A Privacy-preserving Method to Optimize Distributed Resource Allocation
Olivier Beaude, Pascal Benchimol, Stéphane Gaubert, Paulin Jacquot, Nadia Oudjane

To cite this version:
Olivier Beaude, Pascal Benchimol, Stéphane Gaubert, Paulin Jacquot, Nadia Oudjane. A Privacy-preserving Method to Optimize Distributed Resource Allocation. 2019. hal-02262271

HAL Id: hal-02262271
https://hal.archives-ouvertes.fr/hal-02262271
Submitted on 12 Aug 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
A PRIVACY-PRESERVING METHOD TO OPTIMIZE DISTRIBUTED RESOURCE ALLOCATION

OLIVIER BEAUDE †, PASCAL BENCHIMOL‡, STEPHANE GAUBERT §, PAULIN JACQUOT†§, AND NADIA OUDJANE†

Abstract. We consider a resource allocation problem involving a large number of agents with individual constraints subject to privacy, and a central operator whose objective is to optimize a global, possibly nonconvex, cost while satisfying the agents’ constraints, for instance an energy operator in charge of the management of energy consumption flexibilities of many individual consumers. We provide a privacy-preserving algorithm that does compute the optimal allocation of resources, avoiding each agent to reveal her private information (constraints and individual solution profile) neither to the central operator nor to a third party. Our method relies on an aggregation procedure: we compute iteratively a global allocation of resources, and gradually ensure existence of a disaggregation, that is individual profiles satisfying agents’ private constraints, by a protocol involving the generation of polyhedral cuts and secure multiparty computations (SMC). To obtain these cuts, we use an alternate projection method, which is implemented locally by each agent, preserving her privacy needs. We address especially the case in which the local and global constraints define a transportation polytope. Then, we provide theoretical convergence estimates together with numerical results, showing that the algorithm can be effectively used to solve the allocation problem in high dimension, while addressing privacy issues.

1. Introduction.

1.1. Motivation. Consider an operator of an electricity microgrid optimizing the joint production schedules of renewable and thermal power plants in order to satisfy, at each time period, the consumption constraints of its consumers. To optimize power generation or market costs and the integration of renewable energies, this operator relies on demand response techniques, that is, taking advantage of the flexibilities of some of the consumers electric appliances—those which can be controlled without impacting the consumer’s comfort, as electric vehicles or water heaters [18]. However, for privacy reasons, consumers are not willing to provide neither their consumption constraints nor their consumption profiles to a central operator or any third party, as this information could be used to infer private information such as their presence at home.

The global problem of the operator is to find an allocation of power (aggregate consumption) \( p = (p_t) \) at each time period (resource) \( t \in \mathcal{T} \), such that \( p \in \mathcal{P} \) (feasibility constraints of power allocation, induced by the power plants constraints). Besides, this aggregate allocation has to match an individual consumption profile \( x_n = (x_{n,t})_{t \in \mathcal{T}} \) for each of the consumer (agent) \( n \in \mathcal{N} \) considered. The problem can be written as follows:

\[
\begin{align*}
\text{(1.1a)} & \quad \min_{x \in \mathbb{R}^{N \times T}, \; p \in \mathcal{P}} f(p) \\
\text{(1.1b)} & \quad x_n \in \mathcal{X}_n, \quad \forall n \in \mathcal{N} \\
\text{(1.1c)} & \quad \sum_{n \in \mathcal{N}} x_{n,t} = p_t, \quad \forall t \in \mathcal{T},
\end{align*}
\]

The (aggregate) allocation \( p \) can be made public, that is, revealed to all agents. However, the individual constraint set \( \mathcal{X}_n \) and individual profiles \( x_n \) constitute private information of agent \( n \), and should not be revealed to the operator or any third party.

It will be helpful to think of problem (1.1) as the combination of two interdependent subproblems:

i) given an aggregate allocation \( p \), the disaggregation problem consists in finding, for each agent \( n \), an individual profile \( x_n \) satisfying her individual constraint (1.1b), so that constraint (1.1c) is satisfied,
or equivalently, solve the problem:

\[
\begin{align*}
(1.2a) & \quad \text{Find } x \in \mathcal{Y}_p \cap \mathcal{X} \\
(1.2b) & \quad \text{where } \mathcal{Y}_p \overset{\text{def}}{=} \{ y \in \mathbb{R}^{NT} \mid y^\top 1_N = p \} \text{ and } \mathcal{X} \overset{\text{def}}{=} \prod_{n \in \mathcal{N}} \mathcal{X}_n .
\end{align*}
\]

When (1.2) has a solution, we say that a \textit{disaggregation} exists for \( p \):

ii) For a given subset \( Q \subset \mathcal{P} \), we define the \textit{master problem},

\[
(1.3) \quad \min_{p \in Q} f(p) .
\]

When \( Q \) is precisely the set of aggregate allocations for which a disaggregation exists, the optimal solutions of the master problem correspond to the optimal solutions of (1.1).

Aside from the example above, \textit{resource allocation problems} (optimizing common resources shared by multiple agents) with the same structure as (1.1), find many applications in energy [25, 18], logistics [22], distributed computing [24], health care [29] and telecommunications [37]. In these applications, several entities or agents (e.g. consumers, stores, tasks) share a common resource (energy, products, CPU time, broadband) which has a global cost for the system. For large systems composed of multiple agents, the dimension of the overall problem can be prohibitive: a solution is to rely on decomposition and distributed approaches [7, 28, 32]. Besides, agents’ individual constraints are often subject to privacy issues [17]. These considerations have paved the way to the development of privacy-preserving, or non-intrusive methods and algorithms, e.g. [36, 20].

In this work, except in Section 4, we consider that each agent \( n \in \mathcal{N} \) has a global demand constraint (e.g. energy demand or product quantity), which confers to the disaggregation problem the particular structure of a transportation polytope [8]: the sum over the agents is fixed by the aggregate solution \( p \), while the sum over the \( T \) resources are fixed by the agent global demand constraint. Besides, individual constraints can also include minimal and maximal levels on each resource, as for instance electricity consumers require, through their appliances, a minimal and maximal power at each time period.

1.2. Main Results. The main contribution of the paper is to provide a non-intrusive and distributed algorithm (Algorithm 3.4) that computes an aggregate resource allocation \( p \), optimal solution of the—possibly nonconvex—optimization problem (1.1), along with feasible individual profiles \( x \) for agents, without revealing the individual constraints of each agent to a third party, either another agent or a central operator. The algorithm solves iteratively instances of \textit{master problems} \( \min_{p \in \mathcal{P}(s)} f(p) \) by constructing successive approximations \( \mathcal{P}^{(s)} \subset \mathcal{P} \) of the aggregate feasible set of (1.1) for which a disaggregation exists, by adding to the set \( \mathcal{P}^{(s)} \) a new constraint on \( p \) (i.e. a cutting plane), before solving the next master problem. We shall see that this cutting plane can computed and added to the master problem without revealing any individual information on the agents.

More precisely, to identify whether or not disaggregation (1.2) is feasible and to add a new constraint in the latter case, our algorithm relies on the alternate projections method (APM) [31, 14] for finding a point in the intersection of convex sets. Here, we consider the two following sets: on the one hand, the affine space of profiles \( x \in \mathbb{R}^{NT} \) aggregating to a given resource allocation \( p \), and on the other hand, the set defined by all agents individual constraints (demands and bounds). As the latter is defined as a Cartesian product of each agent’s feasibility set, APM can operate in a distributed fashion. The sequence constructed by the APM converges to a single point if the intersection of the convex sets is nonempty, and it converges to a periodic orbit of length 2 otherwise. If the APM converges to a periodic orbit, meaning that the disaggregation is not feasible, we construct from this orbit a polyhedral \textit{cut}, i.e. a linear inequality satisfied by all feasible solutions \( p \) of the global problem (1.1), but violated from the current resource allocation (Theorem 3.3). Adding this cut to the \textit{master problem} (1.3) by updating \( Q \) to a specific subset, we can recompute a new resource allocation and repeat this procedure until disaggregation is possible. At this stage, the use of a cryptographic protocol, secure multiparty computation, allows us to preserve the privacy of agents. Another main result stated in
this paper is the explicit upper bound on the convergence speed of APM in our framework (Theorem 3.2), which is obtained by spectral graph theory methods, exploiting also geometric properties of transportation polytopes. This explicit speed shows a linear impact of the number of agents, which is a strong argument for the applicability of the method in large distributed systems.

1.3. Related Work. A standard approach (e.g. [28, 33, 30]) to solve resource allocation problems in a distributed way is to rely on a a Lagrangian based decomposition technique: for instance dual subgradient methods [6, Ch. 6] or ADMM [13]. Such techniques are generally used to decompose a large problem into several subproblems of small dimension. However, those methods often require global convexity hypothesis, which are not satisfied in many practical problems (e.g. MILP). We refer the reader to [6, Chapter 6] for more background. On the contrary, our method can be used when the allocation problem (1.1) is not convex.

As developed in Section 4, the method proposed here can be related to Bender’s decomposition [5]. The difference with Bender’s approach is in the way of generating a new cut to add in the master problem: instead of solving linear programs, we use APM and our theoretical results, which provides a decentralized, privacy-preserving and scalable procedure. In contrast, at each stage, Benders’ algorithm requires to solve a linear program requiring the knowledge of the private constraints of each individual agent (see Subsection 4.1 for more details).

The problem of the aggregation of constraints has been studied in the field of energy, in the framework of smart grids [25, 2]. In [25], the authors study the management of energy flexibilities and propose to approximate individual constraints by zonotopic sets to obtain an aggregate feasible set. A centralized aggregated problem is solved via a subgradient method, and a disaggregation procedure of a solution computes individual profiles. In [2], the authors propose to solve the economic power dispatch of a microgrid, subject to several agents private constraints, by using a Dantzig-Wolfe decomposition method.

The APM has been the subject of several works in itself [14, 3, 4]. The authors of [9] provide general results on the convergence rate of APM for semi-algebraic sets. They show that the convergence is geometric for polyhedra. However, it is generally hard to compute explicitly the geometric convergence rate of APM, as this requires to bound the singular values of certain matrices arising from the polyhedral constraints. A remarkable example where an explicit convergence rate for APM has been established is in [27], where the authors consider a different class of polyhedra arising in submodular optimization. A common point with our results is the use of spectral graph theory arguments to estimate singular values.

1.4. Structure. Section 2 describes the class of resource allocation problems we address in this paper, and formulate the idea of the decomposition with the disaggregation subproblems. In Section 3, we focus on APM, the subroutine used to solve the disaggregation subproblems. After stating results on the convergence of APM, In Subsection 3.1, we show the key result on which relies the proposed decomposition: how to generate a new cut to add in the master problem, from the output of APM. In Subsection 3.2, we show how to improve the privacy of the proposed procedure by using secure multiparty computation techniques. In Subsection 3.3, we prove an explicit upper bound on the rate of convergence of APM in our case. In Section 4, we generalize part of our results and propose a modified algorithm in the case where agents constraints are polyhedral. Finally, in Section 5, we propose numerical examples of the method: Subsection 5.1 gives an illustrative toy example in dimension $T = 4$, while in Subsection 5.2, we consider a larger scale, nonconvex example, coming from the microgrid application exposed at the beginning of the introduction.

Notation. In the remaining of the paper, bold font $\mathbf{x}$ is used to denote a vector, while $x$ refers to a scalar quantity. $v^\top$ denotes the transpose of vector $v$. Calligraphic letters such as $\mathcal{T}, \mathcal{X}, \mathcal{X}$ are used to denote sets, and if $\mathcal{T}_0 \subset \mathcal{T}$, the set $\mathcal{T}_0^c = \{t \in \mathcal{T} \setminus \mathcal{T}_0\}$ denotes the complementary set of $\mathcal{T}_0$. The notation $\mathcal{U}(\{a, b\})$ stands for the uniform distribution on $[a, b]$. The notation $P_C(.)$ refers to the Euclidean projection onto the convex set $C$. For $d \in \mathbb{N}$, $1_d$ denotes the vector of ones $(1 \ldots 1)^\top \in \mathbb{R}^d$.
2.1. A Decomposition based on Disaggregation. As stated in the introduction, we consider a centralized entity (e.g., an energy operator) interested in minimizing a possibly nonconvex cost function $p \mapsto f(p)$, where $p \in \mathbb{R}^T$ is the aggregate allocation of $T$ dimensional resources (for example power production over $T$ time periods). This resource allocation $p$ is to be shared between a set $\mathcal{N}$ of $N$ individual agents, each agent obtaining a part $x_n \in X_n$, where $X_n$ denotes the individual feasibility set of agent $n$.

The global problem the operator wants to solve is described in (1.1). The idea behind the results of this paper is that, in problem (1.1), the constraints set $X_n$ and individual profile $x_n$ are confidential to agent $n$ and should not be disclosed to the central operator or to another agent.

Let us define the set $\mathcal{P}_D$ of feasible aggregate allocations that are disaggregable as:

\begin{equation}
\mathcal{P}_D \overset{\text{def}}{=} \{ p \in \mathcal{P} \mid \exists x \in X : p = \sum_n x_n \}.
\end{equation}

Feasibility of problem (1.1) is equivalent to having $\mathcal{P}_D$ not empty.

Constraints for each agent are composed of a global demand over the resources and lower and upper bounds over each resource, as given below:

**Assumption 1.** For each $n \in \mathcal{N}$, there exists $E_n > 0$, $\underline{x}_n \in \mathbb{R}^T$, $\overline{x}_n \in \mathbb{R}^T$ such that:

\begin{equation}
X_n = \{ x \in \mathbb{R}^T : \sum_{t \in T} x_{n,t} = E_n \text{ and } \underline{x}_{n,t} \leq x_{n,t} \leq \overline{x}_{n,t} \} \neq \emptyset.
\end{equation}

In particular, $X_n$ is convex and compact. Given an allocation $p$, the structure obtained on the matrix $(x_{n,t})_{n,t}$, where sums of coefficients along columns and along rows are fixed, is often referred to as transportation problem which has many applications (see e.g. [1, 26]). We focus on this case in Sections 2 and 3, while in Section 4, we shall give a generalization of some of our results in the general case where $X_n$ is a polyhedron.

