Conception of energy management systems for residential photovoltaic self-consumption installations

Loris Amabile

PSL UNIVERSITÉ PARIS

THÈSE DE DOCTORAT

DE L'UNIVERSITÉ PSL

Préparée à MINES ParisTech

Conception de gestionnaires d'installations photovoltaïques résidentielles en autoconsommation

Conception of energy management systems for residential photovoltaic self-consumption installations

Soutenue par

Loris AMABILE

Le 10 décembre 2021

École doctorale n°621

Ingénierie des Systèmes, Matériaux, Mécanique, Energétique (ISMME)

Spécialité

Mathématique et Automatique

Composition du jury :

Marc PETIT Professeur, CentraleSupélec	*Président*
Mazen ALAMIR Directeur de Recherche CNRS, GIPSA-lab	*Rapporteur*
Scott MOURA Associate Professor, UC Berkeley	*Examinateur*
Florian HUBLER Directeur Technique, EDF ENR	*Examinateur*
Gilles PLESSIS Ingénieur de recherche, EDF R&D	*Examinateur*
Delphine BRESCH-PIETRI Maître-Assistante, MINES ParisTech	*Examinatrice*
Nicolas PETIT Professeur, MINES ParisTech	*Directeur de thèse*

MINES ParisTech | PSL

Remerciements

Je souhaite tout d'abord remercier tous les membres du jury de thèse pour leur participation à la soutenance et leurs remarques fructueuses : Scott Moura, Florian Hubler et Gilles Plessis pour leur rôle d'examinateurs, et en particulier Marc Petit et Mazen Alamir pour avoir accepté de rapporter ce travail de thèse.

Je souhaite exprimer ma profonde gratitude à mon directeur de thèse, Nicolas Petit, pour m'avoir permis de me lancer dans cette aventure, et pour avoir régulièrement éclairé mon chemin au cours de celle-ci. Encore plus que les propositions d'idées originales, les explications de notions complexes ou les clarifications de directions de travail qui étaient autant de repères rassurants, je retiendrai les cafés du matin à discuter de tout le reste, et qui furent finalement trop peu nombreux.

Cette thèse n'aurait également jamais été ce qu'elle est sans le travail inestimable de Delphine Bresch-Pietri, que je ne remercierai jamais assez. Son implication quotidienne dans l'intégralité des développements mathématiques comme dans les innombrables étapes de rédaction ont fait la qualité de cette thèse. Par sa patience, sa disponibilité, sa réactivité face à mes nombreuses sollicitations, elle a été une encadrante idéale et a toute ma reconnaissance.

Je remercie également mes encadrants d'EDF Gilbert El Hajje et Sébastien Labbé, qui m'ont judicieusement guidé avec confiance dans la majeure partie de ces trois années, en clarifiant les différents buts à atteindre tout en me poussant à voir plus loin.

Merci également aux autres collègues du groupe R38 d'EDF pour leur accueil bienveillant aux Renardières. Je salue particulièrement mes compagnons de thèse Nikola, Sylvain et Baptiste que je remercie pour le partage des joies et difficultés communes.

Malgré ma présence écourtée au laboratoire, je remercie tout le CAS de m'avoir permis de faire partie d'une communauté pendant ces trois années. Pour les multiples discussions sur leurs travaux ou sur la vie de l'école, je remercie chaleureusement Aradhana, Aurélien, Dilshad, Florent, Hubert, Louise, Matthieu, Maxime, Mona, Naveen, Nils, Pauline, Pierre-Cyril et Sijia.

Les amis qui m'entourent sont inestimables. Pour la joie et le soutien qu'ils m'apportent depuis la première année aux Mines, merci à Antonin, Alexia, Alexis, Martin, Matthieu, Rémy, Solène, Thibault et William. Ils sont le meilleur de ma vie étudiante parisienne.

Un remerciement spécial aux deux colocataires ultimes Michel et Maxime, ceux qui m'ont entouré au quotidien, étaient les premiers à m'apporter du positif quand j'hésitais, et les premiers à célébrer avec moi quand je réussissais. Vivre ensemble était vraiment génial. Cette thèse était peut-être une dernière tentative d'essayer d'égaler leur aisance scientifique, sans succès. A Maxime, qui aura soutenu à bout de bras l'ensemble de mon parcours en mathématiques, du projet d'automatique fait pour nous deux au cours accéléré d'optimisation pour décrocher l'offre de thèse : merci pour toute cette aide.

A la Dream MiaMines Team des résultats sportifs et des week-ends mémorables, merci.

A tous mes amis de Marseille et des Mines que je ne vois pas assez souvent mais qui comptent tant, merci.

Merci à ma famille étendue, trop nombreuse pour en faire une liste détaillée, mais tellement chère à mes yeux. Quelle chance de vous avoir.

Merci aussi à Nicolas qui m'avait aidé à faire mon tout premier pas dans le monde de la recherche.

Merci à Arno et Aurélien, frères qualité triple A, pour les rires, le soutien, les discussions interminables, l'ouverture à d'autres mondes, les références communes.

A Céline, qui est arrivée sans prévenir pour rendre ma vie meilleure, qui ne cesse de m'émerveiller, qui me fait grandir et me pousse à m'améliorer, qui m'a accompagné dans cette dernière année de rédaction pas toujours facile avec énergie et bienveillance, merci infiniment. J'ai hâte de réaliser à deux tous ces nouveaux projets.

Mes derniers remerciements vont à ma soeur et à mes parents.

A ma soeur Celia, modèle depuis tant d'années qui a pavé la voie à chaque étape de son parcours, m'a montré ce qui était atteignable et ce que ça exigeait. Son expérience, son écoute attentive, sa capacité à motiver et ses bons conseils me permettent de suivre le même chemin escarpé avec le confort d'une piste balisée.

A mes parents, qui m'ont donné depuis toujours amour, confort et confiance nécessaires pour atteindre mes objectifs. Qui m'ont ouvert toutes les portes et laissé le choix. Qui sont un modèle de réussite familiale et professionnelle. Et qui savent encore aujourd'hui me faire voir ce qui est important.

Merci.

Contents

Abbreviations and nomenclature

Abbreviations

DERs	Distributed Energy Resources
DR	Demand Response
DSO	Distribution System Operator
EMS	Energy Management System
EV	Electric Vehicle
EWH	Electric Water Heater
GHG	Greenhouse Gases
MAS	Multi-agent system
PV	Photovoltaic
V2X	Vehicle-to-X

Nomenclature

Table 1: EWH - Nomenclature

Notation	Description	Unit
E	Thermal energy of the water tank	Wh
E_{clear}	EWH energy level giving clearance to resume heating	Wh
E_{sat}	EWH saturation energy level	Wh
E_{PV}	Energy produced by the PV arrays	Wh
k	Thermal loss coefficient of the water tank	h^{-1}
P_{ewh}	EWH power consumption	W
$\overline{P}_{\text{ewh}}$	EWH power rating	W
P_{PV}	Household h produced PV power	W
$\widehat{P}_{\text{PV}}$	Household h surplus of produced PV power	W
Q	Power drained from the EWH by hot-water consumptions	W
T_{ewh}	Mean water temperature in the water tank	°C
t_{ewh}	EWH starting time of the heating command	h
t_{clear}	EWH clearance time	h
t_{sat}	EWH saturation time	h
t_{lim}	Latest acceptable starting time of the heating range	h
τ_{ewh}	EWH heating authorization ending time	h
SC	Self-consumption over one period	Wh

Table 2: Collective optimization - Nomenclature

Notation	Description	Unit
B	Electricity bill	€
c_{d}	Cost for deficit energy bought by one level to the outer level	€/Wh
c_{s}	Cost for surplus energy sold from one level to the outer level	€/Wh
C	Battery capacity	Wh
δ_{C}	Batteries binary charging state variable	
δ_{D}	Batteries binary discharging state variable	
$\delta_{h,\mathrm{s}}$	Household h binary surplus state variable	
$\delta_{h,\mathrm{d}}$	Household h binary deficit state variable	
$\delta_{\mathrm{s,agg}}$	Aggregated agents binary surplus state variable	
$\delta_{\mathrm{d,agg}}$	Aggregated agents binary deficit state variable	
H	Number of timesteps in the optimization horizon	
$P_{h,\mathrm{d}}$	Household h net deficit power	W
$P_{h,\mathrm{s}}$	Household h net surplus power	W
P_{gb}	Charging power from the grid to the battery	W
P_{ab}	Net charging power from the aggregated agents to the battery	W
P_{bg}	Discharging power from the battery to the grid	W
P_{ba}	Net discharging power from the aggregated agents to the agents	W
$P_{h,\mathrm{PV}}$	Household h PV production	W
$P_{h,\mathrm{unc}}$	Household h uncontrolled power consumption	W
P_{gm}	Power imports from outer grid to microgrid	W
P_{mg}	Power exports from microgrid to outer grid	W
$P_{\mathrm{s,agg}}$	Aggregated agents net surplus power	W
$P_{\mathrm{d,agg}}$	Aggregated agents net deficit power	W
t	Timestep index	

Table 3: Privacy preserving communication - Nomenclature

Notation	Description	Unit
β	Sum of all data items	
ξ	Number of assigned agents	
d_h	Data item of household h	
γ_i	Coefficient of order i of Newton polynomial	
p	Prime number for Finite Field $\mathbb{Z}/p\mathbb{Z}$	
$\Pi(x)$	Newton polynomial	
S_i	Power sum of order i	
u_i	Shared values in Algorithm 4.4	
$y_{h,\ldots}$	Random values drawn by household h	
z_g	Sum of random values received by household g	

Introduction (Français)

1 Contexte

La crise climatique menace les conditions de vie sur Terre : l'augmentation prévue du nombre et de l'intensité des vagues de chaleur, des fortes pluies, des sécheresses et des feux de forêts en résultant, et des inondations des zones côtières accroîtra les risques pesant sur les systèmes écologiques et humains. La crise climatique est causée par les activités humaines responsables d'une augmentation sans précédent des émissions de gaz à effet de serre (GES) depuis le début de l'ère industrielle [IPCC, 2014b]. La part estimée de la production d'électricité et de chaleur dans le total des émissions mondiales de GES va de 17% à 30% (données de 2010) [IPCC, 2014a]. La part de l'énergie utilisée dans les bâtiments résidentiels dans le total des émissions mondiales de GES va de 10,9% (données de 2016, [Ritchie and Roser, 2020]) à 12% (données de 2010, [IPCC, 2014a]).

Réduire les émissions de GES tout en soutenant le développement des sociétés humaines nécessite l'adoption de comportements durables dans les pays développés, et de remplacer mondialement les énergies fossiles (charbon, pétrole, gaz) utilisées pour la production d'électricité et pour d'autres usages par des alternatives « bas-carbone ». La production d'électricité d'origine nucléaire et renouvelable est considérée comme bas-carbone puisque les émissions de GES lors du fonctionnement et du cycle de vie complet de ces centrales sont estimées être largement plus faibles que celles des centrales fonctionnant aux énergies fossiles.

C'est une des raisons pour lesquelles les incitations gouvernementales au développement des énergies renouvelables se sont développées sous diverses formes dans les dernières décennies. Les autres raisons sont le besoin d'alternatives face à la raréfaction pétrolière, les inquiétudes liées à la sécurité des centrales nucléaires après l'accident de Fukushima, et le problème sanitaire lié à la pollution de l'air due aux centrales thermiques à énergies fossiles.

Ces politiques de soutien, couplées à des avancées soutenues de recherche et développement, ont mené à des réductions de coût de production exceptionnelles dans les 10 dernières années, permettant une croissance rapide de la capacité installée et en projet de production d'électricité renouvelable [SolarPower Europe, 2018].

Par ailleurs, les unités de production d'électricité renouvelable de suffisamment petite puissance peuvent être raccordées au réseau de basse ou moyenne tension, contrairement aux grandes centrales connectées au réseau de transport d'électricité de haute tension. Dans de nombreux endroits, il est estimé que l'installation de ces sources d'énergie distribuées pourra limiter le coût et les défaillances du système électrique associés à l'augmentation de la consommation d'électricité et à l'extension des réseaux électriques.

Ces deux tendances expliquent la croissance rapide dans le monde des capacités installées de production d'électricité photovoltaïque en particulier, celle-ci étant la forme la moins chère et la plus simple à installer parmi les sources d'électricité renouvelables distribuées. Un avantage supplémentaire de l'électricité photovoltaïque est la variété des tailles d'installations possibles, des centrales gigantesques situées dans les régions désertiques

jusqu'aux installations individuelles composées d'un seul panneau solaire placé sur le toit d'un petit bâtiment.

Dans les pays à revenu faible et intermédiaire [World Bank Data Help Desk, 2021] où les réseaux de transport et de distribution d'électricité n'ont pas encore atteint la taille et la résilience nécessaires pour couvrir la quasi-totalité des lieux habités, la production distribuée offre l'opportunité de fournir de l'électricité dans des lieux reculés, au lieu de devoir attendre l'extension du réseau électrique et la connexion avec des centrales de production distantes.

En somme, la part grandissante des ressources distribuées dans la production d'électricité mondiale permet d'anticiper un changement de paradigme, qui est le passage d'une organisation centralisée du système électrique où la production s'adaptait aux besoins de consommation, vers une organisation hybride, où la demande d'électricité est également pilotée afin de l'adapter à l'intermittence des ressources distribuées qui engendre une plus grande variabilité de la production. Ainsi, les flexibilités de production et de consommation sont de plus en plus valorisées pour les gestionnaires de réseaux de transport et de distribution d'électricité, en tant qu'options alternatives à un renforcement significatif des investissements financiers dans les réseaux [IEA, 2018].

La flexibilité de consommation est la possibilité de déplacer dans le temps ou de modifier la consommation électrique de certains usages. L'autoconsommation, quant à elle, est un cadre réglementaire visant à encourager ces flexibilités de consommation à l'échelle d'un bâtiment, en promouvant la consommation locale plutôt que l'export vers le réseau électrique. L'électricité en question peut être produite localement par tout type de ressource distribuée, mais les panneaux photovoltaïques (PV) sont principalement utilisés étant donné la taille modulable de leurs installations, et leur coût avantageux. L'autoconsommation solaire contribue à la stabilité du réseau de distribution électrique en évitant des hausses de tension lors des pics de production solaire, et aide à augmenter la part du PV dans le mix électrique [Nousdilis et al., 2018]. Grâce à des simplifications de procédures administratives [EU Directive, 2018] et de nouvelles grilles tarifaires qui ont fait de l'autoconsommation électrique une option rentable, ces installations continuent d'être encouragées et sont devenues populaires.

Les installations en autoconsommation PV favorisent le pilotage intelligent des consommations des usages électriques. En effet, tirer pleinement parti des incitations financières liées à l'autoconsommation nécessite un pilotage optimal des consommations du foyer. De plus, la valeur des gestionnaires d'énergie utilisés pour surveiller, contrôler et optimiser les consommations du foyer est accrue par le montant croissant des factures d'électricité des consommateurs. Cette tendance devrait se poursuivre en vue de l'électrification des consommations d'énergie habituellement couvertes par des combustibles fossiles.

2 Objectifs de la thèse

L'objectif de cette thèse est d'avancer sur la voie d'un gestionnaire d'énergie apportant des bénéfices à la fois au consommateur et au gestionnaire du réseau de distribution. Le premier cherche à diminuer sa facture totale d'énergie, tandis que le second désire le plus haut taux d'autoconsommation possible.

Ce travail se concentre sur l'optimisation de l'autoconsommation photovoltaïque des bâtiments résidentiels, étant donné que les bâtiments commerciaux ont des courbes de charge et des contraintes spécifiques et recouvrent de nombreuses catégories différentes. Dans le secteur résidentiel, les foyers peuvent avoir des compositions variées et habiter des appartements ou des maisons individuelles. Les installations en autoconsommation considérées ici ne couvrent que cette dernière catégorie. Dans la première partie de ce travail, l'échelle d'une maison individuelle est considérée, tandis que le cadre d'étude est

élargi à un groupement de maisons individuelles dans la seconde partie.

Avec 12% de la consommation électrique domestique d'Europe dédiée à l'eau chaude en 2018 [Eurostat, 2018], les équipements électriques de chauffage de l'eau (avec et sans stockage) représentent une part importante dans les factures d'électricité des bâtiments résidentiels. Les ballons d'eau chaude, ou chauffe-eau Joule (CEJ) sont des équipements particulièrement intéressants puisqu'ils fournissent un moyen de stockage de l'énergie sous forme thermique. D'un coût initial faible, les CEJ ont l'intérêt d'être déjà répandus dans plusieurs pays : les 57 millions d'unités installées en 2014 en Europe représentent 23% du parc total d'équipement principal de chauffage de l'eau en Europe [VHK, 2019], et cette part de marché atteint 45% en France avec 11 millions d'unités installées [MSI, 2019]. De plus, par leur consommation d'énergie et leur puissance nominale élevées, les CEJ représentent un équipement électrique particulièrement propice au pilotage dans le cadre d'une optimisation de l'autoconsommation du photovoltaïque [Cao et al., 2013, Lefort et al., 2013, Sossan et al., 2013, Heleno et al., 2015, Beeker et al., 2016, Pacaud, 2018]. La première partie de cette thèse porte exclusivement sur le contrôle des CEJ, tandis que l'éventail des usages électriques pilotés est étendu aux usages blancs et aux véhicules électriques dans la seconde partie.

2.1 Définition des objectifs concernant l'autoconsommation

Plusieurs critères permettent d'évaluer le fonctionnement d'une installation en autoconsommation.

Tout d'abord, l'énergie autoconsommée, notée SC dans ce manuscrit, est définie comme la part de la production PV locale qui est consommée localement. Mathématiquement, elle est définie comme l'intégrale sur la période $[0, \tau]$ du minimum entre la production PV locale et la consommation électrique totale :

$$SC = \int_0^\tau \min(C(t), P_{\mathrm{PV}}(t))dt \tag{1}$$

où $C(t)$ est la consommation électrique totale du foyer au temps t, et $P_{\mathrm{PV}}(t)$ est la puissance totale produite par les panneaux PV au temps t.

Ensuite, le taux d'autoconsommation est le ratio de l'énergie autoconsommée SC sur l'énergie totale localement produite E_{PV}, sur la même période.

$$SC_{\%} = \frac{SC}{E_{\mathrm{PV}}} \tag{2}$$

où

$$E_{\mathrm{PV}} = \int_0^\tau P_{\mathrm{PV}}(t)dt \tag{3}$$

Un taux d'autoconsommation de 100% indique que la production PV locale est totalement consommée sur place. E_{PV} étant indépendant des stratégies de pilotage, maximiser SC revient à maximiser $SC_{\%}$.

De même, le taux d'autoproduction est le ratio de l'énergie autoconsommée SC sur l'énergie totale localement consommée E_{C}, sur la même période.

$$SP_{\%} = \frac{SC}{E_{\mathrm{C}}} \tag{4}$$

où

$$E_{\mathrm{C}} = \int_0^\tau C(t)dt \tag{5}$$

Un taux d'autoproduction de 100% indique que la consommation locale est totalement couverte par la production locale, qui peut même être en surplus. E_{C} dépend des stratégies de pilotage des consommations, donc maximiser SC ne maximise pas forcément $SP_{\%}$.

Par ailleurs, un critère important pour les propriétaires d'une installation en autoconsommation est la facture globale d'énergie. Celle-ci dépend des profils et des volumes de consommation et de production, de leur concomitance, des prix de revente ou subventions accordées pour l'injection de puissance dans le réseau, et des prix horaires de l'électricité. Par toutes ces composantes, l'optimisation de l'énergie autoconsommée peut ne pas coïncider avec l'optimisation de la facture globale d'énergie.

Le point de vue d'un gestionnaire de réseau de distribution mènerait à vouloir minimiser le volume d'énergie réinjectée dans le réseau, défini par

$$\int_0^\tau \max(P_{\mathrm{PV}}(t) - C(t), 0)dt \tag{6}$$

Une installation en autoconsommation connectée au réseau mais cherchant à atteindre une autonomie énergétique pourrait souhaiter minimiser le volume total d'énergie importée, défini par

$$\int_0^\tau \max(C(t) - P_{\mathrm{PV}}(t), 0)dt \tag{7}$$

Enfin, des critères composites peuvent être définis afin de prendre en compte d'autres aspects tels quel des mesures de confort, en plus d'un ou plusieurs des critères présentés.

Les objectifs habituellement retenus sont différents selon la configuration de l'installation en autoconsommation. On distingue généralement trois configurations :

1. Un micro-réseau isolé : dans ce cas, le fonctionnement doit garantir la continuité de l'approvisionnement électrique à partir d'unités de production locales renouvelables, soutenues par d'autres sources d'énergie fossiles pilotables, ou un volume conséquent de stockage. Le taux d'autoconsommation est nécessairement de 100%.
2. Le réseau de distribution local est soumis à de fortes contraintes. Etant donné qu'historiquement, les réseaux de distribution étaient conçus pour satisfaire un certain niveau de consommation, des contraintes peuvent apparaître lorsque la production décentralisée renouvelable est trop importante et dépasse la consommation locale. Si trop de puissance est réinjectée, des hausses de tension peuvent apparaître et endommager le réseau. Des sanctions financières dissuasives peuvent être appliquées en cas de réinjection de puissance dépassant un certain seuil. Dans ce cas, le gestionnaire cherchera à avoir des taux d'autoconsommations aussi élevés que possible.
3. L'installation en autoconsommation individuelle ou collective est connectée au réseau de distribution, sans contraintes de hausses de tension. Aucune contrainte ne porte alors sur les taux d'autoconsommation ou d'autoproduction. L'objectif sera alors habituellement de minimiser la facture d'énergie.

Dans cette thèse, aucune contrainte d'autonomie ni de réinjection n'est considérée, ce qui correspond au troisième cas présenté. Dans les deux premiers chapitres, l'objectif est d'évaluer le taux d'autoconsommation maximal atteignable par une installation individuelle standard, afin d'obtenir un outil de comparaison. Les deux derniers chapitres, eux, portent sur la minimisation de la facture d'énergie, pour considérer un objectif plus aligné avec les préoccupations des propriétaires de ces systèmes.

2.2 Objectifs de robustesse face aux incertitudes

La production PV dépend de l'intensité de l'irradiance solaire, qui peut être affectée par plusieurs phénomènes : variations saisonnières, variabilité journalière, couverture nuageuse, ombre de nuages passant. Ces phénomènes ne sont pas parfaitement prévisibles. Ainsi la variabilité de la météo impacte la fiabilité des prévisions de météo et de production [Inman

et al., 2013]. Plus la prédiction concerne un horizon de temps lointain, plus l'incertitude est grande.

Il n'est pas rare de présenter les performances d'un gestionnaire d'énergie en autoconsommation testé dans un cadre négligeant les incertitudes de production PV. Cela est valable si l'on se concentre sur les performances *a posteriori* d'un gestionnaire d'énergie, où l'on mesure la performance réelle et passée, différente de la performance qui était prévue par le gestionnaire au moment de la prise de décision en anticipation. Néanmoins une évaluation valide des performances *a priori* se doit de prendre les incertitudes en compte. Cela nécessite de considérer que les prédictions de production PV utilisées pour décider d'une stratégie de contrôle peuvent ne pas se matérialiser exactement. Une évaluation rigoureuse des performances de la stratégie de pilotage choisie peut être faite en testant la stratégie de pilotage face à un grand nombre de scénarios de production simulés, et en mesurant la performance moyenne. Chaque scénario de cet ensemble doit être réaliste individuellement, et les variations de tout l'ensemble doivent être représentatives du niveau d'incertitudes associé à la prévision initiale. Le Chapitre 2 propose une méthode pour construire un ensemble de scénarios de production PV cohérent de ce type, et l'utilise pour fournir une évaluation correcte et *a priori* de la performance de l'algorithme de contrôle déterministe développé dans le Chapitre 1, face aux incertitudes de prévisions de production. Dans la deuxième partie de cette thèse, l'approche par horizon glissant de la commande prédictive est adoptée, permettant d'adapter les stratégies de contrôle aux informations mises à jour progressivement (Chapitres 3 et 4).

2.3 Objectifs de déploiement sur les micro-réseaux avec échanges d'énergie

Un micro-réseau est un réseau électrique peu étendu, composé de plusieurs bâtiments et de sources d'énergie distribuées produisant localement une partie de l'énergie requise. Les bâtiments peuvent être tertiaires ou résidentiels ; les sources d'énergie distribuées peuvent être des panneaux PV, de petites éoliennes, des centrales de cogénération, des batteries électriques ; les sources d'énergie distribuées peuvent appartenir aux agents individuellement ou collectivement, ou appartenir à une entité tierce (comme le gestionnaire du réseau de distribution, ou un fournisseur d'énergie) ; le micro-réseau peut être complètement isolé, isolé sur demande, ou bien complètement connecté au réseau électrique principal ; un micro-réseau peut couvrir des zones de tailles variées, allant d'un seul immeuble à l'échelle d'une petite île.

Puisqu'un micro-réseau est constitué de plusieurs bâtiments, de quelques usages électriques pilotables et de sources d'énergie distribuées, ces installations peuvent être décrites comme des communautés en autoconsommation collective, si les échanges d'énergie entre les participants sont permis. Dans ce cas, la définition initiale de l'autoconsommation peut être étendue au périmètre de la communauté, en incluant toute l'énergie produite et consommée au sein du micro-réseau. Les échanges d'énergie locaux correspondent principalement à un principe comptable permettant de partager la production d'énergie locale entre les participants de la communauté énergétique, afin d'augmenter la part d'énergie localement produite et consommée par chaque entité[1]. La seule contrainte physique à respecter pour la stabilité du réseau est d'assurer que les pics de tension dus aux réinjections locales de puissance ne dépassent pas le niveau que peut supporter le réseau de distribution.

[1] Physiquement, les électrons suivent toujours le chemin le plus court (ou chemin de moindre résistance), il est donc impossible de garantir l'origine physique de l'énergie consommée à un endroit d'un réseau connecté et maillé.

2.4 Objectifs de protection des données personnelles

Dans les Chapitres 3 et 4, des échanges d'énergie sont permis entre des membres d'un micro-réseau, un groupement d'habitations individuelles s'accordant pour se vendre et s'acheter de l'énergie, pour partager l'usage d'une ressource d'énergie distribuée (telle une batterie), et pour chercher à atteindre un objectif commun, la minimisation du coût de fourniture de l'énergie pour la communauté dans son ensemble et pour chaque participant particulier. Dans ce cadre, l'augmentation de l'énergie autoconsommée à l'échelle du micro-réseau peut servir leur objectif. Dans cette communauté énergétique, les gestionnaires d'énergie de chaque foyer doivent communiquer entre eux et avec l'entité centrale, ce qui soulève des problèmes de divulgation des informations privées des participants.

Dans les expériences présentées dans [Huberman et al., 2005], des individus devaient établir la valeur (en dollars) de certaines informations personnelles. Un des résultats notables est que le pourcentage de personnes demandant une somme « infinie » (plus de 100$ dans l'expérience) pour révéler des informations telles que le salaire et la situation conjugale était bien plus élevé que le pourcentage de personnes demandant une telle somme pour d'autres types d'informations personnelles telles que l'âge et le poids. Ceci montre l'existence d'un attachement à un certain niveau de protection de certaines informations personnelles.

Cependant, il se trouve que certaines de ces informations peuvent être déduites des données de consommations électriques. En effet, [Molina-Markham et al., 2010] a montré que des traitements statistiques simples des consommations enregistrées chaque seconde pendant deux mois permettaient d'identifier les équipements électriques d'un foyer, et peut-être même la routine journalière en cas de disponibilité de larges jeux de données. Au sein d'un enregistrement des consommations électriques de 180 foyers sur 14 jours au pas de temps horaire, [Buchmann et al., 2013] a réussi à identifier une combinaison unique de paramètres (comme le premier pic de demande dans la journée, ou la consommation totale journalière) permettant d'identifier précisément 68 % des foyers.

Il est donc possible de déduire des informations personnelles telles que la richesse, le statut professionnel, les départs en vacances ou le nombre de personnes vivant dans une maison. C'est pourquoi le droit à la protection des données personnelles est protégé par la réglementation européenne en matière de protection de la vie privée [EU Regulation, 2016], bien connue sous le nom de Règlement Général sur la Protection des Données (RGPD). Ce règlement établit que « la protection des personnes physiques à l'égard du traitement des données à caractère personnel est un droit fondamental ». Concrètement, il est déclaré que « les données à caractère personnel doivent être collectées pour des finalités déterminées, explicites et légitimes, et ne pas être traitées ultérieurement d'une manière incompatible avec ces finalités ».

Ainsi la protection des données personnelles est un enjeu majeur pour la conception de futures communautés locales d'énergie supposées bénéfiques pour les consommateurs et le gestionnaire du réseau de distribution, puisque :

- le gestionnaire du réseau de distribution et les tierces parties doivent se conformer au RGPD ;
- une certaine garantie quant au niveau de protection de ces données pourrait rassurer les nouveaux participants potentiels d'une telle communauté.

Il est à noter que malgré ces règles et les aspirations à la protection des données personnelles, les auteurs de [Huberman et al., 2005] rappellent que les individus sont prêts à faire des compromis sur ces aspects pour obtenir certains services. Cet équilibre est également rappelé dans le règlement européen [EU Regulation, 2016], qui déclare que « le traitement des données à caractère personnel devrait être conçu pour servir l'humanité. Le droit à la protection des données à caractère personnel n'est pas un droit absolu ; il

doit être considéré par rapport à sa fonction dans la société et être mis en balance avec d'autres droits fondamentaux, conformément au principe de proportionnalité. »

3 Organisation du manuscrit

3.1 Résumé du problème considéré dans la thèse

Le problème considéré dans cette thèse peut être décrit comme le pilotage d'usages électriques d'une certaine entité (foyer ou micro-réseau) afin de maximiser la consommation de production locale d'énergie photovoltaïque, tout en respectant des contraintes de confort et de fonctionnement.

3.2 Contenu

Le Chapitre 1 de ce manuscrit présente la formulation mathématique du problème d'optimisation de l'autoconsommation du photovoltaïque dans un bâtiment résidentiel. Le problème d'optimisation a été formulé sous une forme sans contraintes incluant une prévision de production PV. Les propriétés mathématiques du problème sont étudiées (non-concavité, différentiabilité). On modélise le CEJ comme un volume d'eau à température homogène équipé d'un thermostat. Les équations dynamiques qui en régissent le fonctionnement peuvent être explicitement résolues, menant à un algorithme de pilotage optimal simple et efficace. Moins générale que des approches par Model Predictive Control (MPC), qui pourraient avoir été utilisées dans ce contexte, cette stratégie requière moins de puissance de calcul et ne nécessite pas d'ajuster de paramètres. Le temps de calcul extrêmement faible de l'algorithme proposé permet de résoudre plusieurs instances différentes du même problème dans un temps court, et le rend compatible avec le temps réel.

Dans le Chapitre 2, les performances de l'algorithme développé sont étudiées en considérant les incertitudes des prévisions de production. La prévision de production PV considérée comme la prévision la plus probable est celle retenue nominalement, et est associée à un certain niveau de confiance. A partir de ces deux éléments, une méthode novatrice est décrite pour générer des ensembles de scénarios de production PV. Chaque scénario individuel peut dévier de la prévision la plus probable. L'ensemble des scénarios suit globalement la prévision la plus probable, et s'en éloigne en fonction du niveau de confiance initial. La performance de l'algorithme de pilotage précédemment développé peut ainsi être évaluée dans des conditions plus réalistes, en confrontant la stratégie de pilotage adoptée grâce à la prévision initiale la plus probable avec l'ensemble des scénarios possibles associés. Cette étude est possible grâce au faible temps de calcul requis par l'algorithme de pilotage. Ce chapitre montre que le gain de performance obtenu par l'algorithme proposé par rapport à une heuristique de gestionnaire disponible commercialement dépasse le gain qui serait hypothétiquement obtenu grâce à une prévision de la production PV au pas de temps de 30-minutes plus précise que la prévision nominale actuelle, toujours appliquée à l'heuristique de référence.

Le Chapitre 3 étend le cadre d'étude au contrôle de plusieurs autres équipements électriques dont une modélisation retenue utilise des variables binaires (ex. : usages blancs, voitures électriques, batteries) et à l'optimisation de l'autoconsommation d'un ensemble de foyers en coopération. Ce cadre mène à la description d'un problème d'optimisation linéaire mixte (MILP), pour lequel l'algorithme précédemment présenté n'est pas adapté. Les dimensions du problème MILP à considérer sont précisées, et l'intérêt de la coopération est démontré sur un exemple simple.

Le Chapitre 4 aborde le défi de la protection des données personnelles des participants au micro-réseau, un enjeu rarement considéré dans les études d'optimisation collective des consommations d'énergie. Des méthodes de télécommunication et de cryptographie

classiques sont adaptées et appliquées à une étude de cas, montrant la faisabilité de la coopération efficace sous la contrainte de l'anonymat.

Publications

Ces travaux de thèse ont donné lieu à deux publications.

Article de conférence : Loris Amabile, Delphine Bresch-Pietri, Gilbert El Hajje, Sébastien Labbé, and Nicolas Petit. An optimization methodology for self-consumption of residential photovoltaic energy. In IFAC World Congress 2020 *IFAC-PapersOnLine*, volume 53, pages 13196–13203, 2020. doi :10.1016/j.ifacol.2020.12.145. URL `https://linkinghub.elsevier.com/retrieve/pii/S2405896320304031`. [Amabile et al., 2020]

Article de revue : Loris Amabile, Delphine Bresch-Pietri, Gilbert El Hajje, Sébastien Labbé, and Nicolas Petit. Optimizing the self-consumption of residential photovoltaic energy and quantification of the impact of production forecast uncertainties. *Advances in Applied Energy*, 2:100020, May 2021. ISSN 26667924. doi :10.1016/j.adapen.2021.100020. URL `https://linkinghub.elsevier.com/retrieve/pii/S2666792421000135`. [Amabile et al., 2021]

Introduction

1 Context

The climate crisis is threatening the living conditions on Earth: expected increases of heat waves, heavy rain, drought and associated wildfires, and coastal flooding will further put ecological and human systems at risk. The climate crisis is caused by human activities responsible with the unprecedented increase of greenhouse gases (GHG) emissions since the begining of the industrial era [IPCC, 2014b]. The estimated share of the "electricity and heat production" sector in the total GHG world emissions ranges from 17% to 30% (2010 data) [IPCC, 2014a]. The estimated share of the "residential building" end-use in the total GHG world emissions ranges from 10.9% (2016 data, [Ritchie and Roser, 2020]) to 12% (2010 data, [IPCC, 2014a]).

Reducing GHG emissions while supporting human development requires to adopt sustainable behaviors in developed countries and to replace fossil fuel power generation and other fossil-fuel consumptions by "low-carbon" alternatives worldwide. Nuclear and renewable power plants are considered to be low-carbon as their operation GHG emissions as well as life-cycle estimated GHG emissions are far lower than those of fossil fuel power plants.

This is one of the reasons why incentives for renewable energy have grown worldwide under various forms of government supports in the past decades. Other reasons are the need for alternatives to oil (whose conventional sources of declining productivity let anticipate a rising price of the resource), the concerns regarding the safety of nuclear power plants after the Fukushima accident, and the air pollution sanitary issues due to fossil fuel power plants.

These supporting policies, coupled with sustained research and development breakthroughs, have led to unprecedented production cost reductions in the past decade, enabling a rapid growth of worldwide installed capacities and installation projects of renewable energy [SolarPower Europe, 2018].

Moreover, if the capacity of the renewable energy generation unit is sufficiently small, it can be connected to the distribution system, contrary to the conventional centralized power stations connected to the high-voltage electricity grid. In many places, such distributed energy resources (DERs) are expected to contain the costs and faults associated with increasing power consumption and extending electrical grids.

Both these trends explain the rapid growth of worldwide installed capacities of solar power in particular, which is the cheapest and easiest-to-install distributed renewable energy source. Another advantage of solar power is the variety of installations sizes, from huge power plants in desertic regions down to individual installations composed of a single photovoltaic (PV) panel fitting on the roof of a small building.

In the low and middle-income countries [World Bank Data Help Desk, 2021] where the transmission and distribution networks have not reached the size and robustness required to cover the vast majority of the occupied land, DERs offer the opportunity to supply electricity in remote locations by the installation of a few of these decentralized production

units, instead of waiting on an extension of the network and get connected to a central and distant power plant.

All in all, the growing share of DERs in the installed capacity worldwide lets us envision a shift of paradigm, switching from a centralized organization of the power grid where production is adapted to consumption, towards a hybrid vision, where the demand side is managed as well in order to adapt to the intermittent production of DERs that induces higher variability in supply. Hence, production and consumption flexibilities become increasingly valuable for the transmission and distribution system operators (DSOs) as substitutes for substantial grid investments [IEA, 2018].

Consumption flexibility is the ability to temporally shift or to modify the power consumption of electrical appliances. Self-consumption is one regulatory framework intended to bolster these consumption flexibilities at the building scale, by promoting local consumption over export to the main grid. Power can be locally produced by any kind of DERs, but PV is the main technology used due to its modular size and advantageous cost. PV self-consumption contributes to the distribution grid stability by avoiding voltage rise during peak PV generation periods and helps to reach higher shares of PV generation in the electricity mix [Nousdilis et al., 2018]. As simplified administrative procedures [EU Directive, 2018] and new tariffs have made self-consumption of generated power a profitable option, these installations are increasingly encouraged and have become popular.

PV self-consumption installations foster local smart management of electrical appliances because fully benefiting from the self-consumption financial regulatory incentives requires optimal management of household consumptions. Moreover, the value of these Energy Management Systems (EMSs) used to monitor, control, and optimize the household consumptions is enhanced by the increasing electricity bills of retail consumers. This trend can be expected to continue due to the electrification of energy consumptions traditionally supplied by fossil fuels.

2 Objectives of the thesis

The aim of this work is to shape the way towards an energy management system (EMS) benefiting both stakeholders, the consumer and the DSO. The first is concerned with the final energy bill, while the second desires the highest self-consumption rates possible.

This work focuses on optimizing the PV self-consumption of residential buildings, as commercial buildings have specific load curves and constraints and are subdivided in a number of categories. In the residential sector, households can have various compositions and inhabit apartments or individual houses. The self-consumption installations considered in this work are only of the latter category. In the first part of this work, the scale of an individual house is considered, whereas the scope is enlarged up to a neighborhood composed of several individual houses in the second part.

With 12% of all European domestic electric consumption dedicated to water heating in 2018 [Eurostat, 2018], electric water heating appliances (with and without storage) represent a major share of residential electricity bills. A type of appliance of particular interest is Electric Water Heaters (EWHs), which provide storage capability in thermal form. With their low investment costs, EWHs have the benefit of being already widespread in various countries: the 57 million units installed in 2014 in Europe represent 23% of the total European primary water heater stock [VHK, 2019], and this market share goes up to 45% in France with 11 million units installed [MSI, 2019]. Moreover, considering their high energy consumption and power rating, they represent a highly suitable appliance to control in the context of PV self-consumption optimization [Cao et al., 2013, Lefort et al., 2013, Sossan et al., 2013, Heleno et al., 2015, Beeker et al., 2016, Pacaud, 2018]. The first part of this thesis focuses exclusively on the control of EWHs, whereas the scope of

controlled appliances is enlarged to white goods[2] and electric vehicles (EVs) in the second part.

PV production depends on the intensity of solar irradiance, which can be affected by several phenomena: seasonal variations, intra-day variability, cloud coverage, fast passing clouds. These phenomena are not perfectly predictable. Thus weather variability impacts the reliability of weather and production forecasts [Inman et al., 2013]. In the first part of this thesis, a deterministic framework is developed for the optimization of PV self-consumption (Chapter 1), and its robustness is tested in face of the PV production uncertainties (Chapter 2). In the last part of the thesis, the receding horizon MPC approach can be applied, allowing to adapt the control strategy to updated information (Chapter 3).

2.1 Objectives concerning self-consumption installations

Several metrics can be of interest for the evaluation of PV self-consumption installations.

First, the self-consumed energy, simply called self-consumption and noted SC in the remainder of this thesis, is defined as the part of local PV production that is locally consumed to meet electric consumption. Mathematically, it is defined as the integral over the period $[0, \tau]$ of the minimum between the local PV production and the total electric consumption:

$$SC = \int_0^\tau \min(C(t), P_{\mathrm{PV}}(t))dt \tag{8}$$

where $C(t)$ is the total power consumption of the household at time t, and $P_{\mathrm{PV}}(t)$ is the total power produced by the PV arrays at time t.

Second, the self-consumption rate is the ratio of the self-consumed energy SC to the total local energy production E_{PV}, over the period.

$$SC_{\%} = \frac{SC}{E_{\mathrm{PV}}} \tag{9}$$

where

$$E_{\mathrm{PV}} = \int_0^\tau P_{\mathrm{PV}}(t)dt \tag{10}$$

A 100 % self-consumption rate indicates that the local PV production is totally consumed locally. E_{PV} being independent of the control strategies, an optimization performed over SC also maximizes $SC_{\%}$.

Similarly, the self-production rate is the ratio of the self-consumed energy SC to the total local energy consumption E_{C}, over the period.

$$SP_{\%} = \frac{SC}{E_{\mathrm{C}}} \tag{11}$$

where

$$E_{\mathrm{C}} = \int_0^\tau C(t)dt \tag{12}$$

A 100 % self-production rate indicates that the local energy consumption is totally covered by the local production, which can be in surplus. E_{C} depends on the control strategies, hence maximizing SC does not necessarily maximize $SP_{\%}$.

Then, an important criterion for the inhabitants owning a self-consumption installation is the overall energy bill. It is impacted by the consumption and production patterns and volumes, their co-occurrence, the selling price or subsidies accorded for the injection of

[2] Large electrical appliances used for routine housekeeping tasks (such as cooking, cleaning, washing laundry, or food preservation) which were traditionally available only in white.

power on the electricity grid, and the electricity hourly tariffs. Due to all these features, the optimization of the self-consumed energy might not coincide with the optimization of the overall energy bill.

A DSO point-of-view could lead to try and minimize the volume of injected power, defined by

$$\int_0^\tau \max(P_{\mathrm{PV}}(t) - C(t), 0)dt \tag{13}$$

A self-consumption installation connected to the grid but wishing to reach an energetic autonomy could be interested in minimizing the volume of imported power, defined by

$$\int_0^\tau \max(C(t) - P_{\mathrm{PV}}(t), 0)dt \tag{14}$$

Finally, composite metrics could be formulated in order to take one of the previous criteria into account as well as other aspects such as a measure of comfort.

Three different configurations can be described as self-consumption installations. The goals regarding the self-consumption and self-production metrics will be different:

1. An isolated microgrid: in this case, operations must guarantee the continuity of power supply from a mix of renewable and non-renewable production units, or from a mix of renewable production units backed with a consequent volume of storage. The self-production rate is necessarily of 100%.
2. The local network experiences a high level of constraints. Given that distribution networks were historically sized according to the expected consumption, pressure can come from a high number of renewable generation units generating power at the same time. If production is not consumed locally, a large volume of power can be injected, voltage rises can then appear and damage the grid. Deterrent penalties can be applied in case of upstream power injections above a threshold. In this case, the self-consumption rate has to be as high as possible.
3. The individual or collective self-consumption installation is connected to an external grid without voltage surges constraints: no objective nor constraint are linked to the self-consumption or self-production rates. The objective is usually to minimize the energy bill.

In this work, no constraints on autonomy nor power injection are considered, thus the third configuration is the only one retained. In the first chapter, the aim is to assess the maximal self-consumption rate attainable by a standard individual installation. It can serve as a benchmark for other self-consumption EMSs. Thus the criterion retained is the self-consumed energy volume SC. Maximizing it also maximizes $SC_{\%}$, the self-consumption rate. In the second chapter, the method developed allows to assess the statistical expectancy of this same metric, for a realistic presentation of *a priori* performances. Once these investigations are terminated, the focus can be shifted to a more realistic use-case, hence the last two chapters focus on minimizing the energy bill of the households equipped with PV self-consumption installations.

2.2 Objectives concerning robustness to uncertainties

As PV production depends on meteorological conditions, the production prediction is uncertain. The further in time the prediction, the larger the uncertainty.

It is not uncommon to present the performance of an self-consumption EMS tested in a setting disregarding the uncertainty of PV production. This is acceptable when assessing the *a posteriori* performances of an EMS, when measuring the past and real performance, different from the one predicted by the EMS at the time of decision-making,

in anticipation. However a valid *a priori* performance assessment should always take this uncertainty into account. Doing so requires to consider that the PV production prediction used to choose a control schedule might not materialize exactly. A rigorous performance assessment of the chosen control schedule can be done by testing the schedule in face of a large number of simulated production scenarios, and retaining the mean performance. Each scenario of this set has to be realistic, and the set as a whose has to be representative of the degree of uncertainty associated with the initial prevision. Chapter 2 proposes a method to build such a coherent set of scenarios, and uses it to assess the correct *a priori* performance of the control algorithms.

2.3 Objectives for deployment on microgrids with energy sharing

A microgrid is a small-sized electric network composed of several buildings and DERs producing locally a part of the needed energy. The buildings can be tertiary or residential; the DERs can be PV panels, small wind turbines, combined heat and power plants, electric batteries; the DERs can be owned individually or collectively by the agents, or owned by a third-party (e.g. the DSO or a utility); the microgrid can be completely or optionally isolated, or fully connected to the main grid; the geographical area covered by the microgrid can vary from the scale of a building to the scale of a tiny island.

As the microgrid is composed of several buildings, of some controllable appliances and of DERs, the setup can conveniently be described as a collective self-consumption community if power exchanges between the participants are allowed. In this case, the initial definition of self-consumption is then extended to the community, including all the locally consumed and produced power within the microgrid. The local power exchanges are mainly an accounting principle allowing local power production to be shared between participants of the energy community, in order to increase the share of power locally produced and consumed by each entity[3]. The only physical constraint is to ensure that the voltage increases due to the local power input at the distribution level of the grid will not surpass the grid flexibility.

2.4 Objectives concerning privacy-preserving cooperation

In Chapters 3 and 4, energy exchanges are allowed among the members of a microgrid, i.e. a groupment of individual households agreeing to exchange energy with each other, to share the use of a DER (e.g. a battery), and to pursue the objective of lowering the energy procurement cost for the whole community and for each participant. In this framework, increasing the volume of self-consumed energy at the scale of the microgrid can help achieve this goal. In this energy community, the individual households EMSs have to communicate with each other and with the microgrid EMS, thus raising concerns of disclosure of private information to other participants.

[Huberman et al., 2005] conducted experiments leading individuals to state the value (in US dollars) of private information. Among other results, it highlighted that the percentage of individuals demanding an "infinite" (more than US$100 in the experiment) amount of money to reveal information such as salary and spousal salary was way higher than the percentage of individuals demanding such amount of money for other kind of private information such as weight and age. This result demonstrates the wish of a certain level of protection of some private data.

However, it happens that some elements of private data can be inferred from power consumption data. Indeed [Molina-Markham et al., 2010] showed that simple statisti-

[3]Physically, the electrons always travel through the shortest path (or path of least resistance), hence it is impossible to guarantee the physical origin of the consumed power in one location of a connected and meshed grid.

cal treatment allowed to identify appliances in the household monitored at a one-second timestep during two months, and possibly the daily routine in case of availability of large amounts of data. Among a dataset of 180 households monitored for 14 days at a one-hour timestep, [Buchmann et al., 2013] succeeded in finding a unique combination of features (such as the first peak of demand in the morning, or the aggregated consumption per day) identifying 68 % of the households.

The consumption patterns hence allow one to infer personal information like wealth, employment status, vacation timing or the number of people that live in a household. This is why the right to the protection of personal data is protected by the European privacy regulation [EU Regulation, 2016], well known as the General Data Protection Regulation (GDPR). This regulation states that "the protection of natural persons in relation to the processing of personal data is a fundamental right". In practical terms it is stated that "personal data shall be collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes".

Hence preservation of privacy is a key concern for the design of future energy communities that are deemed beneficial for both the consumers and the DSO, as

- the DSO and concerned third parties must abide to the constraining GDPR;
- a guaranteed level of privacy of personal data is assumed to be a reassuring feature to enroll new potential participants.

