal decision one would make if being rational (having complete information, enough time and computing power, etc.). The characteristics of that deviation may depend on the type of bounded rationality e.g. in terms of bias and variance of such deviations, temporal and population-based structure, etc. It is intuitively expected that it leads to some form of sub-optimality in terms of overall social welfare – hence, higher electricity procurement costs overall for our demand response problem.

(b) Bounded rationality in the decentralised mean-field model predictive control

There are several models proposed by behavioural economists that describe bounded rationality in decision-making. However, it becomes challenging to adapt them to engineering applications such as demand response, for them to be included into decentralised control problems like the one considered here. Our approach to account for bounded rationality hence does not consist in explicitly modeling it in its different forms, but instead in perturbing optimal agent strategies. More advanced approaches should be considered in the future, e.g., by re-thinking the local optimization problem (2.5) using prospect theory (see [34] and references therein) or similar.

We can assume that if agents are bounded rational they will deviate from their optimal response by a small amount ξ. We use this ξ to simulate the gap between the optimal strategy obtained by solving the local optimization problem (2.5) and the actual response of the agents. Focusing on a given time step w of the simulation period $\mathcal{W}$, based on the outcome of (2.5), the actual implemented control $\tilde{u}_{i|w}^0$ for the heat pump is

$$\tilde{u}_{i|w}^0 = u_{i|w}^{\star,0} + \xi_{i|w}, \quad \forall i \in \mathcal{N}, \tag{4.1}$$

where $u_{i|w}^{\star,0}$ is the optimal power consumption of the heat pump and $\xi_{i|w}$ is the small deviation applied by the bounded-rational agent. This small deviation $\xi_{i|w}$ may be fixed – to be more general, it is considered here as the realization from a random variable $\Xi_{i|w}$ with distribution $F_{i|w}$, $\Xi_{i|w} \sim F_{i|w}$. For instance, in the result section, Gaussian distributions will be employed for simplicity. Other distributions may be used, possibly asymmetric, to reflect different views on bounded rationality. In addition, such distributions may be defined for the population as a whole while being time-invariant. However, in a more general case, such distributions may be defined on an individual agent basis, or more likely for groups of agents, while also varying in time e.g. with the time of day, to reflect that bounded rationality may be linked to the type of activity of the agents. An alternative approach to perturbing the decisions of the agents could consist in defining a function that transforms the prices originally broadcast by the central coordinator into "perceived" prices used as input to the local optimization problems.

This small deviation is only added for $t = 0$ of the planning horizon since it is the only decision that the model predictive control actually implements anyway. Therefore, when each agent computes her optimal strategy, at Step 9 in Algorithm 1, we add a value $\xi_{i|w}$ drawn

from $F_{i|w}$ to $u_{i|w}^{\star,0}$ and re-optimise fixing $u_{i|w}^0$ to $\tilde{u}_{i|w}^0$ in order to verify that the bounded-rational strategy is within the feasible space K_i. Consequently, the bounded-rational decision implemented at $t=0$ actually propagates throughout the planning period $\mathcal{T}$. One eventually obtains $\tilde{\boldsymbol{x}}_{i|w}(\boldsymbol{p}_{|w}^k) = [\tilde{\boldsymbol{s}}_{i|w}(\boldsymbol{p}_{|w}^k); \tilde{\boldsymbol{u}}_{i|w}(\boldsymbol{p}_{|w}^k)]$ the bounded-rational strategy of the agent.

The updated version of the Picard-Banach decentralised model predictive control, for the case of bounded-rational agents, is described in Algorithm 2.

Algorithm 2: Picard-Banach decentralised model predictive control with bounded-rational agents

