Set-Membership Dual Predictive Control for Thermoelectric Flexible Loads

Alessandro Del Duca[1,3], Erick Meza[2], Andres Cordoba-Pacheco[1]
Carlos Ocampo-Martinez[2], *Senior Member, IEEE* and Fredy Ruiz[1], *Senior Member, IEEE*

***Abstract*—Flexible loads, such as Thermostatically Controlled Loads, are emerging as key elements for power balancing in smart grids, particularly under increasing renewable energy penetration. However, exploiting their flexibility requires knowledge of their aggregated dynamics, which are difficult to derive analytically. This challenge naturally leads to a dual control problem, where the controller must learn the system dynamics while simultaneously regulating the power consumption, to provide flexibility service to the grid. This paper presents a Set-Membership Dual Model Predictive Control (MPC) framework for the aggregated control of flexible loads that performs simultaneous system identification and optimal control of a set of loads, using Autoregressive with Exogenous Inputs models. The proposed approach actively reduces model uncertainty by minimising the volume of the Feasible Parameter Set, providing an uncertainty-aware dual control strategy. The performance of the method is evaluated through a comparative study against a passive adaptive MPC scheme in a demand response scenario. Simulation results show that the proposed approach reduces the Integral Square Error by 70.5% and the Integral Absolute Error by 57.9%, relative to the benchmark, demonstrating the benefits of incorporating active learning into the control design.**

***Index Terms*—Model Predictive Control, Set-Membership estimation, Dual control, Flexible loads.**

I. Introduction

The technological shift towards renewable energy has transformed the paradigm of the power grid, moving from generator-driven systems to innovative frameworks in which loads adapt their consumption to match available generation. In this context, consumers participate in energy dispatch through Demand Response (DR) programs, enhancing grid flexibility [1], [2]. These services rely on flexible resources capable of adjusting their consumption while ensuring operational reliability, with Flexible Loads (FL) playing a key role [3].

As the electrification of Heating, Ventilation, and Air Conditioning (HVAC) systems grows, devices such as refrigerators, boilers, space heaters, and heat pumps are becoming increasingly relevant in flexible energy resources [4] [5]. When properly coordinated, these systems can serve as valuable assets for DR initiatives.

Thermostatically controlled loads (TCLs) have gained particular attention among the different types of FL. These devices regulate the temperature through electrical consumption and, due to their thermal inertia, the temperature setpoint can be temporarily adjusted without compromising user comfort. Given their small individual capacity, TCLs are typically aggregated to provide flexibility services that can be offered to the Transmission System Operator (TSO), subject to coordination with the Distribution System Operator (DSO). This process is usually handled by an aggregator, a central unit that coordinates the collective response by adjusting individual setpoints [6], [7]. However, achieving this objective requires an adequate understanding of the dynamic behaviour of individual loads. Deriving analytical models that accurately capture these dynamics is often challenging, as they involve complex, nonlinear, and time-varying processes that are both difficult and time-consuming to be modelled [8]. Learning these dynamics is essential for the aggregator to coordinate multiple FL devices effectively.

Given the inherent complexity and heterogeneity of individual systems, learning aggregate dynamics in real time can be an efficient alternative to bottom-up modeling, at the cost of reduced guarantees on disaggregation feasibility. This fact introduces a fundamental trade-off between control performance and actively identifying system dynamics, all while respecting operational constraints such as limits on power ramp-up and ramp-down. This challenge, characterised by the need to handle model uncertainties through active learning, is known in literature as the *Dual Control problem* [9]. Despite extensive research, such a problem remains challenging due to its inherent computational complexity [10].

In this context, the Set Membership Estimation (SME) framework emerges as an effective approach to handle model uncertainty deterministically by representing it as sets [11], [12] induced by Unknown-but-Bounded (UBB) noises. This perspective enables the use of set-theoretic tools to derive computable dual control algorithms, offering advantages over Bayesian approaches, including improved computational tractability and independence from prior assumptions on underlying probability distributions [13].

The SME framework has been formulated within Model Predictive Control (MPC) schemes, demonstrating suitable performance in terms of robustness to disturbances, constraint satisfaction, and recursive feasibility, where the model un-

This work has been supported by the Italian Ministry of University and Research under grant ‘Learning-based Model Predictive Control by Exploration and Exploitation in Uncertain Environments” (PRIN PNRR 2022 fund, ID P2022EXP2W) and the Spanish project SEAMLESS: Sustainable learning-based Management of Multi-resource Large-scale Systems” (ref. PID2023-148840OB-I00), funded by MCIN/AEI/10.13039/501100011033/FEDER, UE.