Given a particular allocation $p \in \mathcal{P}$, the operator will be interested to know if this allocation is disaggregable, that is, if there exists individual profiles $(x_n)_{n \in \mathcal{N}} \in \prod_n X_n$ summing to $p$, or equivalently if the disaggregation problem (1.2) has a solution.

Following (1.2), the **disaggregate** profile refers to $x$, while the **aggregate** profile refers to the allocation $p$. Problem (1.2) may not always be feasible. Some necessary conditions for a disaggregation to exist, obtained by summing the individual constraints on $\mathcal{N}$, are the following aggregate constraints:

\begin{align}
(2.3a) & \quad p^\top 1_T = E^\top 1_N \\
(2.3b) & \quad \underline{x}^\top 1_N \leq p \leq \overline{x}^\top 1_N.
\end{align}

Those conditions are not sufficient in general, as explained in the following section.

2.2. An equivalent flow problem and Hoffman conditions. The particular structure of the problem we consider implies that we can write it as a flow problem in a graph, as stated in Proposition 2.2. We refer the reader to the book [11, Chapter 3] for the terminology.

**Definition 2.1.** Consider a directed graph $G = (V, E)$ with vertices $V$ and edges $E \subset V \times V$, and demands $d : V \to \mathbb{R}$ (where $d_v < 0$ means that $v$ is a production node), edge lower capacities $\ell : E \to \mathbb{R}_+$ and upper capacities $u : E \to \mathbb{R}_+$. A flow on $G$ is a function $x : E \to \mathbb{R}_+$ such that $x$ satisfies the capacity constraints, that is $\forall e \in E, \ell_e \leq x_e \leq u_e$, and Kirchoff’s law, that is, $\forall v \in V, \sum_{e \in \delta^+_v} x_e = d_v + \sum_{e \in \delta^-_v} x_e$, where $\delta^+_v$ (resp. $\delta^-_v$) is the set of edges ending at (resp. departing from) vertex $v$.

The following proposition is immediate:

**Proposition 2.2.** Consider the bipartite graph $G$ with vertices $V = T \cup N$ and edges $E = \{(t, n)\}_{t \in T, n \in N}$. Define demands on nodes $T$ by $d_t = -p_t$ and demands on nodes $N$ by $d_n = E_n$. Assign to each edge $(t, n)$ an upper capacity $u_{n,t} = \overline{x}_{n,t}$ and lower capacity $\ell_{n,t} = \underline{x}_{n,t}$. Then, finding a solution $x$ to (1.2) is equivalent to finding a feasible flow in $G$. 

Hoffman [16] gave a necessary and sufficient condition for the flow problem to be feasible. This generalizes a result of Gale (1957). The stated condition is intuitive: there cannot be a subset of nodes whose demand exceeds its “import capacity”.

**Theorem 2.3 (16).** Given a digraph $G = (V, E)$ with demand $d \in \mathbb{R}^V$ such that $d(V) = 0$ and capacities $\ell \in (\mathbb{R} \cup \{-\infty\})^E$ and $u \in (\mathbb{R} \cup \infty)^E$ with $\ell \leq u$, there exists a feasible flow $x \in E \rightarrow \mathbb{R}^+$ on $G$ if and only if:

\[
\forall A \subset V, \sum_{e \in \delta^-(A)} u_e - \sum_{v \in A} d_v \geq \sum_{e \in \delta^+(A^c)} \ell_e,
\]

where $\delta_+(A) \overset{\text{def}}{=} \{(u, v) \in E | u \in A^c, v \in A\}$ is the set of edges coming to set $A$ and $A^c \overset{\text{def}}{=} V \setminus A$.

The following Proposition 2.4 translates Theorem 2.3 in our framework:

**Proposition 2.4.** Disaggregation is possible iff:

\[
\forall T_0 \subset T, \forall N_0 \subset N, \sum_{t \in T_0} p_t - \sum_{n \in N_0} E_n + \sum_{t \notin T_0, n \in N_0} E_n - \sum_{t \in T_0, n \notin N_0} E_n \leq \sum_{t \in T_0, n \notin N_0} D_n
\]

**Proof.** We apply (2.4) with $A \overset{\text{def}}{=} \mathcal{T}_0^c \cup \mathcal{N}_0^c$ and use the equality $d(V) = 0 = \sum_{v \in \mathcal{A}} d_v + \sum_{v \in \mathcal{A}^c} d_v$.

From Theorem 2.3 or Proposition 2.4 above, one can see that the aggregate constraints (2.3) are in general not sufficient to ensure that the disaggregation problem has a solution.

For a given set $T_0$, there is a choice of $N_0$ which leads to the strongest inequality (2.5), namely:

\[
\sum_{t \in T_0} p_t \leq \min_{N_0 \subset \mathcal{N}} \left\{ \sum_{n \in N_0} E_n - \sum_{t \notin T_0, n \in N_0} E_n + \sum_{t \in T_0, n \notin N_0} D_n \right\}
\]

In this way, we get $2^T - 2$ inequalities corresponding to the proper subsets $T_0 \subset T$. Moreover, in general, these $2^T - 2$ inequalities are not redundant. Although this is not stated in [16], this is a classical result whose proof is elementary.

3. Disaggregation based on APM.

3.1. Generation of Hoffman’s constraints with APM. In this section, we propose an algorithm that solves (1.1) while preserving the privacy of each agent constraints $X_n$ and individual profile $x_n \in \mathbb{R}^T$. To do this, the proposed algorithm is implemented in a decentralized fashion and relies on the method of alternate projections method (APM) to solve the disaggregation problem (1.2).

Let us consider the polyhedron enforcing the agents constraints:

$X \overset{\text{def}}{=} X_1 \times \cdots \times X_N$ where $X_n \overset{\text{def}}{=} \{ x_n \in \mathbb{R}^T_+ | \sum_{t \in T} x_{n,t} = E_n \text{ and } \forall t, \; x_{n,t} \leq \bar{x}_{n,t} \leq \bar{x}_{n,t} \}$.
Besides, given an allocation $p \in \mathcal{P}$, we consider the set of profiles aggregating to $p$:

$$\mathcal{Y}_p \overset{\text{def}}{=} \{ x \in \mathbb{R}^{NT} \mid \forall t \in T, \sum_{n \in \mathcal{N}} x_{n,t} = p_t \}.$$ 

Note that $\mathcal{Y}_p$ is an affine subspace of $\mathbb{R}^{NT}$ (to be distinguished from $\mathcal{P}$ which is a subset of $\mathbb{R}^T$), and that $\mathcal{Y}_p \cap \mathcal{X}$ is empty iff $p \notin \mathcal{P}_D$, according to the definition of $\mathcal{P}_D$ in (2.1). The idea of the proposed algorithm is to build a finite sequence of decreasing subsets $(\mathcal{P}(s))_{0 \leq s \leq S}$ such that:

$$\mathcal{P} = \mathcal{P}(0) \supset \mathcal{P}(1) \supset \cdots \supset \mathcal{P}(S) \supset \mathcal{P}_D.$$ 

At each iteration, a new aggregate resource allocation $p^{(s)}$ is obtained by solving an instance of the master problem introduced in (1.3) with $Q = \mathcal{P}(s)$:

\begin{align*}
(3.1a) & \quad \min_{p \in \mathbb{R}^T} f(p) \\
(3.1b) & \quad \text{s.t. } p \in \mathcal{P}(s). 
\end{align*}

In the remaining of the paper, we will refer to (3.1) as an instance of master problem. Our procedure relies on the following immediate observation:

**Proposition 3.1.** If $p^{(s)}$ is a solution of (3.1), and $\mathcal{Y}_{p^{(s)}} \cap \mathcal{X} \neq \emptyset$ and $x \in \mathcal{Y}_{p^{(s)}} \cap \mathcal{X}$, then $(p^{(s)}, x)$ is an optimal solution of the initial problem (1.1).

Having in hands a solution $p^{(s)}$, we can check if $\mathcal{Y}_{p^{(s)}} \cap \mathcal{X} \neq \emptyset$ using APM on $\mathcal{X}$ and $\mathcal{Y}_{p^{(s)}}$, as described in Algorithm 3.1 below (where $\mathcal{Y} = \mathcal{Y}_p$).

**Algorithm 3.1** Alternate Projections Method (APM)

**Require:** Start with $y^{(0)}$, $k = 0$; $\varepsilon_{cvg}$, a norm $\| \|$ on $\mathbb{R}^{NT}$.

1. repeat
2. \hspace{1em} $x^{(k+1)} \leftarrow P_X(y^{(k)})$
3. \hspace{1em} $y^{(k+1)} \leftarrow P_Y(x^{(k+1)})$
4. \hspace{1em} $k \leftarrow k + 1$
5. until $\|x^{(k)} - x^{(k-1)}\| < \varepsilon_{cvg}$

The idea of using cyclic projections to compute a point in the intersection of two sets comes from von Neumann [31], where the idea was applied for affine subspaces. Convergence of APM is proved by Theorem 3.2:

**Theorem 3.2** ([14]). Let $\mathcal{X}$ and $\mathcal{Y}$ be two closed convex sets with $\mathcal{X}$ bounded, and let $(x^{(k)})_k$ and $(y^{(k)})_k$ be the two infinite sequences generated by APM on $\mathcal{X}$ and $\mathcal{Y}$ (Algorithm 3.1) with $\varepsilon_{cvg} = 0$. Then there exists $x^\infty \in \mathcal{X}$ and $y^\infty \in \mathcal{Y}$ such that:

\begin{align*}
(3.2a) & \quad x^{(k)} \xrightarrow{k \to \infty} x^\infty, \quad y^{(k)} \xrightarrow{k \to \infty} y^\infty; \\
(3.2b) & \quad \|x^\infty - y^\infty\|_2 = \min_{x \in \mathcal{X}, y \in \mathcal{Y}} \|x - y\|_2. 
\end{align*}

In particular, if $\mathcal{X} \cap \mathcal{Y} \neq \emptyset$, then $(x^{(k)})_k$ and $(y^{(k)})_k$ converge to a same point $x^\infty \in \mathcal{X} \cap \mathcal{Y}$.

The convergence theorem is illustrated in Figure 3.1 in the case where $\mathcal{X} \cap \mathcal{Y} = \emptyset$, that is, when the disaggregation problem (1.2) is not feasible. The idea of the algorithm proposed in this paper is, in the case where $\mathcal{Y}_{p^{(s)}} \cap \mathcal{X} = \emptyset$, to use the resulting vectors $x^\infty$ and $y^\infty$ to construct a new subset $\mathcal{P}(s+1)$ by adding a constraint of type (2.5) to $\mathcal{P}(s)$; indeed, from Proposition 2.4, we know that, if $\mathcal{Y}_{p^{(s)}} \cap \mathcal{X} = \emptyset$, there exists at least one inequality (2.5) violated.

The difficulty is to guess one of this violated inequality among the set of $2^T$ possible inequalities. It turns out that, using the output of APM, we can build such an inequality.
Once (2.6) can be rewritten as:

\[
\text{min} \frac{1}{2} \|x - y\|_2^2 \quad \forall n \in \mathcal{N}, \sum_{t \in T} x_{n,t} = E_n \quad (\lambda_n)  
\]

(3.5) defines a valid inequality for the disaggregation problem violated by the current allocation following optimization problem:

\[
\text{min} \sum_{n,t} \frac{1}{2} \|x_{n,t} - y_{n,t}\|_2^2 - \lambda^T (\sum_{t \in T} x_{n,t} - E_n)_n - \mu^T (x_n - \bar{x}) - \nu^T (\sum_{n} y_n - p)  
\]

We notice that the stationarity condition of the Lagrangian with respect to the variable \(y_{n,t}\) yields:

\[
\forall n \in \mathcal{N}, \forall t \in T, \nu_t = x_{n,t} - y_{n,t} .  
\]

Let us consider the sets \(T_0 \subset T\) and \(N_0 \subset \mathcal{N}\) defined from the output of APM on \(X\) and \(Y_p\) as:

\[
T_0 \overset{\text{def}}{=} \{ t \in T \mid \exists n \in \mathcal{N}, y_{n,t} > x_{n,t} \} \quad \text{and} \quad N_0 \overset{\text{def}}{=} \{ n \in \mathcal{N} \mid E_n - \sum_{t \notin T_0} x_{n,t} - \sum_{t \in T_0} x_{n,t} < 0 \}.  
\]

In Theorem 3.3 below, we show that applying the inequality (2.5) with the sets \(T_0\) and \(N_0\) defined in (3.5) defines a valid inequality for the disaggregation problem violated by the current allocation \(p\).

The intuition behind the definition of \(T_0\) and \(N_0\) in (3.5) is the following: \(T_0\) is the subset of resources for which there is an over supply (which overcomes the upper bound for at least one agent). Once \(T_0\) is defined, \(N_0\) is the associated subset of \(N\) minimizing the right hand side of (2.6). Indeed, (2.6) can be rewritten as:

\[
\sum_{t \in T_0} p_t \leq \min_{N_0 \subset \mathcal{N}} \left\{ \sum_{n \in N_0} (E_n - \sum_{t \notin T_0} x_{n,t} - \sum_{t \in T_0} x_{n,t}) \right\} + \sum_{t \in T_0} x_{n,t} .  
\]

The following Theorem 3.3 is the key result on which relies the algorithm proposed in this paper.