---

1   Choose ϵ_{stop}
2   Initialise $s_{i|0}^0, \forall i \in \mathcal{N}$
3   **for** $w = 1, \ldots, n(\mathcal{W})$ **do**
4   Initialise $\epsilon > \epsilon_{\text{stop}}$
5   Set $k = 1$
6   Draw $\xi_{i|w}$ from $F_{i|w}$
7   Calculate $\boldsymbol{p}_{|w}^k$ with $\hat{\boldsymbol{d}}_{|w}$ and $\overline{\boldsymbol{u}}_{|w}^k = \frac{1}{N}\sum_{i=1}^N \boldsymbol{u}_{i|w}^k$
8   **while** $\epsilon > \epsilon_{stop}$ **do**
9   Broadcast $\boldsymbol{p}_{|w}^k$
10   Compute $\boldsymbol{x}_{i|w}^\star(\boldsymbol{p}_{|w}^k) = [\boldsymbol{s}_{i|w}^\star(\boldsymbol{p}_{|w}^k); \boldsymbol{u}_{i|w}^\star(\boldsymbol{p}_{|w}^k)]$ by solving (2.5), $\forall i \in \mathcal{N}$
11   Obtain $\tilde{u}_{i|w}^0$ based on $\xi_{i|w}$ using (4.1), $\forall i \in \mathcal{N}$
12   Obtain $\tilde{\boldsymbol{x}}_{i|w}(\boldsymbol{p}_{|w}^k) = [\tilde{\boldsymbol{s}}_{i|w}(\boldsymbol{p}_{|w}^k); \tilde{\boldsymbol{u}}_{i|w}(\boldsymbol{p}_{|w}^k)]$ by re-optimizing through the planning horizon, $\forall i \in \mathcal{N}$
13   Compute $\overline{\boldsymbol{u}}_{|w}^{k+1} = \frac{1}{N}\sum_{i=1}^N \tilde{\boldsymbol{u}}_{i|w}^\star(\boldsymbol{p}_{|w}^{\boldsymbol{k}})$
14   Update $\boldsymbol{p}_{|w}^{k+1}$ using (2.10)
15   Update $\epsilon \leftarrow \left\| \overline{\boldsymbol{u}}_{|w}^{k+1} - \overline{\boldsymbol{u}}_{|w}^k \right\|_2$
16          $k \leftarrow k+1$
17   Implement $\tilde{u}_{i|w}^0$ for time step w, $\forall i \in \mathcal{N}$
18   Compute $s_{i|w}^1$ as in (2.1) and set $s_{i|w+1}^0 = s_{i|w}^1$, $\forall i \in \mathcal{N}$

This approach to account for bounded rationality in the decentralised model predictive control alters the original mean-field game that was obtained in Section 2. It is not straightforward to see, based on theoretical arguments, whether the existence and uniqueness of the Nash equilibrium is preserved or not. Extensive sets of simulations showed that it seems to be the case though. The necessary theoretical treatment of this modified mean-field game is left for future work.

5. Application and case-study

The aim of this section is firstly to illustrate the workings of the decentralised model predictive control for a large population of heat pumps. In parallel, it allows illustrating the impact of forecast uncertainty and information updates. Finally, the impact of bounded rationality is assessed, both at the agent and system levels.

(a) General setup

Different case studies allow showing how forecast accuracy and bounded rationality affect the Nash equilibrium. In all the simulations we consider that the number of agents is $N = 10^6$. We base our studies on the Danish power system, with an installed capacity of $1.2 \ 10^7$kW and a typical inflexible load curve for March 2018 depicted in Figure 2. This results in the total

generation capacity scaled by the population size $\kappa = 12$kW, as well as the d_t the inflexible demand scaled by the population size as shown in the figure. In the perfect information case, the forecast for the inflexible load is the same of the observation.

The planning horizon considered in the optimization problem of each agent is $\tau = 13$h, also using a rolling window approach. It means that, at each time step and based on the information updates, agents consistently solve their local optimization problems for these τ hours. Note we will also focus on a given period of 13 hours in following when analysing the results, for which updates where computed at each time step based on local optimization problems solved for the following τ hours. The linear coefficient used in the price function is set to $\beta = 0.12$ following [7]. In the simulations we consider that the population of agents is heterogeneous and is characterised by different input parameters, e.g. heat-pump capacity, ramping limits, building thermal characteristics, solar irradiance and ambient temperature profiles, as well as tracking temperature and comfort bounds. Parameter values for individual agents are drawn independently from a Gaussian distribution. An overview of the mean and standard deviation for each parameter distribution can be seen in Table 1.

Table 1. Characterization of the heterogeneous population of agents considered, based on the mean and standard deviation of their parameters.

	c_i(kW)	$\underline{\delta u_i}$ (kW)	$\overline{\delta u_i}$ (kW)	$\tilde{s}_i$(°C)	$\underline{s}_i$(°C)	$\overline{s}_i$(°C)
μ	3	-2	2	21	17	25
σ	0.3	0.2	0.2	0.5	0.5	0.5

The data used to simulate the building heat dynamics is taken from [25] and the solar irradiance and ambient temperature profiles are obtained from [35]. These sources are used as the mean of a normal distribution from which we sample the building parameters and weather time series for each agent. The temperature penalization factor was set to $\alpha = 100$. This value was chosen after doing some exploratory tests: lower values were not enough to track the reference temperature while higher values resulted in indifference to electricity prices. The predicted inelastic demand of the system is generated with adding a multivariate Gaussian deviation with an exponentially decreasing covariance structure to the inflexible load measurements. The covariance matrix $\boldsymbol{\Sigma} \in \mathbb{R}^{\tau \times \tau}$ is based on a correlation matrix $\mathbf{C} \in \mathbb{R}^{\tau \times \tau}$ and a model for the standard deviation $\boldsymbol{\sigma} \in \mathbb{R}^{\tau}$ as a function of the lead time. This writes