[1]Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133, Milano, Italy; (e-mail: {alessandro.delduca, andres.cordoba, fredy.ruiz}@polimi.it).

[2]Automatic Control Department, Universitat Politècnica de Catalunya - BarcelonaTech, Barcelona, Spain; (e-mail: erick.manuel.meza@estudiantat.upc.edu, carlos.ocampo@upc.edu).

[3]Ricerca sul Sistema Energetico - RSE S.P.A. Milano, Italy; (e-mail: Alessandro.DelDuca@rse-web.it)

certainty is quantified through the size of predicted state tubes [14]. Alternatively, a scenario-based MPC formulation that combines SME uncertainty sets with MPC strategies has been proposed in [15], [16], aiming to reduce computational requirements.

However, these approaches do not explicitly address the active reduction of model uncertainty during control operation. To overcome this limitation, this work proposes an approach to actively reduce the uncertainty by minimising the volume of the Feasible Parameter Set (FPS), based on the results developed in [15], [16].

Specifically, a novel SME Dual MPC framework is introduced that simultaneously performs system identification and control tasks by using Autoregressive with Exogenous Inputs (ARX) structures, with a point-wise error bound defining the model class. Therefore, the main contribution of this paper lies in formulating the FPS volume minimisation as a control objective, enabling an explicit trade-off between uncertainty reduction and tracking performance while ensuring constraint satisfaction. Finally, the effectiveness of the proposed method is demonstrated through the control and identification of the aggregated dynamics related to a group of TCLs for DR applications.

The remainder of this paper is organized as follows. In Section II, the TCL aggregation problem is defined. Section III presents the proposed SME Dual MPC algorithm and its properties. Section IV shows the simulation results where the algorithm is applied to solve the TCL aggregation problem. Finally, Section V concludes the paper with some final remarks and future work.

II. Problem Definition

Large populations of small TCLs can provide fast, reliable balancing services when coordinated by an aggregator. Thermo-Electric Refrigerators (TERs) are particularly appealing for DR problems due to their precise temperature control and the ability to shift steady-state power by adjusting set points. In practical applications, aggregators must comply with strict operational requirements imposed by grid services. These requirements translate into constraints on the aggregate power response that must be satisfied by the low-level controllers. The modelling and aggregation strategy of TERs considered in this work follow the methods reported in [6].

The dynamics of each TER can be represented by using a continuous-time nonlinear state-space model, with t being the continuous time variable, of the form

$$\dot{x}_i(t) = f_i(x_i(t), \beta_i(t)), \tag{1a}$$

$$P_i(t) = g_i(x_i(t), \beta_i(t)), \tag{1b}$$

where $x_i(t) \in \mathbb{R}^n$ represents the n TER's states, such as temperature, currents and voltages; locally, each TER is regulated by means of an internal controller, to track a reference temperature selected among a set of three reference signals, encoded by the exogenous signal $\beta_i(t) \in \{-1, 0, 1\}$.; the mapping function $f_i(\cdot) : \mathbb{R}^n \times \{-1, 0, 1\} \to \mathbb{R}^n$ describes the nonlinear dynamics of the $i-th$ thermoelectric system, including thermal loads, head flows, and Peltier module transduction, among others; the map $g_i(\cdot) : \mathbb{R}^n \times \{-1, 0, 1\} \to \mathbb{R}$ describe the relation between the system state $x_i(t)$ and exogenous signal $\beta_i(t)$ and the instantaneous power consumed by the TER $P_i(t)$, inferred by current and voltage measures, to maintain a reference temperature $T_{set}(\beta_i(t))$, selected as follows:

$$\beta_i(t) = \begin{cases} 0, & \text{Nominal set-point } T_{\text{set}} = T_{\text{nom}} \in \{3, ..7\}\,^\circ\text{C}, \\ 1, & \text{High set-point } T_{\text{set}} = 8\,^\circ\text{C}, \\ -1, & \text{Low set-point } T_{\text{set}} = 2\,^\circ\text{C}, \end{cases}$$

where T_{nom} represents a nominal temperature set-point, corresponding to a nominal power consumption P_i^{nom}.

The aggregator's role is to shape the aggregate power consumption and provide DR services to the DSO, by coordinating the input signals β_i across multiple TER units.

A. Aggregated dynamics

An aggregation of m heterogeneous TERs capable of providing upward and downward reserve services to the DSO is considered. The aggregator is a discrete-time controller, working with a sampling time $T_s = 1s$, considering the slow temperature dynamics and the constraints on the requested power settling time of $30s$.