Note that despite these protections and desires for privacy, the authors of [Huberman et al., 2005] recall that individuals are ready to make privacy trade-offs to gain access to specific services. This balance is also recalled in the EU regulation [EU Regulation, 2016], stating that "the processing of personal data should be designed to serve mankind. The right to the protection of personal data is not an absolute right; it must be considered in relation to its function in society and be balanced against other fundamental rights, in accordance with the principle of proportionality."

3 Manuscript organization

Problem statement

The problem at hand can be described as scheduling controllable appliances of a given entity (be it household or microgrid) in order to consume the maximum PV power produced locally and respecting comfort or functioning constraints.

Content

In Chapter 1 of this thesis, the aim is to build the optimization algorithm at the core of a residential EMS, showcasing the potential optimal performance attainable by a PV self-consumption installation. Doing so will provide a standard for the evaluation of other self-consumption optimization algorithms. The first accomplishment is to lay the mathematical formulation of the problem of optimizing the PV self-consumption of a residential building. The optimization problem is formulated in an unconstrained form involving a PV production prediction. The mathematical properties of the problem are studied (non-concavity, differentiability). Modeling the EWH a an homogeneous-temperature tank equipped with a thermostat, the dynamics at stake can be explicitly resolved, leading to a simple and efficient optimal EWH control algorithm. This strategy is less general than Model Predictive Control (MPC) approaches, for example, which could have been used in this context, but reveals less computationally intensive and requires no parameter tuning. Hence, because the computing time of the proposed algorithm is extremely low, various instances of the same problem can be solved in a short time.

In Chapter 2, the developed EMS performances are studied considering the uncertainties of production forecasts. The PV production prediction considered as a most likely prediction is the one nominally selected, associated with a certain level of confidence. From these two elements, a novel method is described to generate sets of PV production scenarios. Each individual scenario can deviate from the most likely prediction. As a whole, the scenarios set follows the most likely prediction, and its range corresponds to the given level of uncertainty. The performances of the proposed tailored EMS of Chapter 1 can thus be evaluated in a more realistic setting, where the most likely prediction used to chose the control strategy does not necessarily materialize exactly, and where this strategy is evaluated with the ensemble of possible scenarios. This study is possible thanks to the low computational cost of the proposed algorithm. This chapter will show that the performance gain attained by the proposed algorithm compared to a commercially available heuristic dwarfs the gain that would supposedly procure a 30-minute timestep PV production prediction with increased accuracy over the nominal current prevision.

Chapter 3 extends the scope to the control of several other appliances chosen to be modeled using binary variables (e.g. white goods, EVs, batteries) and to the cooperative optimization of a group of households. These elements leads to formulating a mixed-integer linear problem (MILP). The dimension of the MILP problem is detailed, and this benefit of cooperation is established with a simple example.

Chapter 4 is the opportunity to address the challenge of individual data privacy preservation, rarely considered in studies on collective optimization of energy resources. Classical methods of telecommunication and cryptography are adapted and applied to the case study at hand, showing the feasibility of efficient cooperation under an anonymity requirement.

Publications

The results of this thesis have been presented in two publications.

Conference proceedings: Loris Amabile, Delphine Bresch-Pietri, Gilbert El Hajje, Sébastien Labbé, and Nicolas Petit. An optimization methodology for self-consumption of residential photovoltaic energy. In IFAC World Congress 2020 *IFAC-PapersOnLine*, volume 53, pages 13196–13203, 2020. doi:10.1016/j.ifacol.2020.12.145. URL `https://linkinghub.elsevier.com/retrieve/pii/S2405896320304031`. [Amabile et al., 2020]

Journal paper: Loris Amabile, Delphine Bresch-Pietri, Gilbert El Hajje, Sébastien Labbé, and Nicolas Petit. Optimizing the self-consumption of residential photovoltaic energy and quantification of the impact of production forecast uncertainties. *Advances in Applied Energy*, 2:100020, May 2021. ISSN 26667924. doi:10.1016/j.adapen.2021.100020. URL `https://linkinghub.elsevier.com/retrieve/pii/S2666792421000135`. [Amabile et al., 2021]

Chapter 1

A self-consumption optimization algorithm

Un algorithme d'optimisation de l'autoconsommation. Dans ce chapitre, on considère une maison équipée de panneaux solaires dont l'unique consommation électrique contrôlable est celle du chauffe-eau Joule (CEJ). Le chapitre présente le modèle physique du CEJ considéré comme un volume d'eau chaude à température homogène, la définition du volume d'énergie autoconsommée retenu comme critère d'optimisation, la formalisation du pilotage optimal de ce CEJ sous la forme d'un problème sans contraintes, ainsi qu'un algorithme original de résolution. La performance de l'algorithme sur une année complète est évaluée par simulations numériques et comparée à celle de deux autres méthodes de pilotage. Dans l'exemple considéré, la méthode présentée permet d'augmenter le volume d'énergie auto-consommée de 28 % par rapport à une méthode de référence disponible commercialement. Un avantage manifeste de l'algorithme proposé est sa rapidité d'exécution.

Remark The content of this chapter corresponds to an in-depth description of the method presented in the first half of [Amabile et al., 2021]. The use-case considered here resembles the one presented in [Amabile et al., 2020], but the algorithm presented in this thesis is a improved version of the one presented in the article, thus resulting in better performances. To this extent, the numerical results presented in Section 1.3 are different from those described in the article.

1.1 Introduction

The first part of this thesis (Chapters 1 and 2) focuses on a simplified set-up where the only regulated appliance is the electric water heater (EWH), while the rest of the appliances represents an uncontrollable load. The objective is to design an optimization algorithm at the core of a residential EMS, aiming to maximize the self-consumption of the photovoltaic (PV) arrays (a key evaluation criterion for such installations), and able to update frequently the optimization strategy. Reformulating the EWH dynamics and constraints, the problem under consideration is transformed into an unconstrained optimization problem. The corresponding objective function is proven to be continuously differentiable almost everywhere and an explicit expression of its critical points and derivative discontinuity points is provided. From it, an optimization algorithm is deduced (using a much more computationally efficient numerical routine than other classical techniques (such as

scheduling algorithms, MPC, (stochastic) optimal control)). The performance of this optimization methodology is then compared to an industrial heuristic using a high-fidelity simulation platform.

1.2 Problem statement

1.2.1 Setup

The general setup under consideration is pictured in Figure 1.1 and consists of an individual house connected to the electrical grid. It is equipped with PV arrays, heating and cooling equipment, an EWH, and other electrical appliances such as lights, a refrigerator, and a dishwasher.

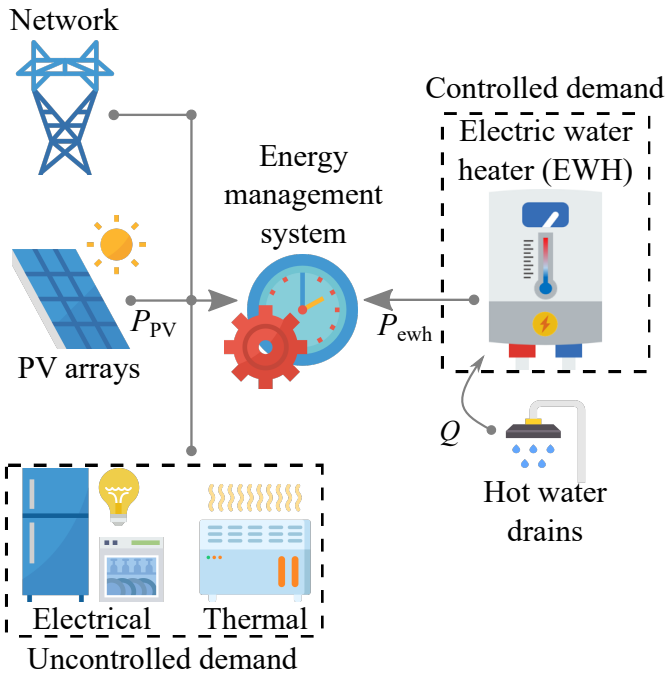

Figure 1.1: Setup under consideration: the Energy Management System regulates the EWH.

The heating and cooling systems are assumed to be entirely dedicated to maintaining house thermal equilibrium. Therefore, these appliances are considered as uncontrolled and are not associated to any decision variables in the remainder of this study. In this chapter and the following, the only controllable appliance to be considered is the EWH. To solve the optimization problem, the proposed algorithm assumes to have an exact prediction of the following elements:

- PV production curve;
- Electrical consumption profile of uncontrollable loads;
- Hot-water drains from the EWH.

Note that in Chapter 2, the impact of the imperfect nature of PV production forecasts on the proposed EMS performance will be studied.

1.2.2 Electric water heater modeling and control

The EWH is modeled as a homogeneous temperature volume. Leaving aside the modeling of the temperature stratification inside the tank implies a loss of precision, but simplifies the control design [Beeker-Adda, 2016]. Considering an initial energy state at time t_0, thermal losses, the power input from resistive heating, and the hot-water consumption during the given period, the energy stored at a time t_2 in the EWH, knowing the energy $E_1 = E(t_1)$, is given by

$$E(t_2) = E_1 \, \mathrm{e}^{-\mathrm{k}(t_2 - t_1)} + \int_{t_1}^{t_2} \mathrm{e}^{-\mathrm{k}(t_2 - t)} \left[P_{\mathrm{ewh}}(s, t) - Q(t) \right] dt \tag{1.1}$$

where

- $E(t)$ is the energy stored in the EWH at time t;
- k is the thermal loss coefficient [Beeker-Adda, 2016];
- $P_{\mathrm{ewh}}(s, t)$ is the electric power input of the EWH at time t, following the decision variable s relative to the heating strategy (detailed in the sequel);
- $Q(t)$ is the energy used at time t by a hot-water drain.

Here, the EWH is assumed to be equipped with an internal thermostat that is adjusted on a setpoint temperature T_{set} and a deadband of width $2M$. Hence, the internal thermostat stops any heating command when the measured temperature reaches the saturation temperature $T_{\mathrm{set}} + M$, and enables heating to resume when the temperature drops below the clearance temperature $T_{\mathrm{set}} - M$. Because the temperature inside the EWH is considered homogeneous, these specific temperatures $T_{\mathrm{set}} + M$ and $T_{\mathrm{set}} - M$ correspond to energy levels, denoted as E_{sat} and E_{clear}.

Correspondingly, the power input of the EWH can be written as

$$P_{\mathrm{ewh}}(s, t) = (1 - \delta_{\mathrm{sat}}(t)) P^{\circ}_{\mathrm{ewh}}(s, t) \tag{1.2}$$

where s is the decision variable relative to the control strategy, t is the time, P°_{ewh} is the chosen authorized heating power, and δ_{sat} is a hysteresis function describing the saturation state of the tank, equal to 1 if saturation is occurring and 0 otherwise. It is

$$\delta_{\mathrm{sat}}(t) = \begin{cases} 0 & \text{if } E(t) < E_{\mathrm{clear}} \\ 1 & \text{if } E(t) \geq E_{\mathrm{sat}} \\ \delta_{\mathrm{sat}}(t^-) & \text{if } E(t) \in [E_{\mathrm{clear}}, E_{\mathrm{sat}}[\end{cases} \tag{1.3}$$

where t^- is the left one-sided limit of time t.

The variable δ_{sat} in Eqs. (1.2)-(1.3) describes the behavior of the EWH internal thermostat, allowing heating to remain activated if the water temperature has not yet reached the saturation threshold or to resume if it has dropped below the clearance temperature.

The function P°_{ewh} is defined as follows. The EWH under consideration has a power consumption of either 0 or $\overline{P}_{\mathrm{ewh}}$, its constant power rating, triggered by an On/Off authorization mechanism. In this work, to simplify the heating control strategy, this authorization is chosen to be On only during a unique and continuous period $[t_{\mathrm{ewh}}, \tau_{\mathrm{ewh}}]$ determined by the EMS, which corresponds to the decision variable $s = (t_{\mathrm{ewh}}, \tau_{\mathrm{ewh}}) \in \mathbb{R}^2$. In detail,

$$P^{\circ}_{\mathrm{ewh}}(s, t) = \overline{P}_{\mathrm{ewh}} \mathbb{1}_{[t_{\mathrm{ewh}}, \tau_{\mathrm{ewh}}]}(t) \tag{1.4}$$

in which $\mathbb{1}_{\mathrm{I}}$ is the indicator function of the time interval I. Within this authorized period, the EWH internal thermostat might start and stop power activation according to Eq. (1.3).

The full description of the EWH behavior requires the following key assumption:

Assumption 1. *The saturation energy level and the hot water drains are such that* $-\,\mathrm{k}\,E_{\mathrm{sat}} + \overline{P}_{\mathrm{ewh}} - Q(t) \geq \eta > 0$ *and* $Q(t) \leq \overline{Q}$ *for certain constants* $\eta > 0$ *and* $\overline{Q} > 0$ *and for all time* t.

Additionally, note that obviously $0 < E_{\mathrm{clear}} < E_{\mathrm{sat}}$ and that $E_0 < E_{\mathrm{sat}}$. For a typical setting (such as the scenario considered in Section 1.3), Assumption 1 is not restrictive as the energy input from the resistive element of the EWH, with its high nominal power ($\overline{P}_{\mathrm{ewh}}$), largely surpasses the energy decrease due to standard hot-water consumptions ($Q(t)$) and energy losses ($\mathrm{k}\,E_{\mathrm{sat}}$).

Assumption 1 together with Eq.(1.1) ensure that, for $t \in [t_{\mathrm{ewh}}, \tau_{\mathrm{ewh}}]$, the EWH energy remains in $[E_{\mathrm{clear}}, E_{\mathrm{sat}}]$ once it is reached.

1.2.3 Optimization problem formulation

Objective function and decision variables

Self-consumption (SC) is usually considered as a proxy for reducing electricity bills, and high SC levels can be critical in locations where the state of the distribution grid cannot allow large volumes of PV surplus to be injected. Mathematically, it is defined as the integral over the period $[0, \tau]$ of the minimum between the local PV production and the total electric consumption:

$$SC(s, P_{\mathrm{PV}}) = \int_0^{\tau} \min(C(s,t), P_{\mathrm{PV}}(t))dt \tag{1.5}$$

where C is the total power consumption of the household depending on s and t, and P_{PV} is the total power produced by the PV arrays. Figure 1.2 gives a visual representation of it.

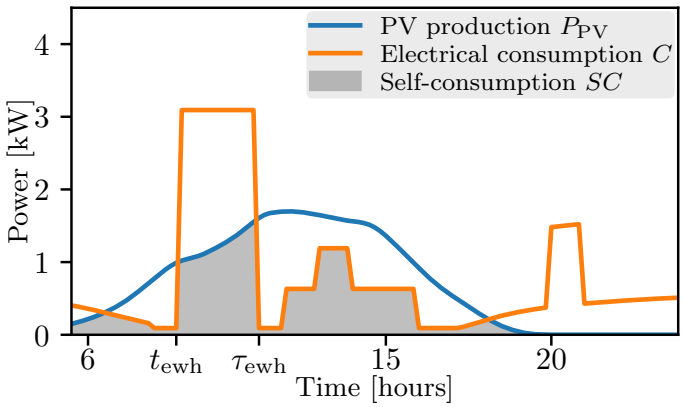

Figure 1.2: The self-consumption SC corresponds to the power that is simultaneously produced and consumed locally.

This definition is rewritten by substracting the uncontrollable electric consumption (heating and cooling, lights, white goods) from the PV production, and considering only the positive part of the remaining PV production. This remaining PV production available for consumption is called the PV surplus, and noted $\widehat{P}_{\mathrm{PV}}$. Observing from Eqs. (1.2)-(1.4) that $P_{\mathrm{ewh}}(s,t)$ is null when $t \notin [t_{\mathrm{ewh}}, \tau_{\mathrm{ewh}}]$, the mathematical definition of the SC becomes the integral over a period $[t_{\mathrm{ewh}}, \tau_{\mathrm{ewh}}]$ of the minimum between the PV surplus and the controllable load:

$$SC(s, P_{\mathrm{PV}}) = \int_{t_{\mathrm{ewh}}}^{\tau_{\mathrm{ewh}}} \min(P_{\mathrm{ewh}}(s,t), \widehat{P}_{\mathrm{PV}}(t))dt + \text{Constant} \triangleq SC(s, \widehat{P}_{\mathrm{PV}}) + \text{Constant} \tag{1.6}$$

where $s = (t_{\text{ewh}}, \tau_{\text{ewh}}) \in \mathbb{R}^2$ is the decision variable and $\widehat{P}_{\text{PV}}(t) \geq 0, \forall t$, is the positive part of the surplus of local production, over which the controllable load can be optimally placed.

For the sake of simplicity, the notation SC will be used in the remainder of the manuscript to refer to the controllable part of Eq. (1.6) (illustrated in Figure 1.3).

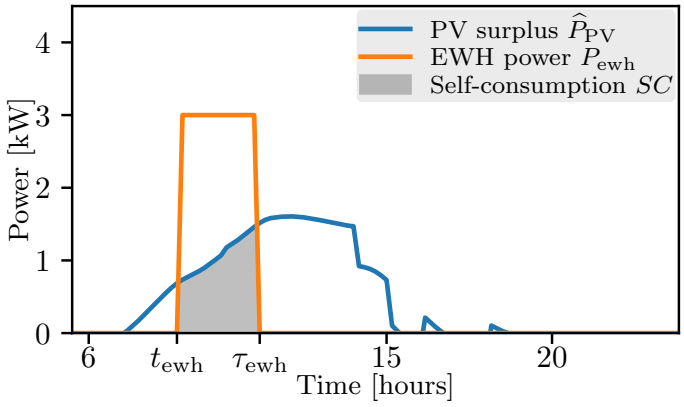

Figure 1.3: Optimizing SC, which is the controllable part of the self-consumption, corresponds to maximizing the size of the gray area.

Constraints

The optimal control problem considers the following boundary and control constraints:

$$E(t_0) = E_0 \tag{1.7}$$

$$E(t_{\text{f}}) \in [E_{\text{clear}}, E_{\text{sat}}] \tag{1.8}$$

where t_0 is a generic starting time at which the energy level is known.

Eq. (1.8) is a comfort constraint: the control shall guarantee that the EWH reaches an energy interval at a final time t_{f}. Here, this energy interval is chosen to be centered around the energy level corresponding to a hot-water tank and the final time is the beginning of the evening. Enforcing this constraint ensures that the tank has stored enough energy to cover the hot-water consumption likely to happen at the end of the day.

Reducing the number of variables

Due to the hysteresis function described in Eq. (1.3) and Assumption 1, once the energy reaches the interval $[E_{\text{clear}}, E_{\text{sat}}]$, it remains in it as long as heating is allowed by Eq. (1.4). Hence, a simplifying choice to ensure that constraint (1.8) is met is

$$\tau_{\text{ewh}} = +\infty \tag{1.9}$$

even though the control strategies developed will only be observed on the interval $[t_{\text{ewh}}, t_{\text{f}}]$. Hence, τ_{ewh} is no longer a decision variable, and this makes it possible to redefine s as the starting time alone:

$$s = t_{\text{ewh}} \tag{1.10}$$

Furthermore, taking Eq. (1.7) into account, one obtains the existence of an upper bound t_{lim} for t_{ewh}. According to (1.1)–(1.4), it is defined as the unique solution of the equation

$$E_{\text{clear}} = \mathrm{e}^{-\mathrm{k}(t_f - t_0)}\, E_0 + \int_{t_{\text{lim}}}^{t_f} \mathrm{e}^{-\mathrm{k}(t_f - t)}\, \overline{P}_{\text{ewh}} dt - \int_{t_0}^{t_f} \mathrm{e}^{-\mathrm{k}(t_f - t)}\, Q(t) dt \tag{1.11}$$

which is

$$t_{\text{lim}} = t_{\text{f}} + \frac{1}{\text{k}} \log\left(1 + \frac{\text{k}}{\overline{P}_{\text{ewh}}}\left(e^{-\text{k}(t_f - t_0)}E_0 - E_{\text{clear}} - \int_{t_0}^{t_f} \text{e}^{-\text{k}(t_f - t)}\, Q(t)dt\right)\right) \tag{1.12}$$

that is, the latest acceptable heating starting time to reach E_{clear} at t_{f} and thus satisfy constraint (1.8). Hence constraint (1.8) can be replaced by the two simple equations $\tau_{\text{ewh}} = +\infty$ and $t_{\text{ewh}} \leq t_{\text{lim}}$.

Figure 1.4 gives an example of the evolution of the EWH energy and its oscillation inside the interval $[E_{\text{clear}}, E_{\text{sat}}]$ for one day and a specific starting time.

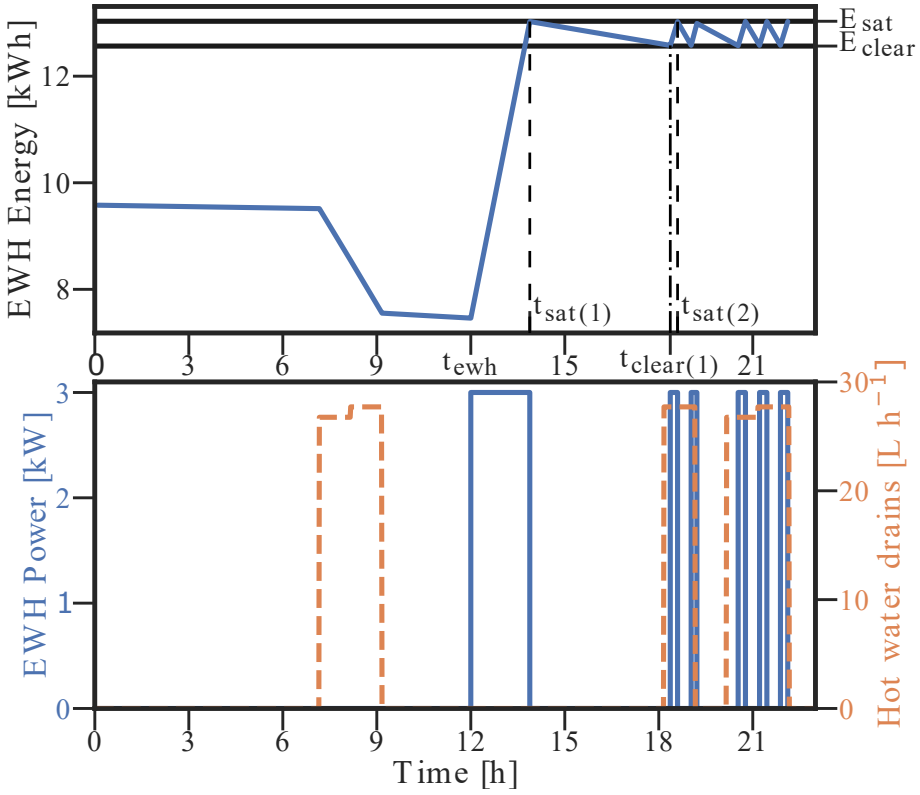

Figure 1.4: Hot-water drains (lower figure, dashed curve) decrease the EWH energy (top figure). EWH heating intervals $[t_{\text{clear}(n)}, t_{\text{sat}(n+1)}]$ (lower figure, plain curve), calculated from a chosen t_{ewh}, increase the EWH energy and keep it oscillating between the deadband $[E_{\text{clear}}, E_{\text{sat}}]$.

This behavior defines (i) saturation times, when the EWH internal thermostat detects a temperature reaching the allowed upper limit (the saturation temperature) and halts heating, and (ii) clearance times, when thermal losses or hot-water consumption have led the internal temperature to drop below the lower hysteresis value (the clearance temperature), thus making heating resume. These times are respectively denoted by $t_{\text{sat}(n+1)}$ and $t_{\text{clear}(n)}$, $n \in \mathbb{N}$. They are defined mathematically as $t_{\text{clear}(0)} = t_{\text{ewh}}$ and as the unique

solutions on $[t_0, +\infty)$ (see Proposition 1 in the sequel) to

$$\begin{aligned} E_{\text{sat}} &= \mathrm{e}^{-\mathrm{k}(t_{\text{sat}(1)}-t_0)}\, E_0 + \int_{t_0}^{t_{\text{sat}(1)}} \mathrm{e}^{-\mathrm{k}(t_{\text{sat}(1)}-t)}[P_{\text{ewh}}(t_{\text{ewh}}, t) - Q(t)]dt \\ &= \mathrm{e}^{-\mathrm{k}(t_{\text{sat}(1)}-t_0)}\, E_0 - \int_{t_0}^{t_{\text{sat}(1)}} \mathrm{e}^{-\mathrm{k}(t_{\text{sat}(1)}-t)}\, Q(t)dt + \int_{t_{\text{ewh}}}^{t_{\text{sat}(1)}} \mathrm{e}^{-\mathrm{k}(t_{\text{sat}(1)}-t)}\, \overline{P}_{\text{ewh}} dt \\ &= \mathrm{e}^{-\mathrm{k}(t_{\text{sat}(1)}-t_0)}\, E_0 - \int_{t_0}^{t_{\text{sat}(1)}} \mathrm{e}^{-\mathrm{k}(t_{\text{sat}(1)}-t)}\, Q(t)dt + \frac{\overline{P}_{\text{ewh}}}{\mathrm{k}} \left(1 - \mathrm{e}^{-\mathrm{k}(t_{\text{sat}(1)}-t_{\text{ewh}})}\right) \end{aligned} \tag{1.13}$$

for $n = 0$, and for $n \geq 1$,

$$\begin{aligned} E_{\text{sat}} &= \mathrm{e}^{-\mathrm{k}(t_{\text{sat}(n+1)}-t_{\text{clear}(n)})}\, E_{\text{clear}} + \int_{t_{\text{clear}(n)}}^{t_{\text{sat}(n+1)}} \mathrm{e}^{-\mathrm{k}(t_{\text{sat}(n+1)}-t)}[\overline{P}_{\text{ewh}} - Q(t)]dt \\ &= \mathrm{e}^{-\mathrm{k}(t_{\text{sat}(n+1)}-t_{\text{clear}(n)})} \left[E_{\text{clear}} - \frac{\overline{P}_{\text{ewh}}}{\mathrm{k}}\right] - \int_{t_{\text{clear}(n)}}^{t_{\text{sat}(n+1)}} \mathrm{e}^{-\mathrm{k}(t_{\text{sat}(n+1)}-t)}\, Q(t)dt + \frac{\overline{P}_{\text{ewh}}}{\mathrm{k}} \end{aligned} \tag{1.14}$$

and

$$E_{\text{clear}} = \mathrm{e}^{-\mathrm{k}(t_{\text{clear}(n)}-t_{\text{sat}(n)})}\, E_{\text{sat}} - \int_{t_{\text{sat}(n)}}^{t_{\text{clear}(n)}} \mathrm{e}^{-\mathrm{k}(t_{\text{clear}(n)}-t)}\, Q(t)dt \tag{1.15}$$

In the following, N denotes the index of the last clearance time $t_{\text{clear}(n)}$ before t_{f}, that is, such that $t_{\text{clear}(N)} \leq t_{\text{f}} < t_{\text{clear}(N+1)}$.

Therefore, the EWH consumption curve can be represented by a sum of non-overlapping boxcar functions corresponding to the intervals where $\delta_{\text{sat}}(t) = 0$: the EWH is On between $[t_{\text{ewh}}, t_{\text{sat}(1)}]$ and between each $[t_{\text{clear}(n)}, t_{\text{sat}(n+1)}]$ pair, with the last interval potentially shortened to $[t_{\text{clear}(N)}, t_{\text{f}}]$ if $t_{\text{f}} < t_{\text{sat}(N+1)}$.

Note that the hot-water drains depicted by orange dashed lines in the lower part of Figure 1.4 have a visible impact on the internal EWH energy (upper part).

Problem statement

From the previous considerations, the problem of optimizing self-consumption can be summarized as:

Problem 1.1. Given a PV production surplus curve $\widehat{P}_{\text{PV}}$, a power rating $\overline{P}_{\text{ewh}}$, an initial energy level E_0, saturation and clearance energy levels E_{sat} and E_{clear}, a hot-water consumption curve Q, an initial time t_0 and a final time t_{f}, solve for each day

$$\begin{aligned} &\max_{t_{\text{ewh}}} \int_{t_{\text{ewh}}}^{t_{\text{f}}} \min(P_{\text{ewh}}(t_{\text{ewh}}, t), \widehat{P}_{\text{PV}}(t))dt \\ &\text{s.t. } t_{\text{ewh}} \leq t_{\text{lim}} \end{aligned} \tag{1.16}$$

where t_{lim} is defined in (1.12) and P_{ewh} is defined as the sum of boxcar functions

$$P_{\text{ewh}}(t_{\text{ewh}}, t) = \overline{P}_{\text{ewh}} \sum_{n=0}^{N} \mathbb{1}_{[t_{\text{clear}(n)}, \min(t_{\text{sat}(n+1)}, t_{\text{f}})]}(t) \tag{1.17}$$

where $t_{\text{clear}(0)} = t_{\text{ewh}}$, $t_{\text{clear}(n)}$ and $t_{\text{sat}(n+1)}$ depend on t_{ewh} and are defined in (1.14)–(1.15) along with N.

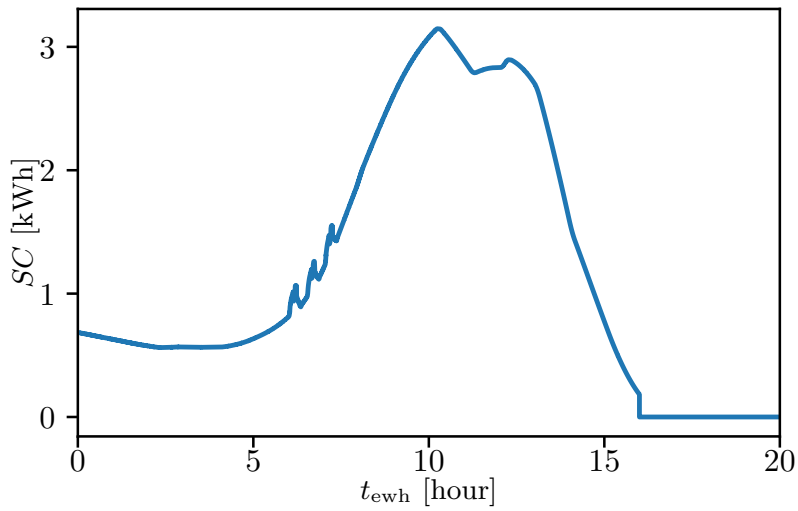

Figure 1.5: The objective function is not concave.

1.2.4 Optimization algorithm

In all generality, Problem 1.1 is non-concave, as multiple local maxima can be observed in the example shown in Figure 1.5. To solve it, this study proposes to identify the critical and derivative discontinuity points of its objective function.

Smoothness analysis of the objective function

To provide a suitable optimization solution, the nature of the objective function is analyzed.

Proposition 1. *Assume that* $\widehat{P}_{\text{PV}}$ *and* Q *are continuous positive-valued functions such that Assumption 1 holds. Then, the solutions* $t_{\text{sat}(n)}$ *and* $t_{\text{clear}(n)}$ *to* (1.13)–(1.15) *are unique and the corresponding functions*

$$\varphi_n : t_{\text{ewh}} \in [t_0, +\infty) \mapsto t_{\text{sat}(n)} \ (n \geq 1) \quad \text{and} \quad \psi_n : t_{\text{ewh}} \in [t_0, +\infty) \mapsto t_{\text{clear}(n)} \ (n \in \mathbb{N}) \tag{1.18}$$

are differentiable, increasing and bijective between $\mathbb{R}$ *and* $\mathbb{R}$.

Besides, SC defined in (1.6) *is almost everywhere continuously differentiable with respect to* t_{ewh}*, and its derivative is given, for* $N \in \mathbb{N}^*$*, by*

$$\frac{dSC(t_{\text{ewh}}, \widehat{P}_{\text{PV}})}{dt_{\text{ewh}}} = \tag{1.19}$$

$$\begin{cases} \displaystyle\sum_{n=0}^{N-1} \left[\frac{dt_{\text{sat}(n+1)}}{dt_{\text{ewh}}} \min(\overline{P}_{\text{ewh}}, \widehat{P}_{\text{PV}}(t_{\text{sat}(n+1)})) - \frac{dt_{\text{clear}(n)}}{dt_{\text{ewh}}} \min(\overline{P}_{\text{ewh}}, \widehat{P}_{\text{PV}}(t_{\text{clear}(n)})) \right] \\ \qquad - \dfrac{dt_{\text{clear}(N)}}{dt_{\text{ewh}}} \min(\overline{P}_{\text{ewh}}, \widehat{P}_{\text{PV}}(t_{\text{clear}(N)})), \text{ if } \quad t_{\text{ewh}} \in (\psi_N^{-1}(t_{\text{f}}), \varphi_{N+1}^{-1}(t_{\text{f}})) \\ \displaystyle\sum_{n=0}^{N} \left[\frac{dt_{\text{sat}(n+1)}}{dt_{\text{ewh}}} \min(\overline{P}_{\text{ewh}}, \widehat{P}_{\text{PV}}(t_{\text{sat}(n+1)})) - \frac{dt_{\text{clear}(n)}}{dt_{\text{ewh}}} \min(\overline{P}_{\text{ewh}}, \widehat{P}_{\text{PV}}(t_{\text{clear}(n)})) \right], \\ \qquad\qquad \text{if } \quad t_{\text{ewh}} \in (\varphi_{N+1}^{-1}(t_{\text{f}}), \psi_{N+1}^{-1}(t_{\text{f}})) \end{cases}$$

in which

$$\frac{dt_{\text{sat(n)}}}{dt_{\text{ewh}}} = \begin{cases} \dfrac{\overline{P}_{\text{ewh}}\, \mathrm{e}^{-\,\mathrm{k}(t_{\text{sat(n)}} - t_{\text{ewh}})}}{\overline{P}_{\text{ewh}} - \mathrm{k}\, E_{\text{sat}} - Q(t_{\text{sat(n)}})}\,, & n = 1 \\ \dfrac{\left[\mathrm{k}\, E_{\text{clear}} - \overline{P}_{\text{ewh}} + Q(t_{\text{clear}(n-1)})\right] \mathrm{e}^{-\,\mathrm{k}(t_{\text{sat}(n)} - t_{\text{clear}(n-1)})}}{\mathrm{k}\, E_{\text{sat}} - \overline{P}_{\text{ewh}} + Q(t_{\text{sat}(n)})} \cdot \dfrac{dt_{\text{clear}(n-1)}}{dt_{\text{ewh}}}\,, & n \geq 2 \end{cases} \tag{1.20}$$

$$\frac{dt_{\text{clear}(n)}}{dt_{\text{ewh}}} = \begin{cases} 1\,, & n = 0 \\ \dfrac{\left[\mathrm{k}\, E_{\text{sat}} + Q(t_{\text{sat}(n)})\right] \mathrm{e}^{-\,\mathrm{k}(t_{\text{clear}(n)} - t_{\text{sat}(n)})}}{kE_{\text{clear}} + Q(t_{\text{clear}(n)})} \cdot \dfrac{dt_{\text{sat}(n)}}{dt_{\text{ewh}}}\,, & n \geq 1 \end{cases} \tag{1.21}$$

Furthermore, the number of discontinuity points of the derivative of SC on $[t_0, t_f]$ *is finite.*

Notice that $\varphi_{N+1}^{-1}(t_{\text{f}})$ is the value of t_{ewh} such that $t_{\text{sat}(N+1)} = t_{\text{f}}$ and that $\psi_{N+1}^{-1}(t_{\text{f}}))$ is the value of t_{ewh} such that $t_{\text{clear}(N+1)} = t_{\text{f}}$.

Hence, Proposition 1 identifies two sources of discontinuities:

1. when $t_{\text{clear}(N)}$ becomes equal to t_{f}, prompting a switch from final index N to $N+1$;
2. when $t_{\text{sat}(N)}$ becomes equal to t_{f}, prompting a switch in the result of $\min(t_{\text{f}}, t_{\text{sat}(N)})$.

Appendix A explores another source of discontinuity, in the case of two controllable appliances. The number of additional discontinuities in that case will depend on the shape of the appliances load curves, and impacts the scalability of the optimal control research method followed in the present chapter.

Proof. The differential form of (1.1) is

$$\frac{dE}{dt} = -kE + P_{\text{ewh}}(t_{\text{ewh}}, t) - Q(t) \tag{1.22}$$

Due to the EWH power actuation law (1.2)–(1.4) and (1.9), $P_{\text{ewh}}(t_{\text{ewh}}, t) = 0$ if $t \in [t_0, t_{\text{ewh}})$ and therefore, from (1.22), $E(t_{\text{ewh}}) \leq E_0 < E_{\text{sat}}$. The first saturation time is defined as $t_{\text{sat}(1)} = \min\{t \mid t \geq t_{\text{ewh}} \text{ and } E(t) = E_{\text{sat}}\}$ if this set is non-empty and $t_{\text{sat}(1)} = +\infty$ otherwise. By definition of $t_{\text{sat}(1)}$, $E(t) < E_{\text{sat}}$ for $t \in [t_{\text{ewh}}, t_{\text{sat}(1)})$ and thus, from the EWH power actuation law (1.2)–(1.4) and (1.9), $P_{\text{ewh}}(t_{\text{ewh}}, t) = \overline{P}_{\text{ewh}}$ for $t \in [t_{\text{ewh}}, t_{\text{sat}(1)})$. Consequently, from Assumption 1 and using Gronwall's lemma on the differential inequalities resulting from Eq. (1.22),

$$E(t_{\text{ewh}}) + \eta(t - t_{\text{ewh}}) \leq E(t) \leq E(t_{\text{ewh}}) + \overline{P}_{\text{ewh}}(t - t_{\text{ewh}}) \quad \text{for} \quad t \in [t_{\text{ewh}}, t_{\text{sat}(1)}) \tag{1.23}$$

This implies that $t_{\text{sat}(1)} < +\infty$ as $\eta > 0$ and that

$$t_{\text{sat}(1)} - t_{\text{ewh}} \in \left[\frac{E_{\text{sat}} - E(t_{\text{ewh}})}{\overline{P}_{\text{ewh}}}, \frac{E_{\text{sat}} - E(t_{\text{ewh}})}{\eta}\right] \tag{1.24}$$

Hence the duration between t_{ewh} and $t_{\text{sat}(1)}$ is strictly upper- and lower-bounded. Furthermore, this also implies that $t_{\text{sat}(1)}$ is equivalently defined as the solution to (1.13), which is thus unique.

The first clearance time after t_{ewh} is defined as

$$t_{\text{clear}(1)} = \min\left\{t \mid t \geq t_{\text{sat}(1)} \text{ and } E(t) = E_{\text{clear}}\right\} \tag{1.25}$$

if this set is non-empty and $t_{\text{clear}(1)} = +\infty$ otherwise. From the EWH power actuation law (1.2)–(1.4) and (1.9), $P_{\text{ewh}}(t_{\text{ewh}}, t) = 0$ for $t \in [t_{\text{sat}(1)}, t_{\text{clear}(1)})$, and from (1.22) and using Gronwall's Lemma,

$$E_{\text{sat}} + \left(\mathrm{k}\, E_{\text{sat}} + \overline{Q}\right)(t_{\text{sat}(1)} - t) \leq E(t) \leq E_{\text{sat}} + \mathrm{k}\, E_{\text{clear}}(t_{\text{sat}(1)} - t) \text{ for } t \in [t_{\text{sat}(1)}, t_{\text{clear}(1)}) \tag{1.26}$$

This implies that $t_{\text{clear}(1)} < +\infty$, and that

$$t_{\text{clear}(1)} - t_{\text{sat}(1)} \in \left[\frac{E_{\text{sat}} - E_{\text{clear}}}{\mathrm{k}\, E_{\text{sat}} + \overline{Q}}, \frac{E_{\text{sat}} - E_{\text{clear}}}{\mathrm{k}\, E_{\text{clear}}}\right] \tag{1.27}$$

Hence the duration between $t_{\text{sat}(1)}$ and $t_{\text{clear}(1)}$ is non-null.

Similarly, one can define by induction $t_{\text{sat}(n)}$ and $t_{\text{clear}(n)}$ for $n \geq 2$ and conclude that these instants are finite and uniquely defined as the solutions to (1.14) and (1.15) respectively. In addition, one also gets by induction

$$t_{\text{sat}(n+1)} - t_{\text{clear}(n)} \in \left[\frac{E_{\text{sat}} - E_{\text{clear}}}{\overline{P}_{\text{ewh}}}, \frac{E_{\text{sat}} - E_{\text{clear}}}{\eta}\right], \quad n \in \mathbb{N} \tag{1.28}$$

$$t_{\text{clear}(n)} - t_{\text{sat}(n)} \in \left[\frac{E_{\text{sat}} - E_{\text{clear}}}{\mathrm{k}\, E_{\text{sat}} + \overline{Q}}, \frac{E_{\text{sat}} - E_{\text{clear}}}{\mathrm{k}\, E_{\text{clear}}}\right], \quad n \in \mathbb{N} \tag{1.29}$$

.

Now, one can define

$$\begin{aligned} f_1\left(t_{\text{ewh}}, t_{\text{sat}(1)}\right) \triangleq\ & \mathrm{e}^{-\mathrm{k}(t_{\text{sat}(1)} - t_0)}\, E_0 - \int_{t_0}^{t_{\text{sat}(1)}} \mathrm{e}^{-\mathrm{k}(t_{\text{sat}(1)} - t)}\, Q(t)dt \\ & + \frac{\overline{P}_{\text{ewh}}}{\mathrm{k}}\left(1 - \mathrm{e}^{-\mathrm{k}(t_{\text{sat}(1)} - t_{\text{ewh}})}\right) - E_{\text{sat}} \end{aligned} \tag{1.30}$$

$$\begin{aligned} f_2\left(t_{\text{clear}(n)}, t_{\text{sat}(n+1)}\right) \triangleq\ & \mathrm{e}^{-\mathrm{k}(t_{\text{sat}(n+1)} - t_{\text{clear}(n)})}\, E_{\text{clear}} - \int_{t_{\text{clear}(n)}}^{t_{\text{sat}(n+1)}} \mathrm{e}^{-\mathrm{k}(t_{\text{sat}(n+1)} - t)}\, Q(t)dt \\ & + \frac{\overline{P}_{\text{ewh}}}{\mathrm{k}}\left(1 - \mathrm{e}^{-\mathrm{k}(t_{\text{sat}(\mathrm{n}+1)} - t_{\text{clear}(\mathrm{n})})}\right) - E_{\text{sat}} \end{aligned} \tag{1.31}$$

$$h\left(t_{\text{sat}(n)}, t_{\text{clear}(n)}\right) \triangleq \mathrm{e}^{-\mathrm{k}(t_{\text{clear}(n)} - t_{\text{sat}(n)})}\, E_{\text{sat}} - \int_{t_{\text{sat}(n)}}^{t_{\text{clear}(n)}} \mathrm{e}^{-\mathrm{k}(t_{\text{clear}(n)} - t)}\, Q(t)dt - E_{\text{clear}} \tag{1.32}$$

to rewrite (1.13)–(1.15) as

$$f_1\left(t_{\text{ewh}}, t_{\text{sat}(1)}\right) = 0 \tag{1.33}$$

$$f_2\left(t_{\text{clear}(n)}, t_{\text{sat}(n+1)}\right) = 0 \tag{1.34}$$

$$h\left(t_{\text{sat}(n)}, t_{\text{clear}(n)}\right) = 0 \tag{1.35}$$

with f_1, f_2 and h continuously differentiable functions of their arguments as Q is continuous. These equations imply that an implicit relation exists between the two variables of each of these three equations, but the explicit expression of one variable as a function of the other has to be proved. Applying Leibniz integral rules for differentiation under the integral sign, straightforward computations yield

$$\begin{aligned} \frac{\partial f_1}{\partial t_{\text{sat}(1)}} =\ & -\mathrm{k}\, E_0\, \mathrm{e}^{-\mathrm{k}(t_{\text{sat}(1)} - t_0)} - Q(t_{\text{sat}(1)}) + \mathrm{k} \int_{t_0}^{t_{\text{sat}(1)}} \mathrm{e}^{-\mathrm{k}(t_{\text{sat}(1)} - t)}\, Q(t)dt \\ & + \overline{P}_{\text{ewh}}\, \mathrm{e}^{-\mathrm{k}(t_{\text{sat}(1)} - t_{\text{ewh}})} \\ =\ & \overline{P}_{\text{ewh}} - \mathrm{k}\, E_{\text{sat}} - Q(t_{\text{sat}(1)}) \end{aligned} \tag{1.36}$$

in which the last equality follows from (1.33). Similarly, one obtains

$$\begin{aligned}\frac{\partial f_2}{\partial t_{\mathrm{sat(n+1)}}} &= -\,\mathrm{k}\,E_{\mathrm{clear}}\,\mathrm{e}^{-\,\mathrm{k}(t_{\mathrm{sat}(n+1)}-t_{\mathrm{clear}(n)})} - \left[Q(t_{\mathrm{sat}(n+1)}) - \mathrm{k}\int_{t_{\mathrm{clear}(n)}}^{t_{\mathrm{sat}(n+1)}} \mathrm{e}^{-\,\mathrm{k}(t_{\mathrm{sat}(n+1)}-t)}\,Q(t)\right] \\ &\quad + \overline{P}_{\mathrm{ewh}}\,\mathrm{e}^{-\,\mathrm{k}(t_{\mathrm{sat}(n+1)}-t_{\mathrm{clear}(n)})} \\ &= -\,\mathrm{k}\,E_{\mathrm{sat}} + \overline{P}_{\mathrm{ewh}} - Q(t_{\mathrm{sat}(n)}) \end{aligned} \tag{1.37}$$

and

$$\begin{aligned}\frac{\partial h}{\partial t_{\mathrm{clear(n)}}} &= -\,\mathrm{k}\,E_{\mathrm{sat}}\,\mathrm{e}^{-\,\mathrm{k}(t_{\mathrm{clear}(n)}-t_{\mathrm{sat}(n)})} - \left[Q(t_{\mathrm{clear}(n)}) - \mathrm{k}\int_{t_{\mathrm{sat}(n)}}^{t_{\mathrm{clear}(n)}} \mathrm{e}^{-\,\mathrm{k}(t_{\mathrm{clear}(n)}-t)}\,Q(t)\right] \\ &= -\,\mathrm{k}\,E_{\mathrm{clear}} - Q(t_{\mathrm{clear}(n)}) \end{aligned} \tag{1.38}$$

From Assumption 1, the right-hand side expressions in (1.36)-(1.38) are positive. Consequently, using the implicit function theorem, there exists a continuously differentiable function φ_1 such that $t_{\mathrm{sat}(1)} = \varphi_1(t_{\mathrm{ewh}})$, at least locally. Using again the implicit function theorem, and a direct induction argument, the existence of ψ_n $(n \geq 1)$ and φ_n $(n \geq 2)$ as defined in (1.18) follows. The implicit function theorem also implies that

$$\varphi_1'(t_{\mathrm{ewh}}) = -\left(\frac{\partial f_1}{\partial t_{\mathrm{sat}(1)}}(t_{\mathrm{ewh}}, \varphi_1(t_{\mathrm{ewh}}))\right)^{-1} \frac{\partial f_1}{\partial t_{\mathrm{ewh}}}(t_{\mathrm{ewh}}, \varphi_1(t_{\mathrm{ewh}})) \tag{1.39}$$

in which, taking a partial derivative of (1.33),

$$\frac{\partial f_1}{\partial t_{\mathrm{ewh}}}(t_{\mathrm{ewh}}, \varphi_1(t_{\mathrm{ewh}})) = -\overline{P}_{\mathrm{ewh}}\,\mathrm{e}^{-\,\mathrm{k}(t_{\mathrm{sat}(1)}-t_{\mathrm{ewh}})} \tag{1.40}$$

Gathering (1.39)–(1.40) and (1.36) gives the first expression in (1.20). Similarly, the implicit function theorem implies

$$\varphi_{n+1}'(t_{\mathrm{ewh}}) = -\left(\frac{\partial f_2}{\partial t_{\mathrm{sat}(n+1)}}\right)^{-1} \frac{\partial f_2}{\partial t_{\mathrm{clear}(n)}}\,\psi_n'(t_{\mathrm{ewh}}) \tag{1.41}$$

in which, taking a partial derivative of (1.34),

$$\begin{aligned}\frac{\partial f_2}{\partial t_{\mathrm{clear}(n)}} &= \mathrm{k}\,E_{\mathrm{clear}}\,\mathrm{e}^{-\,\mathrm{k}(t_{\mathrm{sat}(n+1)}-t_{\mathrm{clear}(n)})} + \mathrm{e}^{-\,\mathrm{k}(t_{\mathrm{sat}(n+1)}-t_{\mathrm{clear}(n)})}\,Q(t_{\mathrm{clear}(n)}) \\ &\quad - \overline{P}_{\mathrm{ewh}}\,\mathrm{e}^{-\,\mathrm{k}(t_{\mathrm{sat}(n+1)}-t_{\mathrm{clear}(n)})} \end{aligned} \tag{1.42}$$

and

$$\psi_n'(t_{\mathrm{ewh}}) = -\left(\frac{\partial h}{\partial t_{\mathrm{clear}(n)}}\right)^{-1} \frac{\partial h}{\partial t_{\mathrm{sat}(n)}}\,\varphi_n'(t_{\mathrm{ewh}}) \tag{1.43}$$

in which, taking a partial derivative of (1.35),

$$\frac{\partial h}{\partial t_{\mathrm{sat}(n)}} = \mathrm{k}\,E_{\mathrm{sat}}\,\mathrm{e}^{-\,\mathrm{k}(t_{\mathrm{clear}(n)}-t_{\mathrm{sat}(n)})} + \mathrm{e}^{-\,\mathrm{k}(t_{\mathrm{clear}(n)}-t_{\mathrm{sat}(n)})}\,Q(t_{\mathrm{sat}(n)}) \tag{1.44}$$

The second expression in (1.20) follows from (1.41)–(1.42) and (1.37), while (1.21) follows from (1.43)–(1.44) and (1.38).