**Theorem 3.3.** Consider the sequence of iterates \((x^{(k)}, y^{(k)})_{k \in \mathbb{N}}\) generated by the APM on \(X\) and \(Y_p\) (see Algorithm 3.1). Then one of the following holds:

(i) if \(X \cap Y_p \neq \emptyset\), then \(x^{(k)}, y^{(k)} \rightarrow x^\infty \in X \cap Y_p\);
(ii) else, if $\mathcal{X} \cap \mathcal{Y}_p = \emptyset$, then $x^{(k)} \xrightarrow[k \to \infty]{} x^\infty \in \mathcal{X}$ and $y^{(k)} \xrightarrow[k \to \infty]{} y^\infty \in \mathcal{Y}_p$. Then, considering the sets $\mathcal{T}_0$ and $\mathcal{N}_0$ as defined in (3.3) gives an inequality of Hoffman (2.5) violated by $p$, that is:

$$
\sum_{n \in \mathcal{N}_0} E_n - \sum_{t \in \mathcal{T}_0} p_t + \sum_{t \notin \mathcal{T}_0, n \notin \mathcal{N}_0} x_{n,t} - \sum_{t \notin \mathcal{T}_0, n \in \mathcal{N}_0} x_{n,t} < 0.
$$

Moreover, this Hoffman inequality can be written as a function of $x^\infty$ as:

$$
A_{\mathcal{T}_0}(x^\infty) < \sum_{t \in \mathcal{T}_0} p_t \text{ with } A_{\mathcal{T}_0}(x^\infty) \overset{\text{def}}{=} \sum_{t \in \mathcal{T}_0} \sum_{n \in \mathcal{N}} x_{n,t}.
$$

Before giving the proof of Theorem 3.3, we need to show some technical properties on the sets $\mathcal{T}_0, \mathcal{N}_0$. For simplicity of notations, we use $x$ and $y$ to denote $x^\infty$ and $y^\infty$ in Proposition 3.4 and the proof of Theorem 3.3.

**Proposition 3.4.** With $x \neq y$ solutions of problem (3.3) (outputs of the APM on $\mathcal{X}$ and $\mathcal{Y}_p$ with $\varepsilon_{cvg} = 0$),

(i) $\forall t \in \mathcal{T}_0, \forall n \notin \mathcal{N}_0, y_{n,t} \geq x_{n,t}$ and $x_{n,t} = x_{n,t}$;

(ii) $\mathcal{T}_0 = \{ t \mid \nu_t > 0 \} = \{ t \mid p_t > \sum_n x_{n,t} \}$, where $\nu_t$ is the optimal Lagrangian multiplier associated to (3.3d);

(iii) $\forall n \in \mathcal{N}_0, \lambda_n < 0$;

(iv) $\forall t \notin \mathcal{T}_0, \forall n \notin \mathcal{N}_0, x_{n,t} = x_{n,t}$;

(v) the sets $\mathcal{T}_0, \mathcal{T}_0^c, \mathcal{N}_0$ and $\mathcal{N}_0^c$ are nonempty.

The proof of Proposition 3.4 is technical and given in Appendix A. With Proposition 3.4, we are now ready to prove Theorem 3.3.

**Proof of Theorem 3.3.** We have:

$$
\sum_{n \in \mathcal{N}_0} E_n + \sum_{t \in \mathcal{T}_0, n \notin \mathcal{N}_0} x_{n,t} + \sum_{t \notin \mathcal{T}_0, n \in \mathcal{N}_0} x_{n,t} - \sum_{t \notin \mathcal{T}_0} x_{n,t} = \sum_{n \in \mathcal{N}_0} \sum_{t \in \mathcal{T}_0} x_{n,t} + \sum_{n \notin \mathcal{N}_0} x_{n,t} - \sum_{t \notin \mathcal{T}_0, n \in \mathcal{N}_0} x_{n,t} - \sum_{t \notin \mathcal{T}_0} x_{n,t} \quad \text{(from (3.3b))}
$$

$$
= \sum_{n \in \mathcal{N}_0} \left( \sum_{t \in \mathcal{T}_0} x_{n,t} + \sum_{n \notin \mathcal{N}_0} x_{n,t} \right) + \sum_{t \notin \mathcal{T}_0, n \notin \mathcal{N}_0} x_{n,t} - \sum_{t \notin \mathcal{T}_0, n \in \mathcal{N}_0} x_{n,t} - \sum_{t \notin \mathcal{T}_0} x_{n,t} \quad \text{(from Prop.3.4 (i) and (iv))}
$$

$$
= \sum_{t \in \mathcal{T}_0} \left( \sum_{n \in \mathcal{N}_0} x_{n,t} + \sum_{n \notin \mathcal{N}_0} x_{n,t} \right) - \sum_{t \notin \mathcal{T}_0} p_t = \sum_{t \in \mathcal{T}_0} \left( \sum_{n \in \mathcal{N}} x_{n,t} - \sum_{n \notin \mathcal{N}} y_{n,t} \right)
$$

$$
= \sum_{t \in \mathcal{T}_0} \left( \sum_{n \notin \mathcal{N}} \nu_t \right) = \sum_{t \in \mathcal{T}_0} \left( - \sum_{n \in \mathcal{N}} |x_{n,t} - y_{n,t}| \right)
$$

using the stationarity conditions (3.4) and for all $t \in \mathcal{T}_0, \nu_t > 0$ by Prop.3.4 (ii). Moreover, using:

$$
\sum_{t \in \mathcal{T}_0, n \in \mathcal{N}} (x_{n,t} - y_{n,t}) = \sum_{n \in \mathcal{N}} E_n - \sum_{t \in \mathcal{T}} p_t = 0,
$$

we see that:

$$
\sum_{t \notin \mathcal{T}_0} \left( - \sum_{n \in \mathcal{N}} |x_{n,t} - y_{n,t}| \right) = -(\|x - y\|_1)/2 < 0,
$$

\[\Box\]
which shows (3.6). We now show that inequality (3.7) is obtained by a rewriting of (3.6), indeed:

\[
\begin{align*}
\sum_{n \in N_0} E_n + & \sum_{n \in N_0} \sum_{t \in T, n \notin N_0} x_{n,t} - \sum_{n \in N_0} \sum_{t \notin T, n \notin N_0} x_{n,t} \\
= & \sum_{n \in N_0} \sum_{t \in T} x_{n,t} + \sum_{n \in N_0} \sum_{t \notin T, n \notin N_0} x_{n,t} - \sum_{n \in N_0} \sum_{t \notin T, n \notin N_0} x_{n,t} \\
= & \sum_{t \in T, n \in N_0} x_{n,t} + \sum_{t \notin T, n \notin N_0} x_{n,t} = \sum_{n \in N_0} \sum_{t \in T} x_{n,t} = A_{T_0}(x).
\end{align*}
\]

(from Prop.3.4 (i) and (iv))

Suppose, as before, that the two sequences generated by the APM on \(X\) and \(Y\) converge to two distinct points \(x^\infty\) and \(y^\infty\). Then, at each round \(k\), we can define from (A.1) and considering any \(n \in N\), the multiplier \(\nu(k) = y_n(k) - x_n(k)\) tends to \(y^\infty - x_n^\infty \equiv \nu^\infty\). The set \(T_0\) of Theorem 3.3 is:

\[
T_0^\infty = \{ t \in T \mid 0 < \nu(t) \},
\]

which raises an issue for practical computation, as \(\nu^\infty\) is only obtained \textit{ultimately} by APM, possibly in infinite time. To have access to \(T_0^\infty\) in \textit{finite} time, that is, from one of the iterates \((\nu(k))_k\), we consider the set:

\[
T_0^{(K)} = \{ t \in T \mid B \nu(t) < \nu(t) \},
\]

where \(\nu_{\text{cvg}}\) is the tolerance for convergence of APM as defined in Algorithm 3.1, \(B > 0\) is a constant, and \(K\) (depending on \(\nu_{\text{cvg}}\)) is the first integer such that \(\|x^{(K)} - x^{(K-1)}\| < \nu_{\text{cvg}}\).

We next show that we can choose \(B\) to ensure that \(T_0^{(K)} = T_0^\infty\) for \(\nu_{\text{cvg}}\) small enough. We rely on the geometric convergence rate of APM on polyhedra [9, 27]:

**Proposition 3.5.** [27] If \(X\) and \(Y\) are polyhedra, there exists \(\rho \in (0, 1)\) such that the sequence \((x^{(k)})_k\) and \((y^{(k)})_k\) generated by APM verify for all \(k \geq 1\):

\[
\|x^{(k+1)} - x^{(k)}\|_2 \leq \rho \|x^{(k)} - x^{(k-1)}\|_2 \quad \text{and} \quad \|y^{(k+1)} - y^{(k)}\|_2 \leq \rho \|y^{(k)} - y^{(k-1)}\|_2.
\]

Proposition 3.5 applies to any polyhedra \(X\) and \(Y\). In Subsection 3.3 we shall give an explicit upper bound on the constant \(\rho\) in the specific transportation case given by (1.1c) and (2.2).

From the previous proposition, we can quantify the distance to the limits in terms of \(\rho\):

**Lemma 3.6.** Consider an integer \(K\) such that the sequence \((x^{(k)})_k\) generated by APM satisfies \(\|x^{(K)} - x^{(K-1)}\| \leq \nu_{\text{cvg}}\), then we have for any \(K' \geq K - 1\):

\[
\|x^\infty - x^{(K')}\| \leq \frac{\nu_{\text{cvg}}}{1-\rho}.
\]

**Proof.** From Proposition 3.5, we have for any \(k \geq K\):

\[
\|x^{(k)} - x^{(K')}\| \leq \sum_{s=0}^{k-K'} \|x^{(K+s+1)} - x^{(K+s)}\| \leq \sum_{s=0}^{k-K'} \rho^s \|x^{(K'+1)} - x^{(K')}\| \leq \frac{1}{1-\rho} \nu_{\text{cvg}},
\]

so that, by taking the limit \(k \to \infty\), one obtains \(\|x^\infty - x^{(K')}\| \leq \frac{\nu_{\text{cvg}}}{1-\rho}\).

With this previous lemma, we can state the condition on \(B\) ensuring the desired property:

**Proposition 3.7.** Define \(\nu \equiv \min\{\|\nu(t)\| > 0\}\) (least nonzero element of \(\nu^\infty\)). If the constants \(B\) and \(\nu_{\text{cvg}} > 0\) are chosen such that \(B > \frac{1}{1-\rho}\) and \(\nu_{\text{cvg}} \times 2B < \nu\), and Algorithm 3.1 stops at iteration \(K\), then we have:

\[
T_0^{(K)} = T_0^\infty.
\]
Proof. Let \( t \in T_0^\infty \), that is \( \nu_t^\infty > 0 \) which is equivalent to \( \nu_t^\infty \geq \nu \) by definition of \( \nu \). We have:

\[
\nu_t^{(K)} = \frac{1}{N} (p_t - \sum_n x_{n,t}^{(K)}) = \frac{1}{N} (p_t - \sum_n x_{n,t}^{\infty}) + \frac{1}{N} (\sum_n x_{n,t}^{\infty} - \sum_n x_{n,t}^{(K)}) > \nu_t^{\infty} - \frac{\varepsilon \nu_t^{\infty}}{1-p} \geq \nu - \frac{\varepsilon \nu_t^{\infty}}{1-p} > \varepsilon_{cvg} (2B - \frac{1}{1-p}) ,
\]

and this last quantity is greater than \( B\varepsilon_{cvg} \) as soon as \( B \geq \frac{1}{1-p} \), thus \( t \in T_0^{(K)} \).

Conversely, if \( t \in T_0^{(K)} \), then:

\[
\nu_t^{\infty} = \frac{1}{N} (p_t - \sum_n x_{n,t}^{(K)}) = \frac{1}{N} (p_t - \sum_n x_{n,t}^{(K)}) - \frac{1}{N} (\sum_n x_{n,t}^{\infty} - \sum_n x_{n,t}^{(K)}) \geq \nu_t^{(K)} - \frac{B}{1-p} > \nu_t^{(K)} - B\varepsilon_{cvg} \geq (B - B)\varepsilon_{cvg} \geq 0 ,
\]

so that \( t \in T_0^\infty \). Furthermore, the “approximated” cut \( \sum_{t \in T_0} (\sum_{n \in N} x_{n,t}^{(K)} - p_t) \geq 0 \) is violated by the current value of \( p \) (or \( p^{(s)} \)) at iteration \( s \) in the algorithm as:

\[
\sum_{t \in T_0} \left( \sum_{n \in N} x_{n,t}^{(K)} - p_t \right) \leq \sum_{t \in T_0} \left( \sum_{n \in N} x_{n,t}^{(K)} - x_{n,t}^{\infty} \right) + \sum_{t \in T_0} \left( \sum_{n \in N} x_{n,t}^{\infty} - p_t \right) \leq \left\| x^{(K)} - x^{\infty} \right\|_1 - \frac{1}{2} \left\| x^{\infty} - y^{\infty} \right\|_1
\]

using (3.8) and (3.9). This last quantity is negative as soon as \( \left\| x^{(K)} - x^{\infty} \right\|_1 < \frac{1}{2} \left\| x^{\infty} - y^{\infty} \right\|_1 \), which holds in particular if \( B\varepsilon_{cvg} < \frac{1}{2} \left\| x^{\infty} - y^{\infty} \right\|_1 \).

This second proposition shows a surprising result: even if we do not have access to the limit \( x^{\infty} \), we can compute in finite time the exact left hand side term \( A_{T_0}(x^{\infty}) \) of the cut (3.7):

**Proposition 3.8.** Under the hypotheses of Proposition 3.4, we have:

\[
A_{T_0}(x^{(K)}) = \sum_{t \in T_0} \sum_{n \in N} x_{n,t}^{(K)} = A_{T_0}(x^{\infty}) .
\]

**Proof.** We start by showing some technical properties similar to Proposition 3.4:

**Lemma 3.9.** The iterate \( x^{(K)} \) satisfies the following properties:

(i) \( \forall t \in T_0, \forall n \notin N_0, x_{n,t}^{(K)} = x_{n,t}^{\infty} = x_{n,t} \);

(ii) \( \forall t \notin T_0, \forall n \notin N_0, x_{n,t}^{(K)} = x_{n,t}^{\infty} = x_{n,t} \).