To provide the DR service, the aggregator computes a high-level control input $u_k = \sum_{i=1}^{m} \beta_{i,k}$, with $u_k \in [-m, m]$, which represents the number of units required to modify their temperature setpoints T_{set} at the sampled time $k \in \mathbb{Z}^+$, i.e. $\beta_{i,k} = \beta_i(k \cdot T_s)$.

A scheduler then determines which of the m TER units should be activated, ensuring compliance with each TER's turn-on/turn-off and idle constraints, as detailed in [6], assigning to each TER the corresponding vector $\beta_{i,k} \in \{0, 1, -1\}^m$, indicating whether the i-th TER must adjust its reference signal during the interval $t \in [k \cdot T_s, (k+1) \cdot T_s]$.

Through this coordination mechanism, the aggregator regulates the aggregate power consumption y_k, defined as

$$y_k = \sum_{i=1}^{m} \left(P_{i,k}^{\text{nom}} - P_{i,k} \right), \tag{2}$$

to increase or decrease power demand as needed, while remaining within operational constraints.

The overall control architecture is shown in Fig. 1. The scheme consists of three main components: the Dual MPC aggregator, which computes the control input u_k according to the required flexibility y_k^{ref}; the TER scheduler, which assigns the control inputs $\beta_{i,k}$ to each TER unit; and finally, the set of TERs, which tracks the assigned temperature setpoints and adjusts their power consumption $P_{i,k}$ accordingly.

Note that the aggregator control level is solely responsible for determining the number of units that will change the setpoint $T_{set}(\beta_{i,k})$. Then, a primary controller decides which of the m TERs must change its reference signal during the time interval $t \in [k \cdot T_s, (k+1) \cdot T_s]$ by assigning the corresponding activation vector $\beta = [\beta_1 \cdots \beta_m]^\top$.

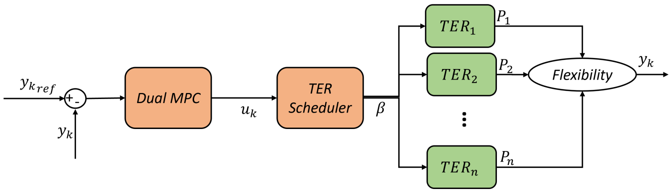

Fig. 1: TER control architecture

The TER aggregation problem can be formulated as a receding-horizon discrete-time optimization problem related to an MPC strategy, namely

$$\min_{u_{k|k},\dots,u_{N-1|k}} \sum_{j=0}^{N-1} Q(y_{k+j|k} - y^{\text{ref}}_{k+j|k})^2 + Ru^2_{k+j|k} \tag{3a}$$

s.t.

$$\text{System dynamics (1a)}, \quad \forall i = 1,\dots,m, \tag{3b}$$

$$\text{Power deviation (2)}, \tag{3c}$$

$$y_{k+j+1|k} \in [y_{min}, y_{max}], \quad \forall j = 1,\dots,N-1, \tag{3d}$$

$$u_{k+j|k} \in [-m, m] \quad \forall j = 0,\dots,N-1, \tag{3e}$$

where N is the prediction horizon and the notation $v(k+i|k)$ denotes the prediction of the variable v at time $k+i$, made at time instant k.

The main challenge faced by the aggregator is that the aggregated dynamics are unknown a priori and cannot be identified offline without disrupting operations. This fact motivates the use of a learning MPC framework to regulate the aggregated power tracking tasks. To address this challenge, a Dual MPC approach is adopted, in which the controller actively explores the aggregated dynamics to learn an accurate model of the system from limited prior information, while robustly controlling the reserve service.

III. Set-Membership Dual MPC

In this section, a Dual MPC strategy with active learning is presented, capable of optimally managing the exploration/exploitation trade-off required to learn the aggregated TERs dynamics while properly tracking power deviation requests.

A. Black-box aggregated dynamics

Let the aggregated dynamics of power consumption be represented as an ARX model of order (n_a, n_b), i.e.,

$$y_{k+1} = a_1 y_k + \cdots a_{n_a} y_{k-n_a+1} + b_1 u_k + \cdots b_{n_b} u_{k-n_b+1} + e_k,$$

where u_k and y_k denote the system input (number of TERs with modified temperature set point) and output (net power consumption variation) at time instant k, respectively. This model can be compactly expressed as

$$y_{k+1} = \theta^{\star\top} \varphi_k + e_k, \tag{4}$$

where

- $\varphi_k = [y_k, \dots \ y_{k-n_a+1}, u_k, \dots \ u_{k-n_b+1}]^\top$ is the regressor vector, and
- $\theta^* = [a^\star_1, \dots \ a^\star_{n_a}, b^\star_1, \dots \ b^\star_{n_b}]^\top \in \mathbb{R}^n$ is the unknown parameter vector, containing the true $n = n_a + n_b$ model parameters.