One can observe that the right-hand side of (1.20)–(1.21) is positive, which implies that $t_{\mathrm{sat}(n)}$ and $t_{\mathrm{clear}(n)}$ are increasing functions of t_{ewh}. As $t_{\mathrm{clear}(0)} = t_{\mathrm{ewh}}$, one concludes from the lower bounds in (1.28)–(1.29) that $t_{\mathrm{sat}(n)}$ and $t_{\mathrm{clear}(n)}$ are bijections.

Finally, from (1.17), note that the integrand in the objective function SC (1.6) is non-null only during the intervals $[t_{\text{clear(n)}}, t_{\text{sat}(n+1)}]$. For $t_{\text{ewh}} \in \left[\psi_N^{-1}(t_f), \psi_{N+1}^{-1}(t_f)\right)$, the objective function therefore rewrites

$$\begin{aligned}
SC(t_{\text{ewh}}, \widehat{P}_{\text{PV}}) = &\int_{t_{\text{ewh}}}^{t_{\text{sat}(1)}} \min(\overline{P}_{\text{ewh}}, \widehat{P}_{\text{PV}}(t))dt \\
&+ \sum_{n=1}^{N} \left(\underbrace{\int_{t_{\text{sat(n)}}}^{t_{\text{clear(n)}}} \min(0, \widehat{P}_{\text{PV}}(t))dt}_{=0} + \int_{t_{\text{clear(n)}}}^{\min(t_{\text{sat}(n+1)}, t_{\text{f}})} \min(\overline{P}_{\text{ewh}}, \widehat{P}_{\text{PV}}(t))dt \right) \\
&= \sum_{n=0}^{N} \int_{t_{\text{clear(n)}}}^{\min(t_{\text{sat}(n+1)}, t_{\text{f}})} \min(\overline{P}_{\text{ewh}}, \widehat{P}_{\text{PV}}(t))dt \\
&= \sum_{n=0}^{N-1} \int_{t_{\text{clear(n)}}}^{t_{\text{sat(n+1)}}} \min(\overline{P}_{\text{ewh}}, \widehat{P}_{\text{PV}}(t))dt \qquad (1.45) \\
&+ \begin{cases} \displaystyle\int_{t_{\text{clear(N)}}}^{t_{\text{f}}} \min(\overline{P}_{\text{ewh}}, \widehat{P}_{\text{PV}}(t))dt\,, & \text{if} \quad t_{\text{ewh}} \in (\psi_N^{-1}(t_{\text{f}}), \varphi_{N+1}^{-1}(t_{\text{f}})) \\ \displaystyle\int_{t_{\text{clear(N)}}}^{t_{\text{sat(N+1)}}} \min(\overline{P}_{\text{ewh}}, \widehat{P}_{\text{PV}}(t))dt\,, & \text{if} \quad t_{\text{ewh}} \in (\varphi_{N+1}^{-1}(t_{\text{f}}), \psi_{N+1}^{-1}(t_{\text{f}})) \end{cases}
\end{aligned}$$

Hence, Leibniz integral rules for differentiation under the integral sign give (1.19). Observe that the discontinuity points are in finite number on $[t_0, t_f]$ due to the lower bound for the steps defined in (1.28)–(1.29). This concludes the proof. □

Calculations

The specific times $t_{\text{sat}(n+1)}$ and $t_{\text{clear}(n)}$, $n \in \mathbb{N}$, as well as their derivatives, must be computed because they are involved in the expression of the derivative of the objective function in Eq. (1.19), but cannot be explicitly determined from their definitions (1.13)–(1.15). Besides, the derivatives of these times, $\frac{dt_{\text{sat}(n+1)}}{dt_{\text{ewh}}}$ and $\frac{dt_{\text{clear}(n)}}{dt_{\text{ewh}}}$, also appear in (1.19) and involve $t_{\text{sat}(n+1)}$ and $t_{\text{clear}(n)}$. The following discussion will clarify how these terms are evaluated numerically.

Description of the calculation of the t_{sat} times For the EWH saturation times $t_{\text{sat}(n+1)}$, a fixed-point algorithm is used. Indeed, one cannot deduce an explicit expression for $t_{\text{sat}(1)}$ from (1.13) due to the presence of the integral term corresponding to the hot-water drains. However, one can rewrite (1.13) as

$$E_{\text{sat}} = \mathrm{e}^{-\,\mathrm{k}(t_{\text{sat}(1)} - t_0)}\, E_0 - \int_{t_0}^{t_{\text{sat}(1)}} \mathrm{e}^{-\,\mathrm{k}(t_{\text{sat}(1)} - t)}\, Q(t)dt + \frac{\overline{P}_{\text{ewh}}}{\mathrm{k}} \left(1 - \mathrm{e}^{-\,\mathrm{k}(t_{\text{sat}(1)} - t_{\text{ewh}})}\right) \quad (1.46)$$

and transform it into

$$\begin{aligned}
\log\left(\overline{P}_{\text{ewh}} - \mathrm{k}\, E_{\text{sat}} - \mathrm{k} \int_{t_0}^{t_{\text{sat}(1)}} \mathrm{e}^{-\,\mathrm{k}(t_{\text{sat}(1)} - t)}\, Q(t)dt\right) = &\log\left(\overline{P}_{\text{ewh}}\, \mathrm{e}^{kt_{\text{ewh}}} - \mathrm{k}\, \mathrm{e}^{kt_0}\, E_0\right) \\
&+ \log\left(\mathrm{e}^{-kt_{\text{sat}(1)}}\right)
\end{aligned} \quad (1.47)$$

which leads to

$$t_{\text{sat}(1)} = g_1\left(t_{\text{ewh}}, t_{\text{sat}(1)}\right) \quad (1.48)$$

with

$$g_1\left(t_{\text{ewh}}, t_{\text{sat}(1)}\right) \triangleq \frac{1}{\text{k}}\left[\log\left(\overline{P}_{\text{ewh}}\,\text{e}^{\text{k}\,t_{\text{ewh}}} - \text{k}\,E_0\,\text{e}^{\text{k}\,t_0}\right)\right. \\ \left. - \log\left(\overline{P}_{\text{ewh}} - \text{k}\,E_{\text{sat}} - \text{k}\int_{t_0}^{t_{\text{sat}(1)}} \text{e}^{-\,\text{k}(t_{\text{sat}(1)}-t)}\,Q(t)dt\right)\right] \tag{1.49}$$

which can be shown to be a contractive mapping with respect to $t_{\text{sat}(1)}$ for the parameter values under consideration in this study (see Appendix B for more details). Hence, fixed-point iterations were used to estimate $t_{\text{sat}(1)}$ numerically, with successive estimates of $t_{\text{sat}(1)}$ computed through $g_1(t_{\text{ewh}}, t_{\text{sat}(1)})$ and progressively reaching this fixed point. Usually, fewer than one dozen iterations are needed to reach an accurate value.

Similarly, Eq. (1.14) can be rewritten for $n \geq 1$ as

$$E_{\text{sat}} = E_{\text{clear}}\,\text{e}^{-\,\text{k}(t_{\text{sat}(n+1)}-t_{\text{clear}(n)})} \\ + \frac{\overline{P}_{\text{ewh}}}{\text{k}}\left(1 - \text{e}^{-\,\text{k}(t_{\text{sat}(n+1)}-t_{\text{clear}(n)})}\right) - \int_{t_{\text{clear}}(n)}^{t_{\text{sat}(n+1)}} \text{e}^{-\,\text{k}(t_{\text{sat}(n+1)}-t)}\,Q(t)dt \tag{1.50}$$

which can be reorganized as

$$t_{\text{sat}(n+1)} = t_{\text{clear}(n)} + \frac{1}{\text{k}}\left[\log(\overline{P}_{\text{ewh}} - \text{k}\,E_{\text{clear}})\right. \\ \left. - \log\left(\overline{P}_{\text{ewh}} - \text{k}\,E_{\text{sat}} - \text{k}\int_{t_{\text{clear}(n)}}^{t_{\text{sat}(n+1)}} \text{e}^{-\,\text{k}(t_{\text{sat}(n+1)}-t)}\,Q(t)dt\right)\right] \\ t_{\text{sat}(n+1)} \triangleq g_2\left(t_{\text{clear}(n)}, t_{\text{sat}(n+1)}\right) \tag{1.51}$$

in which g_2 is a contractive mapping (see Appendix B for more details). $t_{\text{sat}(n+1)}$ can thus be calculated with fixed-point iterations as well. Note that g_2 requires a prior calculation of $t_{\text{clear}(n)}$, which will now be described in detail.

Description of the calculation of the t_{clear} times The $t_{\text{clear}(n)}$ times satisfy a similar fixed-point equation; however, the resulting mapping is not contractive. Hence, a dichotomy procedure was followed instead to evaluate $t_{\text{clear}(n)}$. Indeed, it can be proven that the energy obtained from Eq. (1.15) between a saturation time $t_{\text{sat}(n)}$ and a clearance time $t_{\text{clear}(n)}$, $n \in \mathbb{N}^*$ is a decreasing function of $t_{\text{clear}(n)}$, ensuring the convergence of the method.

Hence, in a nutshell, the proposed procedure computes alternatively $t_{\text{sat}(n+1)}$ through a fixed-point procedure and $t_{\text{clear}(n)}$ with a dichotomy algorithm and repeats this process iteratively until t_{f} is reached.

Algorithm description

Because the problem is non-concave and not everywhere differentiable, it is necessary to assess the nature of every stationary point and every point of discontinuity of the derivative. Away from discontinuities, the stationary condition can be written as

$$\frac{dSC(t_{\text{ewh}}, \widehat{P}_{\text{PV}})}{dt_{\text{ewh}}} = 0 \tag{1.52}$$

and the discontinuity points belong to

$$S = \left(\left\{\psi_n^{-1}(t_f)\,,\ n \in \mathbb{N}\right\} \cup \left\{\varphi_n^{-1}(t_f)\,,\ n \geq 1\right\}\right) \cap [t_0, t_f] \tag{1.53}$$

in which the functions ψ_n, φ_n are defined in Proposition 1.

Corresponding SC values are then evaluated and compared to solve the optimization problem. Typically, for the examples treated in this thesis, fewer than ten extrema need to be considered.

A key advantage of this approach is that evaluating the stationary condition (1.52) through Eq. (1.19) is much less computationally intensive than exhaustively evaluating the integral in Eq. (1.16).

Algorithm The solutions to the stationary condition being exhaustively determined, along with the discontinuity points of the derivatives, the corresponding values of the objective function are then compared to solve Problem 1.1. This procedure is summarized in Algorithm 1.1. Note that this algorithm is of course intended to be used over a finite time grid. Specifically the *for* loop and the final *argmax* identification consider a finite number of elements, and a solution of Eq. (1.52) or an element of S is identified in the *if* condition when the objective function derivative (1.19) goes from a positive to a non-positive value between two consecutive timesteps. This algorithm can now be applied in a high-fidelity environment.

Algorithm 1.1: Calculate $t^\star_{\text{ewh}}$ solution of Problem 1.1

Input: $\widehat{P}_{\text{PV}}, \overline{P}_{\text{ewh}}, t_0, E_0, t_{\text{f}}, E_{\text{f}}, E_{\text{sat}}, E_{\text{clear}}, Q, I = [0, t_{\text{lim}}]$
for *each timestep* $t \in I$ **do**
  Compute times = [every $t_{\text{clear}(n)}$ and $t_{\text{sat}(n+1)}$ up to t_{f}].
  //necessary to compute Eqs. (1.52) and (1.53)
  if t *is solution of Eq.* (1.52) *or belongs to* S *defined in* (1.53), **then**
    Compute $SC(t, \widehat{P}_{\text{PV}})$
    Update Buffer = [Buffer, $SC(t, \widehat{P}_{\text{PV}})$]
  end
end
Identify $t^\star_{\text{ewh}} = \underset{t}{\text{argmax}}\,\text{Buffer}$

Figure 1.6 displays an example of the result of Algorithm 1.1, with the EWH heating period chosen so that the SC will be maximized.

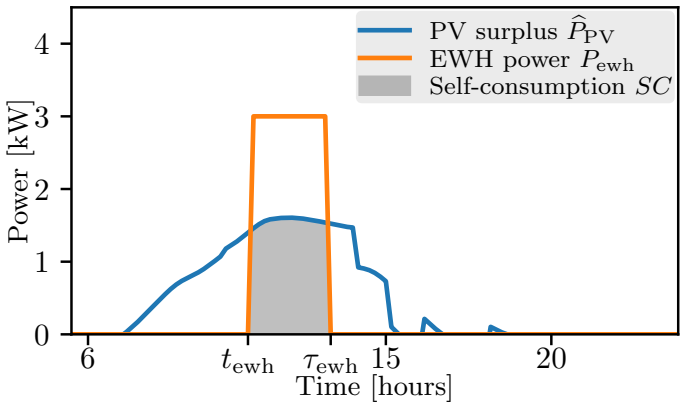

Figure 1.6: The optimized placement of the EWH heating period maximizes the gray area, i.e. the objective SC.

1.3 Numerical experiments

1.3.1 Test setting and hardware specifications

Consider two crystalline silicon PV arrays of 1.5 kW_p[1] each, both inclined at 15° relative to the horizontal and of respective azimuth 15° and 105° relative to the South. An EWH of volume $V = 200\,L$ and power rating $\overline{P}_{ewh} = 3\,kW$ is considered, to cover the consumption of 2 inhabitants. The scenario of outdoor temperature, input cold water, and solar irradiance correspond to a house located in Fontainebleau, in the French region of Ile-de-France. The total yearly PV production amounts to 2.9 MWh, and the total yearly PV surplus available for the EWH consumption is 1.8 MWh. As previously mentioned in Section 1.2.1, a unique deterministic PV production forecast is considered, with perfect prediction: the same load curve is used both for the prediction used by the control strategies to schedule the EWH, and for the *a posteriori* performance evaluation of the chosen schedule on the volume of self-consumed energy. Figure 1.7 displays the total PV production curve and the corresponding PV surplus curve, after substraction of the uncontrolled electric consumption, for a few days.

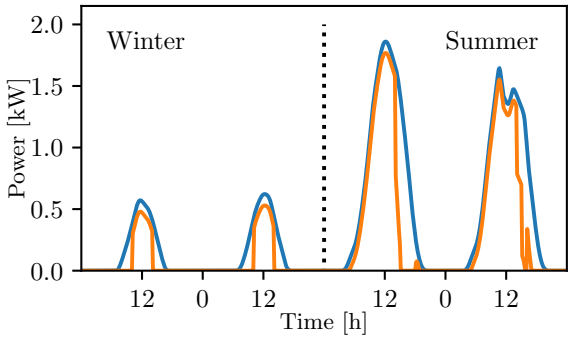

Figure 1.7: Total PV production (blue) and PV surplus (orange) for 2 winter days and 2 summer days.

The hot-water drains scenario is perfectly known and is repeated each day with the same pattern reported in Figure 1.8, with drains in the morning (from 7 a.m. to 9 a.m.) and in the evening (from 6 p.m. to 7 p.m. and from 8 p.m. to 10 p.m), ranging between 26.8 $L\,h^{-1}$ and 27.7 $L\,h^{-1}$, for a total consumption of 956 L per week at 60 °C. Only slight quantity variations occur seasonally, with the same pattern tuned down in summer. The final time t_f is set as 6 p.m., in order to satisfy the main hot-water drains occurring at the end of the day.

Table 1.1: Hardware specifications

PV capacity	3 kW_p
EWH Power rating	3 kW
EWH volume	200 L
Weekly hot-water consumption at 60 °C	956 L

The proposed method is validated through comparisons with two other EWH control strategies:

[1] kW_p stands for kW "peak": it is the maximum electrical power that can be supplied by a PV system, under standardized test conditions of irradiance, temperature and solar spectrum.

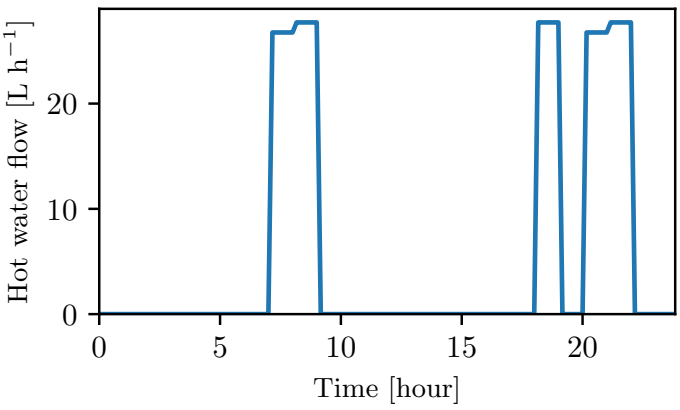

Figure 1.8: Hot-water drains for a winter day.

- a passive heuristic strategy, triggering the heating authorization consistently during [1 a.m.-7 a.m.] and [12 p.m.-2 p.m.] in order to follow the off-peak hours of electricity price;
- an industrial reference heuristic control.

Both methods take advantage of the off-peak hours for the electricity price, between [1 a.m.-7 a.m.] and [12 p.m.-2 p.m.], in order to limit the household electricity bill.

This industrial reference control was developed for commercial use in residential housing equipped with PV arrays with the double aim of maximizing the consumption of local PV production and ensuring a reduced electricity bill. The off-peak hours of electricity price can be put to advantage for the latter objective. It follows priority rules to choose each day whether or not to start the heating authorization, according to the following principles:

- determine when to start the EWH according to a PV surplus threshold set up in advance;
- figure out whether heating the tank without PV production is necessary, according to whether or not the daily heating duration target has been met.

Note that the reference control does not require the hot-water drains sequence to perform an estimation of the next heating duration target, whereas the method proposed here requires exact knowledge of it.

The numerical experiments required the controller presented in this chapter to communicate with a simulation model. It is implemented using Python 3.7.3. The high-fidelity simulation model is implemented using the Dymola 2018 software (based on the Modelica 3.2.2 object-oriented modeling language, see [Wiström, 2013]). It includes a model for the EWH, and readings of simulated datasets for the house uncontrolled electric loads and for the PV power production. The EWH simulation model is more complex and more accurate than the model described in Section 1.2.2. It models the stratification phenomenon, with three layers of water constituting the whole volume for improved representativeness. The thermal energy dissipated by the resistive element, as well as the incoming cold water, are situated in the lower layer of the tank, whereas the hot-water outlet is situated in the upper layer. As described in [Beeker-Adda, 2016], at rest, the layers are mixed only by heat diffusion. The effects of this phenomenon are relatively slow compared to the forced convection and mixing induced by draining. Finally, heat losses at the walls of the tank are more precisely accounted for in this simulation model than in the control model.

1.3.2 Numerical results

Simulations consist of a daily optimization over one year representative of the typical meteorological conditions of the location, starting January 1st. Algorithm 1.1 is used with I restricted to the interval of start times such that the heating authorization would cover the valuable times of high PV production: night hours are not considered as potential start times. The initial and target energy levels E_0 and E_f for the proposed method correspond to the energy levels reached at these times by the EWH under the reference control. Results are extracted with a 30-minute timestep.

A major advantage of the proposed method is its limited computational load: it achieves optimally scheduling these 365 load curves in an average of 40 seconds with Python 3.7.3 and a Core i3 2.4 GHz processor, with 8 Go RAM. The corresponding time of less than 0.11 seconds per curve is short enough to consider repeating the procedure several times for various consumption and production scenarios.

For comparison, the Nelder–Mead optimization method, a derivative-free algorithm based only on objective function evaluations finds the optimal starting time for a single PV production curve in an average of 0.65 seconds. This relatively high computational time could in all likelihood be related to the fixed-point iterations described above and necessary to compute P_{ewh} in Eq. (1.17). Furthermore, this method sometimes converges to a local maximum instead of the global maximum, with a strong dependency on the chosen initialization time. For instance, for the example pictured in Figure 1.5, when initialized at $t = 3$ h, the Nelder–Mead algorithm converges to the local maximum (2.9 h, 0.6 kWh), and initialized between 12 h and 14 h, it converges to the local maximum (1.2 h, 2.9 kWh), which can be observed in Figure 1.5.

Self-consumed energy volumes

The evolution of the cumulative daily SC (cumulative sum starting from January 1st) is displayed in Figure 1.9. The cumulated SC values reached at the end of the year are listed in Table 1.2. These values correspond to the amount of PV surplus energy used for the controllable EWH consumptions.

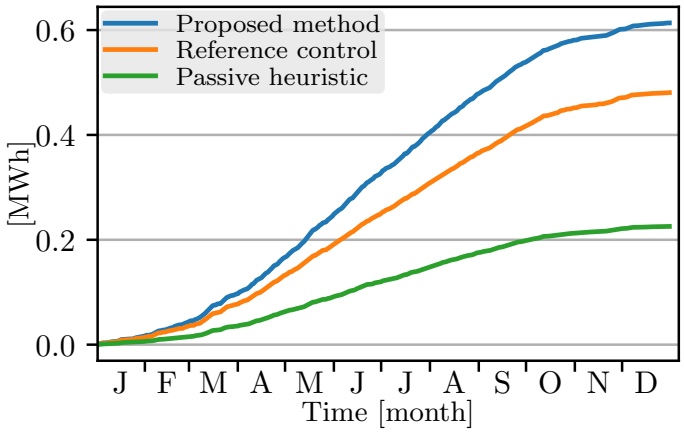

Figure 1.9: The cumulative self-consumed energy varies largely with the chosen strategy.

Table 1.2: Final annual controllable SC volumes for each strategy

Proposed method	**Reference control**	**Passive heuristic**
0.61 MWh	0.48 MWh	0.23 MWh

The final values show that, overall, the proposed method outperforms significantly the reference control, and that both are largely more efficient than the passive strategy (see Table 1.2).

A year-long analysis of the daily SC confirms the interest of the proposed method, both during winter and summer days. This method outputs a higher SC 311 days of the year and a lower SC only for 7 days. The 47 other days results in the same SC value. Over a whole year, the proposed method yields a 28 % increase of the SC compared to the reference control.

Self-consumption rate

An additional perspective is provided by the self-consumption rate, a measure commonly referred to for self-consumption installations. The self-consumption rate relative to the controllable consumptions is the ratio

$$SC_{\%} = \frac{\int_0^\tau \min(P_{\text{ewh}}, \widehat{P}_{\text{PV}})ds}{\int_0^\tau \widehat{P}_{\text{PV}}(s)ds} = \frac{\int_0^\tau \min(P_{\text{ewh}}, \widehat{P}_{\text{PV}})ds}{\widehat{E}_{\text{PV}}} \tag{1.54}$$

$$\sum_{h=1}^{l}\sum_{g=1}^{l} y_{h,g} = \sum_{g=1}^{l}\sum_{h=1}^{l} y_{h,g}$$

where τ can vary from 0 to the end of the year, the numerator is the self-consumed energy considering only the controllable consumptions and the PV surplus, and the denominator is the local PV surplus production over the period.

If time τ is equal to 24 hours, $SC_{\%}$ corresponds to the self-consumption rate relative to the controllable consumptions for the first day, January 1$^{\text{st}}$. If τ is equal to 365 days, $SC_{\%}$ corresponds to this rate calculated for the whole year.

The evolution of the cumulative self-consumption rate relative to the controllable consumptions for each strategy over a one-year simulation is shown in Figure 1.10. The values reached at the end of the year are listed in Table 1.3.

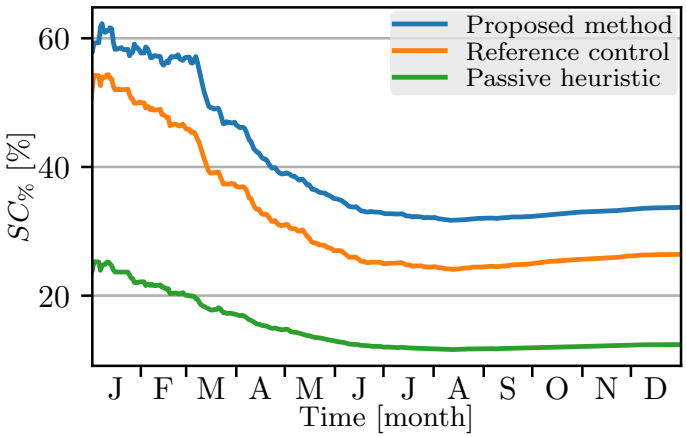

Figure 1.10: The final self-consumption rate varies largely with the chosen strategy.

Table 1.3: Final annual controllable $SC_{\%}$ values for each strategy

Proposed method	**Reference control**	**Passive heuristic**
33.7 %	26.4 %	12.4 %

This proportional perspective confirms the superiority of the proposed algorithm over the reference.

Seasonal analysis

The analysis of Figure 1.9 and 1.10 is now carried on through seasonal observations.

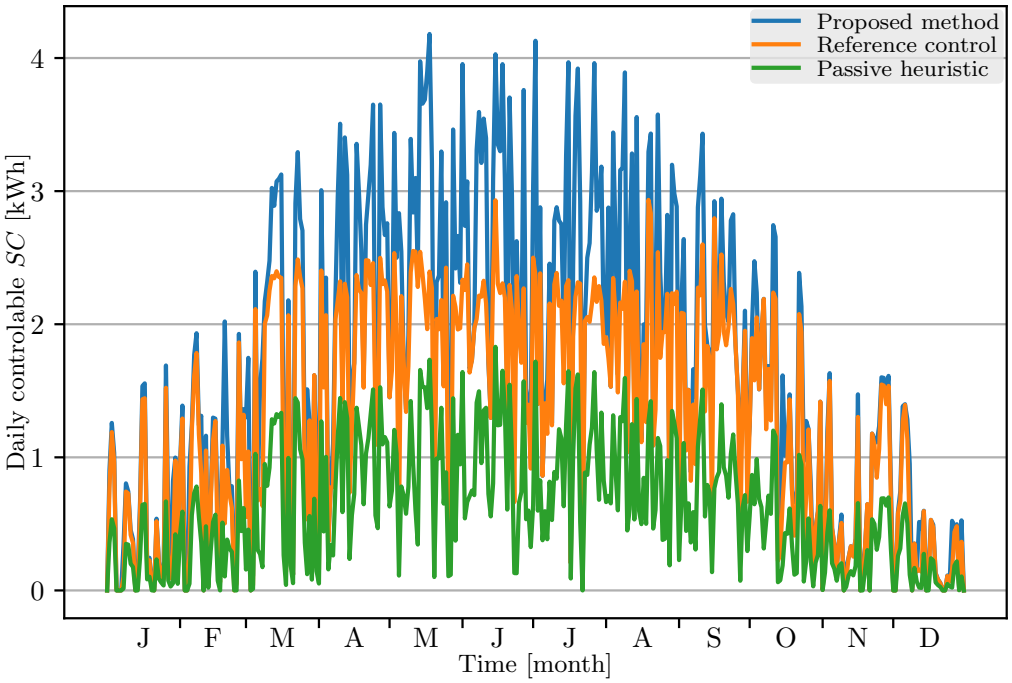

Figure 1.11: The differences in daily self-consumption for each strategy is exacerbated during sunny summer days.

Figure 1.11 highlights the ability of the proposed algorithm to take advantage of large volumes of available PV surpluses. These high stakes days lead to the final difference between each approach witnessed in Figure 1.9 and Table 1.2.

Figure 1.12 displays an example of the EWH consumption curves resulting from the three strategies for a summer day PV surplus. For this particular day, SC for the proposed method is 3.4 kWh, exceeding the 2.3 kWh of the reference control.

Neither the reference control nor the passive heuristic exactly follow their heating command (dashed lines), because the EWH internal thermostat stops the heating when the water temperature reaches the higher limit. This higher limit can be reached early in the day, as displayed in Figure 1.12, because a first heating period usually occurs at night when electricity prices are low, in addition to reduced hot-water consumptions.

The high PV production level triggers the reference control heating authorization early, when the PV surplus is not at its maximum. Thus the simple and efficient working principle of the reference control, based on a PV surplus threshold, prevents it from harnessing the best part of the renewable resource.

During winter days, PV surplus is way more limited. The small extra amounts of PV energy consumed by the EWH controlled by the proposed method over the reference control, visible in Figure 1.11 lead to a massive difference in the daily self-consumption rate $SC_{\%}$. This is visible in Figure 1.13 showing the cumulated self-consumption rate for each strategy over the year.

Figure 1.14 displays the EWH consumption curves resulting from the three strategies and the available PV surplus on a winter day. It highlights the fact that the reference control will not try and consume the very low PV production of winter days and risk yielding an extra cost, but will rather activate according to the off-peak hours of the electricity price. These different approaches result in a typically better SC with the proposed method.

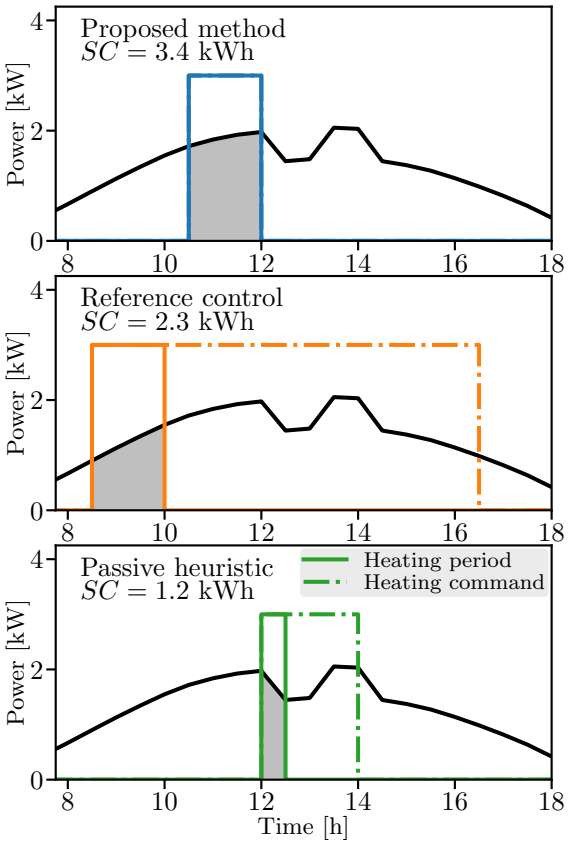

Figure 1.12: The heating signal of the proposed method is the only one capable of waiting and exploiting the summer day PV production (black curves), leading to self-consumption volumes (gray patch) varying from simple to almost triple.

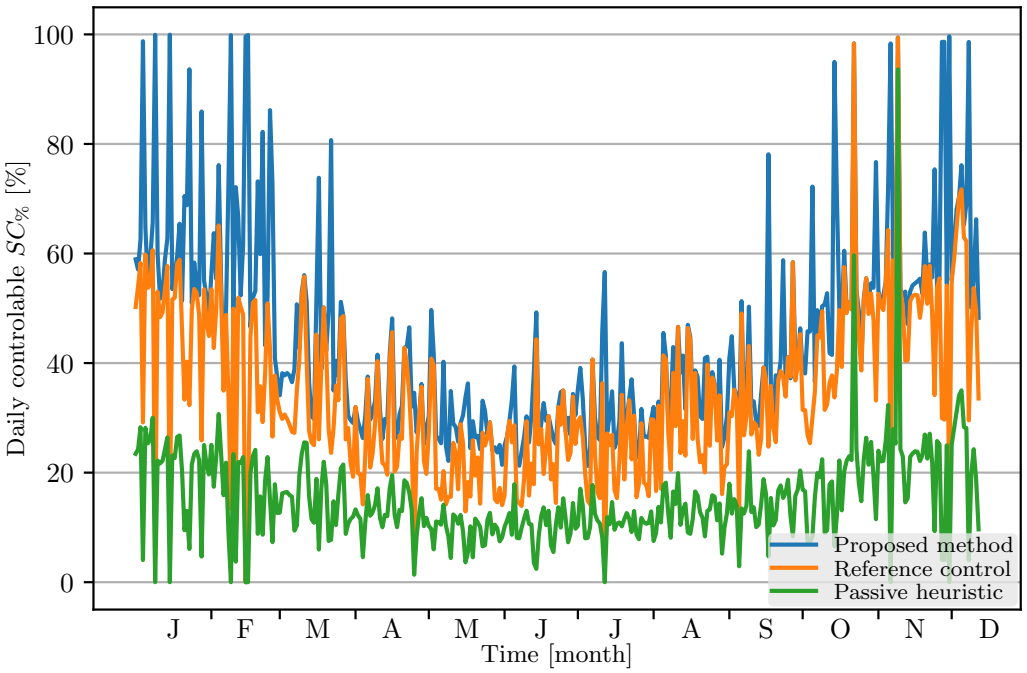

Figure 1.13: Tapping the scarce PV surplus volumes of winter days leads to high daily self-consumption rates for the proposed method.

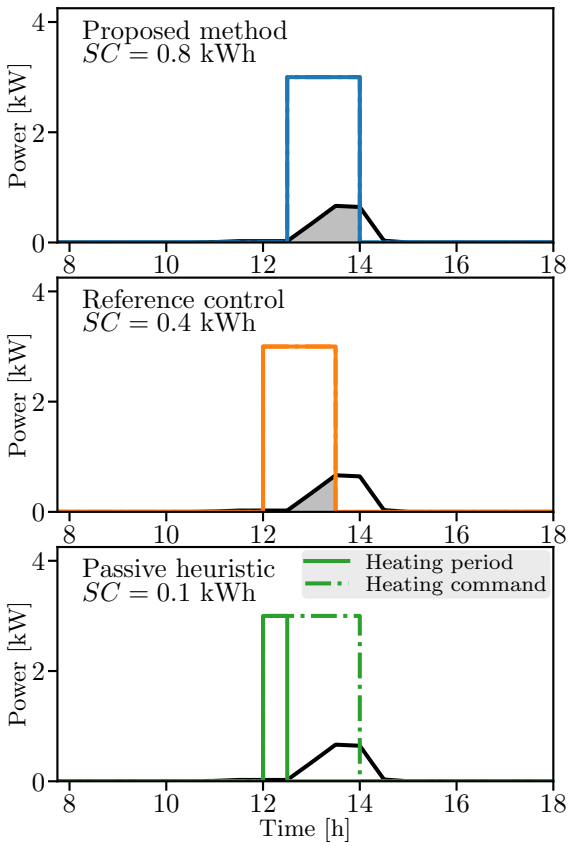

Figure 1.14: The proposed method is the only one correctly detecting the small amounts of PV production surplus (black curves) in the winter, leading to higher self-consumed energy volumes (gray patch).

Whole household self-consumption metrics

As the metrics of self-consumption installations are usually defined at the household level, comparing them for each approach at the household level (i.e. including controllable and non-controllable consumptions) offers another perspective on the comparative impact of each algorithm. Table 1.4 lists the amount of total PV produced energy consumed locally by all the household consumptions, over a whole year. The increase offered by the proposed

Table 1.4: Final total annual self-consumed energy volumes for each strategy, including the non-controllable consumptions

Proposed method	**Reference control**	**Passive heuristic**
1.72 MWh	1.59 MWh	1.33 MWh

method over the reference control is now of only + 8.3 %. Figure 1.15 displays these values and shows the lower difference between each strategy.

Likewise, the differences of daily self-consumption rate at the household level from one approach to the other, visible in Figure 1.16, is largely diminished in comparison with the differences of Figure 1.13. This is especially true for winter days where the extra PV production harnessed by the proposed method amount to a tiny fraction of the total PV produced energy. As the small amounts of PV production occurring during these winter days can be completely used for self-consumption, a few measure points are close to a daily SC rate of 100 %.

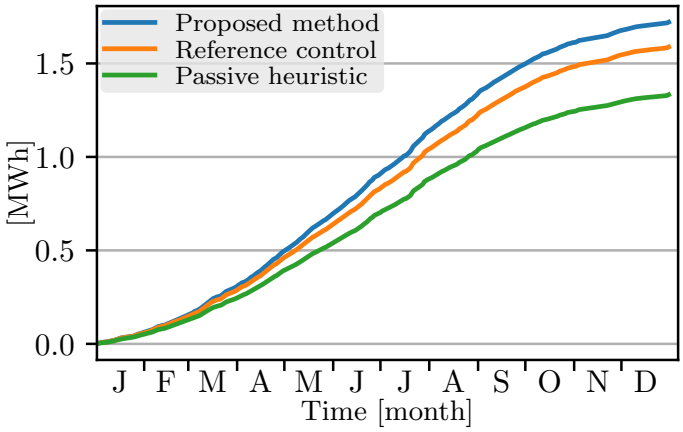

Figure 1.15: Considering the total production and consumption, the final difference in cumulative self-consumed energy between each strategy is lower than when considering only the controllable consumptions (see Figure 1.9).

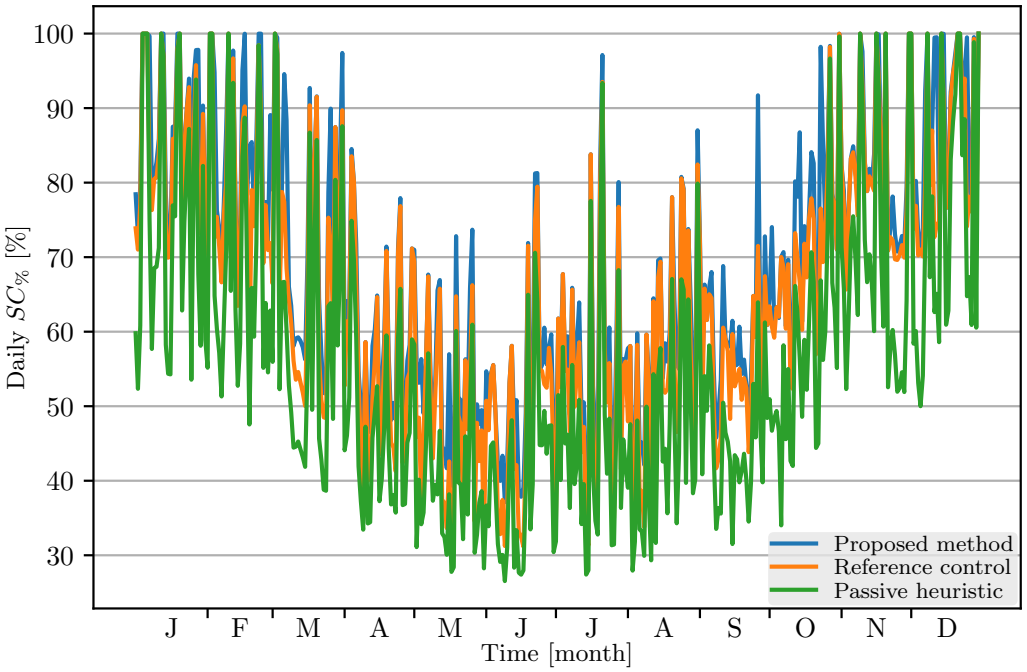

Figure 1.16: The daily total self-consumption rate can be of 100% for the three strategies during winter days, as the total volume of PV production is limited.

The total self-consumption rates evolve accordingly, as Figure 1.17 and Table 1.5 show it.

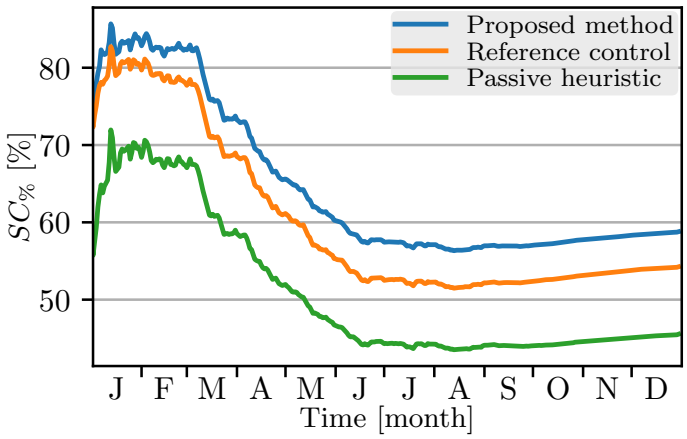

Figure 1.17: Considering the total production and consumption, the final difference in cumulative self-consumption rate between each strategy is lower than when considering only the controllable consumptions (see Figure 1.10).

Table 1.5: Final annual total $SC_{\%}$ values for each strategy

Proposed method	**Reference control**	**Passive heuristic**
58.7 %	54.3 %	45.4 %

1.4 Conclusion and perspectives

In this chapter a methodology has been proposed to optimize the starting time of an EWH in order to maximize the self-consumption rate of a residential PV installation. This method is optimal, simple and has a very low computational cost. Numerical experiments proved the relevance of the method, achieving an improvement of +28 % in yearly self-consumed energy over an industrial energy management solution.

The proposed method could benefit from some functional enhancements: first, implementing an estimation feature for future hot-water consumption, to overcome the unrealistic assumption of perfect knowledge on this matter; then, shifting to the minimization of the total household electricity bill.

Another type of improvement would be to include other controllable home appliances. However, scaling the proposed method to the control of several appliances could be difficult, as Appendix A showed that the overlapping of several load curves can lead to discontinuities in the objective function, and thus greatly increase the number of optimality candidate points and computation time to identify the optimal schedule. For this reason, the optimal control of several appliances will be considered in the second half of this manuscript using Mixed Integer Linear Programming (MILP) formulations (starting Chapter 3).

But first, the lightweight side of the algorithm can be put to advantage in quickly examining several scenarios and thus studying the impact of consumption or production profiles uncertainties. This application of the proposed method is now explored in Chapter 2.

Chapter 2

Impact of production forecast uncertainties

Impact des incertitudes de prévision de production. Ce chapitre présente une méthode de génération d'ensembles de scénarios de production PV autour d'une prévision initiale déterministe, considérée comme la plus probable. Avec ces ensembles de scénarios de production PV et l'algorithme d'optimisation présenté dans le chapitre précédent, il est possible d'étudier la performance du pilotage choisi du chauffe-eau Joule d'un bâtiment résidentiel équipé de panneaux PV, sur un large nombre de scénarios de production. Cette méthode d'évaluation des performances avec prise en compte de l'incertitude de production est plus réaliste que l'évaluation où la production PV est supposée connue d'avance, comme au chapitre précédent. Une étude numérique sur une année indique que le gain de performance apporté par l'algorithme proposé par rapport à l'heuristique de référence est plus important que ce qu'apporterait une méthode de prévision exacte de la production PV utilisée avec l'heuristique de référence.

Remark This chapter is partly extracted from the second half of the paper [Amabile et al., 2021].

2.1 Introduction

2.1.1 Production and consumption uncertainties

The gains obtained with the algorithm presented in Chapter 1 were evaluated in an hypothetical setting, where the photovoltaic (PV) production forecast was considered to materialize exactly.

However, the PV production cannot be predicted perfectly in advance. A more realistic assessment of *a priori* performances (i.e. at the time of decision-making, in anticipation) has to account for the prediction uncertainties.

Numerous self-consumption optimization studies have reported impressive figures. However, they suffer from requiring perfect knowledge of the inhabitants behavior [Sossan et al., 2013] and of the weather forecast (e.g. [Lefort et al., 2013], which uses the same production scenario for the forecast, for the decision algorithm and for performance evaluation of the chosen control), which appears to be an important factor. Indeed, renewable generation is subject to weather variability, leading to a limitation on production forecast precision [Inman et al., 2013]. The robust optimization paradigm [Paridari et al., 2016, Wytock et al., 2017, Pflaum et al., 2018] can be used to tackle these uncertainties,

but these approaches often lead to considering the least favorable scenarios and therefore proposing too-conservative strategies. For instance, [Pflaum et al., 2018] estimates a robust upper bound on the power required by an electric car-sharing charging station during a day, before dividing this available power between the electric vehicles through an ad-hoc heuristic. While the robust approach can satisfy the distribution system operator whose grid stability responsibility is facilitated, it leads to an overestimation of the energy needed by 34 %. Adjusting some of the parameters of the method can lower this overestimation to 23 % but not realistically any further.

Alternatively, Model Predictive Control (MPC) enables weather forecast updates at each prediction step, thus improving performance [Pflaum et al., 2014, Parisio et al., 2015, Wytock et al., 2017, Lefort et al., 2013, Sossan et al., 2013, Parisio et al., 2017]. This improvement is even more significant when stochastic MPC [Oldewurtel et al., 2012, Oldewurtel, 2011] or Stochastic optimal control [Pacaud, 2018] is used, but this comes at the price of potentially burdensome computation times.

2.1.2 Contribution of the chapter

Assessing the *a priori* performance of any energy management system (EMS) requires to test a chosen control schedule with a set of potential PV production scenarios, representative of the estimated uncertainty associated with the initial production prediction.

In order to assess the impact of 30-minute PV production forecast uncertainties on the optimization algorithms considered in Chapter 1, a specific PV production scenarios generation method is developed here.

In this chapter, the reference situation is a rule-based basic heuristic representative of the current state-of-the-art controller in commercial products (already presented in Section 1.3), which relies on a single deterministic PV production forecast, imperfect but considered as the most likely prediction. Considering this class of EMS as a baseline, this chapter compares the performance gains provided by the efficient optimization strategy presented in Chapter 1 over the commercial heuristic, and by a more accurate PV production forecast over the most likely prediction.

To conduct this investigation, an original method is used to generate a vast class of PV power production scenarios. This method is build on [Thorey et al., 2018], which provides a methodology to obtain the quantile forecasts of the PV load factor. The PV load factor is the power production of the PV installation normalized by its nominal installed power. Based on these quantile forecasts, this chapter presents a method to generate PV production scenarios that collectively satisfy a given probabilistic distribution while individually presenting realistic intra-day variability, that is, a realistic correlation factor.

Because the computing time of the algorithm proposed in Chapter 1 is quite low, various instances can be studied in a short time. Hence, the performance of a specific schedule could be quickly assessed as the empirical expectancy of the self-consumed energy reached by the schedule over each instance of the generated set of PV production scenarios. As the method presented in this chapter generates PV production scenarios that collectively satisfy a probabilistic distribution matching the level of uncertainty of the original deterministic forecast, the performance measured for a schedule this way is more realistic than when it is evaluated considering that the PV production prevision is exact.

It is shown in this chapter that, beyond a given level of forecast accuracy, the impact of knowing in advance the exact PV production forecast is negligible compared to the impact of the choice of the EWH control algorithm.

2.1.3 Chapter organization

Section 2.2 presents a methodology to generate a set of correlated realizations of a stochastic process, the PV production scenarios, which are representative of the estimated uncertainties of most likely PV production forecasts. Section 2.3 presents the numerical experiments conducted to validate the overall methodology and then compares the performance gains provided by the optimization strategy presented in Chapter 1 over the commercial heuristic, and by a more accurate PV production forecast over the most likely prediction. Section 2.4 reports conclusions and perspectives.

2.2 Generation of stochastic PV production scenarios

This chapter seeks to evaluate the impact of a discrepancy between the unique deterministic production forecast used in the EMS and the actual realized production. Only a few studies in the literature have numerically evaluated this kind of impact on their EMS, among them [Thomas et al., 2018] and [Rabiee et al., 2016]. However, in both cases, only a few scenarios were considered, which cannot account for potential intra-day variability.