The proof of Lemma 3.9 is similar to Proposition 3.4 and is given in Appendix B. Then, having in mind that \( T_0^{(K)} = T_0^\infty \) from Proposition 3.7, and \( N_0 \) is obtained from \( T_0^\infty \) by (3.5), we obtain:

\[
A_{T_0}(x^{(K)}) = \sum_{n \in N_0} \left( \sum_{t \notin T_0} x_{n,t}^{(K)} + \sum_{t \in T_0} x_{n,t}^{(K)} \right) - \sum_{t \notin T_0, n \in N_0} x_{n,t} - \sum_{t \in T_0, n \notin N_0} x_{n,t}^{(K)} \leq \sum_{n \in N_0} \sum_{t \notin T_0} x_{n,t}^{(K)} - \sum_{t \notin T_0, n \in N_0} x_{n,t} + \sum_{t \in T_0, n \notin N_0} x_{n,t} \quad \text{(from Lemma 3.9)}
\]

which equals to \( A_{T_0}(x^{\infty}) \) as we have \( \sum_{t \in T} x_{n,t}^{(K)} = E_n \) for each \( n \in N \).

Before presenting our algorithm using this last result, we focus on the technique of multiparty secure computation (SMC) which will be used here to ensure the privacy of agent’s constraints and profiles while running the APM.
3.2. Privacy-preserving Projections through SMC. APM, as described in Algorithm 3.1, enables a distributed implementation in our context, by the structure of the algorithm itself: the operator computes the projection on $\mathcal{Y}_p$ while each agent $n$ can compute, possibly in parallel, the projection on $X_n$ of the new profile transmitted by the operator. This enables each agent (as well as the operator) to keep her individual transmitted and not reveal it to the operator or other agents. However, each agent has to transmit back her newly computed individual profile to the operator for the next iteration.

Using a secure multi-party computation (SMC) protocol as introduced by [34], we can avoid this communication of individual profiles and perform APM without revealing the sequence of agent profiles $x$ to the aggregator.

For this, we use the fact that $\mathcal{Y}_p$ is an affine subspace and thus the projection on $\mathcal{Y}_p$ can be obtained explicitly component-wise. Indeed, summing (3.4) on $T$, we immediately obtain:

$$(3.12) \quad \forall n \in \mathcal{N}, \ [P_{\mathcal{Y}_p}(x)]_n = x_n + \frac{1}{N}(p - \sum_{m \in \mathcal{N}} x_m) .$$

Thus, having access to the aggregate profile $S \overset{\text{def}}{=} \sum_{n \in \mathcal{N}} x_n$, each agent can compute locally the component of the projection on $\mathcal{Y}_p$ of her profile, instead of transmitting the profile to the operator for computing the projection in a centralized way.

Using SMC, the sum $S$ can be computed in a non-intrusive manner and by several communications between agents and the operator, as described in Algorithm 3.2. The main idea of SMC is that, instead of sending her profile $x_n$, agent $n$ splits $x_{n,t}$ for each $t$ into $N$ random parts $(s_{n,t,m})_m$, according to an uniform distribution and summing to $x_{n,t}$ (Lines 2-3). Thus, each part $s_{n,t,m}$ taken individually does not reveal any information on $x_n$ nor on $X_n$, and can be sent to agent $m$. Once all exchanges of parts are completed (Line 5), and $n$ has herself received the parts from other agents, agent $n$ computes a new aggregate quantity $\sigma_n$ (Line 7), which does not contain either any information about any of the agents, and sends it to the operator (Line 8). The operator can finally compute the quantity $S = x^T 1_N = \sigma^T 1_N$.

Algorithm 3.2 SMC of Aggregate (SMCA) $\sum_{n \in \mathcal{N}} x_n$

Require: A profile $x_n$ for each agent $n \in \mathcal{N}$
1. for each agent $n \in \mathcal{N}$ do
2. Draw $\forall t, (s_{n,t,m})_m \overset{\text{def}}{=} \mathcal{U}([0,1])^{N-1}$
3. and set $\forall t, s_{n,t,N} \overset{\text{def}}{=} x_{n,t} - \sum_{m=1}^{N-1} s_{n,t,m}$
4. Send $(s_{n,t,m})_{t \in T}$ to agent $m \in \mathcal{N}$
5. done
6. for each agent $n \in \mathcal{N}$ do
7. Compute $\forall t, \sigma_{n,t} = \sum_{m \in \mathcal{N}} s_{m,t,n}$
8. Send $(\sigma_{n,t})_{t \in T}$ to operator
9. done
10. Operator computes $S = \sum_{n \in \mathcal{N}} \sigma_n$ (and broadcasts it to agents)

**Remark 1.** As $\sigma_n$, and $s_n$ are random by construction, an eavesdropper aiming to learn the profile $x_n$ of $n$ has no choice but to intercept all the communications of $n$ to all other agents (to learn $(s_{n,t,m})_{m \neq n}$ and $(s_{m,t,n})_{m \neq n}$) and to the operator (to learn $\sigma_n$).

We sum up in Algorithm 3.3 below the procedure of generating a new constraint as stated in Theorem 3.3 from the output of APM in finite time (see Proposition 3.7) and in a privacy-preserving way using SMC.

To choose $B$ and $\varepsilon_{cvg}$ satisfying the conditions of Proposition 3.7 a priori, one has to know the value of $\nu$. Although a conservative lower bound could be obtained by Diophantine arguments if we consider rationals as inputs of the algorithm, in practice it is easier and more efficient to proceed in an iterative manner for the value of $\varepsilon_{cvg}$. Indeed, one can start with $\varepsilon_{cvg}$ arbitrary large so that
The following Proposition 3.11 shows the correctness of our Algorithm 3.4.

Proposition 3.11. Let $B$ and $\varepsilon_{cvg}$ satisfy the conditions of Proposition 3.7 and 3.8. Then:

- if the problem (1.1) has no solution, Algorithm 3.4 exits at Line 4 after at most $2^T - 2$ iterations;
- else, Algorithm 3.4 computes, after at most $s \leq 2^T - 2$ iterations, an aggregate solution $p^{(s)} \in \mathcal{P}$.
Algorithm 3.4 Non-intrusive Optimal Disaggregation

Require: $s = 0$, $\mathcal{P}^{(0)} = \mathcal{P}$; DISAG = False
1: while Not DISAG do
2:   Solve $\min_{p \in \mathcal{P}(s)} f(p)$
3:   if problem infeasible then
4:      Exit
5:   else
6:      Compute $p^{(s)} = \arg \min_{p \in \mathcal{P}(s)} f(p)$
7:   end
8:   DISAG ← NI-APM($p^{(s)}$) (Algo. 3.3)
9:   if DISAG then
10:      Operator adopts $p^{(s)}$
11:   else
12:      Obtain $\mathcal{T}_0^{(s)}$, $A^{(s)}_{\mathcal{T}_0}$ from NI-APM($p^{(s)}$)
13:      $\mathcal{P}^{(s+1)} \leftarrow \mathcal{P}(s) \cap \{ p | \sum_{t \in \mathcal{T}_0^{(s)}} p_t \leq A^{(s)}_{\mathcal{T}_0} \}$
14:   end
15:   $s \leftarrow s + 1$
16: done

associated to individual profiles $(x^*)_n = NI-APM(p^{(s)})$ such that:

$$p^{(s)} \in \mathcal{P}, \forall n \in \mathcal{N}, x^*_n \in \mathcal{X}_n, \quad \| \sum_{n \in \mathcal{N}} x^*_n - p^{(s)} \| \leq \varepsilon_{dis} \quad \text{and} \quad f(p^{(s)}) \leq f^*,$$

where $f^*$ is the optimal value of problem (1.1).

Proof. The proof is immediate from Theorem 3.3, Proposition 3.7 and Proposition 3.8. □

Remark 2. The upper bound on the number of constraints added has no dependence on $N$ because, as stated in (2.6), once a subset of $\mathcal{T}$ is chosen, the constraint we add in the algorithm is found by taking the minimum by taking a minimum over the subsets of $\mathcal{N}$.

Although there exist some instances with an exponential number of independent constraints, this does not jeopardize the proposed method: in practice, the algorithm stops after a very small number of constraints added. Intuitively, we will only add constraints “supporting” the optimal allocation $p$. Thus, Algorithm 3.4 is a method which enables the operator to compute a resource allocation $p$ and the $N$ agents to adopt profiles $(x^*_n)_n$, such that $(x, p)$ solves the global problem (1.1), and the method ensures that both agent constraints (upper bounds $\mathcal{X}^*_n$, lower bounds $\mathcal{X}^ _n$, demand $\mathcal{E}_n$); and disaggregate (individual) profile $x^*_n$ (as well as the iterates $(x^{(k)})_k$ and $(y^{(k)})_k$ in NI-APM) are kept confidential by agent $n$ and can not be induced by a third party (either the operator or any other agent $m \neq n$).

Remark 3. A natural approach to address problem (1.1) in a distributed way, assuming that both the cost function $p \mapsto f(p)$ and the feasibility set $\mathcal{P}$ are convex, is to rely on Lagrangian based decomposition techniques. Examples of such methods are Dual subgradient methods [6, Chapter 6], auxiliary problem principle method [10], ADMM [13],[35] or bundle methods [23].

One can think of a privacy-preserving implementation of those techniques, where Lagrangian multipliers associated to the (relaxed) aggregation constraint $\sum_n x_n = p$ would be updated using the SMC technique as described in Algorithm 3.2. However, those techniques usually ask for strong convexity hypothesis: for instance, in ADMM, in order to keep the decomposition structure in agent by agent, a possibility is to use multi-blocs ADMM with $N+1$ blocs ($N$ agents and the operator), which is known to converge in the condition that strong convexity of the cost function in at least $N$ of the $N+1$ variables holds [12]. The complete study of privacy-preserving implementations of Lagrangian decomposition methods is left for further work.
The advantage of Algorithm 3.4 proposed in this paper is that convergence is ensured (see Proposition 3.10) even if the cost function \( p \mapsto f(p) \) and the feasibility set \( P \) are not convex, which is the case in many practical situations (see Subsection 5.2).

**Remark 4.** Algorithm 3.4 solves problem (1.1) in a privacy-preserving manner. For this, we use both the results stated in Theorem 3.3 and SMC to securely transmit the aggregate profile to the operator at each step. For the latter point, other techniques could be used instead of SMC such as consensus-based aggregation algorithms [15]. A comparison of the different possible techniques, relying on quantitative privacy indicators, would be interesting and is an avenue for further work.

In the next section, we focus on the convergence rate of APM in the particular case of transportation constraints, precisely on the geometric rate stated in Theorem 3.2.

### 3.3. Complexity Analysis of APM in the Transportation Case

In this section we analyze the speed of convergence of the alternate projections method (APM) described in Algorithm 3.1 on the sets \( \mathcal{X} \) and \( \mathcal{Y}_P \) defined in Section 2.

A general result in [9] gives an upper bound of the sequences generated by APM on subspaces \( \mathcal{X} \) and \( \mathcal{Y} \) if these two sets are semi-algebraic. In particular, it establishes the geometric convergence for polyhedral sets. However, as stated in [27], given two particular polyhedral sets \( \mathcal{X} \) and \( \mathcal{Y} \), it is not straightforward to deduce an explicit rate of convergence from their result.

The authors in [27] establish in a particular case a geometric convergence with an explicit upper bound on the convergence rate. They consider APM on two sets \( P \) and \( Q \), where \( P \) is a linear subspace and \( Q \) is a product of base polytopes of submodular functions.

In this section, we also establish an explicit upper bound on the convergence rate of APM in the transportation case, that is with \( \mathcal{X} \) and \( \mathcal{Y}_P \) defined in (2.2) and (1.2b):

**Theorem 3.12.** For the two sets \( \mathcal{X} \) and \( \mathcal{Y}_P \), the sequence of alternate projections converges to \( x^* \in \mathcal{X}, \ y^* \in \mathcal{X}^P \) satisfying \( \|x^* - y^*\| = \inf_{x \in \mathcal{X}, \ y \in \mathcal{Y}_P} \|x - y\| \), at the geometrical rate:

\[
\left\| x^{(k)} - x^* \right\| \leq 2 \left\| x^{(0)} - x^* \right\| \times \left( 1 - \frac{4}{N(T+1)^2(T-1)} \right)^k,
\]

and the symmetric inequalities hold for \( (y^{(k)})_k \).

For the remaining of this section, we will just use \( \mathcal{Y} \) to denote \( \mathcal{Y}_P \), as \( p \) remains fixed during APM. For the result stated in Theorem 3.12 above, we use several partial results of [27].

**Proof.** First, we use the fact stated in [27] that APM on subspaces \( U \) and \( V \) converge with geometric rate \( c_F(U, V)^2 \), where the rate is given by the square of the cosine of the Friedrichs angle between \( U \) and \( V \), given by:

\[
c_F(U, V) = \sup\{u^T v \mid u \in U \cap (U \cap V)^\perp, \ v \in V \cap (U \cap V)^\perp, \ \|u\| \leq 1, \ \|v\| \leq 1\}.
\]

An intuitive generalization of this result for polyhedra \( \mathcal{X} \) and \( \mathcal{Y} \), considering all affine subspaces supporting the faces of \( \mathcal{X} \) and \( \mathcal{Y} \) is given in [27]:

**Lemma 3.13 ([27]).** For APM on polyhedra \( \mathcal{X} \) and \( \mathcal{Y} \) in \( \mathbb{R}^D \), the convergence is geometric with rate bounded by the square of the maximal cosine of Friedrichs angle between subspaces supporting faces of \( \mathcal{X} \) and \( \mathcal{Y} \):

\[
\max_{x, y} c_F(\text{aff}_0(\mathcal{X}_x), \text{aff}_0(\mathcal{Y}_y)),
\]

where, for any \( x \in \mathbb{R}^D \), \( \mathcal{X}_x \) is the face of \( \mathcal{X} \) generated by direction \( x \) and \( \text{aff}_0(C) = \text{aff}(C) - c \) for some \( c \in C \) denotes the subspace supporting the affine hull of \( C \), for \( C = \mathcal{X}_x \) or \( C = \mathcal{Y}_y \).