The process disturbance $e = \{e_k\}_{k=0}^{\infty}$, governing the mismatch between the actual and the predicted power modifications, is assumed to be an Unknown But Bounded (UBB) signal, i.e., $||e||_\infty \le \bar{e}$, with finite bound $\bar{e} > 0$.

The SME framework enables the characterisation of the set of all models that are compatible with the a priori assumptions on the system and the data generated by the system during normal operation.

Let $\mathscr{K}_0 \subseteq \mathbb{R}^n$ be a closed and convex polytope encoding prior knowledge about the ARX model parameters, such as physical bounds or structural properties. Whenever a new measurement $\{\tilde{y}_k, \varphi_k\}$ is acquired, the uncertainty on the model parameters is reduced. In particular, each measurement defines a strip of unfalsified parameter vectors in the parameter space [12], [13], i.e.,

$$\mathscr{S}_{k+1} := \left\{ \theta \in \mathbb{R}^n \,\middle|\, |\tilde{y}_{k+1} - \varphi_k^\top \theta| \le \bar{e} \right\}. \tag{5}$$

Geometrically, each strip $\mathscr{S}_k$ corresponds to a pair of parallel hyperplanes enclosing a slab in $\mathbb{R}^n$. The set of all models consistent with the prior knowledge of the system and the data up to time $k \in \mathbb{Z}^+$ is called the FPS, denoted by Θ_k, and defined as

$$\Theta_{k+1} = \Theta_k \cap \mathscr{S}_k = (\bigcap_{i=1}^{k} \mathscr{S}_i) \cap \mathscr{K}_0. \tag{6}$$

The FPS adaptation mechanism based on the SME framework ensures that Θ_k is non-increasing over time and is guaranteed to contain the true system parameter vector. Moreover, no assumptions are imposed on the statistical (un)correlation of the input and the noise.

B. Risk-constrained control of aggregated TCLs flexibility

The goal of the control system is to compute, at each time instant k, the optimal control action u_k^* that allows the aggregated TCL population to track a power deviation request as stated in (3). However, the aggregated dynamics are unknown and, at any time instant k, the only information about the system behavior is that it can be described as an ARX model whose parameters belong to Θ_k. To address this challenge, a scenario-based Optimization is adopted [17], where the model uncertainty is represented through a finite set of scenarios randomly extracted from the FPS.

The control strategy is robustified by relying on a scenario approximation of the FPS. Specifically, by randomly selecting K scenarios $\theta_k^{(j)} \in \Theta_k, \quad j = 1,\dots,K$, and imposing the output constraints on the corresponding output trajectory for each scenario. Yet, the cost function follows a certainty-equivalent principle, where the nominal model is the ARX system defined by the Chebyshev center of the FPS, θ_c^k.

The resulting scenario optimisation problem is

$$\min_{u_{k|k},\dots,u_{N-1|k}} \sum_{i=0}^{N-1} Q(y^{(c)}_{k+i|k} - r_{k+i})^2 + Ru^2_{k+i|k} \tag{7a}$$

s.t.

$$y^{(j)}_{k+i+1|k} = \theta_j^\top \varphi^{(j)}_{k+i|k}, \forall i = 0,\dots,N-1, j = 1,\dots,K, \tag{7b}$$

$$y^{(c)}_{k+i+1|k} = \theta_c^{k\top} \varphi^{(c)}_{k+i|k}, \forall i = 0,\dots,N-1, \tag{7c}$$

$$\varphi^{(j)}_{0|k} = \varphi^{(c)}_{0|k} = \varphi_k \quad j = 1,\dots,K, \tag{7d}$$

$$\theta_j \in \Theta_k \quad \forall j = 1,\dots,K, \tag{7e}$$

$$y^{(j)}_{k+i+1|k} \in \mathscr{Y} \quad \forall i = 1,\dots,N-1, j = 1,\dots,K, \tag{7f}$$

$$u_{k+i|k} \in \mathscr{U} \quad \forall i = 0,\dots,N-1. \tag{7g}$$

A limitation of this strategy is that the information gain about the model parameters at each time step is limited by the behavior of the reference signal r_k and the resulting control action u^*_k. In particular, under large uncertainty, the robust control strategy tends to keep the system close to steady-state conditions, avoiding aggressive input maneuvers to guarantee stability.