The following discussion describes the methodology designed to generate a representative set of PV production scenarios.

2.2.1 Available data

Starting from a single deterministic forecast refered to as *most likely*, a set of PV production scenarios is needed to illustrate the uncertainties related to the initial deterministic forecast. The set as a whole shall represent the range of possible production outcomes proportionally to the observed variability of the most likely forecast, and each individual scenario shall present a realistic intra-day variability.

Associated with weather-based PV power production models, ensemble weather forecasts provided by meteorological services could serve this purpose [ECMWF, 2021]. Indeed, they give an indication of the range of possible future states of the atmosphere. However, only a limited number of possible outcomes are produced. In addition, their standard timestep is several hours, hence providing low resolution. Therefore, these data were discarded as a data source for the study at hand.

Instead, the uncertainties of deterministic PV load factor forecasts can be described by quantile forecasts. In [Thorey et al., 2018], quantile regressions are applied to deterministic PV production forecasts to build 21 quantile forecasts (for the quantiles 0.01, 0.99, and from 0.05 to 0.95 with a 0.05 increment). Each deterministic PV production forecast itself is the result of a statistical model fed with the following variables from a deterministic weather forecast: solar irradiance, total cloud cover and 2m temperature[1]. The PV load factor quantile forecasts are the input data to the presented method.

At each 30-minute timestep[2], for each day, following the methodology from [Thorey et al., 2018], an estimation of the uncertainties is available in the form of quantile values. Figure 2.1 reports these quantile curves, with one color for each quantile value. The darkest (and lowest) curve is the 0.01 quantile curve: for each timestep, there is a 1 % chance that the PV load factor will be below the value of this curve. The lightest (and highest) curve is the 0.99 quantile curve: there is a 99 % chance that the PV load factor will be below the value of this curve. The 19 curves in-between correspond to all the quantiles between 0.05 and 0.95, with a 0.05 increment. This quantile generation methodology is *a priori* not climate-dependent, but was tested exclusively for an oceanic temperate climate

[1] 2m temperature refers to air temperature at 2 meters above the surface

[2] Note that PV production forecasts with a 30-minute timestep are here considered to be possibly exact, thus knowingly disregarding faster variations in solar irradiance, for instance due to passing clouds.

(covering the vast majority of the French continental territory).

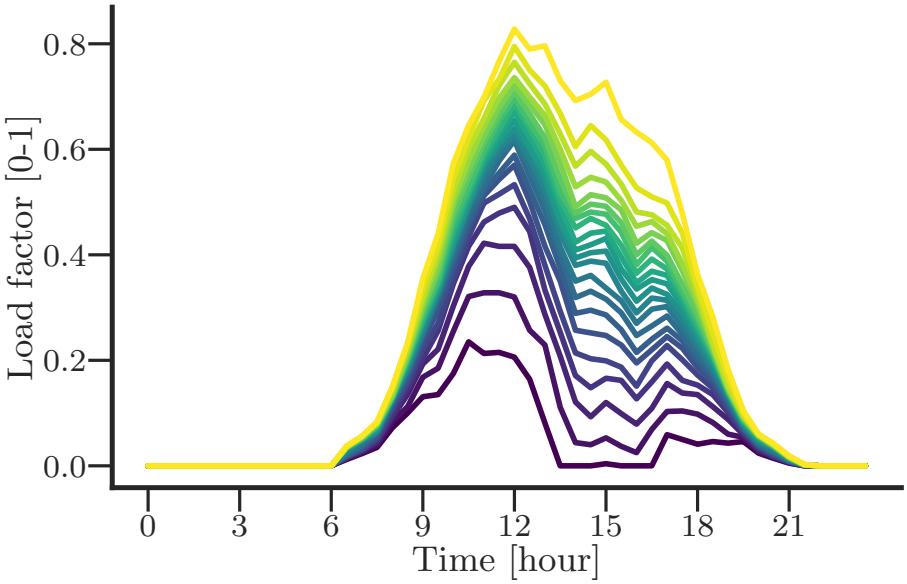

Figure 2.1: Quantile values at each timestep associated with the PV production forecast for one day in June with a cloudy afternoon (timezone: UTC+2). Data generated using the method of [Thorey et al., 2018]

.

Thanks to these quantile values, the cumulative distribution functions of the PV load factor are known for all the consecutive times in a discretized day[3]. With these cumulative distribution functions, the next step is to generate numerous PV load factor scenarios.

2.2.2 Methodology for generating correlated scenarios

Creating the daily scenarios by successively picking one of the values drawn for each timestep according to the cumulative distribution is not a suitable option. Indeed, the generated PV load factor scenarios might then be physically irrealistic, with important changes of values from one timestep to another. It is necessary to introduce some correlation of the PV load factor between two consecutive timesteps.

The cumulative distribution function obtained by the quantile regression inspired by [Thorey et al., 2018] has unfortunately no analytical form and does not correspond to any standard random process *a priori* (see Figure 2.2). Hence, it is necessary to design an *ad hoc* procedure to introduce correlation between two random variables $\mathbf{X_1}$ and $\mathbf{X_2}$ of fixed and known respective cumulative distributions F_1 and F_2.

The solution proposed here is based on the Probability Integral Transform (see e.g. [Angus, 1994]). This result states that if $\mathbf{X}$ is a continuous random variable with cumulative distribution function F_X, then the random variable $\mathbf{Y} = F_X(\mathbf{X})$ has a uniform distribution on $[0, 1]$.

[3]Note that the quantile values resulting from the quantile regression presented by [Thorey et al., 2018] can be greater than the maximum load factor at the given timestamp. It is thus necessary to saturate each cumulative distribution function obtained in proportionality with this daily maximum load factor. This upper limit is obtained assuming a completely clear sky at a given location. Reversely, without a mean to determine the lower bound of the load factor (through a completely cloud-covered sky, for instance), it is simply set to 0 without loss of generality.

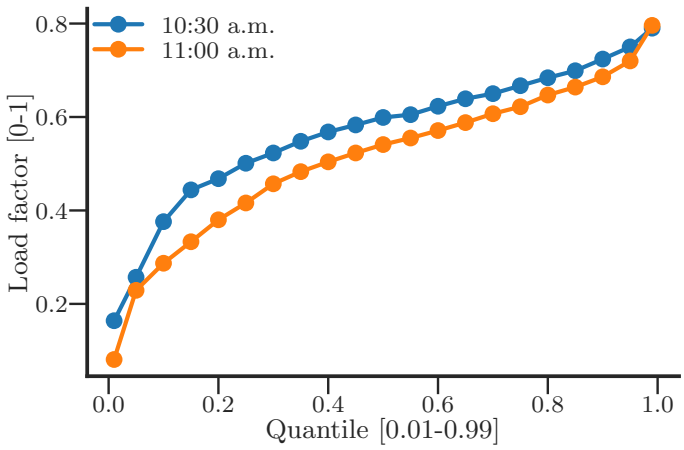

Figure 2.2: The cumulative distribution function for two consecutive timesteps for the same day in June do not follow any standard random process *a priori*. The differences between the consecutive curves demonstrate the changing weather conditions.

To address this correlation issue, consider two random variables $\mathbf{X_1}$ and $\mathbf{X_2}$, of respective cumulative distribution functions F_1 and F_2. Consider $\mathbf{U_1} = F_1(\mathbf{X_1})$ and $\mathbf{U_2}$, two uniformly distributed random variables on [0,1]. Another random variable is built with the following relation:

$$\mathbf{W} = (1-\alpha)\mathbf{U_1} + \alpha\mathbf{U_2} \tag{2.1}$$

with $\alpha \in]0,1[$. The probability density function f of $\mathbf{W}$ can be calculated as

$$f(x) = \begin{cases} \frac{x}{ab} & \text{if } 0 \leq x \leq a \\ \frac{1}{b} & \text{if } a \leq x \leq b \\ \frac{1-x}{ab} & \text{if } b \leq x \leq 1 \end{cases} \tag{2.2}$$

with $a = \min\{\alpha, 1-\alpha\}$ and $b = \max\{\alpha, 1-\alpha\}$. Integrating f yields the cumulative distribution function F of $\mathbf{W}$ as follows

$$F(x) = \begin{cases} \frac{x^2}{2ab} & \text{if } 0 \leq x \leq a \\ \frac{a}{2b} + \frac{x-a}{b} & \text{if } a \leq x \leq b \\ \frac{a}{2b} + \frac{b-a}{b} + \frac{x-b-\frac{x^2-b^2}{2}}{ab} & \text{if } b \leq x \leq 1 \end{cases} \tag{2.3}$$

$F(\mathbf{W})$ is then uniformly distributed according to the Probability Integral Transform. Besides, when the same result is applied again, $F_2^{-1}(F(\mathbf{W}))$ has the same probability density function as $\mathbf{X_2}$. Yet, interestingly, it is correlated with $\mathbf{X_1}$ through Eq. (2.1), which is the desired result.

In a nutshell, denoting by F_k the cumulative distribution function of the random variable $\mathbf{X_k}$, the stochastic process used to generate the scenarios is

$$\begin{cases} x_{k+1} & = F_{k+1}^{-1}(F((1-\alpha)F_k(x_k) + \alpha u_{k+1})), \text{ for } k > 0 \\ x_0 & = 0 \end{cases} \tag{2.4}$$

with $\alpha \in]0,1[$, F defined in Eq. (2.3) and u_{k+1} uniformly distributed over [0,1].

The degree of correlation, given by $\frac{\mathrm{Cov}(x_i,x_j)}{\sigma_{x_i}\sigma_{x_j}}$ (where σ_{x_i} is the standard deviation of x_i), can be tuned by varying the value of α. In the simulations in this study, the value of the parameter α was chosen so that the impact of a past random selection fades after a given number of timesteps. The choice of $\alpha = 0.25$ leads to a steady decrease in the

correlation between two values x_i and x_j separated by an increasing duration, with a median correlation of 0.45 between two values separated by 5 hours.

With 500 values drawn for each timestep, the original estimated uncertainty quantile values and the dispersion of the scenarios generated by this method match almost perfectly. Increasing the number of scenarios would increase the accuracy of the method, but also the time required for computation. Limiting the number of scenarios to 500 seems to be an acceptable compromise between the computational burden and the accuracy of the method.

2.3 Numerical experiments

It is recalled that the modeling of appliances in the considered setup, and the core of the proposed optimization algorithm are presented in Section 1.2. The optimization problem at hand is Problem 1.1.

From the forecast and production data of a PV installation, a set of PV production scenarios was generated according to the method presented in the previous section. Note that the dataset of PV production forecasts is not the same as the one used in Chapter 1. Here, because the data used were not available for two months from the middle of August to the middle of October, the experiments described were conducted over 309 dates only, instead of 365.

Using these scenarios, the impact of PV production uncertainties on the self-consumption performances of two EMSs was investigated in an individual house case study.

2.3.1 Test setting

The test setting considered here is the same as the one presented in Section 1.3.1, regarding the EWH volume and power rating, the number of inhabitants, the scenario of outdoor temperature, input cold water, hot-water consumption and other uncontrollable appliances load curves, the household location, the final time for the comfort constraint.

A difference is that the two crystalline silicon PV arrays of 1.5 $\mathrm{kW_p}$[4] each are now both inclined at 35° relative to the horizontal and facing South.

The scenarios of PV production considered for these numerical experiments are detailed in the subsequent Section 2.3.3.

It is stressed that although this study covers only uncertainties within similar weather conditions (the oceanic temperate climate found in most of north-western Europe), the method is not climate-dependent.

2.3.2 High-fidelity model and hardware specifications

The numerical experiments required the controller presented in this chapter to communicate with a simulation model. The controller corresponds to the optimization algorithm presented in Chapter 1 based on the simplified EWH model. The high-fidelity EWH model and simulation platform considered are the same as those presented in Section 1.3.1.

All computations are run on a Core i3 2.4 GHz processor, with 8 Go RAM and using Python 3.7.3. Running the entire calculations described below (optimization and high-fidelity simulation) took approximately 70 hours for 309 days and 500 scenarios per day.

[4] $\mathrm{kW_p}$ stands for kW "peak": it is the maximum electrical power that can be supplied by a PV system, under standardized test conditions of irradiance, temperature and solar spectrum

2.3.3 Performance evaluation procedure

The performances of two EWH control algorithms was compared under two settings regarding PV production uncertainties, resulting in four configurations.

Strategies Two control algorithms can be considered to compute the heating strategy:

- the industrial reference heuristic control presented in Section 1.3.1;
- the optimization algorithm proposed in Chapter 1.

Both are deterministic, meaning that they consider a unique and supposedly exact PV power prediction to decide on the corresponding optimal heating strategy.

Subsequently, s_{R} denotes a strategy computed with the reference heuristic, whereas s_{P} denotes a strategy computed with the proposed method.

Scenarios The most likely PV production scenario on day j is denoted by S^j_{ML}. The range of outcomes for the PV production on this day j is represented by the set of K distinct PV production scenarios being generated. The k^{th} realization scenario belonging to this set, with $k \in [1, K]$, is denoted by S^j_k.

Two configurations are considered: either the EMS knows only the most likely PV production scenario, or it knows the exact PV power predictions. In the first case, $s_{\mathrm{R}}(S^j_{\mathrm{ML}})$ denotes a strategy computed with the reference heuristic, and $s_{\mathrm{P}}(S^j_{\mathrm{ML}})$ denotes a strategy computed with the proposed algorithm, both optimized over S^j_{ML}, the most likely scenario on day j. In the second case, $s_{\mathrm{R}}(S^j_k)$ denotes a strategy computed with the reference heuristic, and $s_{\mathrm{P}}(S^j_k)$ denotes a strategy computed with the proposed algorithm, both optimized over S^j_k, the realization scenario k of day j. This case corresponds to the availability of what is named a "perfect" PV production forecast.

For a given date j, the performance of the reference control relying on a most likely forecast is the mean SC achieved by this heating strategy evaluated with respect to all the realization scenarios of the generated set. The corresponding mathematical formulation is given in Eq. (2.5). This configuration is hereafter called *Reference - Most Likely*. This empirical average can be considered as an approximation of the expected value of SC for all possible PV production.

Eq. (2.8) shows the mathematical definition of the mean SC score achievable by the proposed algorithm for a given date j, in the case where a "perfect" forecast is available for every realization scenario of the set. This configuration is hereafter named *Proposed - Exact*.

Reference - Exact is the mean SC score attainable by the reference heuristic with knowledge of each production realization scenario (Eq. (2.6)); *Proposed - Most Likely* is the mean SC score achieved by the heating strategy designed by the proposed algorithm relying only on the most likely forecast (Eq. (2.7)).

$$\textit{Reference - Most Likely}: \ \frac{1}{K}\sum_{k=0}^{K} SC\left(s_{\mathrm{R}}(S^j_{\mathrm{ML}}), S^j_k\right) \tag{2.5}$$

$$\textit{Reference - Exact}: \ \frac{1}{K}\sum_{k=0}^{K} SC\left(s_{\mathrm{R}}(S^j_k), S^j_k\right) \tag{2.6}$$

$$\textit{Proposed - Most Likely: } \frac{1}{K}\sum_{k=0}^{K} SC\left(s_{\mathrm{P}}(S^j_{\mathrm{ML}}), S^j_k\right) \tag{2.7}$$

$$\textit{Proposed - Exact: } \frac{1}{K}\sum_{k=0}^{K} SC\left(s_{\mathrm{P}}(S^j_k), S^j_k\right) \tag{2.8}$$

Logically, it should hold that (2.5) < (2.6), and (2.7) < (2.8). However, (2.6) < (2.7) is not guaranteed, and quantifying the relation between these values is one of the purposes of this study.

Table 2.1 details the simulations conducted for the proposed algorithm in this plan of action, with $K = 500$ generated scenarios. Equivalent calculations were conducted for the reference control. Crosses "×" designate combinations that are *not* associated with calculations. Computing the SC score of every single strategy over every single scenario and then computing the mean SC score of each strategy over the complete set of scenarios could lead to identifying the optimal control strategy in presence of production uncertainties, but doing so would require an amount of time so long that performing this computation in order to implement the resulting strategy would not be realistic.

Table 2.1: Scores computed for the proposed algorithm in the conducted numerical simulations.
(Crosses indicate scores that were not evaluated.)

Strategies	PV production scenarios for day j					
	Most Likely, S^j_{ML}	Realization 1, S^j_1	...	Realization k, S^j_k	...	Realization K, S^j_K
$s_{\mathrm{P}}(S^j_{\mathrm{ML}})$ (Most Likely)	$SC\left(s_{\mathrm{P}}(S^j_{\mathrm{ML}}), S^j_{\mathrm{ML}}\right)$	$SC\left(s_{\mathrm{P}}(S^j_{\mathrm{ML}}), S^j_1\right)$	...	$SC\left(s_{\mathrm{P}}(S^j_{\mathrm{ML}}), S^j_k\right)$	...	$SC\left(s_{\mathrm{P}}(S^j_{\mathrm{ML}}), S^j_K\right)$
$s_{\mathrm{P}}(S^j_1)$ (Realization 1)	×	$SC\left(s_{\mathrm{P}}(S^j_1), S^j_1\right)$	×	×	×	×
...	×	×	⋱	×	×	×
$s_{\mathrm{P}}(S^j_k)$ (Realization k)	×	×	×	$SC\left(s_{\mathrm{P}}(S^j_k), S^j_k\right)$	×	×
...	×	×	×	×	⋱	×
$s_{\mathrm{P}}(S^j_K)$ (Realization K)	×	×	×	×	×	$SC\left(s_{\mathrm{P}}(S^j_K), S^j_K\right)$

2.3.4 Results

The impact of the exact PV load factor predictions and of the chosen algorithm can be seen in Figure 2.3. It reports the sum, from January 1st to the end of each month, of the expected daily SC over the set of drawn scenarios $S^j_k, k \in [1, K]$. The control can use either the reference heuristic (orange color) or the proposed tailored algorithm (blue color). The strategy has been optimized either following a most likely PV production forecast (dashed lines) or assuming that the exact PV production forecast is available for each realization (solid lines). The following equation gives the corresponding mathematical formulation with $J \in [\text{Jan.}31^{\text{st}}, ..., \text{Dec.}31^{\text{st}}]$, the date of each end of month:

$$\sum_{j=\text{Jan.}1^{\text{st}}}^{J}\left[\frac{1}{K}\sum_{k=0}^{K} SC\left(X, S^j_k\right)\right] \tag{2.9}$$

with X taking the following values in each case:

$$\textit{Reference - Most Likely: } X = s_{\mathrm{R}}(S^j_{\mathrm{ML}}) \tag{2.10}$$

$$\textit{Reference - Exact: } X = s_{\mathrm{R}}(S^j_k) \tag{2.11}$$

$$\textit{Proposed - Most Likely: } X = s_{\mathrm{P}}(S^j_{\mathrm{ML}}) \tag{2.12}$$

$$\textit{Proposed - Exact: } X = s_{\mathrm{P}}(S^j_k) \tag{2.13}$$

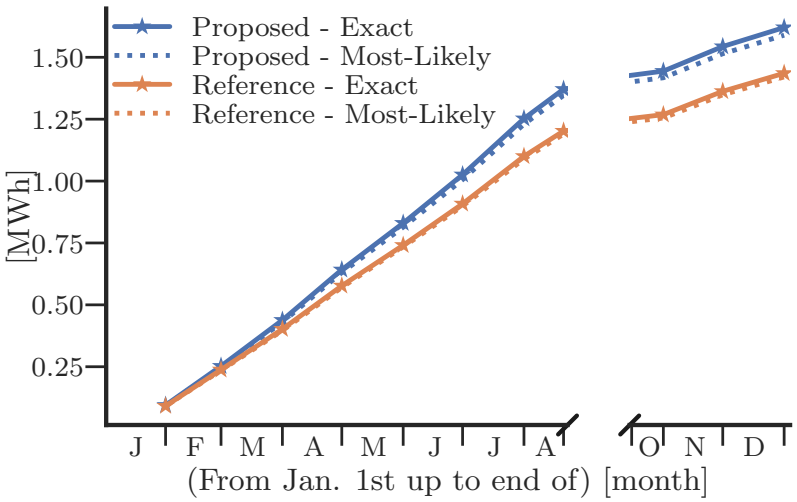

Figure 2.3: The cumulated expected SC is significantly more impacted by the algorithm choice than the forecast precision.

The cumulated SC values reached at the end of the year and the corresponding relative increase compared with the *Reference - Most Likely* case are listed in Table 2.2.

Table 2.2: The cumulated self-consumed energy for 309 days can be increased by 11% with a smarter optimization algorithm.

Strategy	Cumulated Expected SC	Relative increase w.r.t. Reference - Most Likely
Reference - Most Likely (2.10)	1.42 MWh	-
Reference - Exact (2.11)	1.43 MWh	+0.85%
Proposed - Most Likely (2.12)	1.58 MWh	+11%
Proposed - Exact (2.13)	1.61 MWh	+13%

It is clear that, for the studied setup, the impact of a "perfect" PV production forecast is negligible compared with the impact of the choice of the control algorithm. The hypothetical benefit of switching from a most likely PV power prediction to an "exact" one (at a 30-minute timestep) would be slightly greater with the proposed algorithm (+0.03 MWh) than with the Reference heuristic (+0.01 MWh), but remains limited compared with the gains coming from the choice of the control algorithm. Note that the SC values consider both the components due to EWH consumption and those due to uncontrolled demand. It is recalled that the most likely forecast already conveys some information about the PV production curves that will finally occur. Shall a "perfect" PV production forecast be compared with a very low level production forecast (e.g. based on clear-sky solar irradiance only), the impact of the forecast precision would be much greater.

Remark The figures display a gap from the middle of August to the middle of October, because the PV production forecasts were not available for these dates, as mentioned at the beginning of Section 2.3. The "annual" values were hence calculated for 309 days only. As such, the total PV production on the considered period amounts to 3.0 MWh.

Figure 2.4 shows the daily expected self-consumption gains for each configuration compared with the Reference heuristic - Most Likely forecast configuration. For improved clarity, the daily mean *Reference - Most Likely SC* score on day j is denoted as

$$\mathrm{RML}(j) = \frac{1}{K}\sum_{k=0}^{K} SC\left(s_{\mathrm{R}}(S^j_{\mathrm{ML}}), S^j_k\right) \tag{2.14}$$

The curves correspond to the following quantity, with $j \in [\mathrm{Jan.1^{st}}, \mathrm{Dec.31^{st}}]$:

$$\frac{1}{K}\sum_{k=0}^{K} SC\left(X, S^j_k\right) - \mathrm{RML}(j) \tag{2.15}$$

The quantity is evaluated for these three values of X: $s_{\mathrm{R}}(S^j_k)$, $s_{\mathrm{P}}(S^j_{\mathrm{ML}})$, $s_{\mathrm{P}}(S^j_k)$.

The gains are almost always greater with the Proposed algorithm than with the Reference heuristic and Exact PV forecasts. Neither the Reference heuristic nor the Proposed algorithm succeeds in benefiting sensibly from the exact PV forecasts. The ratio of these *SC* gains over the daily score of the *Reference - Most Likely* configuration is not depicted because it follows almost the same pattern, with y-values ranging from -10 % to +30 %.

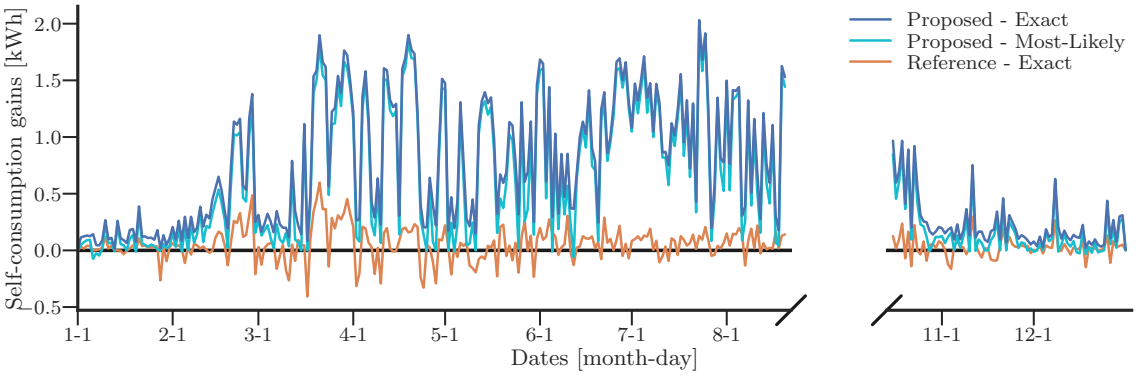

Figure 2.4: The daily *SC* gains for each configuration with respect to the *Reference - Most Likely* configuration are higher during summer days.

Figure 2.4 suggests that the performance difference depends on the weather conditions of the day (e.g., a sunny summer day associated with a high load factor or an intermittently cloudy autumn day associated with a low one). To further analyze these differences, the 309 days of the simulation were partitioned according to their distribution profile. The clustering was made automatically through a K-means method fed with the quantile values of the distribution of the PV load factor values at noon for the $K = 500$ scenarios. Because it seems that the partitioning depends mainly on the median value of this distribution, the partitioning is represented in Figure 2.5 with only the median noon value for each day. The "Low" cohort represented in blue contains 104 days, the "Medium" one, corresponding to the orange dots, contains 86 days, and there are 119 "High" PV scenario days represented by green dots. Unsurprisingly, it can be observed that winter days are mainly grouped in the "Low" cohort, and summer days are in the "High" one.

The clustering allows to break down the relative and absolute expected *SC* gains (w.r.t. the *Reference - Most Likely* configuration) according to the type of day. The lower window of Figure 2.6 represents the variability of the 309 absolute *SC* gains (same data as Figure 2.4), according to the PV scenario types classification. The upper window reports the variability of the 309 relative *SC* gains according to the classification, as detailed in

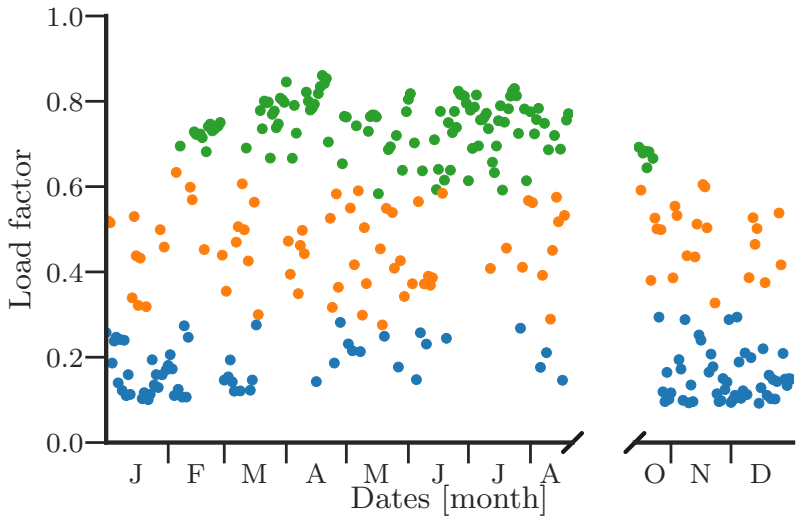

Figure 2.5: PV load factor days clustering.

the following quantities

$$\frac{1}{\mathrm{RML}(j)}\left[\frac{1}{K}\sum_{k=0}^{K} SC\left(X, S_k^j\right) - \mathrm{RML}(j)\right] \tag{2.16}$$

with $\mathrm{RML}(j)$ defined in Eq. (2.14). The equation is valued for the three same values of X as in Eq. (2.15).

It is clear that the gains over the *Reference* - *Most Likely* case, both absolute and relative, are observed during high PV power production days. Because the reference control is activated as soon as the PV production reaches a threshold, it cannot tap into the higher production values later in the day, thus leading to lower self-consumption scores on clear sunny days. Besides, sunny summer days usually have low variability. Hence a most likely forecast is good enough for the proposed algorithm to reach high SC scores. Moreover, it is obvious that a too-basic heuristic control can only poorly benefit from the valuable information carried by an "exact" PV production forecast. This underlines the fact that further research efforts might be better spent on developing better control algorithms than on finding better forecasting methods for the considered time resolution.

2.4 Conclusions

This chapter has studied the impact of PV production forecasts uncertainties on the performance of an EMS optimizing the self-consumption of a residential PV installation. To this end, the optimal EWH scheduling algorithm presented in Chapter 1 has been used with erroneous information. A methodology to generate a set of realistic PV production scenarios was presented. This method ensures that the ensemble of PV power scenarios is representative of the variability associated with most likely weather forecasts. When combined, the efficient optimization algorithm and these sets of PV production scenarios make it possible to assess the performance of such an EMS according to various production scenarios, either assuming the availability of "perfect" PV production predictions or of only the most likely forecasts.

This study has shown that, more globally, self-consumption performance benefits more from an efficient optimal management system like the one proposed in Chapter 1 than from "perfect" PV production predictions at a 30-minute timestep. Indeed, considering a simple rule-based commercial heuristic accessing a *single deterministic most likely forecast* as a

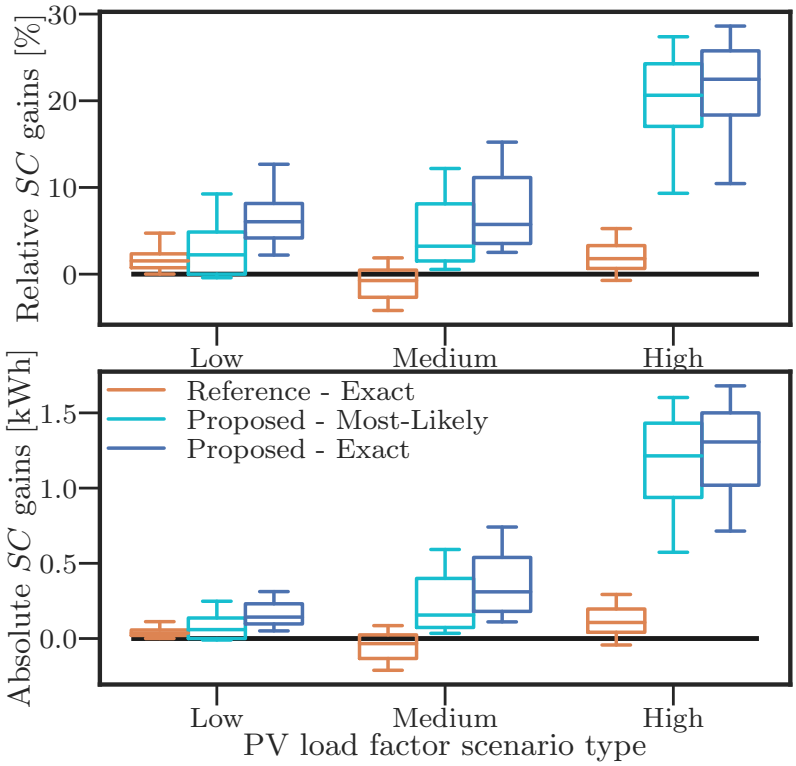

Figure 2.6: The relative and absolute expected *SC* gains w.r.t. *Reference - Most Likely* tend to be higher with the proposed algorithm, and especially during "High" PV production days.

baseline and upgrading to the proposed optimization algorithm yields an 11% increase in annual cumulated expected self-consumed energy, whereas combining this proposed optimization algorithm and a "perfect" PV production forecast at a 30-minute timestep brings only an additional two-percentage-point increase (from +11% to +13%). A detailed analysis shows that the gains yielded by the optimization algorithm choice are concentrated in days of high PV power production.

It is recalled that these findings regarding the volume of self-consumed energy do not necessarily apply to the overall electricity bill of the household, because of the temporal variation of the hourly electricity tariffs.

The PV production scenarios generation method developed for this study is also of interest as it allows to assess more truthfully the *a priori* performance of an EMS.

This study suggests that for these installations:

- the current level of accuracy of PV production forecasts is good enough;
- improving the control algorithms (as it has been proposed in Chapter 1) can be valuable.

This work should be extended to study as well the impact of inhabitants' behavior uncertainties, both for hot-water usage and for uncontrolled power demand.

Assuming that PV production scenarios with a 30-minute timestep can be exact is also an oversimplification, so future research is needed to assess the performance loss of this kind on system due to faster production variations.

Chapter 3

MILP formulation for the microgrid energy bill optimization

Formulation d'un problème d'optimisation linéaire mixte pour l'optimisation de la facture d'énergie d'un micro-réseau. Ce chapitre traite de la minimisation de la facture d'énergie d'un micro-réseau avec autoconsommation, sous forme d'un problème d'optimisation linéaire mixte. Au sein du micro-réseau, plusieurs foyers sont équipés d'usages pilotables, peuvent échanger de l'énergie et accéder à une batterie partagée. Une modélisation discrétisée des usages avec stockage (chauffe-eau Joule, véhicule électrique, batterie) est proposée, inspirée d'une modélisation des usages blancs présente dans la littérature. La formulation du problème d'optimisation à ce niveau implique la concaténation des contraintes de chaque habitation participant au micro-réseau. La complexité du problème centralisé obtenu est explorée. L'intérêt quantitatif de la coopération et le coût de la non-coopération sont présentés au travers d'un cas d'étude simple à deux agents et deux usages.

3.1 Introduction

3.1.1 Context

In the first half of this thesis (Chapter 1 and Chapter 2), an efficient algorithm has been developed and used for the explicit control of an EWH with the aim of maximizing the consumption of the energy locally produced by the residential photovoltaic (PV) self-consumption installation (Chapter 1). This chapter broadens the scope of controlled systems for the optimization of PV self-consumption using several appliances in several households in a local network. An example setup is first introduced before detailing the concepts behind this organization.

Example setup

To visualize the kind of microgrid configuration where appliances control and cooperative energy optimization is desirable, here is detailed an example use-case including several extensions already mentioned in the previous chapters.

Consider a group of neighbor houses (e.g. five) in a residential neighborhood located in a sufficiently sunny city, all equipped with their own PV panels. Each house is assumed to be equipped with a handful of controllable appliances, and with an energy management system (EMS) aiming at minimizing their individual electricity bill. The controllable

appliances of each household can be any selection from the following appliances of interest for control, given their large nominal power and energy consumption and the fact that their activation can be dissociated from the time of use of the service: EWHs, electric vehicles (EVs), dishwashers, washing machines, dryers (the three latter regrouped under the designation "white goods"[1]). A shared battery of substantial capacity is located nearby the houses. Figure 3.1 presents such an example.

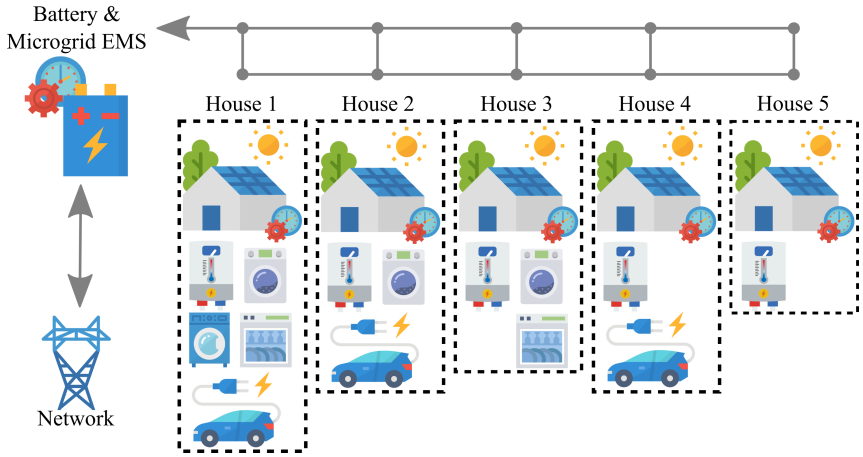

Figure 3.1: The use-case microgrid is composed of households of varied equipment profiles.

Typical technical characteristics of the use-case are the following:

- Installed individual PV power: 9 kW_p[2]
- Maximum power exchange: 9 kV A[3]
- EWH heat pump: variable power [0 W to 600 W] and boost power 2 kW
- EV capacity: 40 kWh
- V2X enabled EVSEs[4], of power: 40 kW
- Shared battery: 100 kWh, 100 kW
- Maximum number of daily battery cycle: 1
- Households composition:
  - House 1 : 2 adults + 1 children;
  - House 2 : 2 adults + 2 children;
  - House 3 : 1 adult;
  - House 4 : 2 retired adults;
  - House 5 : 2 adults.

The cooperative energy management scheme subsequently presented in this chapter can be usefully applied to such a microgrid configuration.

Change of objective function

The first critical extension of this chapter with respect to Chapters 1 and 2 concerns the scope of the objective function. This chapter operates a semantic broadening from the

[1]Large electrical appliances used for routine housekeeping tasks (such as cooking, cleaning, washing laundry, or food preservation) which were traditionally available only in white.

[2]kW_p stands for kW "peak": it is the maximum electrical power that can be supplied by a PV system, under standardized test conditions of irradiance, temperature and solar spectrum

[3]The volt-ampere unit used for the apparent power in an electrical circuit is usually used for analyzing alternating current (AC) circuits.

[4]EV Supply Equipment: individual charging station for an EV

optimizing the *physical criterion* of self-consumption (i.e. the energy simultaneously produced and consumed locally) to optimizing the *financial cost of operation of installations* with self-consumption. Here, the cost of total electricity supply at the microgrid scale is minimized. The amount of self-consumed energy, used as the optimization criterion of the previous chapters, still has to be computed to evaluate the overall electricity bill, and could have been used as the objective function of the problems formulated in this chapter. However, due to the variable cost of electricity through the day, maximizing the self-consumed energy can lead to sub-optimal electricity bills, especially when harnessing the local mid-day production also leads to consuming more energy from the grid at high cost periods. Thus, a more realistic study in the current framework where the households are connected to a main grid experiencing little upstream power constraints is to consider that the criterion minimized by the EMSs is the overall energy cost.

Number and nature of controlled appliances

Other extensions (identified at the end of Chapter 1) regard the number and nature of involved appliances. In this chapter, the range of appliances controlled is thus broadened in order to reach higher levels of self-consumption.

To extend the type of appliances that an EMS could control in the same household, white goods first come to mind as they:

- represent significant shares of the overall residential energy consumption;
- have large power ratings;
- are household appliances whose activation can be desynchronized with the inhabitants presence;
- are already widespread (though not in their controllable "smart" versions).

Then, EVs have to be considered as:

- their sales are expected to increase steadily over the next years as a necessary replacement of internal combustion engine vehicles that contribute to the climate crisis;
- their charging operations will need to be managed to avoid any detrimental effects on both the users energy bills and the electric grid;
- they can provide flexibility services to the grid thanks to Vehicle-to-X capabilities.

Vehicle-to-X (V2X) refers to the possibility for an EV to exchange power back to a place connected to the used EV Supply Equipment. Depending on the considered location, it can correspond to giving power back to the home (V2H), building (V2B) or the grid (V2G). [Gonzalez Venegas et al., 2021] provides a detailed review of potential frameworks and barriers for the provision of flexibility services to distribution grids from EVs. The paper recalls that individual vehicles have long idle periods, and that the limited time required to provide the charge necessary for a usual daily consumption leaves enough margin to provide such services. In the case of distribution grids, EV flexibility can be used to defer or avoid costly infrastructure reinforcements, with great economic savings. These power exchanges require bidirectional chargers and ad-hoc communication protocols.

Modeling these appliances using the framework presented in Chapter 1 would lead to a burdensome combinatorial analysis. Indeed, adapting this approach to only two appliances showed that the specific times where the supports of the load curves were intersecting result in discontinuities in the objective function (see Appendix A). These specific times thus have to be detected in order to find the best times of control in each continuous subset, before comparing them to finally extract the solution. The number of discontinuities would increase in a combinatorial manner for a growing number of controlled appliances, thus rendering this approach unsuitable for the desired use-case.

On the contrary, several works resort to MILP formulations for a simpler, more scalable and more flexible formulation of their operation. Especially, these formulations allow to use binary variables to model On/Off behaviors, conditional activation (e.g. depending on a price signal), or pauses between the various phases of operation. Adapting the previously presented framework to these kind of constraints, for several different appliances would result in a quite complex model. Instead, the path followed in this chapter to model electrical appliances is to resort to classical MILP formulations. As Chapter 2 showed it, self-consumption performance benefits more from an efficient optimal management system based on deterministic most likely PV production forecasts than from "perfect" PV production predictions at a 30-minute timestep. At the light of this result, the MILP formulations followed in this chapter only consider the most likely production forecasts in order to find the optimal appliances schedule at the household or microgrid level.

Microgrid

Another critical extension previously identified is to try and optimize the behavior of a microgrid, whose local production and consumption implies the principle of collective self-consumption but also allows for the collective ownership of distributed energy resources (DERs).

Collective self-consumption Collective self-consumption is the extension of the concept to the totality of energy consumptions and productions in a specified set of buildings, for instance, a few houses located in the same neighborhood.

It has been shown that collective self-consumption can further decrease electricity bills and increase the volume of locally produced, exchanged and consumed energy with respect to individual self-consumption configurations [Mengelkamp et al., 2018]. Moreover, this effect can be reinforced if the share of DERs (e.g. batteries) allotted to each household can be dynamically chosen at each timestep according to all of the microgrid needs, [Dimitrov et al., 2016].

Shared batteries A microgrid can benefit from the addition of batteries to the PV installations to increase the volume of self-consumed energy and decrease electricity bills by performing arbitration in regards to the price of electricity. If multiple individual batteries increase the self-consumption of each agent, [Roberts et al., 2019] has shown that self-consumption is further increased with the installation of shared batteries instead of individual ones. Shared batteries have also the advantage of reducing the investment, maintenance, operational and replacement costs per end-user [Tascikaraoglu et al., 2019], whereas individual batteries still represent an overall cost as long as they do not provide additional services [Quoilin et al., 2016, Goebel et al., 2017, Roberts et al., 2019]). Further, it has been shown by [Koirala et al., 2018] that shared batteries foster social benefits beyond monetary gains, including reinforced social cohesion and local economy, compared to a collection of individually owned and operated batteries. Hence the microgrid considered in this chapter is equipped with a central shared battery. [Dimitrov et al., 2016] has presented an efficient heuristic allowing agents to bid for a portion of the storage capacity allotted to another agent. The algorithm is a distributed control policy, where each user operates independently with the least exchange of information with the other users. It shows that allowing a dynamical allocation of the storage capacity is more profitable than being restricted to multiple private storage capacities.

The microgrid concept calls for the design of collaboration schemes allowing to optimize the collective power use. A direct way to ensure the optimal operation of a collective self-consumption microgrid is to formulate and solve the corresponding optimization problem

at the microgrid level. This yields a lower bound for the cost of supplying energy to the whole community. This is the approach followed in this chapter.

3.1.2 Content of the chapter

The microgrid considered in this chapter, as illustrated in Figure 3.1, is composed of up to a dozen single-family residential buildings, each equipped with their own PV panels and able to store and draw power into and from a shared community battery. The microgrid is connected to the main grid, experiences no voltage constraints (see the Introduction for details), and its size is restricted to a small neighborhood. Energy exchanges between participants are allowed, and even encouraged through the specific tariffs.

The microgrid optimization problem is defined as the problem of minimizing the cost of supplying energy to the microgrid by controlling the shared DERs and all the controllable appliances in the microgrid, and by importing and exporting the remaining deficit and surplus power from and to the external grid operator. Formulating this problem assumes that the microgrid EMS can retrieve all the relevant information from the agents, the DERs and the external grid.

This chapters details the constraints, decision variables and objective function of the microgrid optimization problem, formulated as a MILP.

First, Section 3.2 covers how the microgrid optimization problem can be formulated by aggregating the individual households operational constraints into the definition of the microgrid behavior. The scalability limitations of this approach are then explored in Section 3.3. Finally the value of cooperation at the microgrid level is proved in a small and solvable use-case in Section 3.4.

3.2 Microgrid optimization problem formulation

The microgrid optimization problem under consideration here is the minimization of the overall energy supply cost, under the individual households load constraints, the microgrid coupling constraints and the battery operations constraints.

To ensure that the appliances activation schedules resulting from the microgrid EMS optimization are applied, it is assumed that the individual houses EMSs are fully connected to the household appliances and to the microgrid EMS, that they automatically retrieve the solution from the microgrid EMS and pilot their appliances according to this recommended schedule.

3.2.1 Notations

In the remainder of this chapter, the following notations will be used for matrices and vectors:

The time horizon is uniformly discretized in a finite time grid of H timesteps indexed by the variable t.

For any non-null natural integers n_1, n_2, $\mathbf{0}_{n_1,n_2}$ is the null matrix in $\mathbb{R}^{n_1 \times n_2}$.

For any non-null natural integer n, $\mathbb{1}_n$ is the vector full of ones in $\mathbb{R}^n$ and $\mathbb{0}_n$ the null vector in $\mathbb{R}^n$.

These elements have a simplified notation for the dimension H that often appears:

$\mathbf{I}$ and $\mathbf{0}$ are respectively the identity matrix and the null matrix in $\mathbb{R}^{H \times H}$.

$\mathbb{1}$ is the vector full of ones in $\mathbb{R}^H$ and $\mathbb{0}$ is the null vector in $\mathbb{R}^H$.

3.2.2 White goods modeling: the energy phases concept

[Sou et al., 2011] proposes an advanced modeling framework for controllable appliances

of the white goods category, using mixed integer linear programming (MILP) to model operation characteristics such as duration and peak power consumption, or uninterruptible and sequential operations.

In the proposed scheduling framework for appliances of the white goods category, an appliance operation process is divided into a set of sequential energy phases. Apprehending this concept is fundamental to understand many aspects of the subsequent analysis. Thus a detailed description is given here before presenting the households optimization constraints in the following section.

An energy phase is an uninterruptible sub-task of the appliance operation which uses a pre-specified amount of electric energy, and may have bounds on the instantaneous power consumption and the execution time. The energy phases are uninterruptible in that a sub-task has to be completed in a continuous duration, and cannot be resumed. They are sequential since the next appliance sub-task cannot begin until the previous sub-task is completed (e.g. the washing machine agitator cannot start until the basin is filled with water).

Table 3.1 presents the technical specifications for a dishwasher. These specifications are adapted from [Sou et al., 2011] to work with a 15-minute timestep. Figure 3.2 details the energy phases composing an example load curve for a dishwasher, compliant with the specifications of Table 3.1. A log scale has been used on the second plot (right) to show the low power consumptions of the "drain & dry" energy phase. The number of energy phases of an appliance of the white goods category is thereafter noted J (here $J = 5$ for the considered dishwasher).

Table 3.1: Dishwasher technical specifications from [Sou et al., 2011]

Energy phase	Energy (Wh)	Min power (W)	Max power (W)	Duration (min)
pre-wash	16.0	6.47	140	15
wash	751.2	140.26	2117.8	30
1st rinse & drain	19	10.28	132.4	15
2nd rinse	572.3	187.3	2143	15
drain & dry	1.7	0.2	2.3	60

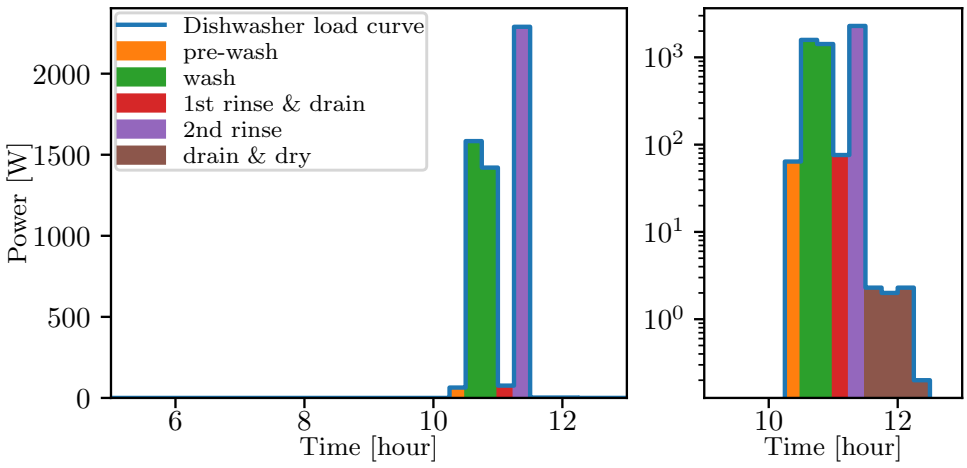

Figure 3.2: Each colored block corresponds to one energy phase, an uninterruptible sub-task of the dishwasher program.