In the remaining of the proof, we bound the quantity (3.13) for our polyhedra \( \mathcal{X} \) and \( \mathcal{Y} \).

For this, we use the space \( \mathbb{R}^{NT} = \mathbb{R}^T \times \cdots \times \mathbb{R}^T \), where the \((n-1)T+1\) to \( nT \) entries correspond to the profile of agent \( n \), for \( 1 \leq n \leq N \). As in [27], we use a result connecting angles between subspaces and the eigenvalues of matrices giving the directions of these spaces:
Lemma 3.14 ([27]). If \( A \) and \( B \) are matrices with orthonormal rows with same number of columns, then:

- if all singular values of \( AB^\top \) are equal to one, then \( c_F(\text{Ker} \, A, \text{Ker} \, B) = 0 \);
- else, \( c_F(\text{Ker} \, A, \text{Ker} \, B) \) is equal to the largest singular value of \( AB^\top \) among those that are smaller than one.

We are left with finding a matricial representation of the faces of polyhedra \( \mathcal{X} \) and \( \mathcal{Y} \) and, then, bounding the corresponding singular values.

In our case, the polyhedra \( \mathcal{Y} \) is an affine subspace \( \mathcal{Y} = \{ \mathbf{x} \in \mathbb{R}^{NT} | A \mathbf{x} = \sqrt{N}^{-1} \mathbf{p} \} \) where:

\[
A \overset{\text{def}}{=} \sqrt{N}^{-1} J_{1,N} \otimes I_T,
\]

where \( \otimes \) denotes the Kronecker product. The matrix \( A \) has orthonormal rows and the linear subspace associated to \( \mathcal{Y} \) is equal to \( \text{Ker}(A) \).

Obtaining a matricial representation of the faces of \( \mathcal{X} \) is more complex. The faces of \( \mathcal{X} \) are obtained by considering, for each \( n \in \mathcal{N} \), subsets of the time periods that are at lower or upper bound (respectively \( \overline{\mathcal{T}}_n \) and \( \underline{\mathcal{T}}_n \), with \( \underline{\mathcal{T}}_n \cap \overline{\mathcal{T}}_n = \emptyset \)). Considering a collection of such subsets, a face of \( \mathcal{X} \) can be written as:

\[
A_{(\mathcal{T}_n,\underline{\mathcal{T}}_n)} \overset{\text{def}}{=} \left\{ (\mathbf{x})_{n,t} | \forall n, \sum_t x_{n,t} = E_n \text{ and } \forall t \in \underline{\mathcal{T}}_n, x_{n,t} = e_{n,t} \text{, and } \forall t \in \mathcal{T}_n, x_{n,t} = x_{n,t} \right\}.
\]

For some particular collection of subsets \( (\overline{\mathcal{T}}_n,\underline{\mathcal{T}}_n) \), the set \( A_{(\mathcal{T}_n,\underline{\mathcal{T}}_n)} \) might be empty. The linear subspace associated to \( A_{(\mathcal{T}_n,\underline{\mathcal{T}}_n)} \) is given by \( \{ \mathbf{x} \in \mathbb{R}^{NT} | B \mathbf{x} = 0 \} = \text{Ker}(B) \), where the \( N \) first rows of \( B \), corresponding to the constraints \( \sum_t x_{n,t} = E_n \), are given before orthonormalization by:

\[
\sqrt{T}^{-1} I_N \otimes J_{1,T},
\]

and the matrix \( B \) has \( b \overset{\text{def}}{=} \sum_n |\mathcal{T}_n| \) more rows, where \( \mathcal{T}_n \overset{\text{def}}{=} \mathcal{T}_n \cup \overline{\mathcal{T}}_n \), corresponding to the saturated bounds. Each of this row is given by the unit vector \( e_{n,t} \in \mathbb{R}^{NT} \) for \( n \in \mathcal{N}, t \in \mathcal{T}_n \), which gives already an orthonormalized family of (unit) vectors. Therefore, a simple orthonormalized matrix \( B \) giving the direction of \( A_{(\mathcal{T}_n,\underline{\mathcal{T}}_n)} \) is given by:

\[
B \overset{\text{def}}{=} \left( \text{diag} \left( \sqrt{T - |\mathcal{T}_1|}^{-1} 1_{\mathcal{T}_1}^\top, \ldots, \sqrt{T - |\mathcal{T}_N|}^{-1} 1_{\mathcal{T}_N}^\top \right) | \text{diag}(B_{\mathcal{T}_1}, \ldots, B_{\mathcal{T}_N}) \right)^\top \in \mathcal{M}_{N+b,NT}(\mathbb{R}) ,
\]

where \( 1_{\mathcal{T}_n} \in \mathbb{R}^{T} \) is the vector where the indices in \( \mathcal{T}_n \) are equal to 1 and 0 otherwise, and \( B_{\mathcal{T}_n} \overset{\text{def}}{=} \sum_{1 \leq k \leq |\mathcal{T}_n|} E_{k,t} \) is the matrix \( |\mathcal{T}_n| \times T \) with indices of \( \mathcal{T}_n \). We obtain the double product:

\[
(AB^\top)(A^\top B) = \frac{1}{N} \left( \sum_n \frac{1_{k \notin \mathcal{T}_n \land \ell \notin \mathcal{T}_n}}{T - |\mathcal{T}_n|} B_{\mathcal{T}_n} \right)_{1 \leq k, \ell \leq T} + \frac{1}{N} \sum_n B_{\mathcal{T}_n}^\top B_{\mathcal{T}_n}
\]

We observe that:

- if \( t_0 \in \cap_{n=1}^N \mathcal{T}_n \), then \( e_{t_0} \) is an eigenvector associated to eigenvalue \( \lambda_{t_0} = 1 \);
- the vector \( 1_{\mathcal{T}} \overset{\text{def}}{=} (1_{t \notin \cap_n \mathcal{T}_n})_{t \in \mathcal{T}} \in \mathbb{R}^T \), where \( \mathcal{T} \overset{\text{def}}{=} \cup_n \mathcal{T}_n^c \), is an eigenvector associated to eigenvalue
\( \lambda = 1 \). Indeed, if we denote by \( \mathcal{N}_\theta = \{ n \in \mathcal{N} \mid \theta \in \mathcal{T}_n \} \), then \( |1_\mathcal{T}|_\theta = 1 \iff \mathcal{N}_\theta \neq \emptyset \), and for each \( \theta \in \mathcal{T} \):

\[
[(AB^\top)(A^\top B)]_\theta 1_\mathcal{T} = \frac{1}{N} \left( \sum_{i \in \mathcal{N}_\theta} \sum_{\ell} \frac{1}{T - |T_n|} \left[ 1_\mathcal{T} \right]_\ell + \sum_n \left[ 1_\theta \in \mathcal{T}_n \right] [1_\mathcal{T}]_\theta \right) = \frac{1}{N} \left( \sum_{i \in \mathcal{N}_\theta} \frac{T - |T_n|}{T - |T_n|} 1 + \sum_{i \in \mathcal{N}_\theta} 1 \times [1_\mathcal{T}]_\theta \right) = \frac{|\mathcal{N}_\theta| + |\mathcal{N}_\theta|[1_\mathcal{T}]_\theta}{N} = [1_\mathcal{T}]_\theta.
\]

To bound the other eigenvalues of the matrix \((AB^\top)(A^\top B)\), we rely on spectral graph theory arguments. Consider the weighted graph \( \mathcal{G} = (\mathcal{T}, \mathcal{E}) \) whose vertices are the time periods \( \mathcal{T} \) and each edge \((k, \ell) \in \mathcal{T} \times \mathcal{T}\) with \( k \neq \ell \) has weight \( S_{k,\ell} = \frac{1}{N} \sum_n \frac{|1_{(k,\ell) \in \mathcal{T}_n}|}{T - |T_n|} \) (if this quantity is zero, then there is no edge between \( k \) and \( \ell \)).

The matrix \( P \overset{\text{def}}{=} I_T - (AB^\top)(A^\top B) \) verifies for each \( k \in \mathcal{T} \):

\[
\sum_{\ell \neq k} -P_{k,\ell} = \sum_{\ell \neq k} \frac{1}{N} \sum_n \frac{|1_{(k,\ell) \in \mathcal{T}_n}|}{T - |T_n|} = \frac{1}{N} \sum_n \frac{1_{k \in \mathcal{T}_n} (T - |T_n| - 1)}{T - |T_n|} = \frac{1}{N} \sum_n (1 - 1_{k \in \mathcal{T}_n}) - \frac{1}{N} \sum_n \frac{1_{k \in \mathcal{T}_n}}{T - |T_n|} = P_{kk},
\]

which shows that \( P \) is the Laplacian matrix of graph \( \mathcal{G} \). As \( \text{Sp}(AB^\top A^\top B) = 1 - \text{Sp}(P) \), we want to have a lower bound on the least eigenvalue of \( P \) greater than 0, that we denote by \( \lambda_1 \).

By rearranging the indices of \( \mathcal{T} \) in two blocks \( \mathcal{T} \) and \( \mathcal{T}^c \), we observe that \( P \) can be written as a block diagonal matrix \( P = \text{diag}(P_{\mathcal{T}}, 0_{\mathcal{T}^c}) \). As we are only interested in the positive eigenvalues of \( P \), we can therefore study the linear application associated to \( P \) restricted to the subspace \( \text{Vect}(e_i)_{i \in \mathcal{T}} \).

As \( 1_\mathcal{T} \) is an eigenvector of \( P \) associated to \( \lambda_0 = 0 \), from the minmax theorem, we have:

\[
\lambda_1 = \min_{u \perp 1_\mathcal{T}, u \neq 0} \frac{u^\top Pu}{u^\top u}.
\]

Let us consider an eigenvector \( u \) realizing (3.16). Let \( u_* \overset{\text{def}}{=} \max_t u_t \) and \( u_* \overset{\text{def}}{=} \min_t u_t \) and let \( d_{s^* \cdot t^*} \) be the distance between \( s^* \) and \( t^* \) in \( \mathcal{G} \), and let \( (s^* \cdot t^*) \) denote a shortest path from \( s^* \) to \( t^* \) in \( \mathcal{G} \). As \( P \) is a Laplacian matrix, we have:

\[
(3.17) \quad u^\top Pu = \frac{1}{2} \sum_{k, \ell \in \mathcal{T}} -P_{k,\ell}(u_k - u_\ell)^2 \geq \frac{1}{2} \sum_{(k, \ell) \in (s^* \cdot t^*)} -P_{k,\ell}(u_k - u_\ell)^2 \geq \min_{k, \ell \in (s^* \cdot t^*)} (-P_{k,\ell}) \frac{(u_{s^*} - u_{t^*})^2}{d_{s^* \cdot t^*}},
\]

where the last inequality is obtained from Cauchy-Schwarz inequality.

Let us write the path \((s^* \cdot t^*) = (t_0, t_1, \ldots, t_d)\). As \((s^* \cdot t^*)\) is a shortest path, for each \( k \in \{0, d - 1\} \), the edge \((t_k, t_{k+1})\) exists so there exists \( n \in \mathcal{N} \) such that \( \{t_k, t_{k+1}\} \subset \mathcal{T}_n \). Moreover, for each \( n \), we have \( \mathcal{T}_n \cap \{t_0, \ldots, t_{k-1}, t_{k+2}, \ldots, t_d\} = \emptyset \), otherwise we could “shortcut” the path \((s^* \cdot t^*)\), thus we have \( |T_n| \geq d - 1 \). We obtain:

\[
-P_{t_k, t_{k+1}} = \frac{1}{N} \sum_n \frac{|1_{(t_k, t_{k+1}) \in \mathcal{T}_n}|}{T - |T_n|} \geq \frac{1}{N(T - d + 1)}.
\]

On the other hand, we have \((u_{s^*} - u_{t^*}) \geq u_{s^*} + \frac{u_{t^*}}{T - 1} u_{t^*} \geq \frac{T}{(T - 1)\sqrt{T}} \|u\|_2 \).

Using these bounds and (3.17), we obtain:

\[
u^\top Pu \geq \frac{(u_{s^*} - u_{t^*})^2}{N(T - d_{s^* \cdot t^*} + 1)d_{s^* \cdot t^*}} \geq \frac{4T}{N(T + 1)^2(T - 1)^2} \|u\|_2^2 \geq \frac{4}{N(T + 1)^2(T - 1)} \|u\|_2^2.
\]
Therefore, $\lambda_1 \geq \frac{4}{N(T+1)(T-1)} \kappa_{N,T}$ and the greatest singular value lower than one of $(AB^T)(A^T B)$ is $1 - \kappa_{N,T}$. We conclude by applying successively Lemma 3.14 and Lemma 3.13, to obtain the convergence rate stated in Theorem 3.12.

Figure 3.2: Evolution of the convergence rate, given as $\lambda_1(P)$ (lowest nonzero eigenvalue of $P$), with $N = 6$ and $T \in \{4, 6, 8, 12, 20, 60\}$. The worst convergence rate is evaluated by taking $100 \times T$ random draws of the sets $T_n \subset T$ for each $n$, and evaluating the eigenvalue of the matrix. The slope is around -0.93, which indicates that in practice the convergence rate is $O(T^{-1})$, faster than the upper bound in $O(T^{-3})$ established in Theorem 3.12.