$$\rho_{i,i} = 1, \quad \rho_{i,j} = \exp\left(\frac{-|i-j|}{\nu}\right), \quad \forall i,j \in \{0,\ldots,\tau-1\}, \tag{5.1a}$$

$$\boldsymbol{\sigma} = \Psi \log(k+1), \quad k \in \{0,\ldots,\tau-1\}, \tag{5.1b}$$

$$\boldsymbol{\Sigma} = \boldsymbol{\sigma}\boldsymbol{\sigma}^\top \circ \mathbf{C}, \tag{5.1c}$$

where $\mathbf{C} = (\rho_{ij})_{ij}$ is a correlation matrix, $\boldsymbol{\sigma}$ a standard deviation vector and $\circ$ the Schur product. The parameter Ψ controls the growth of uncertainty with increasing lead times, and hence the forecast accuracy. In parallel, ν is a range parameter that controls the exponential decay of the covariance. With such an approach, forecasts are inherently unbiased. Relevant biases in the forecasts could be readily accounted for, and their effects additionally studied.

(b) Model predictive control vs. open-loop strategy

We start by showing the difference between the open-loop and the model predictive control strategy. In the first case, the forecasts are not updated, and the profile of heat pump consumption is obtained at once with the available information at 20:00 for the following thirteen hours of the planning horizon. On the contrary, with the model predictive control strategy, forecasts are

updated every hour, allowing the population of agents to adjust their profile to changes in the inelastic demand reflected through the electricity prices. To compare both strategies we simulate a disturbance in the system: at 00:00 the inelastic demand undergoes a step reduction and recovers at 3:00. We differentiate between two cases, perfect and imperfect forecasts. In the former case, the forecast captures perfectly the step reduction and its length, while in the latter case, the forecast step reduction lasts until 8:00.

The mean heat pump consumption across the population obtained with the open-loop strategy is plotted in Figure 2 with dark blue and cyan dots for both perfect and imperfect forecast respectively. Both coincide since the information available at 20:00 is the same in both cases. However, there is a difference with the model predictive control strategy, plotted with dark blue and cyan lines, for perfect and imperfect forecast cases, respectively. In the open-loop strategy, the agents do not know that there is going to be a reduction in the inelastic demand and consequently in the electricity prices, and only adjusts their consumption to general shape for the inelastic demand trajectory over the coming period, with a succession of smooth ramps down (20:00 to 3:00) and up (3:00 to 9:00). Meanwhile, with the model predictive control strategy the agents increase their consumption in the hours when the step reduction is happening. In the perfect forecast case the increase is steeper, since the low electricity prices only last three hours. On the contrary, in the imperfect forecast case the increase happens a bit later, assuming that the electricity prices remain lower for longer time. Obviously, after the disturbance has passed (at 3:00), both MPC with perfect and imperfect forecasts realize that the system is back on the original trajectory. However, the impact of the decisions made during the 3-hour window of the disturbance necessarily propagates due to system dynamics and constraints.
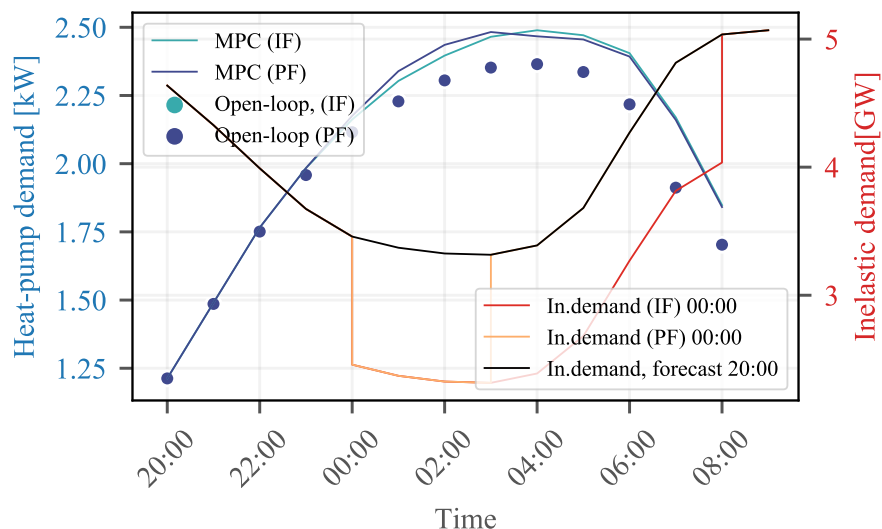

Figure 2. Example of disturbance in the system. Comparison between open-loop and model predictive control (MPC) strategies. (PF) stands for perfect forecast case and (IF) stands for imperfect forecast case.