3.2.3 Households constraints

A framework similar to the one proposed by [Sou et al., 2011] has been developed to model "storage" appliances such as EWHs, batteries and EVs. All these models (both for white goods and storage appliances) are detailed in Appendix C, and can be written with the following equations, for the smart appliance i of an household $h (h \in [1, l])$:

$$A_{hi} x_{hi} \leq b_{hi} \quad \forall\, h, i \tag{3.1}$$

with the content of A_{hi}, x_{hi} and b_{hi} depending on the type of appliance modeled. Note that this inequality formulation can also be used to describe equality equations using two inverted inequalities.

x_{hi} gathers all the variables needed to describe the operation of the appliance i. For "storage loads" such as an EWH or an EV, x_{hi} is composed of the variables defining the stored energy and the instantaneous power flux at every timestep, thus characterizing its load curve. For an EV, additional binary variables help define its status (charging, discharging). For white goods, x_{hi} is composed of variables defining the instantaneous power consumption at every timestep (thus characterizing its load curve), and binary variables indicating the present and past state of each energy phase composing a specific program.

Eq. (3.1) models the behavior of any controllable appliance. The constraints presented in the following paragraphs are necessary to define the overall load curve of an household, uniting the power consumption of its controllable and uncontrollable appliances, and its local power production.

Power exchange limitations

The household EMS takes into account the power exchanges with the rest of the microgrid, which are bounded as follows, for all h, t:

$$0 \leq P^t_{h,\mathrm{s}} \leq P_{h,\max} \delta^t_{h,\mathrm{s}} \tag{3.2}$$

$$0 \leq P^t_{h,\mathrm{d}} \leq P_{h,\max} \delta^t_{h,\mathrm{d}} \tag{3.3}$$

$$\delta^t_{h,\mathrm{d}} + \delta^t_{h,\mathrm{s}} \leq 1 \tag{3.4}$$

Constraint (3.2) describes the limitations on $P_{h,\mathrm{s}}$, the household h net surplus power. $P_{h,\max}$ is the maximum power exchange between household h and the microgrid. The binary decision variable $\delta^t_{h,\mathrm{s}}$ equals 1 when the household load curve is in surplus, 0 otherwise. Constraint (3.3) similarly describes the limitations on $P_{h,\mathrm{d}}$, the household net deficit power, with the binary decision variable $\delta^t_{h,\mathrm{d}}$ indicating when the household load curve is in deficit. Finally, the constraint (3.4) guarantees that the household load curve in not in surplus and in deficit at the same timestep.

Eqs. (3.2)–(3.4) can be formulated as

$$\alpha_{h,\mathrm{def}}\, x_{h,\mathrm{def}} \leq b_{h,\mathrm{def}} \quad \forall\, h \tag{3.5}$$

with

$$\alpha_{h,\mathrm{def}} = \begin{pmatrix} \mathbf{I} & \mathbf{0} & -P_{h,\max}\mathbf{I} & \mathbf{0} \\ -\mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} & -P_{h,\max}\mathbf{I} \\ \mathbf{0} & -\mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{I} \end{pmatrix} \tag{3.6}$$

and with

$$x_{h,\mathrm{def}} = \begin{pmatrix} (P^t_{h,\mathrm{s}})_{t\in[1,H]} \\ (P^t_{h,\mathrm{d}})_{t\in[1,H]} \\ (\delta^t_{h,\mathrm{s}})_{t\in[1,H]} \\ (\delta^t_{h,\mathrm{d}})_{t\in[1,H]} \end{pmatrix}, \quad b_{h,\mathrm{def}} = \begin{pmatrix} \mathbb{0}_{4H} \\ \mathbb{1} \end{pmatrix} \tag{3.7}$$

Supply-demand balance

Power balance at each household level at all times is enforced by the following constraint:

$$P^t_{h,\mathrm{s}} - P^t_{h,\mathrm{d}} = P^t_{h,\mathrm{PV}} - P^t_{h,\mathrm{unc}} - \sum_{i=1}^{m_h} \sum_{j=1}^{J} P^t_{hij} \qquad \forall\, h, t \tag{3.8}$$

where $P_{h,\mathrm{PV}}$ is the PV local production, $P_{h,\mathrm{unc}}$ is the uncontrolled power consumption for household h, and P_{hij} is the power of the energy phase j of appliance i of household h[5].

Eq. (3.8) can be more compactly written as

$$\sum_{i=1}^{m_h} \alpha_{hi}\, x_{hi} + \alpha_{h,\mathrm{join}}\, x_{h,\mathrm{def}} \leq b_{h,\mathrm{join}} \quad \forall\, h \tag{3.9}$$

with

$$\alpha_{hi} = \begin{cases} \begin{pmatrix} & & \mathbb{1}_J^\top & \\ \mathbf{0}_{H,HJ} & \mathbf{0}_{H,HJ} & & \ddots \\ & & & \mathbb{1}_J^\top \\ & & -\mathbb{1}_J^\top & \\ \mathbf{0}_{H,HJ} & \mathbf{0}_{H,HJ} & & \ddots \\ & & & -\mathbb{1}_J^\top \end{pmatrix} & \text{if the appliance is one of the white goods category} \\ \begin{pmatrix} \mathbf{0} & \overline{P}_{\mathrm{ewh}}\mathbf{I} \\ \mathbf{0} & -\overline{P}_{\mathrm{ewh}}\mathbf{I} \end{pmatrix} & \text{if the appliance is an EWH of power rating } \overline{P}_{\mathrm{ewh}} \\ \begin{pmatrix} \mathbf{0} & \mathbf{I} & -\mathbf{I} & \mathbf{0}_{H,2H} \\ \mathbf{0} & -\mathbf{I} & \mathbf{I} & \mathbf{0}_{H,2H} \end{pmatrix} & \text{if the appliance is an EV} \end{cases} \tag{3.10}$$

with spaces full of zeros left empty for better readability, with $x_{h,i}$ defined in Appendix C.2.5, $x_{h,\mathrm{def}}$ defined in Eq. (3.7), and with

$$\alpha_{h,\mathrm{join}} = \begin{pmatrix} \mathbf{I} & -\mathbf{I} & \mathbf{0} & \mathbf{0} \\ -\mathbf{I} & \mathbf{I} & \mathbf{0} & \mathbf{0} \end{pmatrix}, \quad b_{h,\mathrm{join}} = \begin{pmatrix} (P^t_{h,\mathrm{PV}} - P^t_{h,\mathrm{unc}})_{t\in[1,H]} \\ (-P^t_{h,\mathrm{PV}} + P^t_{h,\mathrm{unc}})_{t\in[1,H]} \end{pmatrix} \tag{3.11}$$

Summary

Gathering (3.1), (3.5) and (3.9), the household consumption constraints can be concatenated under the form

$$A_h\, x_h \leq b_h \tag{3.12}$$

with

$$A_h = \begin{pmatrix} A_{h,1} & 0 & \dots & 0 & 0 \\ 0 & A_{h,2} & \ddots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \ddots & A_{h,m_h} & 0 \\ 0 & 0 & \dots & 0 & \alpha_{h,\mathrm{def}} \\ \alpha_{h,1} & \alpha_{h,2} & \dots & \alpha_{h,m_h} & \alpha_{h,\mathrm{join}} \end{pmatrix}, x_h = \begin{pmatrix} x_{h,1} \\ x_{h,2} \\ \vdots \\ x_{h,m_h} \\ x_{h,\mathrm{def}} \end{pmatrix}, b_h = \begin{pmatrix} b_{h,1} \\ b_{h,2} \\ \vdots \\ b_{h,m_h} \\ b_{h,\mathrm{def}} \\ b_{h,\mathrm{join}} \end{pmatrix} \tag{3.13}$$

[5] For "storage" appliances that are not modeled with energy phases, this power variable is not indexed by j.

3.2.4 Microgrid constraints

The microgrid constraints are the concatenation of the constraints of its parts, namely the constraints of all the households and of the shared battery. In its resulting shape, the formulation of the microgrid constraints is thus an extension of the work presented in [Sou et al., 2011], from the household level to the microgrid level.

Aggregated households behavior

The microgrid EMS requires the equations linking the individual load curves to the variables defining the behavior of the aggregated agents, for all t:

$$P^t_{\text{s,agg}} - P^t_{\text{d,agg}} = \sum_{h=1}^{l} (P^t_{h,\text{s}} - P^t_{h,\text{d}}) \tag{3.14}$$

$$0 \leq P^t_{\text{s,agg}} \leq \delta^t_{\text{s,agg}} \sum_{h=1}^{l} P_{h,\max} \tag{3.15}$$

$$0 \leq P^t_{\text{d,agg}} \leq \delta^t_{\text{d,agg}} \sum_{h=1}^{l} P_{h,\max} \tag{3.16}$$

$$\delta^t_{\text{s,agg}} + \delta^t_{\text{d,agg}} \leq 1 \tag{3.17}$$

Constraint (3.14) defines $P_{\text{s,agg}}$ and $P_{\text{d,agg}}$, the two positive variables resulting from an overall surplus load curve or an overall deficit load curve, respectively, for the aggregated households. Constraints (3.15) and (3.16) are enforced to meet the upper and lower limitations on power exchanged from the households when the aggregated power are in surplus or in deficit at each timestep t. Binary decision variables $\delta^t_{\text{s,agg}}$ and $\delta^t_{\text{d,agg}}$ indicate whether the aggregated load curve is in surplus or in deficit. The constraint (3.17) should be satisfied to make sure that the microgrid load curve in not in surplus and in deficit at the same timestep.

Eqs. (3.14)–(3.17) are summarized as

$$\sum_{h=1}^{l} \alpha_h \, x_h + \alpha_{\text{agg}} \, x_{\text{agg}} \leq b_{\text{agg}} \tag{3.18}$$

with the individual household decision variables $x_h \in \mathbb{R}^{n_h} \, \forall h \in [1, l]$, defined in Eq. (3.13), and

$$\alpha_h = \begin{pmatrix} \mathbf{0}_{H,n'_h} & -\mathbf{I}_H & \mathbf{I}_H & \mathbf{0}_{H,2H} \\ \mathbf{0}_{H,n'_h} & \mathbf{I}_H & -\mathbf{I}_H & \mathbf{0}_{H,2H} \\ & & \mathbf{0}_{5H,n_h} & \end{pmatrix} \tag{3.19}$$

where $n'_h = n_h - 4H$ is the dimension of $\begin{pmatrix} x^\top_{h,1} & x^\top_{h,2} & \dots & x^\top_{h,m_h} \end{pmatrix}^\top$ (see Eq. (3.13)), with

$$\alpha_{\text{agg}} = \begin{pmatrix} \mathbf{I} & -\mathbf{I} & \mathbf{0} & \mathbf{0} \\ -\mathbf{I} & \mathbf{I} & \mathbf{0} & \vdots \\ \mathbf{I} & \mathbf{0} & -\sum_{h=1}^{l} P_{h,\max}\mathbf{I} & \vdots \\ -\mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} & -\sum_{h=1}^{l} P_{h,\max}\mathbf{I} \\ \mathbf{0} & -\mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{I} \end{pmatrix} \tag{3.20}$$

and

$$x_{\text{agg}} = \begin{pmatrix} (P^t_{\text{s,agg}})_{t\in[1,H]} \\ (P^t_{\text{d,agg}})_{t\in[1,H]} \\ (\delta^t_{\text{s,agg}})_{t\in[1,H]} \\ (\delta^t_{\text{d,agg}})_{t\in[1,H]} \end{pmatrix}, \quad b_{\text{agg}} = \begin{pmatrix} \mathbb{0}_{6H} \\ \mathbb{1} \end{pmatrix} \tag{3.21}$$

Battery operations

The operation constraints defined in the battery model in Section C.3.3 are summarized as

$$A_{\text{bat,op}}\, x_{\text{bat}} \leq b_{\text{bat,op}} \tag{3.22}$$

where the content of $A_{\text{bat,op}}$ and $b_{\text{bat,op}}$ are detailed in Eqs. (C.38) and (C.39) in Appendix C.3.3 and where x_{bat} gathers the variables defining the behavior of the shared battery and is defined as

$$x_{\text{bat}} = \begin{pmatrix} (P^t_{\text{ba}})_{t\in[1,H]} \\ (P^t_{\text{ab}})_{t\in[1,H]} \\ (P^t_{\text{bg}})_{t\in[1,H]} \\ (P^t_{\text{gb}})_{t\in[1,H]} \\ (\delta^t_{\text{C}})_{t\in[1,H]} \\ (\delta^t_{\text{D}})_{t\in[1,H]} \\ (E^t_{\text{bat}})_{t\in[1,H]} \end{pmatrix} \tag{3.23}$$

P_{ba} is the power discharged from the battery to feed the aggregated agents demand, P_{ab} is the power surplus coming from the agents and charging the battery, P_{bg} is the power discharged from the battery towards the outer grid, P_{gb} is the power imported from the grid to charge the battery. δ_{C} and δ_{D} are binary variables identifying if the battery is respectively charging or discharging, and E_{bat} is the stored energy.

Battery limitations and supply-demand balance

The microgrid EMS considers the limitations on the battery model variables and the overall power balance, for all t:

$$0 \leq P^t_{\text{ba}} \leq P^t_{\text{d,agg}} \tag{3.24}$$
$$0 \leq P^t_{\text{ab}} \leq P^t_{\text{s,agg}} \tag{3.25}$$
$$0 \leq P^t_{\text{bg}} \leq P^t_{\text{mg}} \tag{3.26}$$
$$0 \leq P^t_{\text{gb}} \leq P^t_{\text{gm}} \tag{3.27}$$
$$P^t_{\text{gm}} - P^t_{\text{mg}} = P^t_{\text{d,agg}} - P^t_{\text{s,agg}} + P^t_{\text{gb}} - P^t_{\text{bg}} + P^t_{\text{ab}} - P^t_{\text{ba}} \tag{3.28}$$

Constraints (3.24)-(3.27) define the power exchange between each layer of the operation, and their respective upper and lower limitations. The power discharged from the battery to supply the aggregated households P^t_{ba} cannot be greater than the aggregated load curves deficit (3.24). The power production surplus from the aggregated households has to be greater than the power exchange from the agents to the battery P^t_{ab} (3.25). Constraint (3.26) imposes that the power discharged from the battery towards the outer grid P^t_{bg} be not larger than P^t_{mg}, the overall exported power from the microgrid to the outer grid. Similarly, constraint (3.27) imposes that the charging power coming from the outer grid to the battery P^t_{gb} is not larger than P^t_{gm}, the overall imported power from the outer grid to supply the microgrid. Finally, constraint (3.28) is enforced to satisfy the power balance in the system.

Eqs. (3.24)–(3.27) defining the battery limitations and Eq. (3.28) defining the power balance are summarized as:

$$A_{\text{agg}}\, x_{\text{agg}} + A_{\text{bat,lim}}\, x_{\text{bat}} + A_{\text{m}}\, x_{\text{m}} \leq b_{\text{bat,lim}} \tag{3.29}$$

with x_{agg} defined in Eq. (3.21), x_{bat} defined in Eq. (3.23), with

$$A_{\text{agg}} = \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -\mathbf{I} \\ \mathbf{0} & \mathbf{0} \\ -\mathbf{I} & \mathbf{0} \\ \mathbf{0}_{4H,2H} & \\ \mathbf{I} & -\mathbf{I} \\ -\mathbf{I} & \mathbf{I} \end{pmatrix}, \quad A_{\text{bat,lim}} = \begin{pmatrix} -\mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{I} & \mathbf{0} & \vdots & \vdots \\ \mathbf{0} & -\mathbf{I} & \vdots & \vdots \\ \vdots & \mathbf{I} & \mathbf{0} & \vdots \\ \vdots & \mathbf{0} & -\mathbf{I} & \vdots \\ \vdots & \vdots & \mathbf{I} & \mathbf{0} \\ \vdots & \vdots & \mathbf{0} & -\mathbf{I} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I} \\ \mathbf{I} & -\mathbf{I} & \mathbf{I} & -\mathbf{I} \\ -\mathbf{I} & \mathbf{I} & -\mathbf{I} & \mathbf{I} \end{pmatrix} \mathbf{0}_{10H,2H} \qquad \mathbf{0}_{10H,3H} \tag{3.30}$$

and with

$$A_{\text{m}} = \begin{pmatrix} \mathbf{0}_{5H,2H} & \\ \mathbf{0} & -\mathbf{I} \\ \mathbf{0} & \mathbf{0} \\ -\mathbf{I} & \mathbf{0} \\ \mathbf{I} & -\mathbf{I} \\ -\mathbf{I} & \mathbf{I} \end{pmatrix}, \quad x_{\text{m}} = \begin{pmatrix} (P^t_{\text{gm}})_{t\in[1,H]} \\ (P^t_{\text{mg}})_{t\in[1,H]} \end{pmatrix}, \quad b_{\text{bat,lim}} = \mathbb{0}_{10H} \tag{3.31}$$

3.2.5 Microgrid objective function

The microgrid EMS objective depends on the legal and technical organization of the operation. In principle, two main situations could be considered. If the shared battery and microgrid EMS are installed by will of the residents, the objective of this EMS is to minimize the collective cost of energy procurement. If the battery and EMS installation are chosen and funded by the distribution system operator, the objective of the EMS will more likely be to minimize the volume of imported and exported energy. In this second situation, the objective has to be balanced with the minimization of the deviation between the energy bill when participating to this operation and non participating in it, for each household. With non-constant price rates throughout the day, these two objectives might not coincide, and might even be contradictory.

The first configuration is considered in the remainder of this chapter: the microgrid EMS objective is to minimize the cost of collective energy procurement:

$$\min \sum_{t=1}^{H} \left(c^t_{\text{d}} P^t_{\text{gm}} - c^t_{\text{s}} P^t_{\text{mg}} \right) \Delta t \tag{3.32}$$

with Δt the timestep length, P^t_{gm} the power imports from the outer grid to the microgrid at timestep t, P^t_{mg} the power exports from the microgrid to the outer grid at timestep t, c^t_{d} the cost for deficit energy bought from the outer grid at timestep t, and c^t_{s} the cost for surplus energy sold by the microgrid to the main grid at timestep t.

3.2.6 Microgrid optimization problem

The microgrid EMS minimizes the whole microgrid electricity bill defined in Eq. (3.32) under operations constraints for every controllable appliances including the shared battery, power exchange limitations between each level, and supply-demand balancing at each level and for each household.

According to the previous descriptions, the microgrid problem at stake can be formulated as

Problem 3.1.

$$\begin{aligned} &\min c^\top x \\ &\text{s.t. } A\,x \leq b \end{aligned} \tag{3.33}$$

with the cost vector defined as

$$c = \begin{pmatrix} \mathbb{0}_{n_1}^\top & \dots & \mathbb{0}_{n_l}^\top & \mathbb{0}_{4H}^\top & \mathbb{0}_{7H}^\top & (c_{\mathrm{d}}^t)_{t\in[1,H]}^\top \Delta t & -(c_{\mathrm{s}}^t)_{t\in[1,H]}^\top \Delta t \end{pmatrix}^\top \tag{3.34}$$

and with the decision variables

$$x = \begin{pmatrix} x_1^\top & x_2^\top & \dots & x_l^\top & x_{\mathrm{agg}}^\top & x_{\mathrm{bat}}^\top & x_{\mathrm{m}}^\top \end{pmatrix}^\top \tag{3.35}$$

in which each household decision variables vector $x_h, \forall h \in [1, l]$ is defined in Eq. (3.13), x_{agg} is defined in Eq. (3.21), x_{bat} is defined in Eq. (3.23), and x_{m} is defined in Eq. (3.31).

The constraints matrix A, and right-hand side vector b are defined as

$$A = \begin{pmatrix} A_1 & 0 & \dots & 0 & & & \\ 0 & A_2 & \ddots & 0 & & & \\ \vdots & \ddots & \ddots & \ddots & & 0 & \\ 0 & 0 & \ddots & A_l & & & \\ \alpha_1 & \alpha_2 & \dots & \alpha_l & \alpha_{\mathrm{agg}} & \mathbf{0}_{7H,7H} & \mathbf{0}_{7H,2H} \\ & 0 & & & \mathbf{0}_{n_{\mathrm{bat,op}},4H} & A_{\mathrm{bat,op}} & \mathbf{0}_{n_{\mathrm{bat,op}},2H} \\ & & & & A_{\mathrm{agg}} & A_{\mathrm{bat,lim}} & A_{\mathrm{m}} \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_l \\ b_{\mathrm{agg}} \\ b_{\mathrm{bat,op}} \\ b_{\mathrm{bat,lim}} \end{pmatrix} \tag{3.36}$$

in which $\forall h \in [1, l]$, the matrix A_h corresponds to one household constraints matrix defined in Eq. (3.13), α_h is defined in Eq. (3.19), α_{agg} is defined in Eq. (3.20), $A_{\mathrm{bat,op}}$ is defined in Eq. (C.39), A_{agg} is defined in Eq. (3.30), $A_{\mathrm{bat,lim}}$ is defined in Eq. (3.30), A_{m} is defined in Eq. (3.31), and empty matrices 0 here have the adequate shape, in which $n_{\mathrm{bat,op}} = 7H + 3 + H_{\mathrm{crit,bat}} + 2(H-1)$ is the number of lines of $A_{\mathrm{bat,op}}$, in which b_h is defined in Eq. (3.13), b_{agg} is defined in Eq. (3.21), $b_{\mathrm{bat,op}}$ is defined in Eq. (C.38), and $b_{\mathrm{bat,lim}}$ is defined in Eq. (3.31).

Notice that the same principle applies for the construction of both constraints matrices at the household level (Eq. (3.13)) and at the microgrid level (Eq. (3.36)): these constraints matrices consist of the constraints matrices of their individual constituents plus the equations linking these elements into a broader composition. The formulation of the microgrid optimization problem is thus an extension of the work presented in [Sou et al., 2011], from the household level to the microgrid level.

Remark Numerical experiments were conducted using the optimization software "Pyomo" [Hart et al., 2011] with solver "CPLEX" [IBM, 2020]. For some use-cases, the solver is successful only when modeling equality constraints as double inequality constraints [Pyomo Forum - Google groups, 2018]. This situation happened for the case-study at hand,

hence the choice to model the equality constraint of Eqs. (3.8), (3.14) and (3.28) as double inequality constraints in Eqs. (3.9), (3.18) and (3.29) respectively. Note that if the resolution method used was an Interior Points method, this formulation would not have been adequate as the intersection of the inequality constraints transformed into strict inequalities would be empty.

Now that the theoretical framework of the whole microgrid optimization problem has been described, its tractability conditions have to be examined. The next section explores the complexity and data-burden of the microgrid optimization problem, while the following highlights the benefit of the centralized approach on a small and solvable use-case.

3.3 Problem complexity

3.3.1 Time discretization

Each household of the microgrid can be equipped with several controllable appliances, among which EWHs and EVs, notably. Comfort constraints for EVs usually correspond to a target level of energy that has to be available in the morning for the daily rides. electric water heaters (EWHs) comfort constraint can similarly apply in the morning, but also in the evening, as hot water is required for cooking and bathing.

Having constraints in both the morning and the evening imposes that the time horizon H be longer than one day. If the problem is solved in an iterative MPC manner, with a receding horizon, having a time horizon long enough MPC enables to take these constraints sufficiently in advance to improve performances. Typically, the time horizon will correspond to 36 or 48 hours. [Serale et al., 2018] shows that the majority of energy applications of MPC use time horizons of 24 hours, but some works consider longer time horizons.

A timestep of 15 minutes will be considered, as is usually done in smart appliances and smart grid applications. Note that this duration could profitably be shortened to 5 minutes per slot for a more accurate control of white goods load curves, but this would further increase the number of decision variables and the computational burden. [Sou et al., 2011] showed that the time slot length has a significant impact on computation time, while offering only a moderate improvement of the optimal solution. Indeed, in comparison with a 3-minute timestep, scheduling three appliances with a 10-minute timestep divides by 55 the resolution time, but only decreases by 1 percentage point the maximal saving between the worst case (energy bill maximization) and the best case (energy bill minimization). They conclude that "this justifies the use of lower fidelity optimization models, so long as the temporal constraints (e.g. process time bounds) are reasonably captured."

3.3.2 Complexity calculation

According to the modelization proposed by [Sou et al., 2011] (presented in Section C.2), one appliance of the white goods category modeled with H timesteps and J energy phases requires

$$4J + 4HJ + 2(H-1)J + 3H(J-1) \text{ constraints} \tag{3.37}$$

$$3HJ \text{ variables} \tag{3.38}$$

The EWH and EV MILP modelizations, adapted from the framework proposed in [Sou et al., 2011], are detailed in Section C.3 The EWH MILP modelization requires

$$4H + 2(H-1) + H_{\text{crit,ewh}} + 2 \text{ constraints} \tag{3.39}$$

$$3H \text{ variables} \tag{3.40}$$

where $H_{\text{crit,ewh}}$ is the number of critical timesteps for which a specified level of energy is imposed for the hot-water stock. For instance, the hot-water consumption of an household might require to have a high level of available hot-water at 7 a.m., at noon and at 7 p.m.: this constraint would lead to $H_{\text{crit,ewh}} = 3$.

The EV MILP modelization requires

$$7H + 3 + H_{\text{crit,ev}} + 2(H-1) \text{ constraints} \tag{3.41}$$

$$5H \text{ variables} \tag{3.42}$$

where $H_{\text{crit,ev}}$ is the number of critical timesteps for which a specified level of energy is imposed for the EV battery (e.g. the battery must be full every morning for a round-trip).

To these number of variables have to be added the slack variables and constraints necessary to describe the overall household behavior described in Eqs. (3.2)-(3.4) and (3.8), which amounts to $7H$ constraints and $4H$ variables.

Given the J energy phases required to model any appliance of the white goods category, the number of variables and equations required for their modelization is almost directly affected by a factor J. Thus, the MILP modelization for an EWH and an EV are much less cumbersome than the modelization of a white good modeled with a few energy phases, which can reach prohibitive sizes.

To get a first idea on the problem complexity, the size of the constraints matrix modeling each type of appliance are listed in Table 3.2. These results are direct computations from Eqs. (3.37)-(3.42), with $H = 96$ timesteps (a classical 15-minute discretization of a 24-hour day), with the specified number of energy phases J when applicable, and the number of critical timesteps for which a specific level of energy has to be met set to $H_{\text{crit,ewh}} = 2$ for the EWH and to $H_{\text{crit,ev}} = 1$ for the EV.

Table 3.2: The appliances model require very large sparse matrices

Appliance	Number of energy phases	Number of constraints	Number of variables	Share of non-zero elements
EWH	\	578	288	0.69 %
EV	\	866	480	0.41 %
Washing machine	7	5774	2016	0.13 %
Dryer	1	578	288	0.92 %
Dishwasher	5	4042	1440	0.18 %

The results listed in Table 3.2 show the high complexity of an optimization problem that would be limited to even a few appliances. For instance, a house equipped with a single controllable appliance of "dishwasher" type, modeled with $J = 5$ energy phases over $H = 96$ timesteps corresponds to a constraints matrix of shape (4714, 1824)[6], resulting in $4714 \times 1824 = 8\,598\,336$ elements, which would require already a considerable amount of memory (see Section 4.2.1 for more details on this). It shows as well the considerably sparse nature of the constraints matrices used to model the smart appliances. This sparse nature of the constraints matrix could suggest to use adequate methods allowing not to store the high volume of null elements, which will be discussed in the next chapter.

Consider again the use-case outlined in Section 3.1.1. A total of five houses is considered, each equipped according to Table 3.3.

Taking into account only the white goods models (as they are the largest in number of elements), the constraints matrix of the most-equipped house according to the use-case described in Table 3.3 would require to store at least 27 000 non-zero elements. The

[6]Direct calculation from Eqs. (3.37)-(3.38), with addition of the household behavior equations described using $7 \times H$ constraints and $4 \times H$ variables.

	House 1	House 2	House 3	House 4	House 5
PV arrays	X	X	X	X	X
EMS	X	X	X	X	X
EWH	X	X	X	X	X
Washing machine	X	X	X		
Dryer	X				
Dishwasher	X		X		
EV	X	X		X	

Table 3.3: Equipment of each of the microgrid households for the use-case detailed in Section 3.1.1

resulting size of the microgrid problem is such that it might not be handled as is by the solver. This limitation confirms the need of decentralized computing approaches, which has already been recommended and is a current research topic in the demand response literature [Vardakas et al., 2015, Howell et al., 2017].

The high complexity of even a single household optimization problem coincides with the limitation observed in [Sou et al., 2011] regarding the resolution time of the formulated optimization problem. The paper highlights the impossibility for CPLEX to solve the problem with 10 appliances modeled with 6 phases and a 10-minute timestep. This use-case results in 10 (5580, 2112) constraints matrices, hence a total of 117 849 600 elements in the constraints matrices requiring 0.9 GiB as float24 data. The paper hence states that this MILP modelization "does not admit scalable solution algorithms, and it should be restricted to the case of a single household with a few appliances (e.g. less than five)." Although CPLEX cannot find the optimal solution in a limited time for such large problems, the paper details that CPLEX can manage to find a feasible solution for up to 20 appliances.

These elements indicate that the tractability of the proposed approach will depend on the time resolution chosen, the level of detail of the load curves modelization (e.g. the value of H and J), and the number of appliances. The precision of the time grid discretization as well as the number of energy phases considered for the modelization of the appliances in the white goods category have a major impact on the complexity of the problem, but reducing these values would result in a coarser modelization and less realistic results.

3.4 The cost of non-cooperation

Despite the limitations mentioned in the previous section, the interest of the proposed approach persists for some cases where this complexity is limited. This section presents a numerical example showing the extent of the gains attained.

Consider a simple microgrid composed of two individual houses ($l = 2$), of similar composition and identical PV installations (and hence production). The first house owns a controllable dishwasher, the second owns a controllable washing machine (see Appendix C for the details of these models proposed by [Sou et al., 2011]). A battery is shared between both houses.

The two households may decide to cooperate or not. The impact of this choice is now explored.

3.4.1 Cooperative setting

If the two households cooperate, the microgrid EMS optimizes its energy bill by scheduling all the controllable appliances of both households, and the shared battery operation.

The cooperative setting corresponds to solving Problem 3.1. Its output is denoted B_{MG}, the overall bill of the microgrid. The decision variable x is defined as in Eq. (3.35) as the concatenation of x_1, the vector of decision variables of household 1, x_2 the one of household 2, and $x_{\text{agg}}, x_{\text{bat}}, x_{\text{m}}$ the decision variables necessary for the description of the aggregated household consumptions, of the battery operations, and of the microgrid overall load curve.

3.4.2 Non-cooperative setting

In the non-cooperative setting on the other hand, the overall cost of supplying energy to the microgrid is the sum of the energy bills of the individual households that have optimized their own consumptions without supervision from the microgrid EMS, plus the cost of optimal operation of the shared battery by the microgrid EMS.

First, each household EMS h controls its appliance under its overall constraints summarized in Eq. (3.12), to minimize its household electricity bill, defined as

$$\sum_{t=1}^{H} \left(c_{\text{d}}^t P_{h,\text{d}}^t - c_{\text{s}}^t P_{h,\text{s}}^t \right) \Delta t \tag{3.43}$$

with c_{d}^t the cost for deficit energy bought from the microgrid at timestep t, and c_{s}^t the cost for surplus energy sold from the household to the microgrid at timestep t.

The optimization problem of household $h \in 1, 2$ is summarized as

$$\begin{aligned} B_h = &\min c_h^\top x_h \\ &\text{s.t. } A_h x_h \leq b_h \end{aligned} \tag{3.44}$$

with

$$c_h = \begin{pmatrix} \mathbb{0}_{n'_h}^\top & -(c_{\text{s}}^t)_{t\in[1,H]}^\top \Delta t & (c_{\text{d}}^t)_{t\in[1,H]}^\top \Delta t & \mathbb{0}_{4H}^\top \end{pmatrix}^\top \tag{3.45}$$

and $x_1^\star$ and $x_2^\star$ denote the respective corresponding solutions.

Then, in this non-cooperative configuration, once each household has computed its optimal appliances schedule ($x_1^\star$ and $x_2^\star$), the microgrid EMS collects these solutions and schedules the shared battery operation according to its operational and limit constraints.

It is assumed that the energy prices when buying and selling, $(c_{\text{d}}^t)_{t\in[1,H]}$ and $(c_{\text{s}}^t)_{t\in[1,H]}$, are the same for both households, and are the same as the ones imposed by the main grid to the microgrid. Hence, the microgrid EMS operates the shared battery and transfers the main grid energy prices without generating a profit for itself. The microgrid EMS solely acts as a support to the microgrid participants, and the cost of operating the battery has to be divided among the participants[7].

In this configuration, the microgrid EMS only controls the battery and considers $x_1^\star$ and $x_2^\star$ as inputs instead of decision variables. Thus, the aggregated households behavior in Eq. (3.18) is rewritten with $l = 2$ households into

$$\alpha_{\text{agg}} x_{\text{agg}} \leq b_{\text{agg}} - \alpha_1 x_1^\star - \alpha_2 x_2^\star \tag{3.46}$$

The optimization problem of the microgrid EMS is finally

$$\begin{aligned} B_{\text{bat}} = &\min_{x_{\text{agg}}, x_{\text{bat}}, x_{\text{m}}} \sum_{t=1}^{H} \left(c_{\text{d}}^t P_{\text{gm}}^t - c_{\text{s}}^t P_{\text{mg}}^t \right) \Delta t - c_1^\top x_1^\star - c_2^\top x_2^\star \\ &\text{s.t. } (3.22), (3.29), (3.46) \end{aligned} \tag{3.47}$$

[7] Another option could have been for the microgrid EMS to act as an independent entity and try to maximize its profits by adding a margin to the prices from the main grid – this is not the case here.

with $x_{\text{agg}}, x_{\text{bat}}$ and x_{m} respectively defined in Eqs. (3.21), (3.23), (3.31). When positive, B_{bat} is the additional cost incurred for operating the battery. When negative, B_{bat} is the profit generated by the battery operation. In either case, this amount has to be distributed among the agents, and is added to the individual agents energy bills B_1 and B_2 to calculate the overall energy cost.

Ultimately, the overall energy cost for the microgrid in the non-cooperative setting is the sum $B_1 + B_2 + B_{\text{bat}}$.

3.4.3 Cost comparison

The cost of non-cooperation corresponds to the difference between the overall energy cost of the two configurations, which can be written using the previous notations as:

$$B_1 + B_2 + B_{\text{bat}} - B_{\text{MG}} \tag{3.48}$$

In this example, the microgrid is located near Fontainebleau, in the French region of Ile-de-France. The considered shared battery capacity is $C = 8\,\text{kWh}$. The price of electricity considered is the current Peak/Off-Peak retail tariff from EDF, the major French utility: 18.2 c€/kWh (Peak) and 13.6 c€/kWh (Off-Peak). The subsidy for PV surplus upstream injection is the current flat rate for small domestic installations, fixed by French regulation: 10 c€/kWh. In this example, these prices are considered for the power exchanges between the households and the microgrid, as well as between the microgrid and the external main grid. Both individual houses are assumed to be equipped with the same $3\,\text{kW}_\text{p}$ PV panels, having the same orientation and exposure, thus exactly the same production. The PV production considered here corresponds to a uniformly cloudy day. The prices as well as the PV production are represented in Figure 3.3.

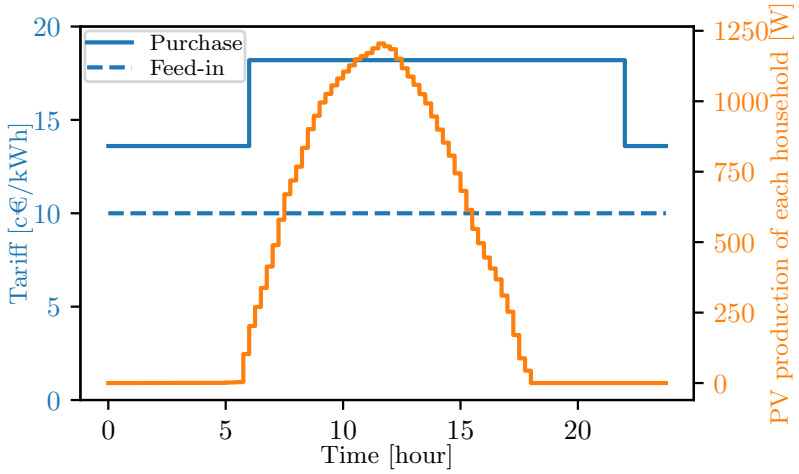

Figure 3.3: Examples of electricity purchase and feed-in tariffs, and PV production of each household for a late winter day. 15-minute timesteps.

The non-controllable power consumption of each household is represented in Figure 3.4. The non-controllable power consumption data is taken from the public "Individual household electric power consumption" dataset [UCI Machine Learning Repository, 2012] containing 47 months of power measurements at a one-minute timestep, from a house located in Sceaux, in the French region of Ile-de-France. Among the measurements are the household total power consumption as well as sub-meterings corresponding to the kitchen, the laundry-room, and the EWH. This allows to select, as the non-controllable load curve for the first household equipped with a controllable dishwasher, the total household power consumption minus the metered load curve of the kitchen. Similarly, the non-controllable

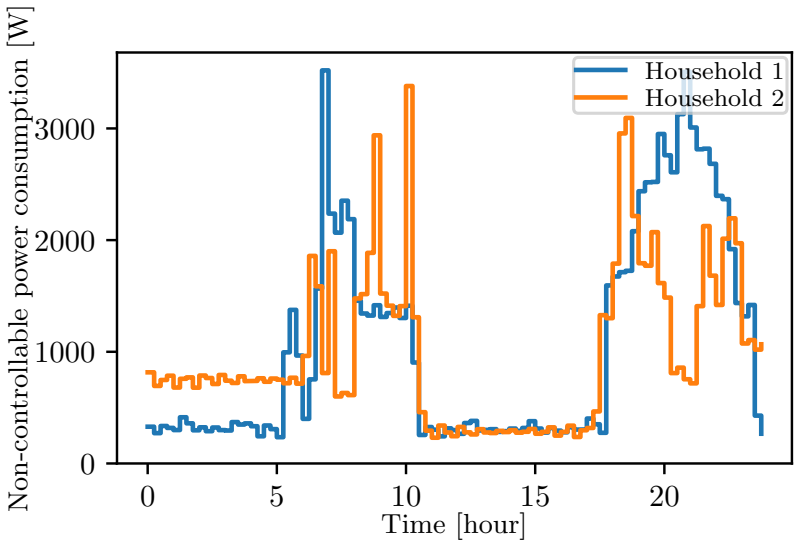

Figure 3.4: Examples of the non-controllable power consumptions for a day.

load curve for the second household equipped with a controllable dishwasher corresponds to the total household power consumption minus the consumptions from the laundry room. The data is re-sampled at a 15-minute timestep.

The addition of the PV production and the non-controllable power consumptions yield a PV surplus production curve available for the controllable appliances to use. The available PV surplus of each household is displayed in Figure 3.5. The benefit of cooperation will come from the concerted use of the shared battery, but also from the pooling of these PV production surpluses.

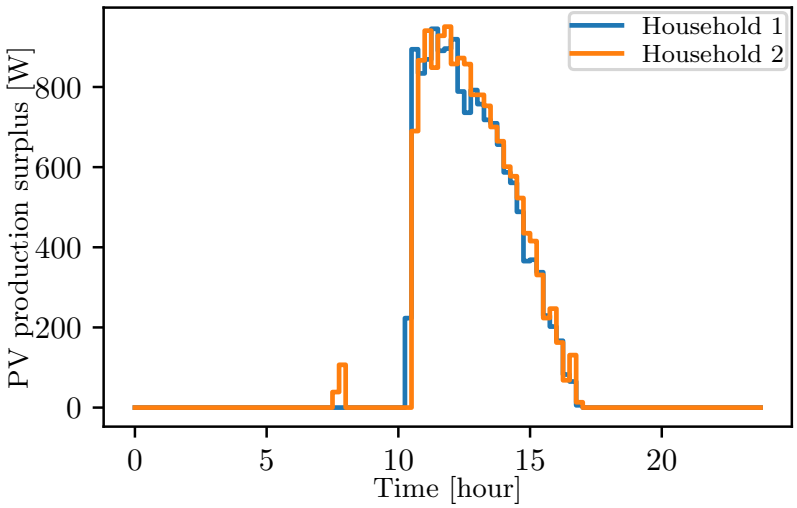

Figure 3.5: The PV production surplus of the each household is available for self-consumption.

Here, the local PV production is supposedly known in advance for both optimization cases. Figure 3.6 reports the load curves or the controllable appliances resulting from optimal scheduling in the two different settings[8]. The upper two figures of Figure 3.6

[8]Notice that the appliances load curves displayed in Figure 3.6 can slightly differ from one configuration to the other, as the MILP formulation of the controllable appliances constraints allows for some flexibility in how is allotted the energy required for each energy phase.

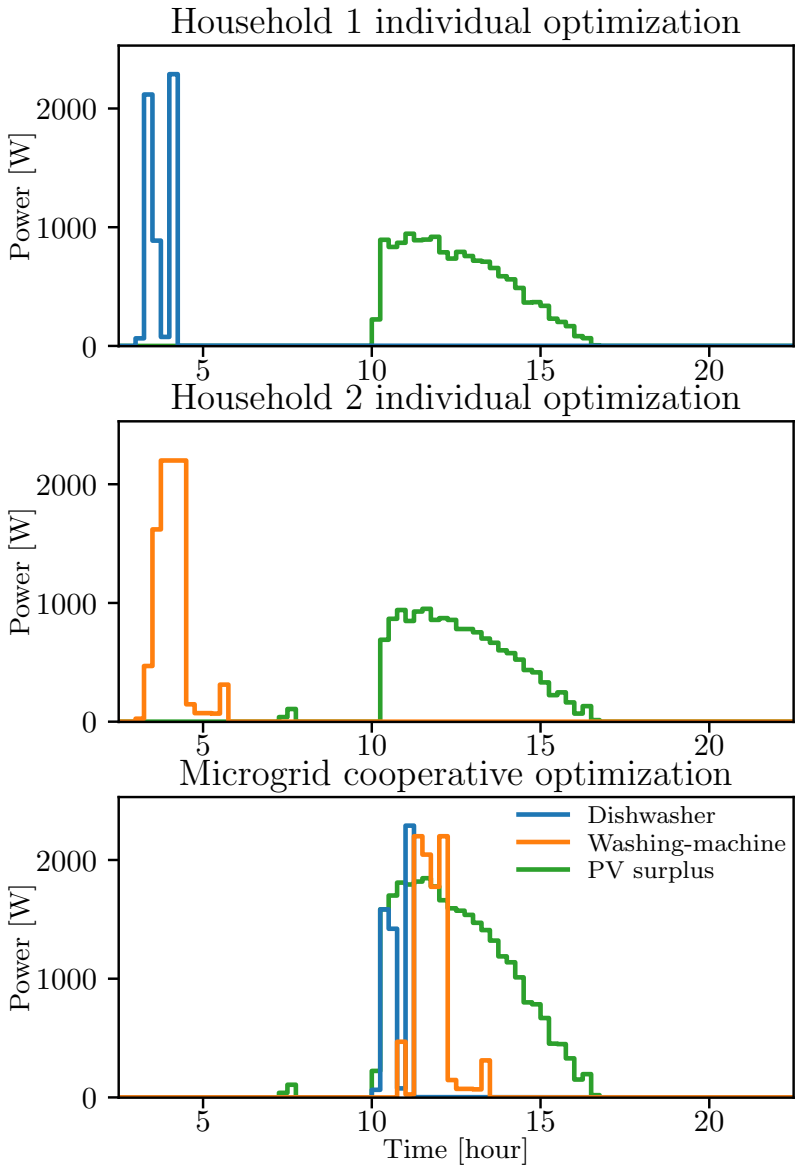

Figure 3.6: Resulting schedule without cooperation (upper two figures) and with cooperation (bottom figure).

correspond to the non-cooperative case, where each household EMS chooses the optimal scheduling of their own appliance, $x_1^\star$ and $x_2^\star$ respectively, according to the supply electricity price and to their own PV production and uncontrolled consumption. Both the dishwasher of the first household and the washing machine of the second household are scheduled at the same time when the cost of electricity is low, because their individual PV surplus is two small for self-consumption to be profitable. The individual costs B_1 and B_2 correspond to these load curves. After each household has chosen its appliance schedule and has transmitted its resulting overall load curve to the microgrid EMS, the battery schedule is chosen leading to the cost B_{bat}.

The resulting overall cost for one day in the non-cooperative case is

$$B_1 + B_2 + B_{\text{bat}} = 5.7996\,€ \tag{3.49}$$

The lower figure of Figure 3.6, on the contrary, corresponds to the cooperative case, where the microgrid EMS collects all the information needed to optimally control both appliances and the shared battery, with the objective of minimizing the total microgrid electricity bill B_{MG}. In the cooperative configuration, the microgrid EMS pools the available PV surpluses, thus allowing self-consumption, and schedules the appliances successively, thus smoothing the overall load.

For this single day, the resulting cost of the microgrid optimization problem is

$$B_{\text{MG}} = 5.7004\,€ \tag{3.50}$$

Hence a cooperation scheme can lead to save up to 1.7 % on the energy cost in this two households case. A more in-depth economical evaluation of this cost of non-cooperation, with respect to a set of hypothesis (on the number of potential microgrids, on their size and composition, on the household profiles, on the prices of energy) was not the subject of the present work, but would be a valuable addition. As the non-cooperative setting already includes some partial cooperation through the use of a shared battery operated with the aim of minimizing the collective cost of energy, the relative gain is quite limited. Yet, other studies available in the literature have shown that the relative gain would be larger if individual batteries where operated for the interest of individual households in the non-cooperative setting [Parisio et al., 2017], or if energy storage was not present at all in both settings [Carli and Dotoli, 2019]. A consideration of the broader cost structure of electricity, including upstream and local grid costs, leads to increased savings attained by the cooperative setting [Hupez et al., 2021]. It can be reasonably expected that a larger degree of freedom, i.e. additional controllable appliances, would increase the attained savings. This cost difference generates the interest for a cooperation protocol yielding the optimal appliances schedule for the microgrid optimization problem.

For more details on the two optimization results, Figure 3.7 presents the aggregated agents load curve, the battery schedule and the microgrid imports and exports from the outer grid, for each configuration. Extra steps would be required to define the individual power bills of both households, as discussed in Section 5.2.

3.5 Conclusion

This chapter has detailed how the microgrid appliances scheduling optimization problem can be formulated as a MILP, and is composed of the elements of the participant households' own optimization problems. Major computational and memory limitations of this approach have been highlighted, but the benefit of this approach has been proven on a small example, where a 1.7 % cost reduction is attained thanks to the cooperative nature of the microgrid optimization. It could be larger when the number of participants is increased.