4. Generalization to Polyhedral Agents Constraints. In this section, we extend our results to a more general framework where for each $n \in \mathcal{N}$, $\mathcal{X}_n$ is an arbitrary polyhedron, instead of having the particular structure given in (2.2). Let us now consider that $(\mathcal{X}_n)_n$ are polyhedra with, for each $n$

\[(4.1) \quad \mathcal{X}_n = \{x_n \in \mathbb{R}^T | A_n x_n \leq b_n \},\]

with $A_n \in \mathcal{M}_{T,k_n}(\mathbb{R})$ with $k_n \in \mathbb{N}$. The disaggregation problem (1.2), with $p \in \mathcal{P}$ fixed, writes:

\[(4.2a) \quad \min_0 0 \]
\[\text{s.t. } A_0 x = B p \quad (\lambda_0)\]
\[A_n x_n \leq b_n, \quad \forall n \in \mathcal{N} \quad (\lambda_n).\]

where $A_0 = J_{1,N} \otimes I_T$, $B = I_T$, (such that (4.2b) corresponds to the aggregation constraint $\sum_n x_n = p$) and $\lambda_0 \in \mathbb{R}^T$, $(\lambda_n)_{n \in \mathcal{N}} \in \mathbb{R}^{\sum_n k_n}$ are the Lagrangian multipliers associated to (4.2b) and (4.2c).

With the polyhedral constraints (4.1), the graph representation of the disaggregation problem, as illustrated in Figure 2.1 is no longer valid. Consequently, one can not directly apply Hoffman’s theorem (Theorem 2.3) to obtain a characterization of disaggregation feasibility by inequalities on $p$. However, using duality theory, Proposition 4.1 below also gives a characterization of disaggregation:

**Proposition 4.1.** A profile $p \in \mathcal{P}$ is disaggregable iff:

\[(4.3) \quad \forall (\lambda_0, \lambda_1, \ldots, \lambda_N) \in \Lambda, \quad \lambda_0^T B p + \sum_{n \in \mathcal{N}} \lambda_n^T b_n \geq 0 ,\]

where $\Lambda \overset{\text{def}}{=} \{\lambda_0 \in \mathbb{R}^{k_0}, \forall n \in \mathcal{N}, \lambda_n \in \mathbb{R}^{k_n} \mid A_0^T \lambda_0 A_0 + A^T (\lambda_n) n = 0\}$, with $A \overset{\text{def}}{=} \text{diag}(A_n)_{n \in \mathcal{N}}$.

**Proof.** From strong duality, we have:

\[(4.4) \quad \min_{x \in \mathbb{R}^N} \max_{\lambda_0 \in \mathbb{R}^{k_0}, \lambda_n \in \mathbb{R}^{k_n}} \lambda_0^T (A_0 x - B p) + \sum_n \lambda_n^T (A_n x_n - b_n)\]
\[= \max_{\lambda_0 \in \mathbb{R}^{k_0}, \lambda_n \in \mathbb{R}^{k_n}} -\lambda_0^T B p - \sum_n \lambda_n^T b_n \]
\[\text{s.t. } \lambda_0^T A_0 + (\lambda_n) A_n = 0 .\]
If the polytope $Y_p \cap \mathcal{X}$ given by the constraints of (4.2) is empty, then there is an infeasibility certificate $\lambda^T = (\lambda_0^T, \lambda_1^T, \ldots, \lambda_N^T) \in \mathbb{R}^T \times \prod_n \mathbb{R}_{+}^{k_n}$ such that:

\begin{equation}
(4.6) \quad \lambda_0^T A_0 + (\lambda_n^T)_n A = 0 \quad \text{and} \quad \lambda_0^T B p + (\lambda_n)_n b < 0 .
\end{equation}

On the other hand, if $Y_p \cap \mathcal{X}$ is nonempty, then a solution to the dual problem (4.5) is bounded, which implies that $\forall \lambda \equiv (\lambda_0, (\lambda_n)_n) \in \Lambda, \quad \lambda_0^T B p + \sum_n \lambda_n^T b_n \geq 0 .
\end{equation}

As opposed to Hoffman circulation’s theorem where disaggregation is characterized by a finite number of inequalities, Proposition 4.1 involves a priori an infinite number of inequalities.

However, we know that the polyhedral cone $\Lambda$ can be represented by a finite number of generators (edges), that is, there exists $\Lambda^* \equiv \{\lambda^{(1)}, \ldots, \lambda^{(d)}\}$ such that:

\begin{equation}
(4.7) \quad \Lambda = \left\{ \sum_{1 \leq i \leq d} \alpha_i \lambda^{(i)} \mid (\alpha_i)_i \in \mathbb{R}_+^d \right\}.
\end{equation}

Thus, we obtain the following corollary to Proposition 4.1:

**Corollary 4.2.** There exists a finite set $\Lambda^* \subset \Lambda$ such that, for any $p \in \mathcal{P}$, $p$ is disaggregable iff:

\begin{equation}
(4.8) \quad \forall (\lambda_0, (\lambda_n)_n) \in \Lambda^*, \quad \lambda_0^T B p + \sum_n \lambda_n^T b_n \geq 0 .
\end{equation}

**Remark 5.** In the transportation case (2.2), we can write each agent constraints in the form $A_n x_n \leq b_n$ (writing the equality $\sum x_{n,t} = E_n$ is written as two inequalities), and Hoffman conditions (2.5) can be written in the form (8). Moreover, Theorem 3.3 ensures that one possibility for $\Lambda^*$ of Corollary 4.2 is to consider the collection of $2^T$ multipliers corresponding to the subsets $\mathcal{T}_0 \subset T$ and $N_0$ minimizing (2.6). We skip the details here for brevity.

As in the first part of the paper, we want to use APM to decompose problem (1.1) and, in the case where disaggregation is not possible, use the result of APM to generate an inequality (4.3) violated by the current profile $p$.

In the case of impossible disaggregation, the APM converges to the orbit $(y^\infty, x^\infty)$, and $\mu = y^\infty - x^\infty$ defines a separating hyperplan $\bar{x} + \mu \perp$, where $\bar{x} = \frac{y^\infty + x^\infty}{2}$, that satisfies, with $a \equiv \bar{x} \mu$ (note that $\bar{x}$ can be replaced by any $y \in [y^\infty, x^\infty]$):

\begin{equation}
(4.9) \quad \forall x \in Y_p, \quad \mu^T x > a \quad \forall x \in \mathcal{X}; \quad a > \mu^T x ,
\end{equation}

which give lower bounds on the linear problems (the second one is decomposed because $A$ is a block-diagonal matrix, but it can also be written in one problem):

\begin{equation}
(4.10) \quad \begin{align*}
\min_{x \in \mathbb{R}^{N_T}} & \quad \mu^T x \\
\text{subject to} & \quad A_0 x = B p (\lambda_0) \quad \text{and} \quad \forall n \in \mathcal{N}, \quad \max_{x \in \mathbb{R}^{k_n}} & \quad \mu_n x_n \\
& \quad A_n x_n \leq b_n (\lambda_n) \quad \text{such that:} \\
& \quad \mu = -A_0^T \lambda_0 \quad \text{and} \quad a < -\lambda_0^T B p \quad \text{and} \quad a > b^T \lambda .
\end{align*}
\end{equation}

Strong duality on these problems implies that there exist $\lambda_0$ and $\lambda$ such that:

\begin{equation}
(4.11) \quad \mu = -A_0^T \lambda_0 \quad \text{and} \quad a < -\lambda_0^T B p \quad \text{and} \quad a > b^T \lambda .
\end{equation}

Thus, we obtain $(\lambda_0, \lambda) \equiv (\lambda^T, \lambda^T) \neq (0, 0)$ that, $\lambda_0^T B p + b^T \lambda < 0$, and we can use this to add a new valid additional inequality on $p$ of form (4.3), that will change the current profile $p$:

\begin{equation}
(4.12) \quad \lambda_0^T B p + \lambda^T b \geq 0 .
\end{equation}

In Algorithm 4.1 below, we summarize the proposed decomposition of problem Equation (1.1). This is a generalization of the decomposition principle used for Algorithm 3.4.
Algorithm 4.1 Non-intrusive optimal disaggregation with polyhedral constraints

Require: Start with $\Lambda^{(0)} = \{\}$, $k = 0$, Disag = false
1: while not Disag do
2: get solution $p^{(k)}$ of problem $\min_{p \in P} \{ f(p) \mid \lambda_0^T B p + \lambda^T b \geq 0, \forall \lambda \in \Lambda^{(k)} \}$
3: get $\mu^{(k)} = y^\infty - x^\infty \leftrightarrow APM(Y_p^{(k)}, X)$
4: if $\mu^{(k)} \neq 0$ then
5: obtain $\lambda_0^{(k)} \leftarrow \max_{\lambda_0 \in \mathbb{R}^n} \{ -\lambda_0^T B p^{(k)} \mid \mu^{(k)} = -A_0^T \lambda_0 \}$
6: obtain for each $n$, $\lambda_n^{(k)} \leftarrow \max_{\lambda_n \geq 0} \{ b_n^T \lambda_n \mid \mu_n^{(k)} = \lambda_n^T A_n \}$
7: add $\Lambda^{(k+1)} = \Lambda^{(k)} \cup \{(\lambda_0^{(k)}, \lambda^{(k)})\}$
8: else
9: Return Disag = true, $p^{(k)}$ as optimal solution
10: end
11: $k \leftarrow k + 1$
12: done

Remark 6. We use the fact that $\mu = y^\infty - x^\infty$ although, as before, we only have an approximation of this quantity. The approximation has to be precise enough to ensure that the solution obtained verifies $\lambda_0^T B p + b^T \lambda < 0$. In practice, one can proceed as in the transportation case and Algorithm 3.3 use a large $\varepsilon_{cvg}$ as stopping criteria in APM, then compute $(\lambda_0, \lambda) \in \Lambda$ and check if $\lambda^T B p + b^T \lambda < 0$. If this is not the case, restart with $\varepsilon_{cvg} = \varepsilon_{cvg}/2$.

Remark 7. When $\mathcal{Y}_p = \{ x \in \mathbb{R}^N \mid A_0 x = B_0 p \} = \{ x \mid \sum_n x_n = p \}$, we can obtain a non-intrusive version of APM on $\mathcal{Y}_p$ and $X$, similar to Algorithm 3.3. In this case, (4.11) ensures that we have $\mu_n = -\lambda_0_n$ for each $n \in N$, and $\lambda_0$ is fixed by $\mu$. The only difference with the transportation case for a non-intrusive APM in the general polyhedral case, is in the way of computing the valid constraint violated by $p$. Thus, Lines 16 to 19 of Algorithm 3.3 have to be replaced by Algorithm 4.2.

Algorithm 4.2 Modification of Lines 16-19 of Algorithm 3.3 for NI-APM with polyhedral constraints
16: for each agent $n \in N$ do
17: compute $M_n$ optimal value of (4.10).
18: done
19: Operator computes $M \leftarrow \text{SMCA}((M_n)_n)$
20: if $-\nu \cdot p + M < 0$ then
21: return Disag $\leftarrow$ False, $-\nu, M$

4.1. Link with Benders’ decomposition. In this generalized case, we obtain an algorithm related to Benders’ decomposition [5] (recall that in our specific case (4.2), the cost function does not involve the variable $x$ but only variable $p$).

The difference between the proposed Algorithm 4.1 and Benders’ decomposition is on the way of generating the new cut. Benders’ decomposition would directly solve the dual problem (4.5): $\max \lambda \{-\lambda_0^T B_0 p - (\lambda_n)_n b \mid \lambda_0 A_0 + (\lambda_n)_n A = 0 \}$ and obtain a cut if it is unbounded. However, this problem involves the constraints of all users (through $A$ and $b$), and it is not straightforward to obtain a method to solve this problem in a decentralized and efficient way.

5. Numerical examples.

5.1. An illustrative example with $T=4$. In this section we illustrate the iterations of the method proposed in this paper on an example with $T = 4$ and $N = 3$. Assuming that we have to satisfy the aggregate constraint $\sum_t p_t = \sum_n E_n$, we can use the projections on this affine space of solutions of master problems $(p^{(s)})_s$, to visualize them in dimension 3.

One can wonder if, in the transportation case, applying Algorithm 3.4 or Algorithm 4.1 will always
lead to the same cuts and solutions: the answer is no, as shown by the instance considered in this
section, for which Algorithm 3.4 converges in 3 iterations and Algorithm 4.1 needs 4 iterations.

We consider the problem (1.1) with agents constraints (2.2) with parameters \( \bar{p} \) and:

\[
\begin{align*}
(5.1) \quad \bar{p} &\equiv \left[ 0.8, 0.2, 0.7, 0.1 \right] \quad E_1 = 1.8 \\
&\equiv \left[ 0.5, 0.1, 0.3, 0.6 \right] \quad E_2 = 0.4, \quad \forall \mathbf{p} \in \mathbb{R}^4, f(\mathbf{p}) \equiv \sum_{1 \leq t \leq 4} 0.8 \times p_t + 0.1 \times p_t^2.
\end{align*}
\]

Considering the aggregate equality constraint \( \sum_{1 \leq t \leq 4} p_t = \sum_{1 \leq n \leq 3} E_n = 3.3 \), we use the canonical projection of 4 dimensional vectors into the 3 dimensional space \((p_1, p_2, p_3)\) to visualize the cuts and solutions. In this example, there exist \( 2^7 - 2 = 14 \) nontrivial Hoffman inequalities characterizing disaggregation from Theorem 2.3. The projection of the obtained polytope \( \mathcal{P}_\text{B} \), as defined in (2.1), is represented in Figure 5.1a. One can remark that this polytope has only 6 facets. Our Algorithm 3.4 applied on this instance with \( \varepsilon_{\text{dis}} = 10^{-3} \) and \( \varepsilon_{\text{cvg}} = 10^{-5} \) converges in 3 iterations, with successive solutions of the master problem (3.1) and cuts added:

\[
\begin{align*}
p^{(1)} &= [1., 0.4, 1., 0.9] \quad \rightarrow \quad p_1 + p_2 + p_4 \leq 1.9 \\
p^{(2)} &= [0.75, 0.4, 1.4, 0.75] \quad \rightarrow \quad p_2 + p_3 + p_4 \leq 2.4 \\
p^{(3)} &= [0.9, 0.4, 1.4, 0.6] .
\end{align*}
\]

Figure 5.1b represents in the projection space the three successive solutions and the two generated cuts (in red for each iteration).