(c) Effect of forecast accuracy

Once we know that the MPC strategy works better than the open-loop one (in the sense that it better adapts to the changes in the environment), we now focus on the impact of forecast accuracy. For that purpose, different forecasts are generated by adding multivariate Gaussian deviations as described in the above. We vary the forecast accuracy by choosing different values of $\Psi = \{0.1, 0.2, 0.3\}$ and perform a Monte Carlo simulation with 10 runs for each value of Ψ. The parameter ν that controls the exponential decay of correlation among lead times is set to $\nu = 13$.

In Figure 3, the mean consumption across the population of heat pumps is depicted for each forecast accuracy and compared to the perfect forecast case. The larger forecast inaccuracies lead to larger deviations from the perfect forecast case. This deviation increases with the lead time, which is in line with the covariance structure, and are more pronounced when the inelastic demand rate of change is higher, e.g. at 22:00. Even if the Nash equilibrium obtained in each case is different it does not have a significant impact on the objective function. The mean objective function across the population is €1649.58 in the perfect forecast case, and €1649.86, €1649.21, €1648.99 for $\Psi = \{0.1, 0.2, 0.3\}$, respectively. Those values are very close, and this result should be interpreted as the fact that changes in forecast uncertainty are mitigated by population effects when it comes to population-average costs. In addition, the distributions have similar shapes and quantiles, therefore we can conclude that the effect of the forecast accuracy is mitigated by the population effects in terms of cost, beyond the mean only, but not in terms of consumption. Recall that the plot shows the mean consumption across the population, but the aggregated effect of this small deviations has a significant impact on the system demand.
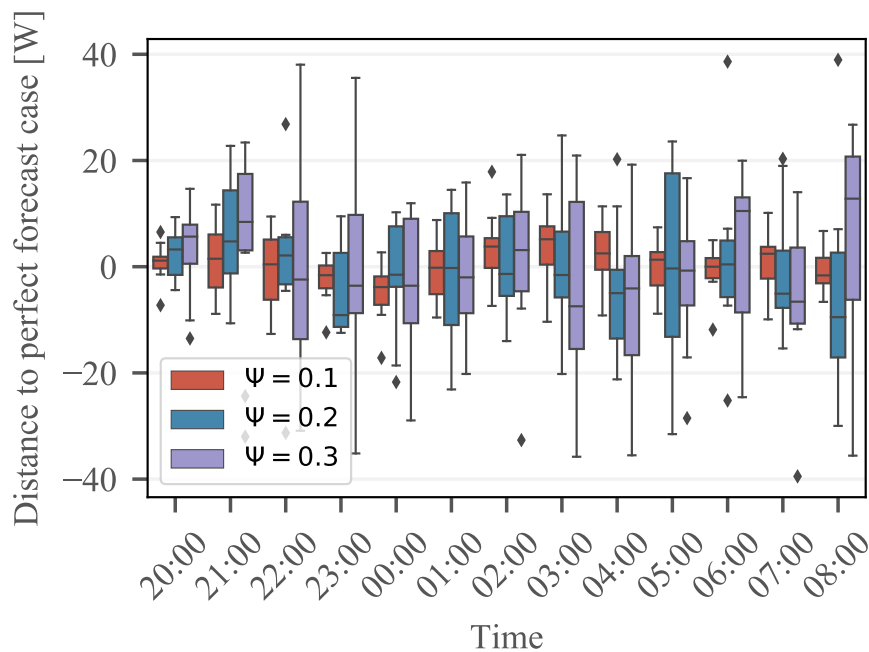

Figure 3. Time-series of mean heat pump consumption across the population. Difference between perfect forecast case and imperfect forecast with $\Psi = \{0.1, 0.2, 0.3\}$.

(d) Effect of bounded rationality

Finally, we relax the rationality assumption and study the impact of bounded rational agents on the system equilibrium. We model bounded rationality following (4.1). In the cases where adding the $\xi_{i|w}$ lead to infeasibility we stick to the rational decision instead. This assumption could fit to reality, i.e., the local controller for the heat pump may check whether the human intervention may lead to infeasibility and if so overrides it. Further work could focus on other alternatives, e.g. projection on the feasible set. Five different setups of bounded rational agents are defined and denoted by "BR 1", "BR 2", "BR 3", "BR 4", and "BR 5":

- In the first three cases ("BR 1", "BR 2", "BR 3"), ξ is drawn from a Normal distribution with $\mu = 0$ and $\sigma = \{83, 167, 250\}$, respectively. Those distributions are the same for all agents and all times. In practice for instance with "BR 1", 99.7% of the population will deviate from the rational choice by $\pm 249W$;
- In the last two cases ("BR 4" and "BR 5"), we cluster the agents into three groups with $\xi_{i|w}$ drawn from three different Normal distributions characterised by $\mu = \{-500, 0, 500\}$ and $\sigma = 83$. In "BR 4" agents are equally distributed between the groups, while in "BR 5" the distribution among groups was randomly chosen. There, the distributions are time-invariant, but vary among groups of agents.