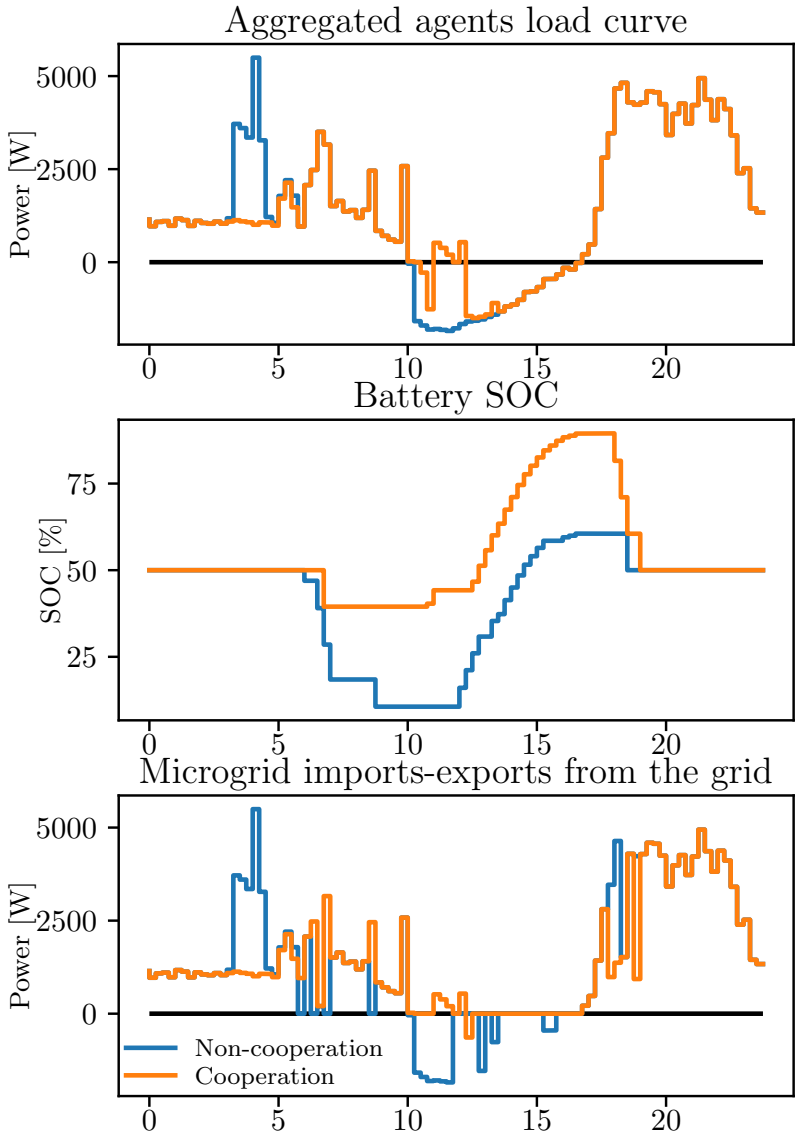

Figure 3.7: In the cooperative case, pooling the PV resources allows the microgrid EMS to import less electricity at the beginning and at the end of the day.

However the cooperative approach generates privacy-protection concerns, as the energy demand of a household contains personal information that agents might not want to disclose to neighbors nor to a third party entity. It seems especially important that the exchanged data for the collaboration scheme does not allow to infer personal details such as wealth, habits, equipment, daily routine, number of occupants.

Hence, while solving the microgrid problem is the optimal collaboration scheme for an energy community, collecting all the information necessary to formulate it requires to build a privacy-preserving communication protocol in order to hide which consumption detail corresponds to which agent.

The next chapter will try and detail a method allowing to transmit the required information from the household EMSs to the microgrid EMS, with the imperative of hiding the private data of a participant to its neighbors, and the identity of the participant owning the data transmitted to the microgrid.

Chapter 4

Privacy-preserving cooperation

Protocoles de communication au sein d'un micro-réseau avec protection des données personnelles. Ce chapitre aborde la question de la coopération au sein d'un micro-réseau avec protection de l'anonymat des agents. En s'inspirant de techniques bien connues de cryptographie, trois protocoles sont proposés pour communiquer les contraintes des gestionnaires d'énergie des habitations jusqu'à celui du micro-réseau. Ces protocoles permettent ainsi de formuler le problème d'optimisation centralisé présenté au chapitre précédent, tout en garantissant l'anonymat des données échangées. Le temps de calcul additionnel nécessaire à chacune de ces trois méthodes est comparé sur un cas d'étude simple à deux habitations et deux usages pilotés. Certaines questions ouvertes par l'anonymisation des participations au sein d'un micro-réseau sont listées.

4.1 Introduction

In the previous chapter, the construction of an energy optimization problem encompassing all the households controllable power consumptions at the microgrid level has been described. Despite the mentioned computational complexity limitations, the interest of the approach on a small scale use-case has been demonstrated.

Implicitly, all data were shared between participants of the grid, and for this reason one can consider that the microgrid EMS had on omniscient view of the situation. In mode details, in its present form, the construction of such an *omniscient* optimization problem requires for the microgrid energy management system (EMS) to collect all data and, in particular, to have knowledge of sensitive information such as the number and type of appliances owned by an household or the time preferences for their operation.

To comply with the privacy protection aspirations and legal requirements on privacy preservation, the challenge of the present chapter is to anonymously transmit the necessary information from the participants to the microgrid EMS in order to formulate a problem equivalent to the omniscient one for reaching the same solution. This property is referred to as "privacy preservation".

4.1.1 A brief review on decentralized cooperation in microgrids

Decentralized optimization approaches can have the benefit of keeping the agents constraints private and letting them perform their own optimization. However such methods were not chosen for the use-case at hand for reasons detailed in the following paragraphs.

Decentralized cooperation in microgrids through game theory approaches

As the cost of energy for one agent in a microgrid can depend on its neighbors decisions as well as its own, the problem of optimizing one's energy bill for an agent in a microgrid can be stated in the framework of game theory. Regarding the privacy of personal data, game-theoretic methods have the benefit of operating in a decentralized manner: each agent can keep its constraints private and perform its own optimization, depending on the cost provided by the microgrid EMS, and the behavior of other agents. Furthermore, these decentralized approaches can scale up to large systems.

In [Jacquot et al., 2018], the demand response framework applied to a microgrid and the associated hourly billing mechanism is formulated under the form of a game. The mechanism is proved to converge to a Nash equilibrium after a limited number of iterations. The final privacy step would be to anonymize the communication of an agent's optimal load curve at the end of an iteration to the central operator. This chapter will present useful tools to perform this additional privacy requirement.

In [Xie et al., 2020], a local energy market is described as a game with buyer and seller coalitions. Secure multiparty computation and encryption protocols are used to preserve the agent's private data. The optimal trading price between both coalitions is computed for each minute and the method is scalable.

In general, the Nash Equilibrium of game-theoretic methods does not achieve collective optimality[1]. Similarly as for the algorithm developed in Chapter 1, the aim is to reach the optimal organization. This is why this manuscript focuses on methods from control theory rather than game theory.

Decentralized cooperation in microgrids through control approaches

[Paridari et al., 2015] proposes a decentralized computation of the optimal load curve of the agents in a microgrid, in cooperation thanks to a central operator communicating a target signal to each agent.

Similarly, in [Jacquot et al., 2019] the agents respond to a target signal to optimize their own load curve in a decentralized manner, but the data required for the aggregator problem formulation is collected through secure multiparty computation, thus protecting the agents' private data. In this paper, the cooperation between the microgrid consumers results from the proposed "Alternate Projections Method", and the preservation of privacy comes from the "Secure Multiparty Computation of Aggregate" method. The convex nature of the admissible set of consumption profiles of each agent allows to perform the projection of the power production distribution desired by the central operator onto theses admissible sets.

In the problem at hand in this manuscript, the detailed modeling of electrical appliances such as white goods, electric water heaters (EWHs) and Electric Vehicles (EVs) is included. It creates non-convex admissible sets for the load curves of the microgrid agents. Thus the non-convexity of the admissible sets of consumption profiles considered here forbids to apply the methods proposed in [Jacquot et al., 2019], but some common methods will be reused such as the "Secure Multiparty Computation of Aggregate" of this paper that corresponds to the "Secure Sum" algorithm from [Yao, 1986] which is presented in the next section.

The approach chosen here to create cooperation among the microgrid participants is to anonymously communicate all the data necessary to control the agents' controllable

[1]Note that some Nash equilibria can be proven to reach the collective optimum, or to have a bounded price of anarchy (as in [Jacquot et al., 2018]), which is the ratio of the optimum collective cost of the system and the collective cost of the worst Nash Equilibrium, providing a measure of the efficiency of a mechanism in game theory.

appliances to the microgrid EMS, responsible for formulating and solving the microgrid optimization problem.

4.1.2 Problem definition

The challenge considered in this chapter is to formulate an optimization problem yielding the same solutions as the omniscient problem of a microgrid, while hiding the correspondence between a participating household and its private information.

The omniscient problem, as defined in Problem 3.1, is

Problem 4.1. [Problem 3.1, Omniscient problem]

$$\begin{aligned} &\min \begin{pmatrix} \mathbb{0}_{n_1}^\top & \dots & \mathbb{0}_{n_l}^\top & c^\top \end{pmatrix}^\top x \\ &\text{s.t. } A\,x \leq b \end{aligned} \tag{4.1}$$

where the decision variable x is defined in Eq. (3.35) as

$$x = \begin{pmatrix} x_1^\top & x_2^\top & \dots & x_l^\top & x_{\text{agg}}^\top & x_{\text{bat}}^\top & x_{\text{m}}^\top \end{pmatrix}^\top \tag{4.2}$$

and the cost vector from Eq. (3.34) is simplified into

$$c = \begin{pmatrix} \mathbb{0}_{4H}^\top & \mathbb{0}_{7H}^\top & (c_{\text{d}}^t)_{t\in[1,H]}^\top \Delta t & -(c_{\text{s}}^t)_{t\in[1,H]}^\top \Delta t \end{pmatrix}^\top \tag{4.3}$$

with c_{d} and c_{s} the costs for deficit energy bought from or surplus energy sold to the outer grid at timestep t, Δt the timestep length, and the constraints matrix A and the right-hand side vector b are defined in Eq. (3.36) as

$$A = \begin{pmatrix} A_1 & 0 & \dots & 0 & & & \\ 0 & A_2 & \ddots & 0 & & 0 & \\ \vdots & \ddots & \ddots & \ddots & & & \\ 0 & 0 & \ddots & A_l & & & \\ \alpha_1 & \alpha_2 & \dots & \alpha_l & \alpha_{\text{agg}} & \mathbf{0}_{7H,7H} & \mathbf{0}_{7H,2H} \\ & 0 & & & \mathbf{0}_{n_{\text{bat,op}},4H} & A_{\text{bat,op}} & \mathbf{0}_{n_{\text{bat,op}},2H} \\ & & & & A_{\text{agg}} & A_{\text{bat,lim}} & A_{\text{m}} \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_l \\ b_{\text{agg}} \\ b_{\text{bat,op}} \\ b_{\text{bat,lim}} \end{pmatrix} \tag{4.4}$$

in which the various matrices and dimensions are defined in Problem 3.1.

The private information owned by the households and necessary for the problem formulation and resolution is composed of a constraints matrix A_h and a right-hand side vector b_h for each household $h \in [1, l]$ (see Eq. (3.36)).

Assume that the households decision variables are standardized (which will be considered in the remainder of this chapter, and detailed in the following section). Then the households decision variables x_h all have the same structure and dimension $n_h = n$, $\forall h \in [1, l]$, and the matrices α_h, $h \in [1, l]$ (described in Eq. (3.19)) are all equal and simply denoted $\alpha \in \mathbb{R}^{7H,n}$.

Under this assumption, one can reformulate Problem 4.1 as follows

Problem 4.2. [Problem 4.1 reformulated]
Solve, for $x = \begin{pmatrix} x_1^\top, \dots, x_l^\top, y^\top \end{pmatrix}^\top$,

$$\begin{aligned} &\min_x c^\top y \\ &\text{s.t. } \begin{cases} A_h\, x_h \leq b_h, \quad h = 1, \dots, l \\ \alpha_{\text{sys}} \displaystyle\sum_{h=1}^{l} x_h + A_{\text{sys}}\, y \leq b_{\text{sys}} \end{cases} \end{aligned} \tag{4.5}$$

with

$$\alpha_{\text{sys}} = \begin{pmatrix} \alpha \\ \mathbf{0}_{n_{\text{bat,op}}+10H,\,n} \end{pmatrix}, \; A_{\text{sys}} = \begin{pmatrix} \alpha_{\text{agg}} & \mathbf{0}_{7H,7H} & \mathbf{0}_{7H,2H} \\ \mathbf{0}_{n_{\text{bat,op}},4H} & A_{\text{bat,op}} & \mathbf{0}_{n_{\text{bat,op}},2H} \\ A_{\text{agg}} & A_{\text{bat,lim}} & A_{\text{m}} \end{pmatrix}, \; b_{\text{sys}} = \begin{pmatrix} b_{\text{agg}} \\ b_{\text{bat,op}} \\ b_{\text{bat,lim}} \end{pmatrix} \tag{4.6}$$

Notice that this problem is invariant by permutation of the decision variables x_h. More precisely, consider the following problem:

Problem 4.3. For a given permutation σ of [1, ..., l] and for $x = \left(x_{\sigma(1)}^\top, \ldots, x_{\sigma(l)}^\top, y^\top\right)^\top$,

$$\begin{aligned} &\min_x \; c^\top y \\ &\text{s.t.} \; \begin{cases} A_{\sigma(h)}\, x_{\sigma(h)} \leq b_{\sigma(h)}, \quad h = 1, \ldots, l \\ \alpha_{\text{sys}} \sum_{h=1}^{l} x_{\sigma(h)} + A_{\text{sys}}\, y \leq b_{\text{sys}} \end{cases} \end{aligned} \tag{4.7}$$

with α_{sys}, A_{sys} and b_{sys} defined in Eq. (4.6).

Problems 4.2 and 4.3 are trivially equivalent. Thus, for the microgrid EMS to solve Problem 4.2, it is sufficient to know A_h and b_h, $h = 1, \ldots, l$ up to a given permutation (notice that the variables y and the matrices α_{sys}, A_{sys}, b_{sys} concern general specifications of the microgrid and are thus known by the microgrid EMS). This inspired the core of the privacy-preserving methodologies proposed in this chapter. Note that it is not necessary to know σ to solve Problem 4.3.

4.1.3 Content

In this chapter, three approaches are considered to ensure identity protection when building the microgrid optimization problem of a microgrid. The first approach, Protocol A (Section 4.3), is to follow an encrypted communication protocol that allows each agent to securely send information to the microgrid EMS, and to provide a masked return address for the microgrid EMS to use. The latter can then communicate the result of the optimization back to the agent, without knowing which agent is addressed. The data are impossible to read and the identities remain unknown. This first approach could be expected to be the most costly one in terms of computing times, due to the encryption of all the information, but allows to keep the result of the optimization secret. The method uses the encrypted communication protocol presented in [Chaum, 1981].

The second approach, Protocol B_1 (Section 4.4.2), is to anonymously assign a unique identifier (ID) to each participant in the microgrid, in such a way that neither an agent nor the microgrid EMS can know the ID of the appliances of its neighbors. The association of the households with the set of IDs forms the permutation σ involved in the microgrid optimization problem 4.3, which can be solved by the microgrid EMS. The result of the microgrid problem (or an encrypted version of it, potentially) is then openly broadcast. All the tools necessary to perform this protocol are presented in [Dunning and Kresman, 2013] and recalled here in Section 4.4.1 for clarity.

Finally, a variant of the previous approach is proposed in Protocol B_2 (Section 4.4.2). It applies the same unique ID assignment and secure multiparty computation, this time to transmit both private data and public encryption keys from the agents to the microgrid EMS. This allows to keep private the solution vector of each agent returned by the microgrid EMS.

In any case, once the information is collected through one of the proposed privacy-preserving communication protocols, the microgrid EMS is responsible for solving the microgrid optimization problem 4.3, including choosing the charging and discharging cycles

of the battery with the aim of minimizing the supply cost of energy of the overall microgrid, and is finally responsible for sending the results to the agents.

Before the communication schemes are presented, the next section details the sparse data compression that can be performed to save significant volumes of data transmission and consequently save significant amounts of computational burden in the protocols.

4.2 Data formats

Notice that the obviously sparse nature of the agents' constraints matrices A_h (see definition (3.13)) and their large dimensions invite to use adequate algorithms for their transfer from the agents to the microgrid EMS and thus save a non-negligible amount of time.

Either a sparse data transfer or a standardized data structure can be adopted. This is exposed below.

4.2.1 Sparse data communication

Sparse matrix transfer consists in communicating three elements (value, row index, column index) per non-zero element, instead of communicating all the elements of the matrix. The non-communicated elements are considered to be null by default (see e.g. [Golub and Van Loan, 2013]).

The interest of such simple data compression is made clear with a simple example. As mentioned in the previous chapter (Section 3.3), a house equipped with a single controllable appliance of "dishwasher" type, modeled with $J = 5$ energy phases[2] over $H = 96$ timesteps (a classical 15-minute discretization of a 24-hour day) ends up with $4714 \times 1824 = 8\,598\,336$ elements. Only 11 783 of these elements are non-zeros (0.14 %). For memory use reduction, it is considered that non-zero elements are stored under the half-precision floating-point format where each element is stored using 2 bytes (B), i.e. 16 bits (instead of using the double-precision floating-point format intended for larger float variables and requiring 8 bytes). It is also considered that rows and columns indices are stores as 16-bit unsigned integers for the same reason. Thus, transmitting this matrix as a sparse matrix made results in exchanging 3 elements (value, row and column indices) of 2 bytes for each of the 11783 non-zero values, thus exchanging $11\,783 \times 3 \times 2 = 70\,698$ B $\simeq 0.07$ MiB, whereas transmitting the whole matrix would correspond to transmitting approximately 16 MiB (the Mi binary prefix denotes 2^{20} B), which is impracticable.

4.2.2 Data standardization

As each household can be composed of any combination of smart appliances (see for instance Table 3.3), each household constraints matrix A_h can be of different size and composition. Using the sparse data communication method, as discussed in the previous section, thus requires to exchange the location (row and column indices) of all non-zero values to exchange a matrix.

On the other hand, a standard composition can be assumed for all households, leading to a standard structure for the constraints matrices. The standard structure being known to all participants in the microgrid, only the values of the elements of interest have to be transmitted, instead of sending also the row and column indices. A value of interest corresponds to an element in the matrices that is not structurally null. This group of values encompasses the above-mentioned non-zero elements, but is not restricted to it, as some zeros will be variables that could have been non-null.

[2] An energy phase is an uninterruptible sub-task of the operation of an appliance in the white goods category, which uses a pre-specified amount of electric energy, see Section C.2 for more details.

In this scheme, the number of elements whose values are transmitted is higher, due to the communicated zeros, but the size of data transmitted for each non-zero value is divided by 3 in comparison with the sparse data communication scheme, thanks to the rows and columns indices not communicated.

This approach requires some assumptions to be validated:

- each household constraints matrix has to host a standardized list of appliances constraints matrices, e.g. one EWH, one EV, one of every type of white goods, in a specific order. If a household does not own a specific appliance, the corresponding matrix block has to be left entirely null and transmitted as such, and cannot be removed.
- the content of the appliances model constraints matrices A_{hi} has to be standardized: each appliance i has to use the standard structure of its type (one structure for each type of the white goods category, one structure for the EWHs, one structure for the EVs), with a specified number of energy phases if applicable, and a standardized order in the decision variables.
- all the household EMSs and the microgrid EMS agree on this overall structure.

These standardization requirements ensure that each household constraints matrix follows the same structure, is composed of the same blocks defining the appliances, with key variables located in the same spots[3].

As defined in Eq. (3.13), each household constraints matrix A_h is composed of standardized blocks of appliances $A_{h,i}$, of common behavior blocks $\alpha_{h,\text{def}}$, $\alpha_{h,i}$ and $\alpha_{h,\text{join}}$, and of large empty blocks, with $i \in [1, m_h]$ and $h \in [1, l]$. The standardization enables to communicate only the values of the variables of interest (not structurally zeros) of each of these blocks, written in vectors denoted $\overline{A}_{h,i}$, $\overline{\alpha}_{h,\text{def}}$, $\overline{\alpha}_{h,i}$ and $\overline{\alpha}_{h,\text{join}}$, with $i \in [1, m_h]$ and $h \in [1, l]$.

These vectors of the non-null elements of the blocks composing A_h are concatenated in a large vector, named the compressed vector

$$\overline{A}_h = \begin{pmatrix} \overline{A}_{h,1}^\top & \overline{A}_{h,2}^\top & \dots & \overline{A}_{h,m_h}^\top & \overline{\alpha}_{h,\text{def}}^\top & \overline{\alpha}_{h,1}^\top & \overline{\alpha}_{h,2}^\top & \dots & \overline{\alpha}_{h,m_h}^\top & \overline{\alpha}_{h,\text{join}}^\top \end{pmatrix}^\top \tag{4.8}$$

The length of this standardized compressed vector is noted $\overline{n}$. For reference, considering that the minimal standard corresponds to a household composed of one EWH, one EV, one dishwasher, one washing-machine and one dryer, the length of $\overline{A}_h$ would be $\overline{n} = 34\,660$.

The data standardization defines unambiguously to which appliance corresponds a given block of variables in $\overline{A}_h$, and which element in the original block matrix corresponds to the communicated variable of interest. Accordingly, these standardization rules yield a vector $\overline{b}_h, \forall h \in [1, l]$ containing the ordered and variables of interest of each of the block vectors composing $b_h, h = 1, \dots, l$. The length of this standardized compressed vector is noted $\overline{m}$.

The data standardization also establishes the order and type of the decision variables. The microgrid EMS is thus able to formulate and solve the problem.

The obvious drawback of this method is that the standardization requires to transmit an important number of potentially null values: if a household is only equipped with one appliance, it still has to transmit a significant amount of zeros for the elements of interest of the matrices of the other appliances[4].

[3] The standardization requires that each type of white goods is modeled using a number of energy phases corresponding to the worst case of their potential operations. If the appliance is used in a way that requires less energy phases, the remaining elements of the unnecessary energy phases can be left empty.

[4] To avoid transmitting too much unnecessary values, it is important to define a minimal standard. For example, if no participant owns a dishwasher, then the standard constraints matrix of a dishwasher and its elements of interest does not have to be considered in the standard household matrices.

The computational burden of the communication methods presented in the two following sections is impacted by the size of the data transmitted. For this reason, it is considered that all these methods are conducted using the compressed vectors $\overline{A}_h$ and $\overline{b}_h$, for $h = 1, \ldots, l$, containing the variables of interest of the constraints matrices A_h and right-hand side vectors b_h.

Now are detailed three methods that could be used to ensure identity protection when transmitting the necessary information from the participants to the microgrid EMS.

4.3 Encryption-based communication (Protocol A)

The first approach presented in this section is based on encrypted communication protocols. [Chaum, 1981] provides the tools to hide the identity of an agent (in the present case, a house) sending encrypted data to a central operator (in the present case, the microgrid EMS), and to return the result of the optimization problem to the corresponding agent without the central operator knowing its identity. This paper is the main source of inspiration in this section.

4.3.1 Definitions

First of all, the whole data of an agent h in its compressed version (see Section 4.2.2), comprised of its constraints matrix $\overline{A}_h$ and its right-hand side vector $\overline{b}_h$, is denoted $\overline{D}_h$.

On one hand, as defined in [Chaum, 1981], a public key cryptosystem uses a pair of keys K and Inv(K). "The public key K is made known to the other users or anyone else who cares to know it; the private key Inv(K) is never divulged." The encryption of a message X with key K is denoted $\mathrm{K}(X)$, and the keys are inverses in the sense that

$$\mathrm{Inv(K)}(\mathrm{K}(X)) = \mathrm{K}(\mathrm{Inv(K)}(X)) = X \tag{4.9}$$

This public/private key pair is also named "asymmetric". On the other hand, a "symmetric" encryption key K is a key that is its own inverse:

$$\mathrm{K}(\mathrm{K}(X)) = X \tag{4.10}$$

In order to forbid a listener to guess that two messages are equal ($Y = X$) by checking whether their encryption is the same ($\mathrm{K}(Y) = \mathrm{K}(X)$), a large string of random bits R is attached to X before encrypting. The encrypted message is then denoted $\mathrm{K}(R,X)$.

The cryptosystem includes two additional intermediary "Bots" that will process each message sent before it is delivered.

The purpose of the Bot of type 1 is to hide the correspondences between the items in its input and those in its output, because it hides the order of arrival by outputting the items in batches. Thus a listener cannot know which message received by the microgrid EMS comes from what agent. It performs the following transformation for a message sent from an agent to the microgrid EMS:

$$\mathrm{K_{m1}}(\mathrm{R_1}, \mathrm{K_{MG}}(\mathrm{R_0}, D), \mathrm{M_{MG}}) \rightarrow \mathrm{K_{MG}}(\mathrm{R_0}, D), \mathrm{M_{MG}} \tag{4.11}$$

where $\mathrm{K_{m1}}$ is the first Bot public key, $\mathrm{K_{MG}}$ is the microgrid EMS public key, $\mathrm{K_{MG}}(\mathrm{R_0}, D)$ is the encrypted message containing the data D, $\mathrm{M_{MG}}$ is the microgrid EMS address in the local communication network, and $\mathrm{R_0}$ and $\mathrm{R_1}$ are two random strings. The Bot decrypts its input with its private key, throws away the random string $\mathrm{R_1}$, outputs the remainder, and finally forwards the sealed messages $\mathrm{K_{MG}}(\mathrm{R_0}, D)$ of the output to $\mathrm{M_{MG}}$ the address of the microgrid EMS who then decrypts the message with their own private key.

The second category of Bot allows the microgrid EMS to respond to the agent while still keeping the identity of the agent secret from the microgrid EMS. The solution is for the agent to form an untraceable return address $\mathrm{K_{m2}}(\mathrm{R}_2, \mathrm{M}_h)$ hiding its real address M_h. Here, R_2 is another key that will also act as a random string for blurring. The masked return address can be sent from the agent to the microgrid EMS as part of the first message.

The second kind of Bot performs this transformation of the input created by the microgrid EMS:

$$\mathrm{K_{m2}}(\mathrm{R}_2, \mathrm{M}_h), \mathrm{K}_h(\mathrm{R}_3, x_h^\star) \rightarrow \mathrm{M}_h, \mathrm{R}_2(\mathrm{K}_h(\mathrm{R}_3, x_h^\star)). \tag{4.12}$$

where $\mathrm{K_{m2}}$ is the second Bot public key, R_2 and R_3 are two random strings, K_h is the agent public key, and $x_h^\star$ is the part of the solution of the optimization problem that involves the agent control variables. This Bot uses the string of bits R_2 that it finds after decrypting the address part $\mathrm{K_{m2}}(\mathrm{R}_2, \mathrm{M}_h)$ as a key to re-encrypt the message part $\mathrm{K}_h(\mathrm{R}_3, x_h^\star)$. Only the addressee h can decrypt the resulting output because h created both R_2 and K_h.

Both these Bots are used as intermediary for the messages transmission system, providing additional protection.

The concepts presented in this section are combined into a communication protocol adapted to the problem at hand.

4.3.2 Encrypted communication protocol

M_h denotes the real address of agent h in the local communication network, and K_h denotes the symmetric encryption key it generates. The asymmetric pair of public/private keys generated by an entity e is noted $\mathrm{K}_e, \mathrm{K'}_e$.

Below are enumerated the successive steps of the proposed encrypted communication protocol. Table 4.1 lists the acquired knowledge by each entity at each step. This protocol is simply referred to as Protocol A thereafter.

0. Agent h knows its real address M_h, its symmetric encryption key K_h, its whole data $\overline{D}_h = (\overline{A}_h, \overline{b}_h)$, three random strings R_0, R_1 and R_2. The type 1 Bot knows its public/private keys pair $\mathrm{K_{m1}}$, $\mathrm{K'_{m1}}$. The type 2 Bot knows its public/private keys pair $\mathrm{K_{m2}}$, $\mathrm{K'_{m2}}$. The microgrid EMS knows its address $\mathrm{M_{MG}}$, its public/private keys pair $\mathrm{K_{MG}}$, $\mathrm{K'_{MG}}$, and a random string R_3.
1. Both Bots as well as the microgrid EMS broadcast their public keys to every participant. The microgrid EMS also broadcasts its address.
2. Agent h formulates the message to send including its data $\overline{D}_h$, its masked return address $\mathrm{K_{m2}}(\mathrm{R}_2, \mathrm{M}_h)$ and its symmetrical encryption key K_h. Encrypted, the message sent is
$$\mathrm{K_{m1}}(\mathrm{R}_0, \mathrm{K_{MG}}(\mathrm{R}_1, [\overline{D}_h, \mathrm{K_{m2}}(\mathrm{R}_2, \mathrm{M}_h), \mathrm{K}_h]), \mathrm{M_{MG}}) \tag{4.13}$$
3. The type 1 Bot receives the message, decrypts it according to the transformation in Eq. (4.11) with its private key $\mathrm{K'_{m1}}$, identifies the elements, and sends the content
$$\mathrm{K_{MG}}(\mathrm{R}_1, [\overline{D}_h, \mathrm{K_{m2}}(\mathrm{R}_2, \mathrm{M}_h), \mathrm{K}_h]) \tag{4.14}$$
to the address $\mathrm{M_{MG}}$ of the microgrid EMS.
4. The microgrid EMS receives the message, decrypts it with its private key $\mathrm{K'_{MG}}$, and identifies the elements.
5. The microgrid EMS solves the microgrid optimization problem, Problem 4.2 using the information of every agent plus the information of the shared battery and obtains the optimal solution $\left(x^{\star\top}, y^{\star\top}\right)^\top$.

6. The microgrid EMS formulates the message

$$\mathrm{K}_{m2}(\mathrm{R}_2, \mathrm{M}_h), \mathrm{K}_h(\mathrm{R}_3, x_h^\star) \tag{4.15}$$

in order to send $x_h^\star$, the agent's share of the solution, in a message encrypted with the agent symmetrical encryption key K_h to the type 2 Bot, addressed to the agent masked return address $\mathrm{K}_{\mathrm{m2}}(\mathrm{R}_2, \mathrm{M}_h)$.

7. The type 2 Bot receives the message, decrypts the first part according to the transformation in Eq. (4.12) with its private key $\mathrm{K'}_{\mathrm{m2}}$, identifies the elements, and sends the doubly encrypted message

$$\mathrm{R}_2(\mathrm{K}_h(\mathrm{R}_3, x_h^\star)) \tag{4.16}$$

to the real address M_h of agent h.

8. The agent decrypts the doubly encrypted message using the two symmetrical keys R_2 and K_h which it is the only one to know, and finally reads $x_h^\star$.

Step	Agent h EMS	Type 1 Bot	Type 2 Bot	Microgrid EMS
0	M_h, K_h, $\overline{D}_h$, R_0, R_1, R/K_2	K_{m1}, $\mathrm{K'}_{\mathrm{m1}}$	K_{m2}, $\mathrm{K'}_{\mathrm{m2}}$	M_{MG}, K_{MG}, $\mathrm{K'}_{\mathrm{MG}}$, R_3
1	K_{m1}, K_{m2}, K_{MG}, M_{MG}	-	-	K_{m1}, K_{m2}
3	-	$\mathrm{K}_{\mathrm{MG}}(\mathrm{R}_1, [\overline{D}_h, \mathrm{K}_{\mathrm{m2}}(\mathrm{R}_2, \mathrm{M}_h), \mathrm{K}_h])$, M_{MG}, R_0	-	-
4	-	-	-	$\overline{D}_h$, $h = 1, \ldots, l$, $\mathrm{K}_{\mathrm{m2}}(\mathrm{R}_2, \mathrm{M}_h)$, K_h, R_1
5	-	-	-	$x_h^\star$, $h = 1, \ldots, l$
7	-	-	R_2, M_h, $\mathrm{K}_h(\mathrm{R}_3, x_h^\star)$	-
8	$x_h^\star$	-	-	-

Table 4.1: Newly acquired knowledge after each step of the communication protocol

This communication protocol can be employed using the GNU Privacy Guard (GnuPG). It is a complete and free implementation of the OpenPGP standard, also known as PGP. As explained in [GnuPG, 2021], "GnuPG allows you to encrypt and sign your data and communications; it features a versatile key management system, along with access modules for all kinds of public key directories." The *python-gnupg* module allows Python programs to make use of the functionality provided by GnuPG [Sajip, 2021]. Figure 4.1 presents the principle of PGP encryption. Figure 4.2 presents the principle of PGP decryption. Note

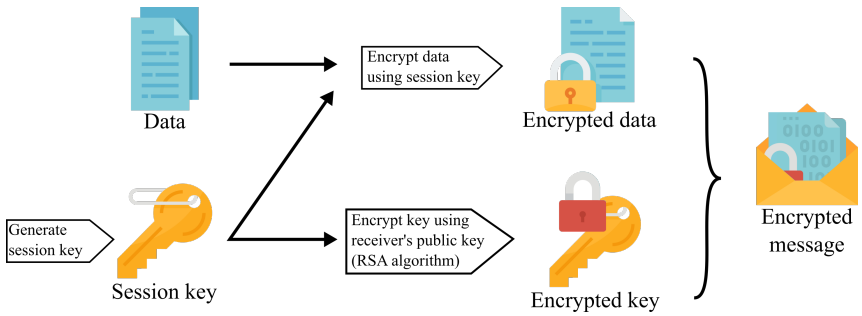

Figure 4.1: Encryption principle of PGP [Zimmermann, 1995] (illustration from [Wikipedia, 2021]).

that a symmetric session key is generated in order to encrypt the actual data, whereas the receiver's public key is only used to encrypt the session key itself. This is due to the limited size of messages encrypted by RSA, the higher efficiency for multi-recipient

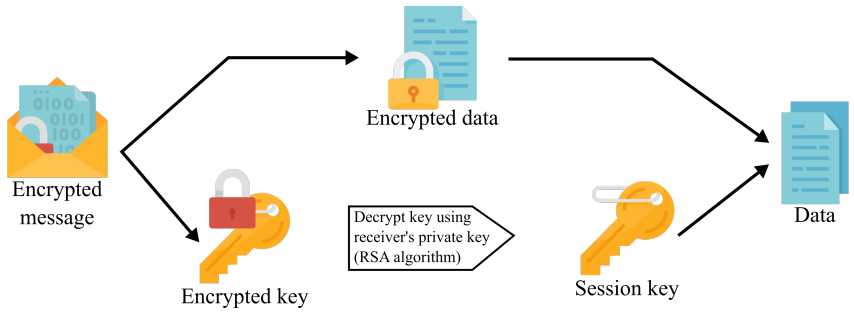

Figure 4.2: Decryption principle of PGP [Zimmermann, 1995] (illustration from [Wikipedia, 2021]).

data exchange, the data-enlargment of RSA encryption, and the slower speed of RSA encryption (see [Information Security Stack Exchange, 2012]).

Importantly, the procedure proposed here allows to keep secret the optimal schedule of the appliances of each agent from all other agents.

4.4 Anonymous ID - based communication (Protocols B_1 and B_2)

The second approach now presented in this section is based on anonymous identifier assignment and secure multiparty computation.

4.4.1 Algorithmic tools

The cornerstone of the methods presented in this section is to assign an anonymous ID to each agent participating in the microgrid, in the form of a unique integer. The ID can then be used to coordinate the anonymous information exchange. In the end, the association of the households with the set of IDs forms the permutation σ involved in the Problem 4.3, which can be solved by the microgrid EMS to reach the same solution as the omniscient problem, as it is equivalent to Problem 4.2, as discussed previously.

[Dunning and Kresman, 2013] is the main source of inspiration for the privacy-preserving communication protocol considered in this section. The article presents the "Find Anonymous ID Assignment (AIDA)" algorithm reproduced here for convenience (see Algorithm 4.4), used to perform the unique ID assignment anonymously, as desired. It is based on two other algorithms, Algorithms 4.1 and 4.3 which are first detailed. Following this article, it is assumed that all agents in the microgrid are *semi-honest*, a concept also known as honest-but-curious, in which each agent can be described as "a legitimate participant in a communication protocol who will not deviate from the defined protocol but will attempt to learn all possible information from legitimately received messages." [Paverd and Martin, 2014]. This property implies that protective measures are not needed to counter malevolent behaviors, and that the protocols will be executed faithfully.

The first tool used in the proposed protocol is Algorithm 4.1, the "Distributed Secure Sum" algorithm. It is used to compute the sum of pieces of data distributed across several agents, without revealing the data of each agent to other agents nor to the microgrid EMS. This classical ingredient of secret communications, is also presented in [Jacquot et al., 2019] under the name "Secure Multiparty Computation of Aggregate".

Algorithm 4.1: Distributed Secure Sum: Share a sum of data items $\beta = \sum_{h=1}^{l} d_h$ among all l agents without revealing the values d_h

Input: l agents each holding a data item d_h from $\mathbb{Z}$ with $h = 1, \ldots, l$.

for *each agent* $h \in [1, l]$ **do**

Agent h draws random values $y_{h,1}, \ldots, y_{h,l-1}$ without any constraint and according to any desired probability distribution.

Agent h then chooses $y_{h,l}$ such that

$$y_{h,l} = d_h - \sum_{g=1}^{l-1} y_{h,g} \tag{4.17}$$

to ensure that the sum of drawn values adds up to the initial data.

for *each agent* $g \in [1, l]$ **do**

Value $y_{h,g}$ is transmitted from agent h to agent g

end

end

for *each agent* $g \in [1, l]$ **do**

Agent g computes the sum of the random values received as

$$z_g = y_{1,g} + \cdots + y_{l,g} \tag{4.18}$$

Agent g broadcasts the sum z_g to all other agents.

end

for *each agent* $h \in [1, l]$ **do**

Agent h adds all the received sums to compute

$$\beta = z_1 + \cdots + z_l \tag{4.19}$$

end

In the end, the algorithm succeeds in transmitting the sum because the data items are parted then reunited

$$\begin{aligned}
\beta &= \quad z_1 \quad + \cdots + \quad z_l \\
&= y_{1,1} + \cdots + y_{l,1} + \cdots + y_{1,l} + \cdots + y_{l,l} \\
&= y_{1,1} + \cdots + y_{1,l} + \cdots + y_{l,1} + \cdots + y_{l,l} \\
&= \quad d_1 \quad + \cdots + \quad d_l
\end{aligned}$$

The "Distributed Secure Sum" (Algorithm 4.1) allows every agent to compute the final sum. Another version of the "Secure Sum" algorithm is the "Centralized Secure Sum" (Algorithm 4.2), where only the microgrid EMS can compute the final sum. The algorithm differs only in the final steps. It is also used in the final protocol considered.

Algorithm 4.2: Centralized Secure Sum: Communicate a sum of data items $\beta = \sum_{h=1}^{l} d_h$ with the microgrid EMS without revealing the values d_h

Input: l agents each holding a data item d_h from $\mathbb{Z}$ with $h = 1, \ldots, l$.

for *each agent* $h \in [1, l]$ **do**

Agent h draws random values $y_{h,1}, \ldots, y_{h,l-1}$ without any constraint and according to any desired probability distribution.

Agent h then chooses $y_{h,l}$ such that

$$y_{h,l} = d_h - \sum_{g=1}^{l-1} y_{h,g} \tag{4.20}$$

to ensure that the sum of drawn values adds up to the initial data.

for *each agent* $g \in [1, l]$ **do**

Value $y_{h,g}$ is transmitted from agent h to agent g

end

end

for *each agent* $g \in [1, l]$ **do**

Agent g computes the sum of the random values received as

$$z_g = y_{1,g} + \cdots + y_{l,g} \tag{4.21}$$

Agent g broadcasts the sum z_g to the microgrid EMS

end

The microgrid EMS adds all the received sums to compute

$$\beta = z_1 + \cdots + z_l \tag{4.22}$$

Notice that broadcasting is not necessarily anonymous in the several communication steps of Algorithms 4.1 and 4.2. With this communication scheme, it is required to eavesdrop on every single information exchange to be able to reconstruct the initial data involved in the final sum.

The second tool presented in [Dunning and Kresman, 2013] is the "Anonymous Data Sharing With Power Sums" algorithm, reproduced here in Algorithm 4.3. It is used to anonymously transmit data $d_h (h = 1, \ldots, l)$ held by each agent to every other participant. The algorithm consists in creating a polynomial of degree l, whose roots on the ring of natural integers modulo p are the data to transmit. The prime number p has to satisfy $p > d_h, \forall h \in [1, l]$. The polynomial coefficients are computed using Newton's Identities, which are collectively computed using the "Secure Sum" algorithm (Algorithm 4.1). Notice that knowing the polynomial coefficients does not enable to deduce a pair (root, agent), because the coefficients of the polynomial are symmetric functions of the roots.

Algorithm 4.3: Anonymous Data Sharing With Power Sums: Make every data item d_h of agent h public to all l agents without revealing the sources

Input: l agents each holding a data item d_h from $\mathbb{Z}$ with $h = 1, \ldots, l$; a prime number p satisfying $p > d_h, \forall h \in [1, l]$.

for *each agent* $h \in [1, l]$ **do**

Compute d_h^n with power values $n = 1, \ldots, l$ over the Finite Field $\mathbb{Z}/p\mathbb{Z}$.

end

Agents use the "Secure Sum" Algorithm 4.1 to share knowledge of the power sums with every agent:

$$S_1 = \sum_{h=1}^{l} d_h, \; S_2 = \sum_{h=1}^{l} d_h^2, \; \ldots, \; S_l = \sum_{h=1}^{l} d_h^l \tag{4.23}$$

for *each agent* $h \in [1, l]$ **do**

Construct the Newton polynomial which has $d_1, \ldots, d_l$ as its roots using Newton's Identities. Representing the Newton polynomial as

$$\Pi(x) = \gamma_l x^l + \ldots + \gamma_1 x + \gamma_0 \tag{4.24}$$

the values $\gamma_0, \ldots, \gamma_l$ are obtained from the equations:

$$\gamma_l = -1 \tag{4.25}$$

$$\gamma_{l-1} = -\frac{1}{1}(\gamma_l S_1) \tag{4.26}$$

$$\gamma_{l-2} = -\frac{1}{2}(\gamma_{l-1} S_1 + \gamma_l S_2) \tag{4.27}$$

$$\gamma_{l-3} = -\frac{1}{3}(\gamma_{l-2} S_1 + \gamma_{l-1} S_2 + \gamma_l S_3) \tag{4.28}$$

$$\ldots$$

$$\gamma_{l-m} = -\frac{1}{m} \sum_{k=1}^{m} \gamma_{l-m+k} S_k \tag{4.29}$$

Determine the roots $d_1, \ldots, d_l$ of the polynomial Π over $\mathbb{Z}/p\mathbb{Z}$

end

This step prevents sharing and computing extremely large numbers, due to the calculation of the values power "up to l". Note that, in the Finite Field $\mathbb{Z}/p\mathbb{Z}$, the calculation of γ_i coefficients (Eqs. (4.25)-(4.29)) requires finding the modular multiplicative inverse $\frac{1}{m}$ for $m = 1, \ldots, l$, which can be computed using the Euclidean algorithm. This algorithm along with a simple example are detailed in Appendix D. An example of the use of Algorithm 4.3 from [Dunning and Kresman, 2013] is reproduced in Appendix E.

Finally, the "Find AIDA" Algorithm 4.4 is the one yielding the unique ID to every agent. At one step, random integers between 1 and S are chosen by each agent. The list of chosen integers is shared using the "Anonymous Data Sharing With Power Sums" Algorithm 4.3. An agent's position will be determined by its position among the chosen integers, provisions being made for collisions. The parameter S should be chosen so that $S \geq l$.

Algorithm 4.4: Find Anonymous ID Assignment: Have the l agents define an anonymous indexing permutation $\sigma : 1, \ldots, l \to 1, \ldots, l$ without any central authority.

Input: l agents.
1) Set the number of assigned agents $\xi = 0$ and $\sigma_h = 0$ for $h = 1, \ldots, l$.
while $\xi < l$ **do**
  for *each agent* $h \in [1, l]$ **do**
    if $\sigma_h = 0$ **then**
      2) Choose a random integer y_h in the range 1 to S.
    else
      2) Choose $y_h = 0$.
    end
  end
  3) Agents share knowledge of y_h for $h = 1, \ldots, l$ with the "Anonymous Data Sharing With Power Sums" Algorithm 4.3. Shared values are denoted q_1, ..., q_l.
  for *each agent* $h \in [1, l]$ **do**
    Remove from $Q = \{q_1, \ldots, q_l\}$ the zero values and all duplicated values to create a new sorted list $U = \{u_1, \ldots, u_f\}$ where f is the number of unique random values.
    if $q_h \in U$ **then**
      4) Pick $\sigma_h = \xi + \text{Card}(u_j : u_j \leq y_h)$, the position of their random number in the revised list as it would appear after being sorted.
    end
    5) Update the number of agents assigned: $\xi = \xi + f$.
  end
end

The 5 steps of the algorithm are identified by indices 1), ..., 5). An example of the use of Algorithm 4.4 from [Dunning and Kresman, 2013] is reproduced in Appendix F.

4.4.2 Secure multiparty computation

Now that the tools from [Dunning and Kresman, 2013] have been detailed, the principle of the privacy-preserving communication scheme that is advocated can be presented.

Protocol B_1

Once an anonymous ID has been assigned to each participant, the compressed vector versions of the data elements $\overline{A}_h$ and $\overline{b}_h, h = 1, \ldots, l$ (see Section 4.2) required for the composition of the microgrid optimization problem 4.2 can be transmitted anonymously.

The idea is to use once again the "Secure Sum" algorithm, but in a different way. Just like the household constraints matrices A_h can be transmitted as vectors of their flattened non-structurally null elements $\overline{A}_h$, the collection of these vectors of interest can

be collected by the microgrid EMS as an ordered vector:

$$V = \begin{pmatrix} \overline{A}_1^\top & \overline{A}_2^\top & \dots & \overline{A}_l^\top \end{pmatrix}^\top \tag{4.30}$$

where $\overline{A}_h$ is the compressed vector of variables of interest of the constraints matrix A_h defined in Section 4.2.2.

The idea proposed in [Dunning and Kresman, 2013] for the microgrid EMS to collect all these compressed vectors is the following: each agent h creates a vector of l blocks of standard compressed vector length $\overline{n}$, and stores its own compressed vector $\overline{A}_h$ in the block of index matching its assigned anonymous ID (see the "Find AIDA" Algorithm 4.4). The created vector $\overline{A}_h^e$ is

$$\overline{A}_h^e = \begin{pmatrix} \mathbb{0}_{\overline{n}}^\top & \dots & \overline{A}_h^\top & \dots & \mathbb{0}_{\overline{n}}^\top \end{pmatrix}^\top \tag{4.31}$$

where $\mathbb{0}_{\overline{n}}$ is the null vector in $\mathbb{R}^{\overline{n}}$. $\overline{A}_h$ is here stored on the block of index i, the anonymous ID assigned to agent h. The same principle applies for the communication of the compressed right-hand side vectors $\overline{b}_h, h = 1, \dots, l$, of standard length $\overline{m}$.

Then, the "Secure Sum" Algorithm 4.2 can be used for the microgrid EMS to collect all these matrices ordered accordingly, the sum of $\overline{A}_h^e$, $h = 1, \dots, l$ resulting in V (Eq. (4.30)), and the sum of $\overline{b}_h^e$, $h = 1, \dots, l$ denoted W.

Upon reception of V and W by the microgrid EMS, it can form the matrices $A_{\sigma(h)}$ and vectors $b_{\sigma(h)}$, $h = 1, \dots, l$ thanks to the knowledge of the data standard which is public. Then, Problem 4.3 can be constructed and solved.

The solution vector is finally openly broadcast. Every agent can read its part of the solution vector, as well as the one of the other participants. This protocol is simply referred to as Protocol B_1 thereafter.

Encrypted response option (Protocol B_2)

For increased security, an option of the presented protocol uses both the secure sum algorithm, and pairs of public-private encryption keys. The secure sum algorithm is now used to transfer the agent public encryption keys K_h (more details on encryption principles are given in Section 4.3), in addition to the compressed vectors $\overline{A}_h$ and $\overline{b}_h$ containing the data of agent h [5].

Then, after solving the microgrid optimization problem, the microgrid EMS can encrypt each part of the solution vector with the corresponding encryption key. The encrypted result is openly broadcast. The agents can read all the solutions parts, but can only decrypt the one corresponding to their own decision variables, with their private encryption key.

Algorithm 6 details this method, which follows the same principle as Eqs. (4.30) and (4.31). This protocol is simply referred to as Protocol B_2 thereafter.

4.4.3 Communication protocol

The successive steps of the protocols based on anonymous ID assignment and secure multiparty communication (Protocols B_1 and B_2) are detailed here:

1. Houses participate in the "Find AIDA" method (Algorithm 4.4) in order to get an AIDA (Anonymous ID Assignment) for each controllable appliance;

[5] This approach assumes that it is possible to work with a hexadecimal or binary version of the encryption key, in order to be able to divide it in several bits that add up to the correct sum in the end of the secure sum algorithm.