On the other hand, applying Algorithm 4.1 with the same precision parameters \( \varepsilon_{\text{dis}}, \varepsilon_{\text{cvg}} \), there are 3 cuts generated and 4 resolutions of master problem needed for convergence, given by (we refer the reader to Remark 6 for the numerical precision obtained in the values):

\[
\begin{align*}
p^{(1)} &= [1., 0.4, 1., 0.9] \quad \rightarrow \quad -0.25p_1 - 0.25p_2 + 1.0p_3 - 0.5p_4 \geq 0.75 \\
p^{(2)} &= [0.8097, 0.4, 1.3984, 0.6919] \quad \rightarrow \quad 1.0p_1 - 0.509p_2 + 0.018p_3 - 0.509p_4 \geq 0.4161 \\
p^{(3)} &= [0.9062, 0.4, 1.3823, 0.6115] \quad \rightarrow \quad -0.333p_1 - 0.333p_2 + 1.0p_3 - 0.333p_4 \geq 0.7666 \\
p^{(4)} &= [0.9, 0.4, 1.4, 0.6] .
\end{align*}
\]

The 4 successive solutions and the 3 added cuts are represented in the three dimensional space on Figure 5.1c: we observe that the last cut needed to obtain the convergence of Algorithm 4.1, corresponds to the first one added with Algorithm 3.4.

Due to the strict convexity of the cost function \( \mathbf{p} \mapsto f(\mathbf{p}) \), the final solution obtained is the same, unique aggregated optimal solution of (1.1). The 4 successive solutions and the 3 added cuts are represented in the three dimensional space on Figure 5.1c: we observe that the last cut needed to obtain the convergence of Algorithm 4.1, corresponds to the first one added with Algorithm 3.4.

### 5.2. A nonconvex example: management of a microgrid

In this section, we illustrate the proposed method on a larger scale practical example from energy. We consider an electricity microgrid [21] composed of N electricity consumers with flexible appliances (such as electric vehicles or water heaters), a photovoltaic (PV) power plant and a conventional generator. The operator of the microgrid aims at satisfying the demand constraints of consumers over a set of time periods \( T = \{1, \ldots, T\} \), while minimizing the energy cost for the community. We have the following characteristics:

- the PV plant generates a nondispatchable power profile \((p^{\text{PV}}_t)_{t \in T}\) at marginal cost zero;
- the conventional generator has a starting cost \( C^\text{cr} \), minimal and maximal power production \( p^\text{m}, p^\text{M} \), and piecewise-linear and continuous generation cost function \( p^\theta \mapsto f(p^\theta) \):

\[
f(p^\theta) = \alpha_k + c_k p^\theta, \quad \text{if } p^\theta \in I_k \equiv [\theta_{k-1}, \theta_k], \quad k = 1 \ldots K,
\]

where \( \theta_0 \equiv 0 \) and \( \theta_K \equiv \bar{p}^\text{c} \);
Figure 5.1: Illustration of the iterations of the proposed decomposition method. The cut \( p_3 \geq 1.4 \), which is added at first for Algorithm 3.4, is only added at the third iteration of Algorithm 4.1.

- each agent \( n \in \mathcal{N} \) has some flexible appliances which require a global energy demand \( E_n \) on \( T \), and has consumption constraints on the total household consumption, on each time period \( t \in T \), that are formulated with \( \mathbf{x}_n \), \( \mathbf{x}_n \). These parameters are confidential because they could for instance contain some information on agent \( n \) habits.

The master problem (3.1) can be written as the following MILP (5.2):

\[
\begin{align*}
\text{(5.2a)} & \quad \min_{\mathbf{p}, \mathbf{p}, \mathbf{p}, \mathbf{b}^{\text{on}}, \mathbf{b}^{\text{st}}, \mathbf{b}^t} \sum_{t \in T} \left( \alpha_1 b_t^{\text{on}} + \sum_k c_k p_k^g t + C^{\text{st}} b_t^{\text{st}} \right) \\
\text{(5.2b)} & \quad p_t^g = \sum_{k=1}^K p_k^g t, \quad \forall t \in T \\
\text{(5.2c)} & \quad b_{k,t}(\theta_k - \theta_{k-1}) \leq p_{k,t}^g \leq b_{k-1,t}(\theta_k - \theta_{k-1}), \quad \forall 1 \leq k \leq K, \quad \forall t \in T \\
\text{(5.2d)} & \quad b_t^{\text{st}} \geq b_t^{\text{on}} - b_t^{\text{st}} - 1, \quad \forall t \in \{2, \ldots, T\} \\
\text{(5.2e)} & \quad p^g b_t^{\text{on}} \leq p_t^g \leq p^g b_t^{\text{on}}, \quad \forall t \in T \\
\text{(5.2f)} & \quad b_t^{\text{on}}, b_t^{\text{st}}, b_{1,t}, \ldots, b_{K-1,t} \in \{0,1\}, \quad \forall t \in T \\
\text{(5.2g)} & \quad \mathbf{p} \leq \mathbf{p}^{\text{pv}} \leq \mathbf{p} \\
\text{(5.2h)} & \quad \mathbf{p}^{\top} \mathbf{1}_T = E^{\top} \mathbf{1}_N \\
\text{(5.2i)} & \quad \mathbf{x}^{\top} \mathbf{1}_N \leq \mathbf{p} \leq \mathbf{x}^{\top} \mathbf{1}_N.
\end{align*}
\]

In this formulation (5.2b-5.2c), where \( b_{0,t} \overset{\text{def}}{=} 1 \) and \( b_{K,t} \overset{\text{def}}{=} 0 \), are a mixed integer formulation of the
generation cost function $f$. One can show that the Boolean variable $b_{k,t}$ is equal to one iff $p_t^k \geq \theta_k$ for each $k \in \{1, \ldots, K - 1\}$. Note that only $\alpha_1$ appears in (5.2a) because of the continuity of $f$.

Constraints (5.2d-5.2c) ensure the on/off and starting constraints of the power plant, (5.2g) ensures that the power allocated to consumption is not above the total production, and (5.2h-5.2i) are the aggregated feasibility conditions already referred to in (2.3). The nonconvexity of (5.2) comes from the existence of starting costs and constraints of minimal power, which makes necessary to use Boolean state variables $b^{st}, b^{on}$.

We simulate the problem described above for different values of $N \in \{2^4, 2^5, 2^6, 2^7, 2^8\}$ and one hundred instances with random parameters for each value of $N$. A scaling factor $\kappa_N = N/20$ is applied on parameters to ensure that production capacity is large enough to meet consumers demand. The parameters are chosen as follows:

- $T = 24$ (hours of a day);
- production costs: $K = 3$, $\theta = [0, 70, 100, 300] \kappa_N, c = [0.2, 0.4, 0.5]$, $p^g = 50 \kappa_N, p^s = 300 \kappa_N$, $\alpha_1 = 4$ and $C^{st} = 15$;
- photovoltaic: $p_t^{pv} = 50(1 - \cos(\frac{t - 6 \pi}{6}))/\kappa_N$ for $t \in \{6, \ldots, 20\}$, $p_t^{pv} = 0$ otherwise;
- consumption parameters are drawn randomly with: $x_{n,t} \sim \mathcal{U}([0, 10]), x_{n,t} \sim \mathcal{U}([0, 5]) + x_{n,t}$ and $E_n \sim \mathcal{U}([1, \frac{1}{2}], 1, \frac{1}{2})$, so that individual feasibility ($X_n \neq \emptyset$) is ensured.

<table>
<thead>
<tr>
<th>$N$</th>
<th>$2^4$</th>
<th>$2^5$</th>
<th>$2^6$</th>
<th>$2^7$</th>
<th>$2^8$</th>
</tr>
</thead>
<tbody>
<tr>
<td># master</td>
<td>193.6</td>
<td>194.1</td>
<td>225.5</td>
<td>210.9</td>
<td>194.0</td>
</tr>
<tr>
<td># projs.</td>
<td>9507</td>
<td>15367</td>
<td>24319</td>
<td>26538</td>
<td>26646</td>
</tr>
</tbody>
</table>

Table 5.1: number of subproblems solved (average on 100 instances)

We implement Algorithm 3.4 using Python 3.5. The MILP (5.2) is solved using Cplex Studio 12.6 and Pyomo interface. Simulations are run on a single core of a cluster at 3GHz. For the convergence criteria (see Lines 11 and 12 of Algorithm 3.3), we use $\varepsilon_{dis} = 0.01$ with the operator norm defined by $\|x\| = \max_{n \in N} \sum_t |x_{n,t}|$ (to avoid the $\sqrt{N}$ factor in the convergence criteria appearing with $\|\cdot\|_2$), and starts with $\varepsilon_{cvx} = 0.1$. The largest instances took around 10 minutes to be solved in this configuration and without parallel implementation. As the CPU time needed depends on the cluster load, it is not a reliable indicator of the influence on $N$ on the complexity of the problems. Moreover, one advantage of the proposed method is that the projections in APM can be computed locally by each agent in parallel, which could not be implemented here for practical reasons.

Table 5.1 gives the number of master problems solved and the total number of projections computed, on average over the hundred instance for each value of $N$.

One observes that the number of master problems (5.2) solved (number of “cuts” added), remains almost constant when $N$ increases. In all instances, this number is way below the upper bound of $2^{24} > 1.6 \times 10^7$ possible constraints (see Proposition 3.10), which suggests that only a limited number of constraints are added in practice. The average total number of projections computed for each instance (total number of iterations of the while loop of Algorithm 3.3, Line 1 over all calls of APM in the instance) increases in a sublinear way which is even better that one could expect from the upper bound given in Theorem 3.12.

6. Conclusion. We provided a non-intrusive algorithm that enables to compute an optimal resource allocation, solution of a –possibly nonconvex–optimization problem, and affect to each agent an individual profile satisfying a global demand and lower and upper bounds constraints. Our method uses local projections and works in a distributed fashion. Hence, the resolution of the problem is still efficient even in the case of a very large number of agents. The method is also privacy-preserving, as agents do not need to reveal any information on their constraints or their individual profile to a third party.

Several extensions and generalizations can be considered for this work. Section 4 generalizes the procedure to arbitrary polyhedral constraints for agents. However, the number of constraints (cuts)
added to the master problem is not proved to be finite as done in the transportation case. Proving that only a finite number of constraints can be added (maybe up to a refinement procedure of the current constraint obtained) will enable to have a termination result for the algorithm in the general polyhedral case. In the transportation case, we showed the geometric convergence of APM with a rate linear in the number of agents. Moreover, the number of cuts added in the procedure is finite but the upper bound that we have remains exponential. In practice however, the number of constraints to consider remains small, as seen in Section 5. A thinner upper bound on the number of cuts added in the algorithm in this case would constitute an interesting result.

**Appendix A. Proof of Proposition 3.4.**

**Proof of Item (i).** Let us write the stationarity conditions associated to problem (3.3):

(A.1) \[ \forall n \in \mathcal{N}, \forall t \in \mathcal{T}, \quad 0 = (x_{n,t} - y_{n,t}) - \lambda_n - \mu_{n,t} + \pi_{n,t} \quad \text{and} \quad y_{n,t} = x_{n,t} + \nu_t. \]

By summing the preceding equalities on \(\mathcal{T}\) and \(\mathcal{N}\), we obtain the three equalities:

(A.2) \[ \sum_t \nu_t = \sum_t y_{n,t} - E_n, \quad \forall n \in \mathcal{N} \quad \quad p_t = \sum_n x_{n,t} + N \nu_t, \quad \forall t \in \mathcal{T} \]

(A.3) \[ |\mathcal{T}_n^\circ| \lambda_n = E_n - \sum_{t \in \mathcal{T}_n^\circ} \bar{x}_{n,t} - \sum_{t \in \mathcal{T}_n^\circ} y_{n,t} - \sum_{t \in \mathcal{T}_n^\circ} \bar{\pi}_{n,t}, \forall n \in \mathcal{N}, \]

where we define for each \(n \in \mathcal{N}\):

\(\mathcal{T}_n^\circ \overset{\text{def}}{=} \{ t \mid \bar{x}_{n,t} < x_{n,t} < \bar{\pi}_{n,t} \}\), \(\mathcal{T}_n = \{ t \mid x_{n,t} = \bar{x}_{n,t} \}\) and \(\mathcal{T}_n = \{ t \mid x_{n,t} = \bar{\pi}_{n,t} \}\).