In all the cases, for simplicity, we consider that there is perfect forecast accuracy. We did not include results where both bounded rationality and forecast accuracy are studied together, since by construction they are not interdependent, and it may then be difficult to isolate the effect of bounded rationality. Obviously in practice both effects will be there at the same time and possibly magnify each other.

Figure 4 shows the difference of the mean consumption across the population between the bounded-rational and the rational case. During the hours with low inelastic demand, the equilibria computed for the bounded-rational cases are systematically lower than for the rational case. In those hours, the heat pumps are running at high load, taking advantage of low electricity prices. Therefore, the bounded-rational agents trying to increase their consumption would hit the heat pump capacity. Meanwhile, the bounded-rational agents who try to reduce power consumption of the heat pumps are still capable to do so, resulting in a lower mean consumption. This effect is inverted when the inelastic demand increases, resulting in a higher mean consumption. The spikes that appear in "BR 4" and "BR 5" are a result of how bounded rationality diffuses in time. In those cases there are groups of bounded-rational agents who systematically decrease their consumption by a larger amount. This makes that for given time steps, either the lower bound of the heat pump or the low temperature bound of the comfort settings is reached, and the system needs to recover in the following time step. Further work could focus on adding complexity to the modelling of the heat pump; e.g. by introducing minimum on-off times, the bounded rational decisions are expected to have a higher diffusion in time. Lower mean consumption translates to lower electricity prices, specially in the hours of low inelastic demand as we discussed before. The time-series of electricity prices can be observed in Figure 5.

The bounded rational decisions also have an impact on indoor temperature and consequently on agent comfort. Figure 6 illustrates the deviations of the indoor temperature from the reference temperature $\bar{s}_i$. The more the agents deviate from the optimal response, the more the temperature deviates from its reference. This translates to higher discomfort and then higher costs.

The total costs including both the cost of electricity and the discomfort term are depicted in Figure 7. The cost function is on average higher for bounded rational agents. This is mainly due to the discomfort term which is $586 - 15761\%$ larger than in the rational case on average. On the other hand, the cost of electricity is $0 - 6\%$ smaller than in the rational case on average, while the 90% quantile shows a cost of electricity that is $0 - 22\%$ larger than in the rational case. Note that the lower electricity costs are linked to the period simulated, as previously discussed with Figure 4.

6. Conclusions

With the increasing deployment of distributed energy resources, it is of utmost relevance to investigate alternative approaches to their control in a market-based environment for various applications, e.g., demand response. We placed emphasis here on the case of population of heat pumps to be coordinated. Model predictive control was chosen as the control framework, to accommodate the temporal inter-dependencies in the heat dynamics of dwellings, and information renewal with regularly updated forecasts. The decentralised control was re-thought as a mean-field game, where the price sequence communicated by the central coordinator
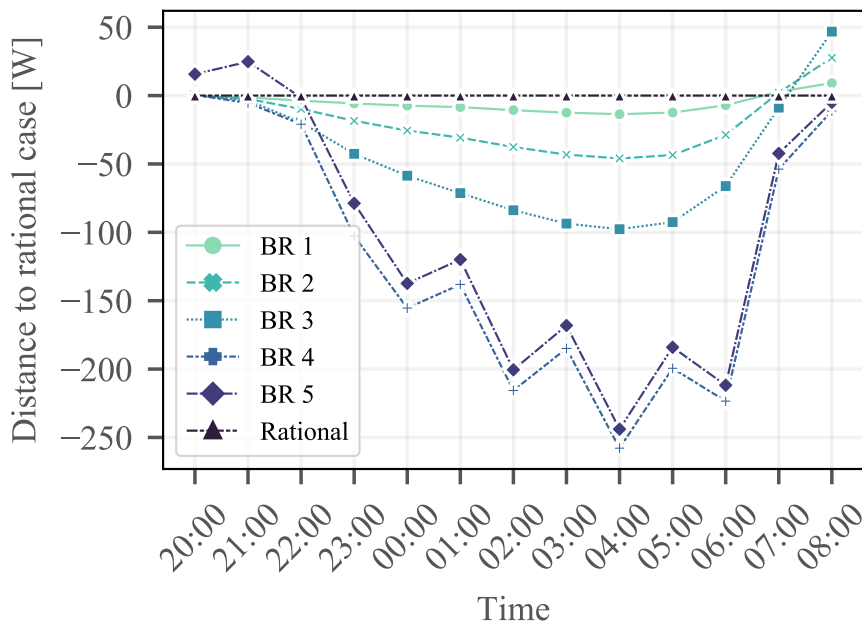
Figure 4. Time-series of mean heat pump consumption across the population. Difference between rational case and bounded rational cases.
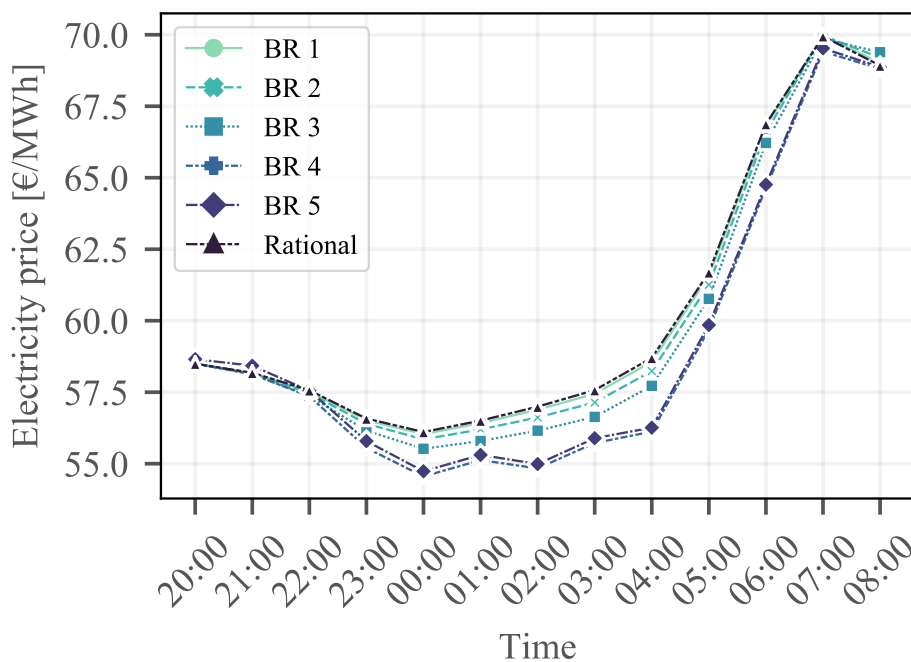