Algorithm 4.5: Secure multiparty transmission of encryption keys and broadcast of encrypted solution (Protocol B_2)

Input: Compressed vectors $\overline{A}_h$ for $h \in [1, l]$, of standard size $\overline{n}$.
Agents pick IDs $n_1, \ldots, n_l$ with Algorithm 4.4.
for *each agent* $h \in [1, l]$ **do**
  Construct $\overline{A}_h^e = \begin{pmatrix} \mathbb{0}_{\overline{n}(n_h-1)}^\top & \overline{A}_h^\top & \mathbb{0}_{\overline{n}(l-n_h)}^\top \end{pmatrix}^\top$
  Construct $\overline{b}_h^e = \begin{pmatrix} \mathbb{0}_{\overline{n}(n_h-1)}^\top & \overline{b}_h^\top & \mathbb{0}_{\overline{n}(l-n_h)}^\top \end{pmatrix}^\top$
  Construct $K_h^e = \begin{pmatrix} \mathbb{0}_{(n_h-1)}^\top & K_h & \mathbb{0}_{(l-n_h)}^\top \end{pmatrix}^\top$
end
Agents share $V = \sum_{h=1}^{l} \overline{A}_h^e$, $W = \sum_{h=1}^{l} \overline{b}_h^e$ and $K = \sum_{h=1}^{l} K_h^e$ with the microgrid EMS using Algorithm 4.2.
The microgrid EMS decompresses V and W to form $\left(A_{\sigma(h)}\right)_{h\in[1,l]}$ and $\left(b_{\sigma(h)}\right)_{h\in[1,l]}$, then computes a solution $\left(x^{\star\top}, y^{\star\top}\right)^\top$ to Problem 4.3
The microgrid EMS ciphers each component $x_h^\star$ of $x^\star$ with the key K_h and broadcasts the corresponding result $x_c^\star$.
for *each agent* $h \in [1, l]$ **do**
  Pick $x_{c,n_h}^\star$ and deciphers it with its private key.
end

2. Secure multiparty computation is used to transmit the necessary data (vector of interest variables of constraints matrix $\overline{A}_h$ (see Section 4.2.2) and constraint vector $\overline{b}_h$) to the microgrid EMS;
3. Optionally, the agents communicate their public encryption keys to the microgrid EMS with the same method (Protocol B_2);
4. The microgrid EMS collects all the data to create the microgrid optimization problem (equivalent to the omniscient problem) and solves it;
5. In the "Encrypted response" option (Protocol B_2), each part of the solution vector is encrypted with the appropriate encryption key;
6. The (potentially encrypted) solution vector is broadcast to the agents;
7. Each individual EMS retrieves its own solution, decrypts it if it has been encrypted, then applies it.

Each agent can read the complete result, and will implement the solution corresponding to its ID. If the "encrypted response" option is implemented, each agent can only decrypt and read its own part of the solution vector. In a microgrid with more than 2 households, as each agent knows only its unique ID, the agents cannot link directly the solution of other IDs (i.e. the control parameters of other agents) to a specific household. In a microgrid restricted to only two households, the "encrypted response" version of this scheme is required, for the agents not to know exactly the final appliances schedule of its neighbor.

These communication protocols can be applied using the *python-FLINT* Python module. It allows to perform number theory operations as required in Algorithm 4.3, such as computing the modular inverse of any natural integer (detailed in Appendix D) or finding the roots of a polynomial modulo a prime.

4.5 Protocols comparison

4.5.1 Security quality

The security of communication corresponds in each protocol to the difficulty for an agent to learn private data of another agent ($\overline{A}_h$ and $\overline{b}_h$). When Protocol A (encryption-based, presented in Section 4.3) is used, learning data would requires to listen to only one communication from one agent to the microgrid EMS, but also to know the encryption key required to decode the message. Indeed, hacking an encrypted message is supposed to be infeasible [Schneier, 1996]. On the contrary, the secure multiparty computation protocols (B_1 and B_2) presented in Section 4.4 require for an eavesdropper to listen to every single communication in order to be able to recompose the data of a single agent. Indeed the microgrid EMS is the only entity knowing the "real" data, recomposed at the end of the "Secure Sum" Algorithm. Regarding the communication of the solution schedule, the use of encryption keys has the obvious advantage of hiding the solution vector (i.e. appliances planning) of any agent from its neighbors. These differences are listed in Table 4.2.

Table 4.2: Summary of the privacy-preserving communication protocols

	Encrypted communications	ID-based communications	
		Without solution encryption	With solution encryption
	(Protocol A)	(Protocol B_1)	(Protocol B_2)
Information originally known by the microgrid EMS	α_{sys}, A_{sys}, b_{sys} c, number of households l, standard matrices structure		
Information known originally only by household h	$\overline{A}_h$, $\overline{b}_h$, K_h, standard matrices structure	$\overline{A}_h$, $\overline{b}_h$, standard matrices structure	$\overline{A}_h$, $\overline{b}_h$, K_h, standard matrices structure
Information learnt by the microgrid EMS during protocol	$\left(\overline{A}_{\sigma(h)}, \overline{b}_{\sigma(h)}\right)_{h\in[1,l]}$ $\left(x^\star_{\sigma(h)}, y^\star_{\sigma(h)}\right)_{h\in[1,l]}$ $\left(K_{\sigma(h)}\right)_{h\in[1,l]}$	$\left(\overline{A}_{\sigma(h)}, \overline{b}_{\sigma(h)}\right)_{h\in[1,l]}$ $\left(x^\star_{\sigma(h)}, y^\star_{\sigma(h)}\right)_{h\in[1,l]}$	$\left(\overline{A}_{\sigma(h)}, \overline{b}_{\sigma(h)}\right)_{h\in[1,l]}$ $\left(x^\star_{\sigma(h)}, y^\star_{\sigma(h)}\right)_{h\in[1,l]}$ $\left(K_{\sigma(h)}\right)_{h\in[1,l]}$
Information learnt by household h at the end of protocol	$x^\star_h$	$x^\star_h$, $x_{\sigma(i)}$ for $\sigma(i) \neq h$	$x^\star_h$
Information necessary for an external agent to learn $\overline{A}_h$ or $\overline{b}_h$	Impossible (PGP protocol)	listening to all channels of exchange to recompose $V = \sum_{h=1}^{l} \overline{A}^e_h$ or $W = \sum_{h=1}^{l} \overline{b}^e_h$	

4.5.2 Runtime

The simple two-participants use-case from Section 3.4 is considered again for a time comparison of the three communication protocols. Only the first phase of the protocols, where the agents communicate their private data to the microgrid EMS, are compared. Only the compressed vector of the constraints matrix is communicated, and not the compressed version of the right-hand side vector of the agents. The time required for the microgrid EMS to sent a response is not accounted for here.

The issue of one agent being able to infer the data of its neighbor in the two-participants case arises. In the case of distributed algorithms, such a limitation is widely known and reported e.g. in [Tassa and Cohen, 2013]. Note that, in the case of the Distributed Secure Sum, it is not a limitation of the protocols, but a natural consequence of the computation even if it would be executed using a trusted third party. As this section focuses on time of execution of the methods, this issue is dismissed here.

The number of timesteps is $H = 96$ (a classical 15-minute discretization of a 24-hour day). The first house is equipped with a dishwasher modeled with $J = 5$ energy phases and the second house is equipped with a washing machine modeled with $J = 7$ energy phases. The constraints matrix of the first house is of shape (4714, 1824), as computed in Section 3.3, and the constraints matrix of the second house is of shape (6446, 2400). However only the variables of interest are communicated, resulting in a vector of 13 031 elements for the first house and of 18 013 elements for the second house.

The tests were conducted on a laptop equipped with Python 3.7.3 and a Core i3 2.4 GHz processor, with 8 Go RAM.

The Encryption protocol (Protocol A) was first applied. It takes on average 9.7 seconds to generate the keys for all the entities involved in the protocol, and an additional 1.9 seconds to perform the data exchanges (limited to the initial communication from the agents to the microgrid EMS). Table 4.3 displays the duration required by each step of this communication protocol. It is recalled that the duration of resolution of the microgrid optimization problem is not accounted for here. Notice that the "Keys generation" and "Keys exchange" steps, significant contributors to the total duration, would be executed only once for several data exchanges, leaving only a fraction of this duration to be repeated at each exchange.

Table 4.3: Protocol A: the duration of encryption, transmission and decryption steps are extremely low.

Step	**Median duration**
Keys generation	9.7 s
Keys exchanges	0.4 s
Agent 1 data encryption and transmission	0.4 s
Agent 2 data encryption and transmission	0.4 s
Type 1 Bot decryption and message handling	0.3 s
Microgrid EMS messages decryption	0.4 s
Total	11.6 s

Consider now Protocol B_1, based on secure multiparty computation. It performs all the data exchange in an average of 3.2 seconds. As with the encrypted communication protocol, only the values of the elements of interest (i.e. elements that are not structurally equal to zero) of the constraints matrices of each agent are exchanged, because the structure of the resulting matrix is standardized and known in advance. This protocol requires that the compressed vectors of the variables of interest of the constraints matrices are included in a null vector (see Eqs. (4.30) and (4.31)), thus resulting in $2\overline{n} \simeq 60\,000$ secure sum calculations.

Finally, consider Protocol B_2, based on secure multiparty computation and with solution encryption. It performs all the data exchanges in an average of 8.37 seconds. Table 4.4 displays the duration required by each step of this communication protocol.

After this private data communication has been performed, the microgrid EMS can formulate and solve the microgrid optimization problem. For reference, the median resolution time is 7 s.

Table 4.4: Protocol B_2: Keys exchange through the Secure Sum is extremely fast.

Step	**Median duration**
Keys generation	4.6 s
Agents data exchange (through Secure Sum)	3.7 s
Agents keys exchange (through Secure Sum)	0.07 s
Total	8.37 s

The very low computational cost of the Encryption communication scheme seems to advocate for this option over the other two, provided that new encryption keys are generated and exchanged only rarely. A thorough comparison with a larger number of participants and including complexity comparison is however needed to compare the three complete schemes (response steps included) over repeated data exchange.

4.6 Conclusion

This chapter has addressed the privacy-preservation challenge in a local energy community.

The aim was to construct the microgrid optimization problem with permutation (Problem 4.3), strictly equivalent to the omniscient one (Problem 3.1), while protecting the privacy of the microgrid agents. To this end, three communication protocols inspired from cryptography and secure multiparty computation have been described. They allow to transmit anonymously (i.e. without the data recipient being able to directly link which data comes from which sender) to the microgrid EMS the individual constraints, consumption and production profiles necessary for the construction of the optimization problem at the microgrid level.

A small-size use-case has allowed to evaluate the time required for communication by the three privacy-preserving protocols.

4.6.1 Possible extensions

The success of both methods in transmitting data in a limited time suggests to try and apply the proposed anonymization method to a larger group of optimization problems of this kind.

Indeed, the solution of any non-linear optimization problem under parametric constraints depending on distributed parameters (belonging to agents in a network) can be found by constructing an equivalent problem after the anonymous collection of the parameters, thanks to the privacy-preserving communication methods proposed in this chapter.

Namely, consider the problem of l agents, each of them having a triplet $(x_h, \theta_h) \in \mathcal{X} \times \Theta$, $h = 1, \dots, l$ satisfying constraints of the form

$$c(x_h, \theta_h) \leq 0\,, \quad h = 1, \dots, l \tag{4.32}$$

with $c : \mathcal{X} \times \Theta \to \mathbb{R}^m$, and wishing to share anonymously these information with a central operator, to construct and solve the mixed-integer nonlinear optimization problem

Problem 4.4.

$$\begin{aligned} &\min_{(x,y)\in\mathcal{X}^l\times\mathcal{Y}} f(x,y,\theta) \\ &\text{s.t.} \begin{cases} c(x_h,\theta_h) \leq 0\,, \quad h = 1,\dots,l \\ c_{op}(x,y,\theta) \leq 0 \end{cases} \end{aligned} \tag{4.33}$$

with the objective function $f : \mathcal{X}^l \times \mathcal{Y} \times \Theta^l \to \mathbb{R}$ and the operator constraint $c_{op} : \mathcal{X}^l \times \mathcal{Y} \times \Theta \to \mathbb{R}^p$, satisfying the following assumption:

Assumption 2. *The cost function f and the operator constraint c_{op} in Problem 4.4 are such that, for any permutation σ of $\{1, \dots, l\}$,*

$$\forall (x, y, \theta) \in \mathcal{X}^l \times \mathcal{Y} \times \Theta^l, \quad \begin{cases} f(x, y, \theta) = f(x_\sigma, y, \theta_\sigma) \\ c_{op}(x, y, \theta) = c_{op}(x_\sigma, y, \theta_\sigma) \end{cases} \tag{4.34}$$

The privacy-preserving protocols presented in this chapter allow to solve Problem 4.4 by exchanging the data θ_h with the central operator.

Furthermore, this study calls for a thorough mapping of the categories of optimization problems that can be anonymized (in addition to the problem at hand in this chapter and to the aforementioned one), of the anonymization techniques applicable to each type of optimization problems, and a comparison of them. For instance, one can wonder if the methods proposed here could apply to the problem studied in [Jacquot et al., 2019], and how their performance would compare to the one used in that paper. Such mappings would be a major addition in generalizing privacy-preserving methods for collective optimization problems.

Finally, a potential extension of the privacy-preserving protocols proposed here could be to have the agents send the technical information (power ratings and limitations, energy requirements, energy phases duration) to the microgrid EMS, and let it perform the discretization of the constraints equations resulting in the constraints matrices formulation. Thus the volume of data transmitted would be significantly lower. This approach was not explored in this thesis.

4.6.2 Open questions for a privacy-preserving collaboration scheme

In view of real world applications, a couple of questions remain to be investigated regarding the final level of privacy, that has to be assessed considering the whole setup. Challenges to a proper level of privacy are the following:

- the monitoring of household consumptions is necessary to establish bills between the distribution system operator and households.
- a fine-grain division of the operation cost of the shared battery requires to track the power exchanges between each household and the battery.

A direction of research to achieve these goals could be to mix the elements proposed in this chapter with game theory approaches.

Chapter 5

Conclusion and perspectives

Conclusion et perspectives. Ce chapitre commence par résumer les principales contributions de cette thèse. Puis il présente certaines directions de travail qui, à la lumière des présents travaux, semblent prometteurs pour les futurs systèmes de gestion de l'énergie. Ces perspectives sont réparties entre celles qui semblent au moins partiellement applicables aux outils présentés dans cette thèse, et celles qui s'inscrivent dans un cadre plus général.

5.1 Conclusion

Here are summarized the main contributions from this thesis. The first two chapters develop useful tools for the performance evaluation of existing PV-equipped smart home control algorithms, while the last two chapters propose a solution to the challenge of enabling cooperation in a microgrid while protecting the privacy of participants.

- In Chapter 1, a method is developed to find efficiently the optimal control of an electric water heater (EWH) in an individual house equipped with photovoltaic (PV) panels, under the following conditions: 1.the assumption of perfect prediction of the hot-water consumptions profile; 2.the modelization of the EWH without the stratification phenomenon; 3.the maximization of a specific criterion, the self-consumed energy. This resolution algorithm can be used as a comparison tool for other EWH control algorithms designed for self-consumption installations, in that it provides the absolute maximum performance and thus informs on the progression margin of the others algorithms.
- Chapter 2 proposed a PV production scenarios generation method to provide a set of individually realistic scenarios that represents as a whole the uncertainties of an initial PV forecast. Using the quick algorithm presented in Chapter 1 on every scenario generated by the present method allows to give the only acceptable *a priori* performance evaluation of a smart home control algorithm with a predictive component. A study conducted in this chapter also allows to assess the hypothetical margin for progress available for an EWH control algorithm that would benefit from perfect PV production predictions.
- Chapter 3 details the formulation of the energy optimization problem at the microgrid level, where the microgrid EMS can act on all the agent's controllable appliances and a shared battery, in order to foster cooperation and to minimize the overall energy cost. The problem is formulated as a MILP. Its interest is proved in a small-size example and its complexity is discussed.
- Chapter 4 draws inspiration from cryptography tools to propose three privacy-preserving communication schemes allowing to formulate the microgrid problem

while anonymizing the data necessary to its construction. For a small-size problem, both schemes are shown to be quick enough with respect to the usual timestep of microgrid energy optimization algorithms.

In order to pave the way for the extension of this work towards the design of fully-implementable PV self-consumption EMSs at the household and microgrid level, some additional notions are presented in the following section.

5.2 Perspectives

A handful of concepts, in addition to the ones used throughout this manuscript, have been proposed in the literature in order to design an ideal energy management system (EMS) both for the household and the microgrid levels. These notions were not included in the present work for better clarity of the work exposed or for lack of time, but could become vital in the design of such ideals EMSs. The present section will present these available building blocks and the corresponding references.

5.2.1 Extensions of the present work

In this section are presented concepts and references that might be applicable to the tools used in this manuscript, or that seem at least like natural extensions of the presented work. These concepts could thus provide valuable additions without shifting entirely the nature of the EMSs developed.

Hot-water consumption estimation Throughout all this thesis, a major hypothesis has been made: the perfect knowledge, in advance, of the hot-water drains. The sequence of time, duration and amplitude of the hot-water drains is supposedly known in advance in both the continuous and discrete-time EWH models used. Future works shall overcome this quite unrealistic assumption.

To this end, inspiration can be drawn from [Arkhangelski et al., 2019] which proposes to forecast load consumption using a Recurrent Neural Network (RNN) method with Long Short Term Memory (LSTM) algorithm. Recurrent Neural Networks (RNN) are said to be more effective in sequential data dynamics modeling than other types of Artificial Neural Networks.

Stochastic MPC Despite potentially large computation times, Stochastic MPC might prove more valuable than other control approaches when dealing with increased levels of uncertainty. Chapter 2 showed that, at a 30-minute timestep, the benefit of anticipating (hypothetically perfectly) the PV production uncertainties, when using a basic control heuristic, was lower than the benefit of switching to an improved control algorithm. However, the benefit of anticipating production uncertainties would probably be increased with a decreased timestep length, and with the occupants behavior uncertainties taken into account.

Peer-to-peer energy trading In a microgrid composed of several households with DERs, if local peer-to-peer energy exchanges are not authorized, each household converses with the microgrid EMS exclusively, all its power surplus and deficit is paid by and to the microgrid EMS, and all agents have this same and unique intermediary. This is the case retained in this manuscript. However, when direct peer-to-peer energy trading is allowed, the trading is then executed at a price in-between the selling and buying prices of the corresponding agents. This configuration should be encouraged in future works at it promotes peer-to-peer power exchanges as they are beneficial to both parties: the seller gets a higher selling

price, and the buyer gets a lower buying price than the respective ones offered by the microgrid EMS.

Cost configuration A question of interest in a microgrid optimization configuration is to know whether the prices for buying and selling power inside the microgrid have to be the same than those from the external grid. In the use-cases considered in the last two chapters of this thesis, these prices were the same at each level of the energy system, the microgrid EMS thus only operating as a neutral entity supporting the agents in the cost-reduction process. Nevertheless, another configuration could have the microgrid EMS apply a profit rate to the prices from the external grid, before exchanging energy with the microgrid participants. The impact of each configuration has to be studied. Moreover, when a price mechanism is designed for the access to a shared battery, the energy consumed at one moment by an agent may have been purchased and stored before by the microgrid EMS through the battery. The price mechanism could account for this time delay in order to establish a fair cost allocation. For instance, [Steriotis et al., 2019] provides a price mechanism ensuring that the price of the energy exiting the shared battery reflects the cost previously paid by the community to make this energy available. The price mechanism increases the fairness of the energy cost division among the agents.

Whatever the cost configuration selected, the evaluation of its impact on the collective and individual costs has to be assessed.

Fair division of collective cost The question of the fair division of the collective cost of operation of the shared resource soon rises in a collective self-consumption operation, as was mentioned in Section 3.4.

A fairness criteria should be used in an extension of the presented work, in order to ensure

- that the battery access price leads to a fairer division of the collective cost of operation of the shared resource;
- and to lead to the best choice of division of the collective cost of operation of the shared resource.

An "effort" criteria could be needed for the fair division of the collective cost. With the aim of finding the fairest distribution of the collective profit, a special retribution could go to the most compliant agents with the demand response signals. Some challenges then arise:

- if a marketplace type of price is considered, how not to pay twice for the same "effort"?
- how to compare the effort made by one type of household to another, with radically different types of households, consumption patterns and constraints (e.g. a retired couple *vs* a family of four with working parents)?
- how to define which behaviors are vertuous, and which are not?

[Steriotis et al., 2019] and [Jacquot et al., 2017] each propose a possible fairness criterion. A billing mechanism is fair if the bill of user n reflects correctly the additional cost it as brought to the collective system.

Despite its difficult evaluation, the fairness of a cost configuration in a microgrid has to be assessed in order to encourage enrollment of new participants.

Discomfort If the comfort constraints are formulated in the optimization problem as hard constraints, better solutions that infringe these rules only by a small margin cannot be proposed. Thus comfort constraints can be softened in order to provide more flexibility to

the system. This is especially interesting when the household participates in DR actions, in frequency regulation actions, or in a microgrid. The common element of these three situations is that a collective objective is acknowledged, beyond the individual household one. In these situations, one can imagine allowing to breach the comfort constraints. But then the resulting discomfort must be taken into account in the problem resolution, and a form of compensation might be required.

For example, [Pflaum et al., 2014] includes a discomfort function in the overall multi-objective function of an energy management problem in a smart district. Figure 5.1 displays the value of the discomfort criterion in function of the state variable y with respect to upper and lower bounds $\overline{y}$ and $\underline{y}$.

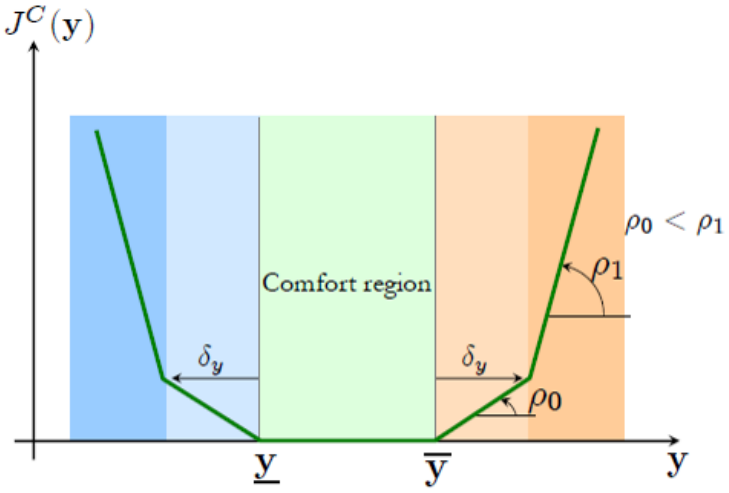

Figure 5.1: Example of a discomfort function from [Pflaum et al., 2014]. Parameters ρ_0, ρ_1, δ_y determine the bounds violations zones.

Such a formulation could be used for the softening of the user preferences constraints in the optimization problem formulated in Chapter 3. In all likelihood, it should translate into improved performances and increase feasibility of the problem at stake.

Multi-objective optimization Similarly, regarding self-consumption installations, multi-objective optimization can be of interest since two different objective functions proved useful throughout this manuscript: the self-consumed volume of energy and the electricity bill. Both these criteria could be incorporated in a single weighted objective in the case of an installation connected to a constrained grid in terms of upstream injection. [Gonzalez Venegas et al., 2021] recalls that several control strategies can apply to EVs, with several potentially aligned or opposed objectives. This is true for any appliance storing energy, be it in a way to provide another principal service (EWHs, EVs) or not (batteries).

Distributed computations The computational burden of solving the microgrid optimization problem in Chapter 3 highlighted the limitations of resorting to a centralized problem for the optimization of energy communities. Similarly, for MPC strategies, [Pflaum et al., 2014] evaluates centralized approaches as unsuitable for microgrids, because of the heavy computational burden, the lack of modularity (each change in one building affecting the centralized formulation), and the data privacy issues.

Distributed computation strategies consist in a decomposition of the central problem into sub-problems processed by the local EMSs which are able to recover the optimal solution of the centralized problem or at least to find a relevant sub-optimal solution, through an iterative communication scheme.

In particular, decomposition-coordination methods [Culioli and Cohen, 1990] have been proposed in order to optimize a common objective function by decentralized computations. In such methods, a centralized optimization problem is split into small-size local optimization problems whose outputs are coordinated dynamically by a central agent so that the overall objective of the system becomes aligned (after a certain number of iterations) with the centralized optimization problem outcome.

Two challenges then face the adoption of distributed computation techniques: the restricted type of problems potentially solved by these techniques, and the protection of the private data exchanged.

Recent research potentially providing insights on privacy preserving energy sharing and optimization, with distributed computing, in order to design scalable solutions: [Ye et al., 2020] proposes a privacy-preserving version of a consensus-based microgrid EMS, based on distributed computation. In this framework, the agents share a common global objective and are willing to collaborate, but refuse to disclose private information to their neighbors. Random weights are used to blur the power supply or demand information or each agent. Contrary to the use-case studied in this thesis, cooperation in this framework requires that all agents share the same global objective, instead of an individual one. Investigations are needed to assess if the privacy guarantees of the method hold when replacing the linear problem solved by the gradient descent of the lagrangian function by a more complex optimization problem (including integer variables for instance).

5.2.2 Towards distributed energy systems

Considering that energy systems will progressively switch from a centralized and scarcely digitalized organization towards a decentralized and highly monitored system, some authors have advocated the use of some concepts in the design of future EMSs, especially for the case of collective energy management.

Multi-agent systems A multi-agent system (MAS) is composed of several agents (e.g. software programs, robots, humans, ants) acting in a given environment according to a given set of rules. The primary characteristics of the agents are the following [Wooldridge, 2002]:

- Autonomy: agents at least partially independent, self-aware, autonomous;
- Local views: no agent has a full global view, or the system is too complex for an agent to exploit such knowledge;
- Decentralization: no agent is designated as controlling (or the system is effectively reduced to a monolithic system).

Inside a MAS, an agent is modeled by expliciting its individual goal (the objective function it will aim at optimizing), the interaction capabilities it has with the other intelligent agents, and the ways it might react to the changes in its environment [Kumar Nunna and Doolla, 2013]. This set of rules defines the behavior of the agents, and thus confer them some degrees of autonomy: plans of actions are put in practice according to their desires and current knowledge [Ionita, 2009].

A MAS is thus a network of intelligent and autonomous controllers. This framework seems adequate to answer the problem of untractable complexity of a centralized resolution of a local energy community optimization problem. Indeed, as the behavior, intelligence and computing power is distributed at the agent level, the evolution of the overall system will emerge as a result of the behaviors of its agents, and the approach is more scalable than centralized control [Howell et al., 2017].

[Logenthiran et al., 2011] points out the advantages of the MAS approach: unit autonomy, lower volume of data manipulation, robustness of the control system, "plug and play" capability, learning of agents.

This "plug and play" capability is especially attractive for the future studies of collective energy management systems, as the framework allows to describe different types of agents and to describe once and for all, in their interaction rules, how each agent will communicate with the different types of agents at reach, despite a potentially evolving environment. This approach could help simulate large systems made of several sub-systems, the decentralized management intelligence being more adaptable and scalable than in centralized approaches.

The JADE (Java Agent DEvelopment) framework is a popular framework for these systems [Jade, 2021].

Introduction to the holonic system approach A natural extension of the cooperation of household EMSs into a local energy community proposed in Chapter 3 is to aim at a modular cooperation between neighboring entities of the same kind, at every scale of the power system.

This idea takes form in the concept of *holonic* systems. [Howell et al., 2017] presents the principle of a holonic organization, where each entity has the dual role of both an aggregator/controller of some sub-level entities, and a part of a larger upper-level entity. Figure 5.2 shows the interactions between the elements of a holonic system. Groupments are meant to be dynamical, evolutionary according to indisponibilities or arrivals and departures of entities. The modularity of this concept is indented to bring resilience.

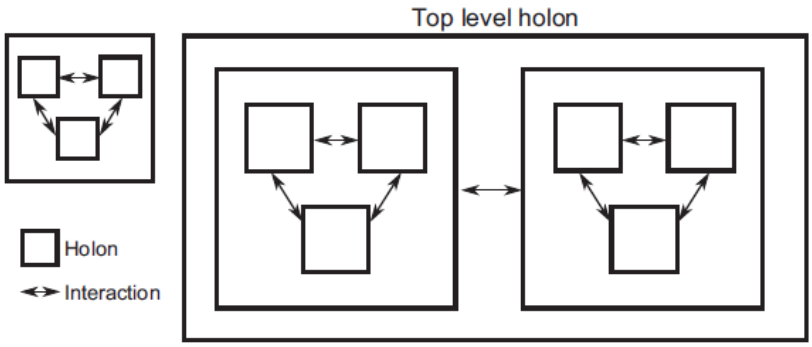

Figure 5.2: In a holonic system, each element is both a part, contained in a larger entity, and a whole, containing groups of smaller entities. Figure from [Howell et al., 2017].

The authors advocate applying this concept in the development of future EMSs, in order to work towards effective and widespread "smart grids" and "microgrids", where a completely distributed and modular organization allows entities of the same nature to communicate and collaborate, but also to be grouped into a larger entity, and this at every possible scale.

An advantage of the concept is its adequation with distributed computations, which are required in order to scale up cooperation and optimization at every level of the power system.

[Ionita, 2009] precises that "the essential feature of the basic holon is its structural self-similarity at different scales", and [Pahwa et al., 2015] explains in other words that an interest of this approach is to "design a core control holon that can be reused from the highest substation level to the lowest levels of the power distribution system".

Thus, the concept of holonic architecture could be beneficial to the design of future EMSs.

Appendix A

Differentiability of the objective function (1.6) for multiple appliances

The objective function considered in the optimization Problem 1.1 has been proven to be almost everywhere continuously differentiable in Section 1.2.4, with a finite number of discontinuity points on any interval.

In the case of several appliances scheduling, discontinuities of the derivative of the objective function (A.1) also appear at times when the support of the two appliances load curves start to overlap, that is, when the ending time of an appliance corresponds to the starting time of another one.

This appendix detailes a simple use-case which allows to illustrate the presence of such discontinuities in the objective function derivative at these specific times.

A.1 Objective function study

Consider two controllable appliances of respective load curves P_1 and P_2 shaped as boxcar functions, with starting times t_1, t_2, ending times $T_1 = t_1 + \Delta_1$, $T_2 = t_2 + \Delta_2$, and power ratings $\overline{P}_1$, $\overline{P}_2$ respectively. Here the appliances load curves are supposed to be of fixed durations Δ_1 and Δ_2 respectively. The decision variable s relative to the control strategy simplifies to (t_1,t_2), and the objective function (1.6) becomes

$$SC(t_1, t_2, \widehat{P}_{\mathrm{PV}}) = \int_0^\tau \min(P_1(t_1, t) + P_2(t_2, t), \widehat{P}_{\mathrm{PV}}(t))dt \tag{A.1}$$

First, consider the case where the supports of two appliances load curves are disjoints, P_1 preceding P_2 without loss of generality ($t_1 < T_1 < t_2 < T_2$):

$$SC(t_1, t_2, \widehat{P}_{\mathrm{PV}}) = \int_{t_1}^{T_1} \min(\overline{P}_1, \widehat{P}_{\mathrm{PV}}(t))dt + \int_{t_2}^{T_2} \min(\overline{P}_2, \widehat{P}_{\mathrm{PV}}(t))dt \tag{A.2}$$

The derivative of the objective function with respect to t_1 is

$$\begin{aligned} \frac{\partial SC(t_1, t_2, \widehat{P}_{\mathrm{PV}})}{\partial t_1} &= \frac{\partial T_1}{\partial t_1} \min(\overline{P}_1, \widehat{P}_{\mathrm{PV}}(T_1)) - \frac{\partial t_1}{\partial t_1} \min(\overline{P}_1, \widehat{P}_{\mathrm{PV}}(t_1)) \\ &= \min(\overline{P}_1, \widehat{P}_{\mathrm{PV}}(T_1)) - \min(\overline{P}_1, \widehat{P}_{\mathrm{PV}}(t_1)) \end{aligned} \tag{A.3}$$

Then, consider the case where the support of the two appliances load curves overlap, that

is, are non-null simultaneously, with P_1 shifted over P_2 ($t_1 < t_2 < T_1 < T_2$):

$$SC(t_1, t_2, \widehat{P}_{\mathrm{PV}}) = \int_{t_1}^{t_2} \min(\overline{P}_1, \widehat{P}_{\mathrm{PV}}(t))dt + \int_{t_2}^{T_1} \min(\overline{P}_1, \max(\widehat{P}_{\mathrm{PV}}(t) - \overline{P}_2, 0))dt + \int_{t_2}^{T_2} \min(\overline{P}_2, \widehat{P}_{\mathrm{PV}}(t))dt \tag{A.4}$$

Hence, in this case, the derivative of the objective function with respect to t_1 is

$$\begin{aligned} \frac{\partial SC(t_1, t_2, \widehat{P}_{\mathrm{PV}})}{\partial t_1} &= -\frac{\partial t_1}{\partial t_1} \min(\overline{P}_1, \widehat{P}_{\mathrm{PV}}(t_1)) \\ &\quad + \frac{\partial T_1}{\partial t_1} \min(\overline{P}_1, \max(\widehat{P}_{\mathrm{PV}}(T_1) - \overline{P}_2, 0)) \\ &= \min(\overline{P}_1, \max(\widehat{P}_{\mathrm{PV}}(T_1) - \overline{P}_2, 0)) - \min(\overline{P}_1, \widehat{P}_{\mathrm{PV}}(t_1)) \end{aligned} \tag{A.5}$$

The fact that P_2 is now non-null at $t = T_1$ leads to a potential difference in the two expressions of Eqs. (A.3) and (A.5). The only case where there is no difference in the two expressions is when $\widehat{P}_{\mathrm{PV}}(T_1) > \overline{P}_1 + \overline{P}_2$, so that $\widehat{P}_{\mathrm{PV}}(T_1) - \overline{P}_2 > \overline{P}_1$. Except in this specific case, two appliances can create discontinuities in the objective function each time the supports of their load curves start to overlap.

A.2 Numerical example

The mapping of the SC score for the two appliances displayed in Figure A.1 shows the non-concave nature of the objective function.

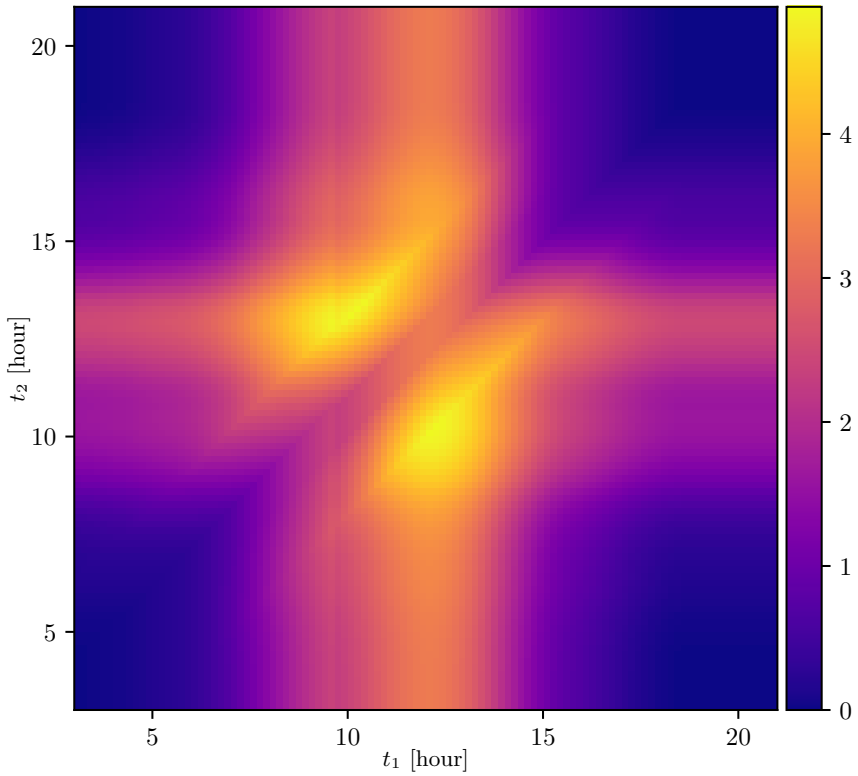

Figure A.1: The mapping of the SC score (in kWh) for two appliances shows the non-concave nature of the objective function.

At a 10-minutes timestep, the appliances scheduling that yields the maximum self-consumed volume $SC = 4.8\,\text{kWh}$ is depicted in Figure A.2 and corresponds to the starting times $t_1^\star = 12\,\text{h}10\,\text{min}$ and $t_2^\star = 10\,\text{h}10\,\text{min}$. The duration of each boxcar is $\Delta_1 = 3\,\text{h}$ and $\Delta_2 = 2\,\text{h}$. The nominal power of the appliances is $\overline{P}_1 = 2.2\,\text{kW}$ and $\overline{P}_2 = 2\,\text{kW}$. The optimal schedule happens to be for load curves of contiguous support.

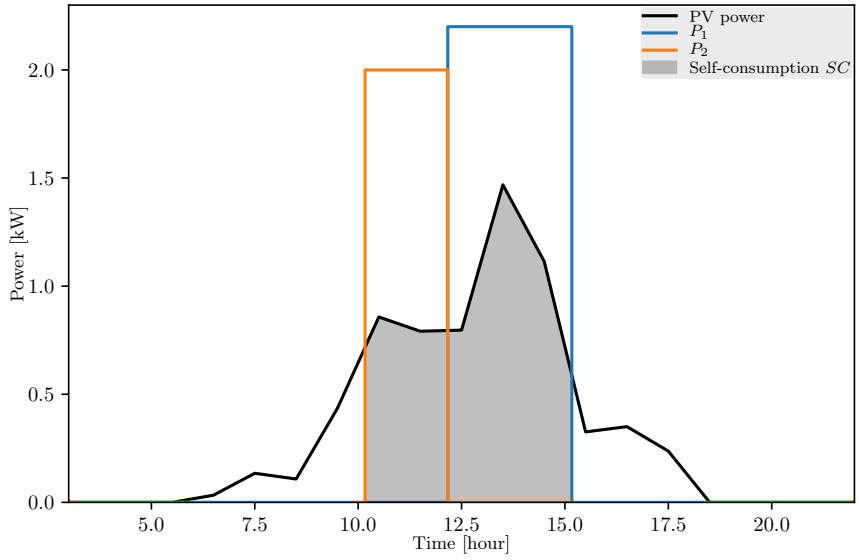

Figure A.2: The optimal scheduling of both appliances covers as much PV power as possible.

The four pairs of starting times corresponding to discontinuities in the objective function derivative can be expressed with the following equations, describing four different configurations of coincident rising or falling edges:

$$t_1 = t_2 \tag{A.6}$$
$$t_2 = t_1 + \Delta_1 \tag{A.7}$$
$$t_1 = t_2 + \Delta_2 \tag{A.8}$$
$$t_1 = t_2 + \Delta_2 - \Delta_1 \tag{A.9}$$

The discontinuities become visible when plotting the objective function partial derivative with respect to variable t_2, $\frac{\partial SC(s,\widehat{P}_{\text{PV}})}{\partial t_2}$ with variable t_1 fixed to $t_1^\star$, as shown in Figure A.3 Figure A.4 identifies the discontinuities defined in Eqs. (A.6)-(A.6) on the SC values mapping (from Figure A.1).

In the case of more complex load curves, with more edges, additional discontinuities would appear.

Hence the objective function of Problem 1.1 presents at least $2n(n-1)$ additional discontinuity points for its derivative, in the case of n controllable appliances. As the resolution method proposed in Section 1.2.4 requires to detect all the discontinuities in the objective function derivative, and to exhaustively compute the objective function on each set described by these discontinuities, its complexity increases quadratically with the number of appliances. A significant increase in the computation time can thus be expected, possibly rendering the proposed method inadequate, or too slow compared to other optimization methods.

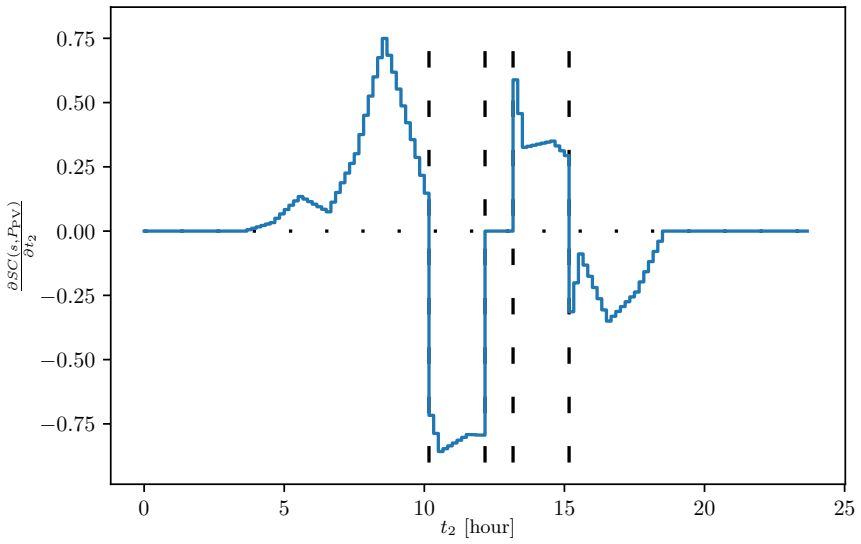

Figure A.3: Discontinuities are visible on the plot of the objective function partial derivative with respect to variable t_2.

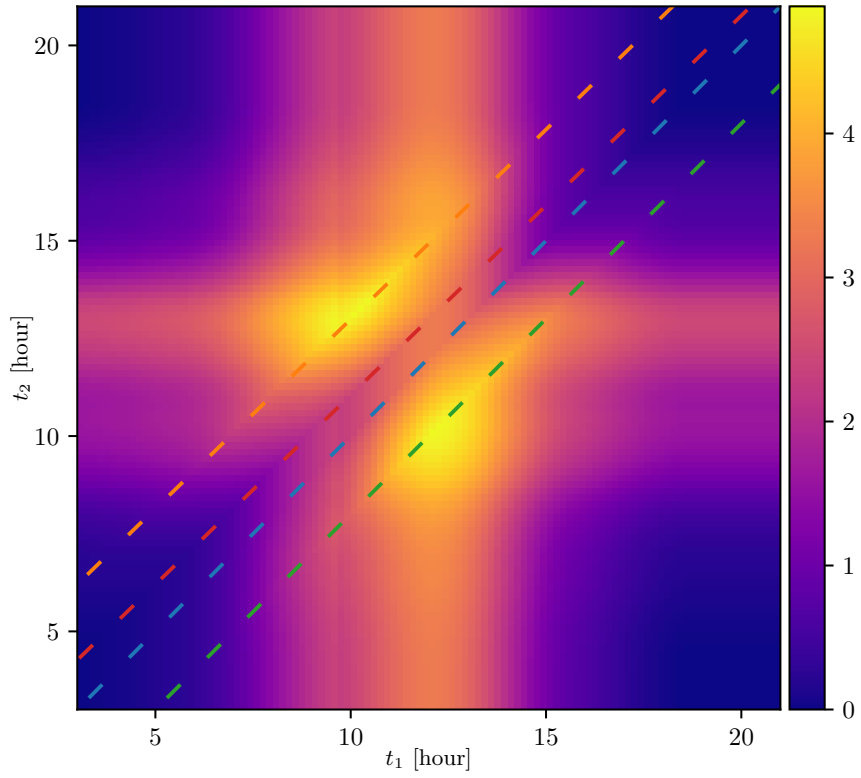

Figure A.4: The discontinuities equations can be identified on the *SC* values mapping.

Appendix B

Details on the computations of the derivative of the objective function of Problem 1.1

Section 1.2.4 presents the calculations necessary to compute the t_{sat} and t_{clear} times of the electric water heater (EWH) control. Additional details are provided here on these calculations.

B.1 Calculation of the EWH saturation times, t_{sat}

Eqs. (1.48) and (1.51) build contractive maps that converge to a fixed point.

This contraction property is used as follows. From a first approximation $\widehat{t}^{1}_{\text{sat}(1)}$ or $\widehat{t}^{1}_{\text{sat}(n+1)}$, the following approximations can be computed as $g_1\left(t_{\text{ewh}}, \widehat{t}^{1}_{\text{sat}(1)}\right) \triangleq \widehat{t}^{1}_{\text{sat}(1)}$ (Eq. (1.48)) or $g_2\left(t_{\text{clear}(n)}, \widehat{t}^{1}_{\text{sat}(n+1)}\right) \triangleq \widehat{t}^{1}_{\text{sat}(n+1)}$ (Eq. (1.51)). The calculation is repeated until the convergence[1] or a maximum number of iterations (e.g. 20) has been reached.

The contraction mapping defined by a function f has at most one fixed point if for any x_1 and x_2, there exists $\kappa \in [0, 1[$ such that:

$$|f(x_1) - f(x_2)| \leq \kappa |x_1 - x_2| \tag{B.1}$$

In the case of a function f being $\mathcal{C}^1$, the fundamental theorem of calculus states that $f(x_2) = f(x_1) + \int_{x_1}^{x_2} f'(s)ds$. It can thus be stated that

$$|f(x_2) - f(x_1)| \leq \sup_{s \in [x_1, x_2]} |f'(s)|(x_2 - x_1) \tag{B.2}$$

Hence the value of κ for the contractive map in the case at hand can be computed through the evaluation of the sup values of their derivatives.

Computing κ numerically for both functions g_1 and g_2 confirms that they form contractive maps.

B.2 Calculation of the EWH resumption times, t_{clear}

The contraction mapping approach cannot work for the calculation of the t_{clear} times however, for which the corresponding functions do not generate a contractive map.

[1] When applied on a discretized time grid, convergence means that the difference between two consecutive iterations is smaller than the timestep length.

Hence another approach is required, through a simple dichotomy procedure.

The dichotomy process is as follows:

- Initialization with $\widehat{t}^{L}_{\text{clear}(n)} = t_{\text{sat}(n)}$, $\widehat{t}^{R}_{\text{clear}(n)} = t_f$ and $\widehat{t}^{M}_{\text{clear}(n)} = \frac{\widehat{t}^{L}_{\text{clear}(n)} + \widehat{t}^{R}_{\text{clear}(n)}}{2}$.
- Calculate $\widehat{E}^{M}_{\text{clear}}$ using Eq. (1.15) and $\widehat{t}^{M}_{\text{clear}(n)}$.
  - If $\widehat{E}^{M}_{\text{clear}} > E_{\text{clear}}$, then $\widehat{t}^{M}_{\text{clear}(n)} < t_{\text{clear}(n)}$, and $\widehat{t}^{L}_{\text{clear}(n)}$ is updated through $\widehat{t}^{L}_{\text{clear}(n)} = \widehat{t}^{M}_{\text{clear}(n)}$ and $\widehat{t}^{M}_{\text{clear}(n)}$ as the mean of both bounds.
  - If $\widehat{E}^{M}_{\text{clear}} < E_{\text{clear}}$, then $\widehat{t}^{M}_{\text{clear}(n)} > t_{\text{clear}(n)}$, and $\widehat{t}^{R}_{\text{clear}(n)}$ is updated through $\widehat{t}^{R}_{\text{clear}(n)} = \widehat{t}^{M}_{\text{clear}(n)}$ and $\widehat{t}^{M}_{\text{clear}(n)}$ as the mean of both bounds.
- Start again until both bounds $\widehat{t}^{L}_{\text{clear}(n)}$ and $\widehat{t}^{R}_{\text{clear}(n)}$ converge[2], or until a maximum number of iterations has been reached (e.g. 20).

[2]When the dichotomy process is applied on a discretized time grid, convergence means that the difference between two consecutive iterations is smaller than the timestep length.

Appendix C

Appliance modeling with mixed integer variables

This appendix details the equations modeling the appliances used in the microgrid optimization problem of Chapter 3. These equations are written for a time horizon H, thus allowing to implement a repeated resolution of the problem in an MPC manner, iteratively.

C.1 Common notations

In this appendix, the following notations will be used for matrices and vectors:

For any natural integers n_1, n_2, $\mathbf{0}_{n_1,n_2}$ is the null matrix in $\mathbb{R}^{n_1 \times n_2}$, $\mathbf{I}_{n_1,n_2}$ is the matrix in $\mathbb{R}^{n_1 \times n_2}$ with ones on the diagonal starting in the upper left corner, and zeros elsewhere, $\mathbf{I}'_{n_1,n_2}$ is the matrix in $\mathbb{R}^{n_1 \times n_2}$ with ones on the diagonal starting in the lower right corner, and zeros elsewhere:

$$\mathbf{I}'_{n_1,n_2} = \begin{pmatrix} 0 & \dots & 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & \dots & 0 & 0 & 1 & \ddots & 0 & 0 \\ \vdots & \dots & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 0 & 0 & \ddots & 1 & 0 \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 & 1 \end{pmatrix} \tag{C.1}$$

and $\tilde{\mathbf{I}}_{n_1,n_2}$ is the substraction of the two previous matrices:

$$\tilde{\mathbf{I}}_{n_1,n_2} = \mathbf{I}_{n_1,n_2} - \mathbf{I}'_{n_1,n_2} \quad \forall n_1, n_2 \tag{C.2}$$

For any natural integer n, $\mathbb{1}_n$ is the vector full of ones in $\mathbb{R}^n$ and $\mathbb{0}_n$ the null vector in $\mathbb{R}^n$.