To overcome this limitation, an Active Learning MPC is proposed. This approach leverages the volume of the FPS, Θ_k, as a measure of parametric uncertainty in the aggregate ARX dynamics, rather than as a guarantee of disaggregation feasibility at the individual load level. The robustness in the disaggregation process, which is described in [6], is ensured by the robustification of the MPC through the scenario-based optimization problem, which balances exploration of the FPS (reducing uncertainty in the aggregation levels) and exploitation of the FPS (generating control actions that allow for robust tracking of the requested flexibility requests). Compared to existing dual control approaches, the proposed formulation offers reduced computational complexity while preserving uncertainty-aware decision making.

The optimization problem behind the Dual Scenario MPC approach maintains the constraints in (7), but it differs in the cost function. The new cost function is augmented with an additional term, $V_{k+i}(y,u)$, which is the volume of the FPS at time $k+i$ and measures the evolution of the uncertainty size. By minimising this term, the controller implicitly aims to reduce model uncertainty while also achieving performance objectives. The optimisation solved by the Dual Scenario MPC becomes:

$$\min_{u_{k|k},\dots,u_{N-1|k}} \sum_{i=0}^{N-1} Q(y^{(c)}_{k+i|k} - r_{k+i})^2 + Ru^2_{k+i|k} + \lambda \cdot V_{k+i}(y_{k+i|k}, u_{k+i|k}), \tag{8a}$$

s.t.

$$(7b), (7c), (7d), (7e), (7f), (7g) \tag{8b}$$

Note that the computation of the volume $V_{k+i|k}(y_{k+i|k}, u_{k+i|k})$ relies on future output measurements $y_{k+i|k}$, which are not available at time k. Since the true parameter $\theta^\star$ that generates the data is unknown, a practical approximation is to evaluate the evolution of the FPS volume under the assumption that the system behaves according to the central model parametrized by θ^k_c. Although this approximation introduces some error in the cost evaluation, it tends to yield an overestimation of the volume, meaning that the realised FPS has a lower volume (i.e. lower uncertainty) than the predicted one. This formulation provides a tractable yet principled approximation of the dual control problem. When uncertainty is large, the volume term has a significant impact on the cost function, encouraging the selection of optimal inputs u^*_k reducing the size of the FPS. By adjusting the weight parameter λ, the trade-off between exploring the parameter space and achieving tracking control objectives can be explicitly balanced.

The overall SME dual MPC scheme is summarized in Algorithm 1.

Algorithm 1 Model Predictive Control with Adaptive Exploration

Require: Initial regressor vector φ_0; initial FPS Θ_0; initial central estimate θ^c_0;

for $k = 0, 1, 2, \dots$ **do**
  Solve the *volume-minimizing* MPC problem
  Compute $u_k \leftarrow u^*_{k|k}$ and apply to the system
  Measure the system output $\tilde{y}_{k+1}$
  Compute the corresponding strip:
  $\mathscr{S}_{k+1} = \left\{ \theta \in \mathbb{R}^n : \left| \tilde{y}_{k+1} - \varphi_k^\top \theta \right| \le \bar{e} \right\}$
  Update the FPS: $\Theta_{k+1} \leftarrow \Theta_k \cap \mathscr{S}_{k+1}$
end for

IV. Dual predictive control of aggregated TCLs

The system comprises $m = 100$ independent TERS, each one with different parameters and dynamics. It was shown in [6] that the aggregated behavior can be represented as a couple of second-order continuous-time linear models with two poles and two zeros. Power increments are described by the transfer function $G_{up}(s)$, and power reductions by $G_{down}(s)$, with

$$G_{up}(s) = \frac{25.95s^2 + 0.017s + 9.3 \times 10^{-6}}{s^2 + 1.1 \times 10^{-3}s + 6.8 \times 10^{-7}}, \tag{9a}$$

$$G_{down}(s) = \frac{-17.35s^2 - 0.018s - 3.6 \times 10^{-6}}{s^2 + 2.3 \times 10^{-3}s + 3.7 \times 10^{-7}}. \tag{9b}$$

These models are used here as a baseline for the true system parameters $\theta^\star$.