Figure 5. Time-series of electricity prices for different cases of bounded rationality.

summarizes the impact of the decisions (as well as constraints and states) of all agents involved. It allowed to give a number of theoretical results related to the existence and uniqueness of a Nash equilibrium for such a game. The Picard-Banach algorithm was used as a solution
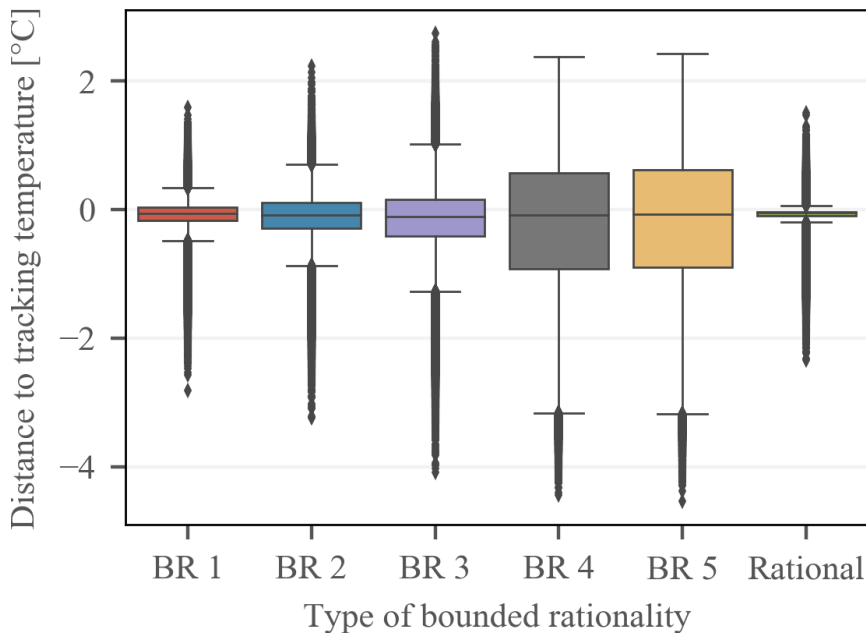
Figure 6. Boxplot of indoor temperature across the population. Distance to tracking temperature for different cases of bounded rationality.
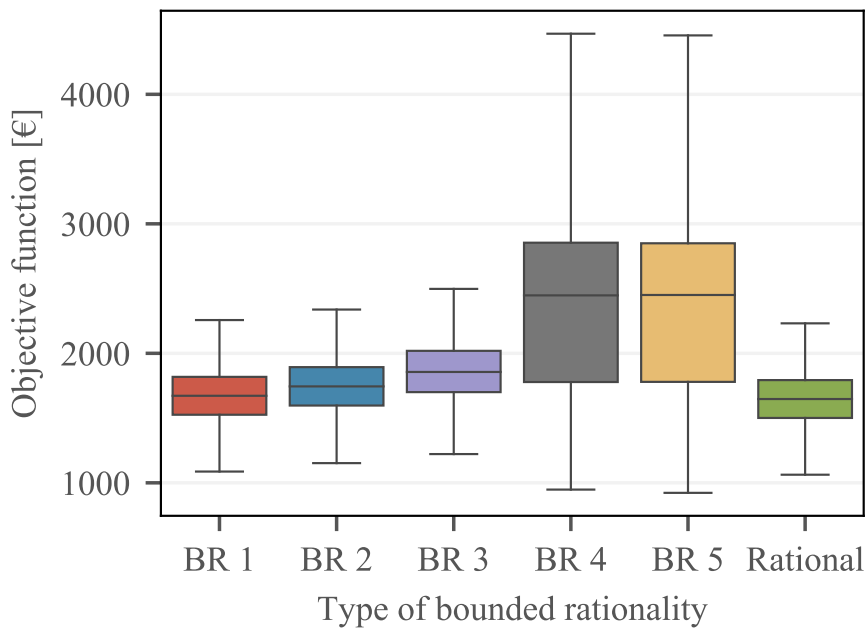