From (A.2) and the aggregate equality \(\sum_n E_n = \sum_t p_t\) we obtain: \(\sum_t \nu_t = 0\) and:

(A.4) \[ \forall n \in \mathcal{N}, \sum_{t \in \mathcal{T}} y_{n,t} = E_n. \]

Suppose that Item (i) is false: there exists \(n \notin \mathcal{N}_0\) and \(t \in \mathcal{T}_0\) such that \(x_{n,t} < \bar{x}_{n,t}\). We have:

\[ x_{n,t} + \nu_t > y_{n,t} + \lambda_n + \nu_t + \lambda_n \implies \lambda_n < 0. \]

Immediately, we have \(\mathcal{T}_n \subset \mathcal{T}_0\): indeed, for \(t \in \mathcal{T}_n\), we have:

\[ y_{n,t} + \lambda_n \geq \bar{\pi}_{n,t} \implies y_{n,t} + \lambda_n > \bar{\pi}_{n,t} \implies t \in \mathcal{T}_0. \]

From the condition (A.4) and from \(\nu_t > 0\) for each \(t \in \mathcal{T}_n\) because \(\mathcal{T}_n \subset \mathcal{T}_0\), we get:

\[ 0 = \sum_{t \in \mathcal{T}} (y_{n,t} - x_{n,t}) = \sum_{t \in \mathcal{T}_n} (y_{n,t} - \bar{x}_{n,t}) + \sum_{t \in \mathcal{T}_n} (-\lambda_n) + \sum_{t \in \mathcal{T}_n} \nu_t \iff \sum_{t \in \mathcal{T}_n} (y_{n,t} - \bar{x}_{n,t}) = \sum_{t \in \mathcal{T}_n} \lambda_n - \sum_{t \in \mathcal{T}_n} \nu_t, \]

which is strictly negative: this implies that there exists \(t' \in \mathcal{T}_n\) such that \(y_{n,t'} < \bar{x}_{n,t'}\). Necessarily, \(t' \notin \mathcal{T}_0\) because \(\nu_{t'} = y_{n,t'} - x_{n,t'} < \bar{x}_{n,t'} - \bar{x}_{n,t'} = 0\). Then, as we have \(\sum_{m \in \mathcal{N}} y_{m,t'} = \bar{\mu}_{t'} \geq \sum_{m \in \mathcal{N}} \bar{x}_{m,t'}\), there exists \(m \in \mathcal{N}\) such that \(y_{m,t'} > \bar{x}_{m,t'}\). If \(\lambda_m \leq 0\), and as \(\bar{x}_{m,t'} = y_{m,t'} - \nu_{t'} > \bar{x}_{m,t'}\), we get:

\[ x_{m,t'} = \min(\bar{x}_{m,t'}, y_{m,t'} + \lambda_m) \leq y_{m,t'} + \lambda_m \leq y_{m,t'} + \nu_{t'} < x_{m,t'} \]

which is impossible, thus \(\lambda_m > 0\). Now, we observe that \(\mathcal{T}_n \subset \mathcal{T}_0\). Indeed, otherwise, if \(t'' \in \mathcal{T}_n \cap \mathcal{T}_0\), we have \(\nu_{t''} = -\lambda_n > 0\) and \(x_{m,t''} = y_{m,t''} - \nu_{t''} < y_{m,t''} - \bar{x}_{m,t''}\), thus we get:

\[ x_{m,t''} = \max(\bar{x}_{m,t'', y_{m,t''} + \lambda_m}) \geq y_{m,t''} + \lambda_n > y_{m,t''} = x_{m,t''} + \nu_{t''} + \lambda_m > x_{m,t''} \]

which is impossible, thus \(\mathcal{T}_n \subset \mathcal{T}_0\).
Finally, since $\mathcal{T}_n \neq \emptyset$, consider $t_0 \in \arg\min_{t \notin \mathcal{T}_n} \{\varphi_{n,t} - y_{n,t}\}$. By (A.3), we obtain:

(A.5) \quad y_{n,t_0} + \lambda_n < \varphi_{n,t_0} \iff E_n - \sum_{t \in \mathcal{T}_n} \varphi_{n,t} - \sum_{t \in \mathcal{T}_n} y_{n,t} - \sum_{t \in \mathcal{T}_n} \varphi_{n,t} < |\mathcal{T}_n^c| (\varphi_{n,t_0} - y_{n,t_0})

and thus:

$$E_n - \sum_{t \in \mathcal{T}_0} \varphi_{n,t} - \sum_{t \in \mathcal{T}_0} \varphi_{n,t} = E_n - \sum_{t \in \mathcal{T}_n} \varphi_{n,t} + \sum_{t \in \mathcal{T}_n \cap \mathcal{T}_n^c} \varphi_{n,t} - \sum_{t \in \mathcal{T}_n \cap \mathcal{T}_n^c} \varphi_{n,t} - \sum_{t \in \mathcal{T}_n \cap \mathcal{T}_n^c} \varphi_{n,t} - \sum_{t \in \mathcal{T}_n \cap \mathcal{T}_n^c} \varphi_{n,t} \quad \text{(as } \mathcal{T}_n \cup \mathcal{T}_n^c \subset \mathcal{T}_0)$$

(A.6) \quad < \sum_{t \in \mathcal{T}_n^c} (\varphi_{n,t_0} - y_{n,t_0}) - (\varphi_{n,t} - y_{n,t}) + \sum_{t \in \mathcal{T}_n \cap \mathcal{T}_n^c} (\varphi_{n,t} - \varphi_{n,t}) \quad \text{(from (A.5))}

(A.7) \quad \leq 0 \quad \text{(from the definition of } t_0 \text{ and } \varphi_{n,t} \leq \varphi_{n,t})

which contradicts $n \notin N_0$ and terminates the proof for Item (i).

Proof of Item (ii). To prove (ii), we see that if $t$ is such that $\nu_t > 0$, then all the facts said before are true for $n \notin N_0$ if we consider $\mathcal{T}_0^t \defeq \{t|\nu_t > 0\}$ instead of $\mathcal{T}_0$. In that case we will have $\lambda_n < 0$. However we cannot have $t^* \in \mathcal{T}_0 \cap \mathcal{T}_n^c$ because this would mean $\nu_{t^*} = -\lambda_n > 0$ but we have $\nu_{t^*} \leq 0$ because $t^* \notin \mathcal{T}_0$. Thus $\mathcal{T}_0^c$ is necessarily empty, and if there is $t \in \mathcal{T}_n \cap T_0^c$, the same sequence of inequalities as (A.6-A.7) show a contradiction. Consequently, for each $t \in \mathcal{T}_0$, $x_{n,t} = \varphi_{n,t}$ and $y_{n,t} = x_{n,t} + \nu_t > \varphi_{n,t}$, thus $t \in \mathcal{T}_0$ and $T_0^c \subset \mathcal{T}_0$. The other inclusion is immediate.

Proof of Item (iii). Suppose on the contrary that there exists $n \in N_0$ such that $\lambda_n \geq 0$. For $t \in \mathcal{T}_n$, we have $\nu_t = -\lambda_n \leq 0$, thus $\mathcal{T}_n \subset \mathcal{T}_0^c$. Then, if $t \in \mathcal{T}_0$ and if $x_{n,t} < \varphi_{n,t}$, we would have:

$$x_{n,t} = \max(\varphi_{n,t}, y_{n,t} + \lambda_n) \geq x_{n,t} + 0 + \nu_t > x_{n,t},$$

which is impossible, thus $x_{n,t} = \varphi_{n,t}$, and $\mathcal{T}_0 \subset \mathcal{T}_n$. As we show independently in Item (v) that $\mathcal{T}_0 \neq \emptyset$, we know $\mathcal{T}_n \neq \emptyset$. Let us consider $t_0 \in \arg\min_{t \notin \mathcal{T}_n} \{y_{n,t} - \varphi_{n,t}\}$. By (A.3), we obtain:

(A.8) \quad y_{n,t_0} + \lambda_n > \varphi_{n,t_0} \iff E_n - \sum_{t \in \mathcal{T}_n} \varphi_{n,t} - \sum_{t \in \mathcal{T}_n} y_{n,t} - \sum_{t \in \mathcal{T}_n} \varphi_{n,t} > |\mathcal{T}_n^c| (\varphi_{n,t_0} - y_{n,t_0})

and thus:

$$E_n - \sum_{t \in \mathcal{T}_0} \varphi_{n,t} - \sum_{t \in \mathcal{T}_0} \varphi_{n,t} = E_n - \sum_{t \in \mathcal{T}_n} \varphi_{n,t} - \sum_{t \in \mathcal{T}_n \cap \mathcal{T}_n^c} \varphi_{n,t} - \sum_{t \in \mathcal{T}_n \cap \mathcal{T}_n^c} \varphi_{n,t} - \sum_{t \in \mathcal{T}_n \cap \mathcal{T}_n^c} \varphi_{n,t} - \sum_{t \in \mathcal{T}_n \cap \mathcal{T}_n^c} \varphi_{n,t} \quad \text{(as } \mathcal{T}_0 \subset \mathcal{T}_n)$$

(A.9) \quad > \sum_{t \in \mathcal{T}_n^c} (y_{n,t} - \varphi_{n,t}) - (y_{n,t} - \varphi_{n,t_0}) + \sum_{t \in \mathcal{T}_n} \varphi_{n,t} - \varphi_{n,t} \quad \text{(from (A.8))}

(A.10) \quad \geq 0 \quad \text{(from the definition of } t_0 \text{ and } \varphi_{n,t} \leq \varphi_{n,t})

which contradicts $n \notin N_0$ and terminates the proof for Item (iii).

Proof of Item (iv). From (ii), we know that $\mathcal{T}_n^c = \{t|\nu_t \leq 0\}$, thus, if $t \notin \mathcal{T}_0$ and $n \in \mathcal{N}_0$, if $x_{n,t} > \varphi_{n,t}$ then we would have $x_{n,t} \leq y_{n,t} - \lambda_n = x_{n,t} + \nu_t + \lambda_n < x_{n,t}$, which is a contradiction.

Proof of Item (v). From $\sum_t \nu_t = 0$, we see that if $\mathcal{T}_0 = \emptyset$, then this means that $\nu_t = 0$ for all $t \in \mathcal{T}$, and thus $y = \varphi$ which is a contradiction. Thus there exists $t_0$ such that $\nu_{t_0} > 0$ and for the same reason, there exists $t_0'$ such that $\nu_{t_0'} < 0$.

If $\mathcal{N}_0 = \emptyset$, then using (i), we would have for all $n$, $y_{n,t_0} > \varphi_{n,t_0}$ and thus $p_{t_0} > \sum_{n \in \mathcal{N}_0} \varphi_{n,t_0}$, which contradicts the aggregate upper bound constraint $\forall t$, $p_t \leq \sum_{n \in \mathcal{N}_0} \varphi_{n,t}$.

If $\mathcal{N}_0^c = \emptyset$, then using (iv), we would have for all $n$, $y_{n,t_0'} < \varphi_{n,t_0'}$ and thus $p_{t_0'} < \sum_{n \in \mathcal{N}_0} \varphi_{n,t_0'}$, which contradicts the aggregate lower bound constraint $\forall t$, $p_t \geq \sum_{n \in \mathcal{N}_0} \varphi_{n,t}$.

Appendix B. Proof of Lemma 3.9.
Proof of Item (i). From $x^{(K)} = P_X(y^{(K-1)})$ and $y^{(K)} = P_Y(x^{(K)})$, we obtain, similarly to (A.1):

$$(B.1) \quad \forall n \in N, \forall t \in T, \quad 0 = (x_{n,t}^{(K)} - y_{n,t}^{(K-1)}) - \lambda_n^{(K)} - \frac{\mu_n^{(K)}}{cvg_n,t} + \nu_{n,t}^{(K)} \quad \text{and} \quad y_{n,t}^{(K)} = x_{n,t}^{(K)} + \nu_{t}^{(K)}.$$ 

where the Lagrangian multipliers $\lambda_n^{(K)}, \frac{\mu_n^{(K)}}{cvg_n,t}, \nu_{n,t}^{(K)}$ (resp. $\nu_{t}^{(K)}$) are associated to the quadratic problem characterizing the projections $P_X(y^{(K-1)})$ (resp. $y^{(K)} = P_Y(x^{(K)})$). We obtain equalities similar to (A.2, A.3). We proceed as for Proposition 3.4(i) and suppose that there exists $n \notin N_0$ and $t \in T_0$ such that $x_{n,t} < \tau_{n,t}$. Then, as $\|y^{(K)} - y^\infty\|^2 \leq \frac{\varepsilon_{cvg}}{1-\rho}$ and $\sum_{t \in T} y_{n,t} = \sum_{t \in T} y_{n,t}^\infty$, we have for each $n \in N, t \in T, |y_{n,t}^{(K)} - y_{n,t}| \leq \frac{\varepsilon_{cvg}}{2(1-\rho)}$, and thus we get:

$$(B.2) \quad \tau_{n,t} \geq x_{n,t}^{(K)} \geq y_{n,t}^{(K-1)} + \lambda_n \geq y_{n,t}^\infty - \frac{\varepsilon_{cvg}}{2(1-\rho)} + \lambda_n^{(K)} = x_{n,t}^\infty + \nu_{t}^{\infty} - \frac{\varepsilon_{cvg}}{2(1-\rho)} + \lambda_n^{(K)}$$

as $\nu_{t}^{\infty} \geq u > 2\varepsilon_{cvg}$. Let us now consider $t' \in \tau_n^{(K)} \cup \tau_n^{(K)}$, then:

$$(B.3) \quad \nu_{t'}^{(K)} = y_{n,t'}^{(K)} - x_{n,t'}^{(K)} \geq y_{n,t'}^{(K)} - y_{n,t'}^{(K-1)} - \lambda_n^{(K)} > -\frac{\varepsilon_{cvg}}{2} + \frac{3}{2}B\varepsilon_{cvg} > B\varepsilon_{cvg} + \frac{\varepsilon_{cvg}}{2}(B - 1) \geq B\varepsilon_{cvg},$$

which shows that $t' \in \tau_n^{(K)} = \tau_n^{(K)}$ and thus $\tau_n^{(K)} \cup \tau_n^{(K)} \subset T_0$. Then, the same sequence of inequalities as (A.5, A.6, A.7) applied to $y^{(K-1)}$ applied to $y^{(K-1)}$ gives a contradiction to $n \notin N_0$.

Proof of Item (ii). The proof of Item (ii) is symmetric to the one of Item (i): if we suppose that there exists $n \in N_0$ and $t \notin T_0$ such that $x_{n,t}^\infty > \tau_{n,t}$, we obtain, symmetrically to (B.2), that $\lambda_n^{(K)} > -\frac{\varepsilon_{cvg}}{2(1-\rho)}$. Then, considering $t' \in \tau_n^{(K)} \cup \tau_n^{(K)}$, we show, symmetrically to (B.3), that $\nu_{t'}^{(K)} < B\varepsilon_{cvg}$ i.e. $t' \notin T_0$ and thus $\tau_n^{(K)} \cup \tau_n^{(K)} \subset T_0$. We conclude by obtaining a contradiction to $n \in N_0$ by the same sequence of inequalities as (A.8, A.9, A.10).

References.