The aggregated dynamics are sampled through a Sampler-and-Hold with the same sampling time $T_s = 1$s; the upward/downward dynamics can therefore be modelled as a linear ARX model

$$A(z^{-1})Y(z^{-1}) \approx B(z^{-1}) \cdot U(z^{-1}) + e(z^{-1}), \tag{10}$$

with $A(z^{-1})$ and $B(z^{-1})$ containing the parameters $\theta \in \mathbb{R}^n$ being the unknown parameter vector ($n = n_a + n_b = 5$), i.e.,

$$\begin{aligned} A(z^{-1}) &= 1 + \theta_1 z^{-1} + \theta_2 z^{-2}, \\ B(z^{-1}) &= \theta_3 z^{-1} + \theta_4 z^{-2} + \theta_5 z^{-3}, \end{aligned}$$

and the signal e_k is unknown but (pointwise) bounded with bound $e_k \le 5$ kW, reflecting the maximum deviation in the aggregated power.

A. Initial Feasible Parameter Set

Considering the second-order ARX model in the causal form as in (10), the characteristic polynomial of the dynamical system is

$$p(z) = z^2 - \theta_1 z - \theta_2. \tag{11}$$

The discrete-time system is *asymptotically stable* (Schur stable) if and only if all roots λ_i of $p(z)$ lie strictly inside the unit disk, i.e.,

$$|\lambda_i| < 1, \qquad i = 1,2. \tag{12}$$

For a quadratic polynomial $z^2 - \theta_1 z - \theta_2$, the classical Schur (Jury) conditions are

$$1 - \theta_1 - \theta_2 > 0, \tag{13}$$
$$1 + \theta_1 - \theta_2 > 0, \tag{14}$$
$$|\theta_2| < 1. \tag{15}$$

For any θ inside this polytope, the characteristic polynomial (11) has both roots inside the unit disk, and the corresponding autonomous second-order system is asymptotically stable. Instead, for parameters θ_i $\forall i = 3,4,5$ the boxes are bounded, i.e.,

$$\theta_i \in [\theta_i^{min}, \theta_i^{max}], \qquad \forall i = 3,4,5. \tag{16}$$

Threrefore, all the prior information can be rewritten as the initial FPS $\mathscr{K}_0 = \{\theta : A_0\theta \le b_0\}$. The initial FPS $\mathscr{K}_0$ is shown in Fig. 3 as the blue area.

B. Simulation setting

The active MPC developed in this work is tested against its passive counterpart, presented in [15], which cannot actively probe the system dynamics; instead, it relies on passively acquiring new data while controlling the system in a closed loop. During real-time operations, the DSO can request either upward or downward reserves activation from the TER aggregation, with a reference signal $y_k^{\text{ref}} \in [-1000, 1000]\,\text{kW}$. To facilitate these bidirectional request, the aggregator is designed to learn and employ two distinct models, G_{down} and G_{up}, depending on the required service type.

For operational safety, the aggregator must ensure that its actual output remains within the admissible range $y_k \in [-500, 500]\,\text{kW}$. Consequently, the requested reserve may sometimes exceeds the maximum deliverable capacity of the system, requiring that the aggregator regulate the system robustly while operating at the boundary of its safe constraints, due to the error e_k.

In Fig. 2, the trajectories of both active and passive closed-loop systems are depicted. In the absence of the DSO service request ($y_{k,ref} = 0$kW), the passive aggregator maintains a zero-power output (0kW), whereas the active aggregator starts probing the system to reduce uncertainty in the model. Therefore, when a service is required, the active aggregator has a better model than the passive one and, consequently, can better follow the step-reference.

At $k = 8$s, the active aggregator begins proving the system to learn its dynamics; then, when the DSO asks for a reserve

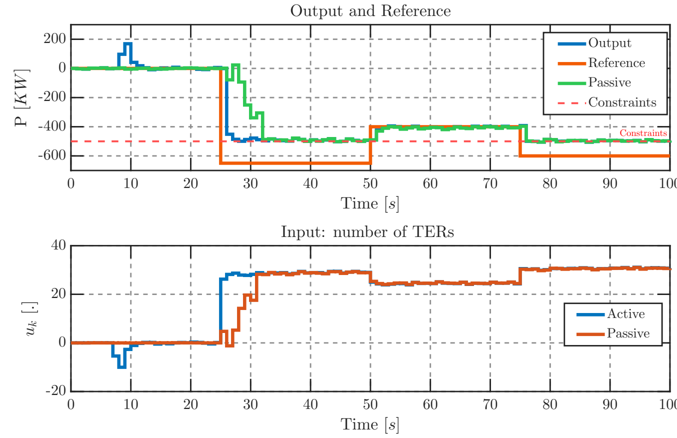

Fig. 2: Top: Output trajectories of the passive (green) and active (blue) closed-loop systems, shown alongside the reference signal (orange) and the constraint (red, dotted). Bottom: Input trajectories of the passive (orange) and active (blue) closed-loop systems.