Figure 7. Boxplot of total objective function across the population for different cases of bounded rationality.

approach to obtain the optimal price sequence and strategies of the agents. Importantly, we relaxed the rational assumptions for the agents in order to obtain a complete framework that allows analysis of the impact of bounded rationality on outcomes at system and agent levels. A Picard-Banach algorithm was proposed for the case of bounded-rational agents. Additional

theoretical considerations are necessary to better characterise the properties of the game with bounded-rational agents and the approach to account for bounded rationality that we proposed.

The numerical simulations showed that the effect of forecast accuracy is mitigated by population effects. In contrast, bounded rationality has a big influence on system equilibria, and consequently on electricity prices, comfort and total costs. Further emphasis should be placed on understanding human decision-making and behaviour when it comes to electricity consumers and their use of distributed energy resources (also including stationary battery storages, electric cars, etc.) in order to optimally utilise and steer available flexibility on the demand side. It is clear that more advanced models for various types of bounded rationality, e.g., anchoring and hyperbolic discounting, ought to be placed in relevant control and game-theoretic frameworks. Eventually, in view of the potential impact of bounded rationality on agent-level and system-level outcomes, hedging strategies should be proposed in order to obtain efficient and fair outcomes.

Data Accessibility. This article has no additional data.

Authors' Contributions. Both authors have contributed to the proposal and analysis of the methods. PP originally brought up the idea for this work. AMR implemented the methods and performed the numerical simulations. Both authors contributed to the writing of the manuscript. Both authors have read and approved the manuscript.

Competing Interests. The authors declare that they have no competing interests.

References

1. Strbac G. 2008 Demand side management: Benefits and challenges.
   Energy Policy **36**, 4419–4426.
2. Kirschen DS. 2003 Demand-side view of electricity markets.
   IEEE Transactions on Power Systems **18**, 520–527.
3. Laskey A, Kavazovic O. 2010 Energy efficiency through behavioral science and technology.
   ACM XRDS **17**, 47–51.
4. O'Connell N, Pinson P, Madsen H, O'Malley M. 2014 Benefits and challenges of electrical demand response: A critical review.
   Renewable and Sustainable Energy Reviews **39**, 686–699.
5. Meyn S, Barooah P, Bušić A, Ehren J. 2013 Ancillary service to the grid from deferrable loads: The case for intelligent pool pumps in Florida.
   In *52nd IEEE Conference on Decision and Control*, pp. 6946–6953.
6. Mohsenian-Rad A, Wong VWS, Jatskevich J, Schober R, Leon-Garcia A. 2010 Autonomous demand-side management based on game-theoretic energy consumption scheduling for the future smart grid.
   IEEE Transactions on Smart Grid **1**, 320–331.
7. Paccagnan D, Kamgarpour M, Lygeros J. 2016 On aggregative and mean field games with applications to electricity markets.
   In *2016 European Control Conference (ECC)*, pp. 196–201.
8. Lasry JM, Lions PL. 2007 Mean field games.
   Japanese Journal of Mathematics **2**, 229–260.
9. Gueant O, Lasry JM, Lions PL. 2010 Mean field games and applications.
   In *Paris-Princeton Lectures on Mathematical Finance 2010* (ed. Carmona AR et al), pp. 205–266.
   Berlin, Heidelberg: Lecture Notes in Mathematics 2003, Springer.
10. Ma Z, Callaway DS, Hiskens IA. 2013 Decentralized charging control of large populations of plug-in electric vehicles.
    IEEE Transactions on Control Systems Technology **21**, 67–78.
11. Chen H, Li Y, Louie RH, Vucetic B. 2014 Autonomous demand side management based on energy consumption scheduling and instantaneous load billing: An aggregative game approach.
    IEEE Transactions on Smart Grid **5**, 1744–1754.