service at $k = 25$s, the active learning aggregator has better information about the system and follows the step signal closely. To compare the active and passive strategies, the initial step responses were compared using Key Performance Indicators (KPIs), specifically the Integral Square Error (ISE) and the Integral Absolute Error (IAE), defined as

$$ISE(k,k_0) = \sum_{i=k_0}^{k} (y_i - y_{i,ref})^2, \tag{17a}$$

$$IAE(k,k_0) = \sum_{i=k_0}^{k} |y_i - y_{i,ref}|. \tag{17b}$$

These values are used to evaluate whether active learning is beneficial for better tracking of the reference, even during the initial learning phase between 8s and 15s. The errors for $k_0 = 0$ and $k = 50$s are computed and presented in Table I.

TABLE I: Performance metrics for the Active and Passive MPC comparison

KPI's	**ISE**	**IAE**	**Solver time**
Active MPC	293,961.8	1,107.2	0.6847 s
Passive MPC	996,898.9	2,632.3	0.2129 s
Reduction [%]	**70.5%**	**57.9%**	**-**

Despite the initial learning phase, the Active MPC significantly improves the closed-loop step response compared to the Passive MPC. Even if the learning transient is included in the KPIs, the overall performance indices of the Active MPC remain substantially lower, demonstrating its superior control efficiency and adaptability.

Figure 3 presents the FPS before and after the learning phase. The system can correctly reduce the distance between the true parameters $\theta^\star$ and the central estimate θ_c^k. With a lower FPS volume, the aggregator has better knowledge

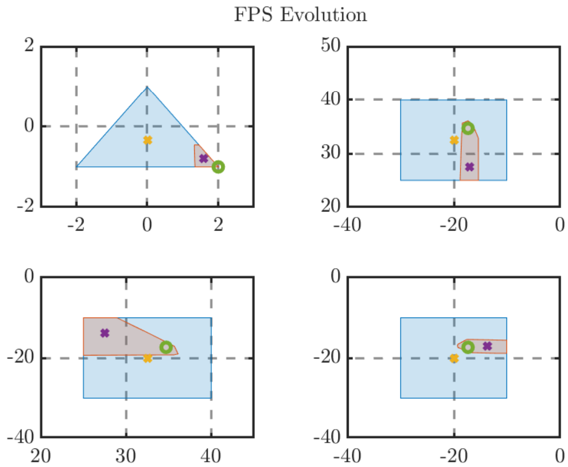

Fig. 3: Initial FPS at $k=1$ (blue area), FPS at $k=25$ (orange area), central estimates θ_c^1 (yellow x) and θ_c^{24} (purple x); real parameter $\theta^\star$ (green circle).

of the underlying system and, therefore, better performance in subsequent references. As a remark, since the passive aggregator does not probe the dynamic, the passive FPS remains equal to the initial FPS Θ_0, meaning that at the time instant in which the step start, i.e. $k=22$s, the passive aggregator propagates the central estimate model θ_c^k, equal to the initial central estimate model θ_c^0, which does not represent the correct system dynamics. However, it is worth noting that, even when the reference signal exceeds the constraints, both active and passive MPC remain within specified limits due to robustification of the scenario.

The average computational time of one step of the Dual MPC is $\tau_c = 0.6847s$ for a model with $n=5$ parameters (over the active-learning windows), which remains compatible with real-time operations, i.e. $\tau_c < 1s$; however, it is worth noting that the computational complexity increases with the ARX model order and the size of the uncertainty set, and a more systematic scalability analysis with respect to these factors is left for future work. The dual control optimisation program is solved with *fmincon* Sequential Quadratic Program (SQP) solver, using Gurobi to solve the Quadratic Program.

V. Conclusions

In this paper, a computationally tractable approximation of the dual control problem is proposed, based on the set membership estimation framework. The presented MPC dual algorithm minimises the volume of the feasible parameter set, serving as a measure of the model-mismatch-induced uncertainty, while guaranteeing robustness through a scenario optimization framework, thereby balancing exploration and performance within the control loop. The algorithm is applied to an aggregated set of thermoelectric loads participating in demand response programs. In this setting, the DSO issues upward or downward reserve requests, which the aggregator meets by adjusting the temperature reference setpoints of individual devices, thereby modulating their power consumption. During periods without reserve requests, the controller actively probes the system's dynamics to improve model accuracy, learning distinct models for each reserve direction. Simulation results demonstrate that the proposed dual MPC achieves superior adaptability and performance compared to a passive adaptive MPC counterpart, particularly during reserve activations and model transitions.