12. Grammatico S, Gentile B, Parise F, Lygeros J. 2015 A mean field control approach for demand side management of large populations of thermostatically controlled loads.
In *2015 European Control Conference (ECC)*, pp. 3548–3553.
13. De Paola A, Angeli D, Strbac G. 2016 Distributed control of micro-storage devices with mean field games.
IEEE Transactions on Smart Grid **7**, 1119–1127.
14. Barrera J, Garcia A. 2015 Dynamic incentives for congestion control.
IEEE Transactions on Automatic Control **60**, 299–310.
15. Kahneman D. 2003 Maps of bounded rationality: Psychology for behavioral economics.
American Economic Review **93**, 1449–1475.
16. Frederiks ER, Stenner K, Hobman EV. 2015 Household energy use: Applying behavioural economics to understand consumer decision-making and behaviour.
Renewable and Sustainable Energy Reviews **41**, 1385 – 1394.
17. Gyamfi S, Krumdieck S, Urmee T. 2013 Residential peak electricity demand response – Highlights of some behavioural issues.
Renewable and Sustainable Energy Reviews **25**, 71–77.
18. Meier A, Peffer T, Pritoni M, Aragon CR. 2010 Thermostat interface and usability: A survey.
Technical report, Lawrence Berkeley National Laboratory, tech. rep. no. LBNL- 4182.
19. Sarran L, Christensen MH, Hviid CA, Marin Radoszynski A, Rode C, Pinson P. 2019 Data-driven study on individual occupant comfort using heating setpoints and window openings in new low-energy apartments - preliminary insights.
volume 111. E3S Web Conference.
Art. no. 04063
20. Mullainathan S, Allcott H. 2010 Behavior and energy policy.
Science **327**, 1204–1205.
21. Pita J, Jain M, Tambe M, Ordóñez F, Kraus S. 2010 Robust solutions to Stackelberg games: Addressing bounded rationality and limited observations in human cognition.
Artificial Intelligence **174**, 1142–1171.
22. Gabaix X. 2014 A sparsity-based model of bounded rationality.
The Quarterly Journal of Economics **129**, 1661–1710.
23. Grammatico S, Parise F, Colombino M, Lygeros J. 2016 Decentralized convergence to Nash equilibria in constrained deterministic mean field control.
IEEE Transactions on Automatic Control **61**, 3315–3329.
24. Moret F, Pinson P. 2019 Energy collectives: A community and fairness based approach to future electricity markets.
IEEE Transactions on Power Systems **34**, 3994–4004.
25. Mahler E. 2016 *Demand response in a market environment*.
Ph.D. thesis, Technical University of Denmark, Department of Electrical Engineering.
26. Bacher P, Madsen H. 2011 Identifying suitable models for the heat dynamics of buildings.
Energy and Buildings **43**, 1511–1522.
27. Parise F, Colombino M, Grammatico S, Lygeros J. 2014 Mean field constrained charging policy for large populations of plug-in electric vehicles.
In *53rd IEEE Conference on Decision and Control*, pp. 5101–5106.
28. Facchinei F, Pang JS. 2007 *Finite-dimensional Variational Inequalities and Complementarity Problems*.
Springer Science & Business Media.
29. Boyd S, Vandenberghe L. 2004 *Convex Optimization*.
Cambridge University Press.
30. Tversky A, Kahneman D. 1974 Judgment under uncertainty: Heuristics and biases.
Science **185**, 1124–1131.
31. Ariely D. 2008 *Predictably Irrational*.
Harper Collins.
32. Gyamfi S, Krumdieck S, Urmee T. 2013 Residential peak electricity demand response – highlights of some behavioural issues.
Renewable and Sustainable Energy Reviews **25**, 71 – 77.
33. Stern PC, Aronson E, Darley JM, Hill DH, Hirst E, Kempton W, Wilbanks TJ. 1986 The effectiveness of incentives for residential energy conservation.
Evaluation Review **10**, 147–176.

34. Barberis NC. 2013 Thirty years of prospect theory in economics: A review and assessment. *Journal of Economic Perspectives* **27**, 173–196.
35. Pfenninger S, Staffell I. 2016 Long-term patterns of european PV output using 30 years of validated hourly reanalysis and satellite data. *Energy* **114**, 1251 – 1265.

22

rsta.royalsocietypublishing.org Phil. Trans. R. Soc. A 0000000