References

[1] M. H. Albadi and E. F. El-Saadany, "Demand response in electricity markets: An overview," in *2007 IEEE Power Engineering Society General Meeting*, 2007, pp. 1–5.

[2] J. Ponoćko and J. Milanović, "The effect of load-follow-generation motivated DSM programme on losses and loadability of a distribution network with renewable generation," in *2019 IEEE PES GTD Grand International Conference and Exposition Asia (GTD Asia)*, 2019, pp. 165–170.

[3] A. Alzahrani, G. Hafeez, G. Rukh, S. Murawwat, F. Iftikhar, S. Ali, S. I. Haider, M. I. Khan, and A. M. Abed, "Demand response for optimal power usage scheduling considering time and power flexibility of load in smart grid," *IEEE Access*, vol. 11, pp. 33 640–33 651, 2023.

[4] Z. Ma, S. Cui, and J. Chen, "Demand response through ventilation and latent load adjustment for commercial buildings in humid climate zones," *Applied Energy*, vol. 373, p. 123940, 2024.

[5] H. Oh, W. Jin, P. Peng, J. A. Winick, D. Sickinger, D. Sartor, Y. Zhang, K. Beckers, K. Kitz, D. Acero-Allard, T. A. Atkinson, and P. Dobson, "Techno-economic performance of reservoir thermal energy storage for data center cooling system," *Applied Energy*, vol. 391, p. 125858, 2025.

[6] C. Diaz-Londono, D. Enescu, F. Ruiz, and A. Mazza, "Experimental modeling and aggregation strategy for thermoelectric refrigeration units as flexible loads," *Applied Energy*, vol. 272, p. 115065, 2020.

[7] A. Cordoba-Pacheco, C. Diaz-Londono, and F. Ruiz, "Data-driven aggregation control for thermoelectric loads in demand response," *IFAC-PapersOnLine*, vol. 55, no. 40, pp. 205–210, 2022, 1st IFAC Workshop on Control of Complex Systems COSY 2022.

[8] S. L. Brunton and J. N. Kutz, *Data-driven science and engineering: Machine learning, dynamical systems, and control*. Cambridge University Press, 2022.

[9] A. A. Feldbâum, "Dual control theory problems," *IFAC Proceedings Volumes*, vol. 1, no. 2, pp. 541–550, 1963, 2nd International IFAC Congress on Automatic and Remote Control: Theory, Basle, Switzerland, 1963.

[10] B. Wittenmark, "Adaptive dual control methods: An overview," *IFAC Proceedings Volumes*, vol. 28, no. 13, pp. 67–72, 1995, 5th IFAC Symposium on Adaptive Systems in Control and Signal Processing 1995, Budapest, Hungary, 14-16 June, 1995.

[11] M. Milanese and A. Vicino, "Optimal estimation theory for dynamic systems with set membership uncertainty: An overview," *Automatica*, vol. 27, no. 6, pp. 997–1009, 1991.

[12] M. Milanese and G. Belforte, "Estimation theory and uncertainty intervals evaluation in presence of unknown but bounded errors: Linear families of models and estimators," *IEEE Transactions on Automatic Control*, vol. 27, no. 2, p. 408 – 414, 1982.

[13] F. Blanchini and S. Miani, "Set-theoretic methods in control," *Systems and Control: Foundations and Applications*, vol. 0, p. 1 – 630, 2015.

[14] A. Parsi, D. Liu, A. Iannelli, and R. S. Smith, "Dual adaptive MPC using an exact set-membership reformulation," *IFAC-PapersOnLine*, vol. 56, no. 2, pp. 8457–8463, 2023, 22nd IFAC World Congress.

[15] E. Meza, A. D. Duca, F. Ruiz, and C. Ocampo-Martinez, "Adaptive scenario-based predictive control: A set membership learning approach," in *2025 IFAC 7th Conference on Intelligent Control and Automation Sciences (ICONS)*, Padova, Italy, 2025.

[16] Q. Shi, A. Cordoba-Pacheco, and F. Ruiz, "Efficient learning control through zonotopic set membership estimation and scenario MPC," in *2025 European Control Conference (ECC)*, 2025, pp. 1–6.

[17] G. Schildbach, L. Fagiano, C. Frei, and M. Morari, "The scenario approach for stochastic model predictive control with bounds on closed-loop constraint violations," *Automatica*, vol. 50, no. 12, p. 3009–3018, Dec. 2014. [Online]. Available: http://dx.doi.org/10.1016/j.automatica.2014.10.035