These elements have a simplified notation for the dimension H that often appears:

$\mathbf{I}$ and $\mathbf{0}$ are respectively the identity matrix and the null matrix in $\mathbb{R}^{H \times H}$.

$\mathbb{1}$ is the vector full of ones in $\mathbb{R}^H$ and $\mathbb{0}$ is the null vector in $\mathbb{R}^H$.

In this framework, the time horizon is discretized in a a finite time grid of H uniform timesteps indexed by $t \in \{1, \dots, H\}$, of length Δt.

C.2 White goods modeling

The mathematical formulation for modeling smart appliances is taken from [Sou et al., 2011] As detailed in Section 3.2.2, in the proposed scheduling framework, an appliance operation process is divided into a set of sequential energy phases.

C.2.1 Specific notations and nomenclature

For any appliance i of the white goods category of household h, the number of uninterruptible energy phases is denoted J[1].

Table C.1 lists the main notations used for the modeling of the appliances of the white goods category.

Table C.1: Appliances' model from [Sou et al., 2011] - Nomenclature

Notation	Description	Unit
E_{hij}	energy requirement for energy phase j in appliance i of household h (fixed for shiftable appliances)	Wh
l	Number of households in the microgrid	
m_h	Number of appliances in each household $h = 1, 2, \ldots, l$	
J	Number of uninterruptible energy phases for each appliance $i = 1, 2, \ldots, m_{\mathrm{h}}$ of each household $h = 1, 2, \ldots, l$	
P^t_{hij}	Power of the energy phase j of appliance i of household h during timestep $[t, t+1[$	W
$\underline{P}^t_{hij}$	lower limit of power assignment to the energy phase j for appliance i of household h during time slot $[t, t+1[$	W
$\overline{P}^t_{hij}$	upper limit of power assignment to the energy phase j for appliance i during time slot $[t, t+1[$	W
$P^t_{h,\max}$	total power upper bound at time slot $[t, t+1[$ for household h	W
$\underline{t}_{hij}$	lower limit of the number of time slots for energy phase j in appliance i of household h to be processed	
$\overline{t}_{hij}$	upper limit of the number of time slots for energy phase j in appliance i of household h to be processed	
$\hat{t}^t_{hi}$	Time preference interval: null if and only if none of the energy phases of appliance i of household h can be processed during time slot $[t, t+1[$	
w^t_{hij}	binary decision variables indicating whether a particular energy phase is being processed or not	
z^t_{hij}	binary decision variable indicating whether energy phase j in appliance i of household h is already finished by time slot $[t, t+1[$	

C.2.2 Problem setup

The number of appliances considered for scheduling is denoted m_h in household h, and the number of uninterruptible energy phases for each appliance i is denoted J for simplicity. Note that in this framework "appliance" and "energy phase" are abstractions. For instance, a single oven used for lunch and dinner can be treated as two separate appliances.

C.2.3 Decision variables

The discretized power profiles are the output of the proposed scheduler and are denoted P^t_{hij}, corresponding to the power assigned to energy phase j of appliance i of household h during the whole period of time slot $[t, t+1[$. Unit for P^t_{hij} is W. The power profiles P^t_{hij} are real (i.e., continuous) decision variables.

[1] Technically, the number of energy phases for an appliance depends on the type and model of appliance i, but is noted simply J for simplicity of the notations.

In addition to P^t_{hij}, auxiliary binary decision variables are required to indicate whether a particular energy phase is being processed or not. These binary decision variables are denoted $w^t_{hij} \in \{0, 1\}$. $w^t_{hij} = 1$ if and only if for appliance i of household h energy phase j is being processed during time slot $[t, t + 1[$. The decision variables w^t_{hij} are required, for instance, to model the energy phase sequential operation constraint described next.

In addition, another set of binary decision variables is needed to model the decision problem. This set of variables contains the $z^t_{hij}, t \in [1, H]$, with a value of one indicating that, in appliance i of household h, energy phase j is already finished by time slot $[t, t+1[$.

More explanation regarding the use of the auxiliary binary decision variables are given below.

C.2.4 Constraints

To ease the description, the constraints are organized in two groups - energy constraints and timing constraints.

Energy Constraints:

Energy phase energy requirement: To make sure that the energy phases fulfill their energy requirements, the following constraint is imposed:

$$\sum_{t=1}^{H} P^t_{hij} \Delta t = E_{hij}, \quad \forall\, h, i, j \tag{C.3}$$

where E_{hij} is the energy requirements for energy phase j of appliance i of household h.

Instantaneous energy phase power assignment bounds: To model whether an energy phase is being processed during time slot $[t, t + 1[$ as well as the lower and upper limits of power assignment to the phase, the following constraint is imposed:

$$\underline{P}^t_{hij} w^t_{hij} \leq P^t_{hij} \leq \overline{P}^t_{hij} w^t_{hij}, \quad \forall\, h, i, j, t \tag{C.4}$$

where $\underline{P}^t_{hij}$ and $\overline{P}^t_{hij}$ are appliance specific data characterizing the lower and upper limits of power assignment to the energy phases, and w^t_{hij} is a binary variable. Note that if $w^t_{hij} = 0$, then the inequalities above collapse into a single equality $P^t_{hij} = 0$.

Timing Constraints:

Energy phase process time limits: To model the limits on energy phase process time, the following constraint is enforced

$$\underline{t}_{hij} \leq \sum_{t=1}^{H} w^t_{hij} \leq \bar{t}_{hij}, \quad \forall\, h, i, j \tag{C.5}$$

where $\underline{t}_{hij}$ and $\bar{t}_{hij}$ are the lower and upper limits of the number of time slots for energy phase j in appliance i of household h to be processed.

As in [Sou et al., 2011], implementation of this constraint assumes that the operation time of the energy phases can be between 80 % and 120 % of the nominal time t^{nom}_{hij}. The bounds $\underline{t}_{hij}$ and $\bar{t}_{hij}$ in Eq. (C.5) are obtained by dividing the operation time limits $[0.8t^{\text{nom}}_{hij}, 1.2t^{\text{nom}}_{hij}]$ by the time slot length, and rounding is performed where necessary (see Eq. (C.15)).

Uninterruptible operation: An energy phase being uninterruptible implies that it cannot be resumed after being stopped. This can be modeled by the property that, for all h, i and $j, w^t_{hij} = 0$ if there exists an earlier time slot $\tilde{t} < t$ such that $w^{\tilde{t}}_{hij} = 1$ and $w^{\tilde{t}+1}_{hij} = 0$. An alternative formulation can be obtained with the aid of auxiliary decision variables z^t_{hij} (introduced in Section C.2.3) as:

$$w^t_{hij} \leq 1 - z^t_{hij} \qquad \forall\, h, i, j, t \tag{C.6}$$

$$w^{t-1}_{hij} - w^t_{hij} \leq z^t_{hij} \qquad \forall\, h, i, j, \forall\, t = 2, 3, \ldots, H \tag{C.7}$$

$$z^{t-1}_{hij} \leq z^t_{hij} \qquad \forall\, h, i, j, \forall\, t = 2, 3, \ldots, H \tag{C.8}$$

In constraint (C.6), if $z^t_{hij} = 1$, then during time slot $[t, t+1[$ energy phase j in appliance i of household h is already finished. Hence, the corresponding w^t_{hij} must be 0. The condition triggering $z^t_{hij} = 1$ is that $w^{t-1}_{hij} = 1$ while $w^t_{hij} = 0$ (i.e., the phase is just finished at time t). This is the situation in (C.7). Then, z^t_{hij} should remain unity, as (C.8) imposes it.

Sequential Processing: Sequential processing of the energy phases of an appliance means that an energy phase cannot be processed unless its preceding phases have finished. This condition can be conveniently described using the auxiliary decision variables z^t_{hij} as follows:

$$w^t_{hij} \leq z^t_{hi(j-1)}, \quad \forall\, h, i, t, \forall\, j = 2, 3, \ldots, J \tag{C.9}$$

Between-phase delay: The constraint enforcing a null delay between energy phases is

$$z^t_{hi(j-1)} = w^t_{hij} + z^t_{hij}, \quad \forall\, h, i, t, \forall\, j = 2, 3, \ldots, J \tag{C.10}$$

Note that $w^t_{hij} + z^t_{hij} \leq 1$ because an energy phase cannot simultaneously be processed and finished (see (C.6)). Hence, the equality in (C.10) is valid.

User time preference: The household user can set up the time preference constraints, specifying the time interval a particular appliance must be finished before. Alternatively, this means that the appliances cannot be run outside of the time preference interval. The constraints are written as

$$w^t_{hij} \leq \hat{t}^t_{hi} \quad \forall\, h, i, j, t \tag{C.11}$$

where $\hat{t}^t_{hi}$ characterizes the time preference interval. That is, $\hat{t}^t_{hi} = 0$ if and only if none of the energy phases of appliance i of household h can be processed during time slot $[t, t+1[$.

Notice that additional constraints could be used to model the power safety and between-phase delay constraints, as well as sequential operations between appliances, as introduced by [Sou et al., 2011]. These equations are not presented here as they are not used in the manuscript.

C.2.5 Matrices

In the thesis, this appliances model serves in a MILP formulation, with a solver requiring equality constraints to be written as two inequality constraints. All previous equations (C.3)-(C.11) for the modelization of appliance i of household h can be summarized as

$$A_{hi}\, x_{hi} \leq b_{hi} \quad \forall\, h, i \tag{C.12}$$

with the vector of decision variables

$$x_{hi} = \begin{pmatrix} \left((w^t_{hij})_{j\in[1,J]}\right)_{t\in[1,H]} \\ \left((z^t_{hij})_{j\in[1,J]}\right)_{t\in[1,H]} \\ \left((P^t_{hij})_{j\in[1,J]}\right)_{t\in[1,H]} \end{pmatrix} \tag{C.13}$$

with the constraints matrix

$$A_{hi} = \begin{pmatrix} A_1^\top & A_2^\top & A_3^\top & A_4^\top & A_5^\top & A_6^\top & A_7^\top & A_8^\top & A_9^\top \end{pmatrix}^\top \tag{C.14}$$

and the right-hand side vector

$$b_{hi} = \begin{pmatrix} (E_{hij})_{j\in[1,J]} \\ (-E_{hij})_{j\in[1,J]} \\ \mathbb{0}_{2HJ} \\ \left(-0.8\frac{t^{\text{nom}}_{hij}}{\Delta t}\right)_{j\in[1,J]} \\ \left(1.2\frac{t^{\text{nom}}_{hij}}{\Delta t}\right)_{j\in[1,J]} \\ \mathbb{1}_{HJ} \\ \mathbb{0}_{(H-1)J} \\ \mathbb{0}_{(H-1)J} \\ \mathbb{0}_{H(J-1)} \\ \mathbb{0}_{2H(J-1)} \\ \left(\hat{t}^t_{hi}\mathbb{1}_J\right)_{t\in[1,H]} \end{pmatrix} \tag{C.15}$$

in which the block matrices are now detailed.

In (C.14), the first block matrix A_1 corresponds to Eq. (C.3)

$$A_1 = \begin{pmatrix} \mathbf{0}_{2J,2HJ} & \Delta t \times \mathbf{I}_{\text{spe}} \end{pmatrix} \text{ with } \mathbf{I}_{\text{spe}} = \begin{pmatrix} \mathbf{I}_{J,J} & \dots & \mathbf{I}_{J,J} \\ -\mathbf{I}_{J,J} & \dots & -\mathbf{I}_{J,J} \end{pmatrix} \in \mathbb{R}^{2J,HJ} \tag{C.16}$$

In (C.14), the second block A_2 corresponds to Eq. (C.4)

$$A_2 = \begin{pmatrix} \underline{P}^1_{hi1} & & & & & & & & & & & \mathbf{0}_{HJ,HJ} & -\mathbf{I}_{HJ,HJ} \\ & \underline{P}^1_{hi2} & & & & & & & & & & & \\ & & \ddots & & & & & & & & & & \\ & & & \underline{P}^1_{hiJ} & & & & & & & & & \\ & & & & \underline{P}^2_{hi1} & & & & & & & & \\ & & & & & \ddots & & & & & & & \\ & & & & & & \underline{P}^H_{hiJ} & & & & & & \\ -\overline{P}^1_{hi1} & & & & & & & & & & & & \\ & -\overline{P}^1_{hi2} & & & & & & & & & & & \\ & & \ddots & & & & & & & & & & \\ & & & -\overline{P}^1_{hiJ} & & & & & & & & \mathbf{0}_{HJ,HJ} & \mathbf{I}_{HJ,HJ} \\ & & & & -\overline{P}^2_{hi1} & & & & & & & & \\ & & & & & \ddots & & & & & & & \\ & & & & & & -\overline{P}^H_{hiJ} & & & & & & \end{pmatrix} \tag{C.17}$$

For readability, the empty spaces denote null matrices of the adequate shape.

In (C.14), A_3 serves for Eq. (C.5).

$$A_3 = \begin{pmatrix} -\mathbf{I}_{\text{spe}} & \mathbf{0}_{2J,2HJ} \end{pmatrix} \text{ with } \mathbf{I}_{\text{spe}} \text{ given in Eq. (C.16).} \tag{C.18}$$

In (C.14), A_4 corresponds to Eq. (C.6)

$$A_4 = \begin{pmatrix} \mathbf{I}_{HJ,HJ} & \mathbf{I}_{HJ,HJ} & \mathbf{0}_{HJ,HJ} \end{pmatrix} \tag{C.19}$$

In (C.14), the fifth block A_5 for Eq. (C.7) is defined as

$$A_5 = \begin{pmatrix} \tilde{\mathbf{I}}_{(H-1)J,HJ} & -\mathbf{I}'_{(H-1)J,HJ} & \mathbf{0}_{(H-1)J,HJ} \end{pmatrix} \tag{C.20}$$

The last constraint (C.8) for the definition of uninterruptible operations is summarized in (C.14) with A_6

$$A_6 = \begin{pmatrix} \mathbf{0}_{(H-1)J,HJ} & \tilde{\mathbf{I}}_{(H-1)J,HJ} & \mathbf{0}_{(H-1)J,HJ} \end{pmatrix} \tag{C.21}$$

In (C.14), A_7 groups the $H \times (J-1)$ equations of Eq. (C.9):

$$A_7 = \begin{pmatrix} \mathbf{I}'_{J-1,J} & & & -\mathbf{I}_{J-1,J} & & \\ & \mathbf{I}'_{J-1,J} & & & -\mathbf{I}_{J-1,J} & \\ & & \ddots & & & \ddots \\ & & \mathbf{I}'_{J-1,J} & & & -\mathbf{I}_{J-1,J} \end{pmatrix} \tag{C.22}$$

where empty spots denote a null matrix of the adequate shape.

Eq. (C.10) is grouped in (C.14) in A_8

$$A_8 = \begin{pmatrix} -\mathbf{I}'_{J-1,J} & & & & \tilde{\mathbf{I}}_{J-1,J} & & & \\ & -\mathbf{I}'_{J-1,J} & & & & \tilde{\mathbf{I}}_{J-1,J} & & \\ & & \ddots & & & & \ddots & \\ & & & -\mathbf{I}'_{J-1,J} & & & & \tilde{\mathbf{I}}_{J-1,J} \\ \mathbf{I}'_{J-1,J} & & & & -\tilde{\mathbf{I}}_{J-1,J} & & & \\ & \mathbf{I}'_{J-1,J} & & & & -\tilde{\mathbf{I}}_{J-1,J} & & \\ & & \ddots & & & & \ddots & \\ & & & \mathbf{I}'_{J-1,J} & & & & -\tilde{\mathbf{I}}_{J-1,J} \end{pmatrix} \tag{C.23}$$

The user-time preference Eq. (C.11) is grouped in (C.14) in A_9:

$$A_9 = \begin{pmatrix} \mathbf{I}_{HJ,HJ} & \mathbf{0}_{HJ,2HJ} \end{pmatrix} \tag{C.24}$$

C.3 Storage loads modeling

C.3.1 Common constraints

In [Parisio et al., 2015], a distinction is made between *thermal loads* and *electrical loads*. Similarly, in [Wytock et al., 2017], a distinction is made between *thermal loads* and *flexible loads*. Here a distinction is made between *white goods*, i.e. appliances which processes are made up of sub-processes called energy phases (see Section C.2), and *storage loads*, i.e. appliances that consist in an energy storage, be it thermal energy as in the EWHs, or electrical energy as in stationary batteries or EVs.

Appliances of this category are supposed to be equipped with a single power element, thus their power profiles are either at 0, or at a fixed constant nominal value.

Whatever the storage load under consideration, bounds are imposed on the level of stored energy E_{ap} with $ap \in$ [ewh, bat, ev], to protect the appliance durability and usability:

$$E_{ap,\min} \leq E_{ap}^{t} \leq E_{ap,\max} \quad \forall\, t, \tag{C.25}$$

Moreover, the energy stored at the first timestep has to match the initialization level:

$$E_{ap}^{1} = E_{ap,\text{init}} \tag{C.26}$$

Finally, $\Delta t_{ap}^{\star}$ is the subset of critical timesteps for which a specified level of energy $E_{\text{ap,fin}}$ for comfort is imposed for the energy stored in the appliance (e.g. battery state of charge at the end of each day must be 50%, EV state of charge in the morning must be above 80%, see Section 3.3.1).

$$E_{ap}^{t} \geq E_{ap,\text{fin}} \quad \forall\, t \in \Delta t_{ap}^{\star} \tag{C.27}$$

The dimension of $\Delta t_{ap}^{\star}$ is noted $H_{\text{crit},ap}$. $\mathbf{I}_{\Delta t_{ap}^{\star}}$ is the matrix in $\mathbb{R}^{H_{\text{crit},ap} \times H}$ composed of a single 1 on each line, located on columns of indices in subset $\Delta t_{ap}^{\star}$. $\mathbb{1}_{\Delta t_{ap}^{\star}}$ is the vector full of ones in $\mathbb{R}^{H_{\text{crit},ap}}$.

The different appliances in this category differ only by the equations describing their dynamical behavior.

C.3.2 Electric Water Heater modeling

The EWH model corresponds to a homogeneous hot-water tank, discarding the stratification phenomenon. The discretization of the model in continuous time in Eq. (1.1) yields

$$E_{\text{ewh}}^{t} = \mathrm{e}^{-\,\mathrm{k}\,\Delta t}\left(E_{\text{ewh}}^{t-1} + (\overline{P}_{\text{ewh}}\delta_{\text{ewh}}^{t-1} - Q^{t-1})\Delta t\right) \quad \forall\, t \in 2, \ldots, H \tag{C.28}$$

where

- E_{ewh}^{t} is the thermal energy of the water stored in the EWH during timestep $[t, t+1[$;
- k is the thermal loss coefficient [Beeker-Adda, 2016];
- $\overline{P}_{\text{ewh}}^{t}$ is the constant power rating during timestep $[t, t+1[$;
- δ_{ewh}^{t} is the binary variable setting the On/Off status during timestep $[t, t+1[$[2];
- Q^{t} is the energy used during timestep $[t, t+1[$ by a hot-water drain.

Eqs. (C.25)-(C.28) can be summarized in the form $A_{\text{ewh}}\, x_{\text{ewh}} \leq b_{\text{ewh}}$ with the decision variable

$$x_{\text{ewh}} = \begin{pmatrix} (E_{\text{ewh}}^{t})_{t\in[1,H]} \\ (\delta_{\text{ewh}}^{t})_{t\in[1,H]} \end{pmatrix} \tag{C.29}$$

with the constraints matrix

$$A_{\text{ewh}} = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{I} & \mathbf{0} \\ \begin{pmatrix} 1 & 0 & \ldots & 0 \end{pmatrix} & \mathbb{0}^{\top} \\ \begin{pmatrix} -1 & 0 & \ldots & 0 \end{pmatrix} & \mathbb{0}^{\top} \\ -\mathbf{I}_{\Delta t_{\text{ewh}}^{\star}} & \mathbf{0}_{H_{\text{crit,ewh}},H} \\ \mathrm{e}^{-\,\mathrm{k}\,\Delta t}\,\mathbf{I}_{H-1,H} - \mathbf{I}'_{H-1,H} & \mathrm{e}^{-\,\mathrm{k}\,\Delta t}\,\overline{P}_{\text{ewh}}\Delta t\mathbf{I}_{H-1,H} \\ -\,\mathrm{e}^{-\,\mathrm{k}\,\Delta t}\,\mathbf{I}_{H-1,H} + \mathbf{I}'_{H-1,H} & -\,\mathrm{e}^{-\,\mathrm{k}\,\Delta t}\,\overline{P}_{\text{ewh}}\Delta t\mathbf{I}_{H-1,H} \end{pmatrix} \tag{C.30}$$

[2] As most EWHs are equipped with a unique resistive element of fixed power rating, the EWH power is either equal to its power rating or to zero.

and with the right-hand side vector

$$b_{\text{ewh}} = \begin{pmatrix} E_{\text{ewh,max}}\mathbb{1} \\ -E_{\text{ewh,min}}\mathbb{1} \\ E_{\text{ewh,init}} \\ -E_{\text{ewh,init}} \\ -E_{\text{ewh,fin}}\mathbb{1}_{H_{\text{crit,ewh}}} \\ \mathrm{e}^{-\mathrm{k}\,\Delta t}\,\Delta t(Q^t)_{t\in[1,H-1]} \\ -\mathrm{e}^{-\mathrm{k}\,\Delta t}\,\Delta t(Q^t)_{t\in[1,H-1]} \end{pmatrix} \tag{C.31}$$

In the constraints matrix, the first two lines correspond to Eq. (C.25), the two following lines correspond to Eq. (C.26), the fifth line corresponds to Eq. (C.27), and the last two lines correspond to Eq. (C.28).

C.3.3 Shared battery modeling

The battery model distinguishes the source and destination of incoming and outgoing power, as is necessary in the implementation.

In the variables used, "b" refers to the shared battery, "g" refers to the grid, and "a" refers to the agents. Thus at timestep t, P^t_{ba} is the power discharged from the battery to supply the aggregated households, P^t_{ab} is the power production surplus from the aggregated households stored in the battery, P^t_{bg} is the power discharged from the battery towards the outer grid, and P^t_{gb} is the charging power coming from the outer grid to the battery.

The dynamic behavior of the stored energy in the battery at a timestep t is described by the following equation

$$E^t_{\text{bat}} = \zeta E^{t-1}_{\text{bat}} + \left((P^{t-1}_{\text{gb}} + P^{t-1}_{\text{ab}})\eta_{\text{C}} - \frac{(P^{t-1}_{\text{bg}} + P^{t-1}_{\text{ba}})}{\eta_{\text{D}}} \right) \Delta t \quad \forall\, t \in 2, \ldots, H \tag{C.32}$$

where ζ is the proportional rate of energy degradation at each timestep, η_{C} and η_{D} are efficiency rates of the charging and discharging operations. For simplicity, the values taken for these parameters are the following: $\zeta = 1$, $\eta_{\text{C}} = 0.98$, $\eta_{\text{D}} = 0.98$.

The following constraints define the operation of the battery during charge and discharge, for all t

$$0 \leq P^t_{\text{gb}} + P^t_{\text{ab}} \leq P_{\text{bat,max}}\delta^t_{\text{C}} \tag{C.33}$$

$$0 \leq P^t_{\text{bg}} + P^t_{\text{ba}} \leq P_{\text{bat,max}}\delta^t_{\text{D}} \tag{C.34}$$

$$\delta^t_{\text{C}} + \delta^t_{\text{D}} \leq 1 \tag{C.35}$$

Constraints (C.33) and (C.34) enforce the upper and lower limitations on the charging and discharging powers at each timestep t. It is assumed that the upper limitation of powers for charging and discharging have the same value $P_{\text{bat,max}}$. To make sure that the battery is not charging and discharging at the same timestep, constraint (C.35) is enforced with binary decision variables δ^t_{C} and δ^t_{D}.

In addition, it holds

$$\sum_{t=1}^{H_{24}} \left((P^t_{\text{gb}} + P^t_{\text{ab}})\eta_{\text{C}} + \frac{(P^t_{\text{bg}} + P^t_{\text{ba}})}{\eta_{\text{D}}} \right) \Delta t \leq C \tag{C.36}$$

Constraint (C.36) ensures that the state of health of the battery is taken into account by limiting the total energy charging and discharging of the battery over a chosen period. The standard limit of the equivalent of the battery capacity C exchanged over a 24-hours period is chosen here.

A cycle is completed by the battery when the sum of the exchanged energy is equal to C, the nominal energy capacity of the battery. According to the previous constraint, only one cycle is allowed per day.

These equations can be summarized in the form $A_{\text{bat,op}}\, x_{\text{bat}} \leq b_{\text{bat,op}}$ using the decision variable x_{bat} defined as

$$x_{\text{bat}} = \begin{pmatrix} (E^t_{\text{bat}})_{t\in[1,H]} \\ (P^t_{\text{gb}})_{t\in[1,H]} \\ (P^t_{\text{ab}})_{t\in[1,H]} \\ (P^t_{\text{bg}})_{t\in[1,H]} \\ (P^t_{\text{ba}})_{t\in[1,H]} \\ (\delta^t_{\text{C}})_{t\in[1,H]} \\ (\delta^t_{\text{D}})_{t\in[1,H]} \end{pmatrix} \tag{C.37}$$

the right-hand side vector

$$b_{\text{bat,op}} = \begin{pmatrix} E_{\text{bat,max}}\mathbb{1} \\ -E_{\text{bat,min}}\mathbb{1} \\ E_{\text{bat,init}} \\ -E_{\text{bat,init}} \\ -E_{\text{bat,fin}}\mathbb{1}_{H_{\text{crit,bat}}} \\ \mathbb{0}_{2(H-1)} \\ \mathbb{0}_{4H} \\ \mathbb{1} \\ C \end{pmatrix} \tag{C.38}$$

and the constraints matrix

$$A_{\text{bat,op}} = \begin{pmatrix} A_{\text{bat,op},1} & A_{\text{bat,op},2} \end{pmatrix} \tag{C.39}$$

with

$$A_{\text{bat,op},1} = \begin{pmatrix} \mathbf{I} & \mathbf{0} & \dots & \dots & \mathbf{0} \\ -\mathbf{I} & \mathbf{0} & \dots & \dots & \mathbf{0} \\ \begin{pmatrix} 1 & 0 & \dots & 0 \end{pmatrix} & \mathbb{0}^\top & \dots & \dots & \mathbb{0}^\top \\ \begin{pmatrix} -1 & 0 & \dots & 0 \end{pmatrix} & \mathbb{0}^\top & \dots & \dots & \mathbb{0}^\top \\ -\mathbf{I}_{\Delta t^\star_{\text{bat}}} & \mathbf{0}_{H_{\text{crit,bat}},H} & \mathbf{0}_{H_{\text{crit,bat}},H} & \mathbf{0}_{H_{\text{crit,bat}},H} & \mathbf{0}_{H_{\text{crit,bat}},H} \\ -\zeta\mathbf{I}_{H-1,H} + \mathbf{I'}_{H-1,H} & -\eta_{\text{C}}\Delta t\mathbf{I}_{H-1,H} & -\eta_{\text{C}}\Delta t\mathbf{I}_{H-1,H} & \frac{\Delta t}{\eta_{\text{D}}}\mathbf{I}_{H-1,H} & \frac{\Delta t}{\eta_{\text{D}}}\mathbf{I}_{H-1,H} \\ \zeta\mathbf{I}_{H-1,H} - \mathbf{I'}_{H-1,H} & \eta_{\text{C}}\Delta t\mathbf{I}_{H-1,H} & \eta_{\text{C}}\Delta t\mathbf{I}_{H-1,H} & -\frac{\Delta t}{\eta_{\text{D}}}\mathbf{I}_{H-1,H} & -\frac{\Delta t}{\eta_{\text{D}}}\mathbf{I}_{H-1,H} \\ \mathbf{0} & \mathbf{I} & \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \vdots & -\mathbf{I} & -\mathbf{I} & \mathbf{0} & \mathbf{0} \\ \vdots & \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{I} \\ \vdots & \vdots & \vdots & -\mathbf{I} & -\mathbf{I} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbb{0}^\top & \eta_{\text{C}}\Delta t\mathbb{1}^\top & \eta_{\text{C}}\Delta t\mathbb{1}^\top & \frac{\Delta t}{\eta_{\text{D}}}\mathbb{1}^\top & \frac{\Delta t}{\eta_{\text{D}}}\mathbb{1}^\top \end{pmatrix} \tag{C.40}$$

$$A_{\text{bat,op,2}} = \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \\ \mathbb{0}^\top & \mathbb{0}^\top \\ \mathbb{0}^\top & \mathbb{0}^\top \\ \mathbf{0}_{H_{\text{crit,bat}},H} & \mathbf{0}_{H_{\text{crit,bat}},H} \\ \mathbf{0}_{H-1,H} & \mathbf{0}_{H-1,H} \\ \mathbf{0}_{H-1,H} & \mathbf{0}_{H-1,H} \\ -P_{\text{bat,max}}\mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -P_{\text{bat,max}}\mathbf{I} \\ \mathbf{0} & \mathbf{0} \\ \mathbf{I} & \mathbf{I} \\ \mathbb{0}^\top & \mathbb{0}^\top \end{pmatrix} \tag{C.41}$$

In the constraints matrix, the first two lines correspond to Eq. (C.25), the two following lines correspond to Eq. (C.26), the fifth line corresponds to Eq. (C.27), the following two lines correspond to Eq. (C.32), then four lines correspond to Eqs. (C.33)-(C.34), the twelfth line corresponds to Eq. (C.35) and the last line corresponds to Eq. (C.36).

C.3.4 Electric Vehicles modeling

The EV can be modeled exactly as a battery, only adding a time-dependent availability, and simplifying the variables. A model of higher precision can include the on-board charger (OBC) efficiency and power rating, and power rating of the EV Supply Equipment. These modeling details are dismissed.

When connected to the household EMS, the EV is necessarily parked. Thus, the dynamic behavior of the stored energy in the EV battery during driving is not modeled. Only dynamic behavior of the stored energy in the EV battery at a timestep t during a vehicle-to-grid operation is modeled by the following equation

$$E_{\text{ev}}^t = \zeta E_{\text{ev}}^{t-1} + \left(P_{\text{ev,c}}^{t-1}\eta_{\text{C}} - \frac{P_{\text{ev,d}}^{t-1}}{\eta_{\text{D}}} \right) \Delta t \quad \forall\, t \in 2, \ldots, H \tag{C.42}$$

where $P_{\text{ev,c}}$ is the incoming power charging the EV battery, $P_{\text{ev,d}}$ is the outgoing power discharging the EV battery, ζ is the proportional rate of energy degradation at each timestep, η_{C} and η_{D} are efficiency rates of the charging and discharging operations.

The EV battery is either charging or discharging, not simultaneously, thus, for all t

$$0 \leq P_{\text{ev,c}}^t \leq P_{\text{ev,max}}\delta_{\text{ev,C}}^t \tag{C.43}$$

$$0 \leq P_{\text{ev,d}}^t \leq P_{\text{ev,max}}\delta_{\text{ev,D}}^t \tag{C.44}$$

$$\delta_{\text{ev,C}}^t + \delta_{\text{ev,D}}^t \leq 1 \tag{C.45}$$

Finally the EV battery state of health is preserved by limiting the total energy exchanged over a chosen period, typically up to the equivalent of the EV battery capacity exchanged over a 24-hours period.

$$\sum_{t=1}^{H_{24}} \left(P_{\text{ev,c}}^t\eta_{\text{C}} + \frac{P_{\text{ev,d}}^t}{\eta_{\text{D}}} \right) \Delta t \leq C \tag{C.46}$$

These equations can be summarized in the form $A_{\text{ev}}\, x_{\text{ev}} \leq b_{\text{ev}}$, using the decision

variable x_{ev} defined as

$$x_{\text{ev}} = \begin{pmatrix} (E_{\text{ev}}^t)_{t\in[1,H]} \\ (P_{\text{ev,c}}^t)_{t\in[1,H]} \\ (P_{\text{ev,d}}^t)_{t\in[1,H]} \\ (\delta_{\text{ev,C}}^t)_{t\in[1,H]} \\ (\delta_{\text{ev,D}}^t)_{t\in[1,H]} \end{pmatrix} \tag{C.47}$$

the right-hand side vector

$$b_{\text{ev}} = \begin{pmatrix} E_{\text{ev,max}}\mathbb{1} \\ -E_{\text{ev,min}}\mathbb{1} \\ E_{\text{ev,init}} \\ -E_{\text{ev,init}} \\ -E_{\text{ev,fin}}\mathbb{1}_{H_{\text{crit,ev}}} \\ \mathbb{0}_{2(H-1)} \\ \mathbb{0}_{4H} \\ \mathbb{1} \\ C \end{pmatrix} \tag{C.48}$$

and the constraints matrix

$$A_{\text{ev}} = \begin{pmatrix} A_{\text{ev},1} & A_{\text{ev},2} \end{pmatrix} \tag{C.49}$$

with

$$A_{\text{ev},1} = \begin{pmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} \\ -\mathbf{I} & \mathbf{0} & \mathbf{0} \\ \begin{pmatrix} 1 & 0 & \dots & 0 \end{pmatrix} & \mathbb{0}^\top & \mathbb{0}^\top \\ \begin{pmatrix} -1 & 0 & \dots & 0 \end{pmatrix} & \mathbb{0}^\top & \mathbb{0}^\top \\ -\mathbf{I}_{\Delta t_{\text{ev}}^\star} & \mathbf{0}_{H_{\text{crit,ev}},H} & \mathbf{0}_{H_{\text{crit,ev}},H} \\ -\zeta\mathbf{I}_{H-1,H} + \mathbf{I}'_{H-1,H} & -\eta_{\text{C}}\Delta t\mathbf{I}_{H-1,H} & \frac{\Delta t}{\eta_{\text{D}}}\mathbf{I}_{H-1,H} \\ \zeta\mathbf{I}_{H-1,H} - \mathbf{I}'_{H-1,H} & \eta_{\text{C}}\Delta t\mathbf{I}_{H-1,H} & -\frac{\Delta t}{\eta_{\text{D}}}\mathbf{I}_{H-1,H} \\ \mathbf{0} & -\mathbf{I} & \mathbf{0} \\ \vdots & \mathbf{I} & \mathbf{0} \\ \vdots & \mathbf{0} & -\mathbf{I} \\ \vdots & \vdots & \mathbf{I} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbb{0}^\top & \eta_{\text{C}}\Delta t\mathbb{1}^\top & \frac{\Delta t}{\eta_{\text{D}}}\mathbb{1}^\top \end{pmatrix} \tag{C.50}$$

$$A_{\text{ev},2} = \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \\ \mathbb{0}^\top & \mathbb{0}^\top \\ \mathbb{0}^\top & \mathbb{0}^\top \\ \mathbf{0}_{H_{\text{crit,ev}},H} & \mathbf{0}_{H_{\text{crit,ev}},H} \\ \mathbf{0}_{H-1,H} & \mathbf{0}_{H-1,H} \\ \mathbf{0}_{H-1,H} & \mathbf{0}_{H-1,H} \\ \mathbf{0} & \mathbf{0} \\ -P_{\text{ev,max}}\mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -P_{\text{ev,max}}\mathbf{I} \\ \mathbf{I} & \mathbf{I} \\ \mathbb{0}^\top & \mathbb{0}^\top \end{pmatrix} \tag{C.51}$$

In the constraints matrix, the first two lines correspond to Eq. (C.25), the two following lines correspond to Eq. (C.26), the fifth line corresponds to Eq. (C.27), the following two

lines correspond to Eq. (C.42), then four lines correspond to Eq. (C.43)-(C.44), the twelfth line corresponds to Eq. (C.45) and the last line corresponds to Eq. (C.46).

Appendix D

Euclidian Algorithm for modular inverses

The following algorithm details the Euclidian Algorithm that can be used to compute u, a modular multiplicative inverse of an integer a in the Finite Field $\mathbb{Z}/p\mathbb{Z}$ (i.e. with modulo value p). It satisfies $au \equiv 1[p]$.

Algorithm D.1: Euclidian Algorithm: Compute the inverse of a modulo p, with a and p integers

Input: Integers a and p
Initialize $(r, u, v, r', u', v') := (a, 1, 0, p, 0, 1)$.
while $r' \neq 0$ **do**
  Compute the quotient q of the Euclidian division of r by r'
  Compute $(r, u, v, r', u', v') := (r', u', v', r - qr', u - qu', v - qv')$
end
Return (r, u, v) //u is the modulo multiplicative inverse of a.

The final rest r has to be 1 in the case considered, as the Algorithm 4.4 has to work with a and p being coprime integers.

Example: $a = 120$ and $p = 23$. Table D.1 shows the successively computed values for r, u, v and q. The algorithm indicates that -9 is the inverse of $a = 120$ modulo $p = 23$.

r	u	v	q
120	1	0	\
23	0	1	\
5	1	- 5	5
3	- 4	21	4
2	5	- 26	1
1	- 9	47	1

Table D.1: The Euclidian algorithm computes the r, u, v and q sequences

If a negative result is obtained but a non-negative one is needed, the value $u + p$ can be chosen.

Appendix E

Example: Anonymous Data Sharing With Power Sums

Here is reproduced an example from [Dunning and Kresman, 2013] of the execution of Algorithm 4.3.

Suppose that $l = 4$ agents h wish to share a data item d_h which takes values from 0 to 10: $d_1 = 6$, $d_2 = 10$, $d_3 = 6$, $d_4 = 2$. Choice of $p = 11$ and Finite Field $\mathbb{Z}/p\mathbb{Z}$ will serve to represent these numbers. The modulus 11 inverses needed will be $1/2 = 6$, $1/3 = 4$, $1/4 = 3$ (see Appendix D for the presentation of the Euclidian Algorithm used to compute these modular inverses). The agents compute the power sums shown in Table E.1.

d_h^n	$n = 1$	$n = 2$	$n = 3$	$n = 4$
$i = 1$	6	3	7	9
$i = 2$	10	1	10	1
$i = 3$	6	3	7	9
$i = 4$	2	4	8	5
$\sum d_h^n$	$S_1 = 2$	$S_2 = 10$	$S_3 = 0$	$S_4 = 2$

Table E.1: Powers of data values d_h chosen by each agent, modulo $p = 11$

Solving each of the Newton identities (4.25)-(4.29) in turn yields $\gamma_4 = -1 = 10$, $\gamma_3 = 2$, $\gamma_2 = 9$, $\gamma_1 = 1$, $\gamma_0 = 6$. And thus the polynomial (4.24) is

$$\Pi(x) = 10x^4 + 2x^3 + 9x^2 + 1x + 6 \pmod{p = 11} \tag{E.1}$$

All the agents receive the values S_1, S_2, S_3 and S_4 and can compute the polynomial and its roots to recover the original data items 2,6,6 and 10, but not their indices.

Appendix F

Example: Find AIDA

Here is reproduced an example from [Dunning and Kresman, 2013] for the execution of Algorithm 4.4.

Suppose that four agents participate in searching for an AIDA. The random integers in the first round y_h can range from 1 to $S = 10$: $y_1 = 6$, $y_2 = 10$, $y_3 = 6$, $y_4 = 2$. All the agents receive the values 2, 6, 6 and 10 with Algorithm 4.3, but not their indices. Agents $h = 2$ and $h = 4$ can pick $\nu_2 = 2$ and $\nu_4 = 1$ by the position of their chosen integer in U, the ordered list of unique non-null random values.

Agents $h = 1$ and $h = 3$ choose 5 and 6 respectively in the second round, while agents $h = 2$ and $h = 4$ choose 0 as they already have indices assigned at that point.

A trace of critical steps in the procedure is shown in Table F.1. Steps are identified by indices 1), …, 5) in Algorithm 4.3.

Round	Step	ξ	y_1	y_2	y_3	y_4	u_1	u_2	u_3	u_4	σ_1	σ_2	σ_3	σ_4
1	2	0	6	10	6	2								
1	3	0	6	10	6	2	2	6	6	10				
1	4	0	6	10	6	2	2	10				2		1
1	5	2										2		1
2	2	2	5	0	6	0						2		1
2	3	2	5	0	6	0	0	0	5	6		2		1
2	4	2	5	0	6	0	5	6			3	2	4	1

Table F.1: Trace of an AIDA algorithm execution

The final AIDA result is then $\sigma_1 = 3$, $\sigma_2 = 2$, $\sigma_3 = 4$ and $\sigma_4 = 1$.

Bibliography

Loris Amabile, Delphine Bresch-Pietri, Gilbert El Hajje, Sébastien Labbé, and Nicolas Petit. An optimization methodology for self-consumption of residential photovoltaic energy. In IFAC World Congress 2020 *IFAC-PapersOnLine*, volume 53, pages 13196–13203, 2020. doi:10.1016/j.ifacol.2020.12.145. URL https://linkinghub.elsevier.com/retrieve/pii/S2405896320304031.

Loris Amabile, Delphine Bresch-Pietri, Gilbert El Hajje, Sébastien Labbé, and Nicolas Petit. Optimizing the self-consumption of residential photovoltaic energy and quantification of the impact of production forecast uncertainties. *Advances in Applied Energy*, 2:100020, May 2021. ISSN 26667924. doi:10.1016/j.adapen.2021.100020. URL https://linkinghub.elsevier.com/retrieve/pii/S2666792421000135.

John E. Angus. The Probability Integral Transform and Related Results. *SIAM Review*, 36(4):652–654, December 1994. ISSN 0036-1445. doi:10.1137/1036146. URL https://epubs.siam.org/doi/10.1137/1036146.

Jura Arkhangelski, Abdou-Tankari Mahamadou, and Gilles Lefebvre. Data forecasting for Optimized Urban Microgrid Energy Management. In *2019 IEEE International Conference on Environment and Electrical Engineering and 2019 IEEE Industrial and Commercial Power Systems Europe (EEEIC / I&CPS Europe)*, Genova, Italy, June 2019. IEEE. ISBN 978-1-72810-653-3. doi:10.1109/EEEIC.2019.8783853. URL https://ieeexplore.ieee.org/document/8783853/.

Nathanael Beeker, Paul Malisani, and Nicolas Petit. Discrete-time optimal control of electric hot water tank. *IFAC-PapersOnLine*, 49(7):882–888, June 2016. ISSN 24058963. doi:10.1016/j.ifacol.2016.07.301. URL https://linkinghub.elsevier.com/retrieve/pii/S2405896316305080.

Nathanaël Beeker-Adda. *Modeling and control of electric hot water tanks: from the single unit to the group*. PhD thesis, PSL Research University, 2016. URL https://pastel.archives-ouvertes.fr/tel-01629304.

Erik Buchmann, Klemens Böhm, Thorben Burghardt, and Stephan Kessler. Re-identification of Smart Meter data. *Personal and Ubiquitous Computing*, 17(4):653–662, April 2013. ISSN 1617-4909, 1617-4917. doi:10.1007/s00779-012-0513-6. URL http://link.springer.com/10.1007/s00779-012-0513-6.

Sunliang Cao, Ala Hasan, and Kai Sirén. Analysis and solution for renewable energy load matching for a single-family house. *Energy and Buildings*, 65:398–411, October 2013. ISSN 03787788. doi:10.1016/j.enbuild.2013.06.013. URL https://linkinghub.elsevier.com/retrieve/pii/S0378778813003514.

Raffaele Carli and Mariagrazia Dotoli. Decentralized control for residential energy management of a smart users microgrid with renewable energy exchange.

IEEE/CAA Journal of Automatica Sinica, 6(3):641–656, May 2019. ISSN 2329-9266, 2329-9274. doi:10.1109/JAS.2019.1911462. URL https://ieeexplore.ieee.org/document/8707104/.

David Chaum. Untraceable Electronic Mail, Return Addresses, and Digital Pseudonyms. *Communications of the ACM*, 24(2):8, February 1981.

Jean-Christophe Culioli and Guy Cohen. Decomposition/Coordination Algorithms in Stochastic Optimization. *SIAM Journal on Control and Optimization*, 28(6):1372–1403, November 1990. ISSN 0363-0129. doi:10.1137/0328072. URL https://epubs.siam.org/doi/10.1137/0328072.

Petio Dimitrov, Luigi Piroddi, and Maria Prandini. Distributed allocation of a shared energy storage system in a microgrid. In *2016 American Control Conference (ACC)*, pages 3551–3556, Boston, MA, USA, July 2016. IEEE. ISBN 978-1-4673-8682-1. doi:10.1109/ACC.2016.7525464. URL http://ieeexplore.ieee.org/document/7525464/.

Larry A. Dunning and Ray Kresman. Privacy Preserving Data Sharing With Anonymous ID Assignment. *IEEE Transactions on Information Forensics and Security*, 8(2):402–413, February 2013. ISSN 1556-6013, 1556-6021. doi:10.1109/TIFS.2012.2235831. URL http://ieeexplore.ieee.org/document/6389771/.

ECMWF. Forecast User Guide: Forecast Ensemble (ENS) - Rationale and Construction, 2021. URL https://confluence.ecmwf.int/display/FUG/5+Forecast+Ensemble+%28ENS%29+-+Rationale+and+Construction.

EU Directive. Directive 2018/2001 of the European Parliament and of the Council of 11 December 2018 on the promotion of the use of energy from renewable sources. *Official Journal of the European Union*, L328:82–209, December 2018. URL http://data.europa.eu/eli/dir/2018/2001/oj.

EU Regulation. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). *Official Journal of the European Union*, L119:1–88, April 2016. URL http://data.europa.eu/eli/reg/2016/679/oj.

Eurostat. Disaggregated final energy consumption in households - dataset, 2018. URL https://ec.europa.eu/eurostat/databrowser/bookmark/fe627fab-b02c-4913-83fe-8cfa16ef44a0?lang=en. code : NRG_D_HHQ.

GnuPG. The GNU Privacy Guard, 2021. URL https://gnupg.org/.

Christoph Goebel, Vicky Cheng, and Hans-Arno Jacobsen. Profitability of Residential Battery Energy Storage Combined with Solar Photovoltaics. *Energies*, 10(7), July 2017. ISSN 1996-1073. doi:10.3390/en10070976. URL http://www.mdpi.com/1996-1073/10/7/976.

Gene H. Golub and Charles F. Van Loan. *Matrix computations*. Johns Hopkins studies in the mathematical sciences. The Johns Hopkins University Press, Baltimore, fourth edition edition, 2013. ISBN 9781421407944.

Felipe Gonzalez Venegas, Marc Petit, and Yannick Perez. Active integration of electric vehicles into distribution grids: Barriers and frameworks for flexibility services. *Renewable and Sustainable Energy Reviews*, 145:111060, July 2021. ISSN 13640321.

doi:10.1016/j.rser.2021.111060. URL https://linkinghub.elsevier.com/retrieve/pii/S1364032121003488.

William E. Hart, Jean-Paul Watson, and David L. Woodruff. Pyomo: modeling and solving mathematical programs in Python. *Mathematical Programming Computation*, 3(3):219–260, September 2011. ISSN 1867-2949, 1867-2957. doi:10.1007/s12532-011-0026-8. URL http://link.springer.com/10.1007/s12532-011-0026-8.

Miguel Heleno, David Rua, Carlos Gouveia, Andre Madureira, Manuel A. Matos, João Peças Lopes, Nuno Silva, and Sergio Salustio. Optimizing PV self-consumption through electric water heater modeling and scheduling. In *2015 IEEE Eindhoven PowerTech*, Eindhoven, Netherlands, June 2015. IEEE. ISBN 978-1-4799-7693-5. doi:10.1109/PTC.2015.7232636. URL http://ieeexplore.ieee.org/document/7232636/.

Shaun Howell, Yacine Rezgui, Jean-Laurent Hippolyte, Bejay Jayan, and Haijiang Li. Towards the next generation of smart grids: Semantic and holonic multi-agent management of distributed energy resources. *Renewable and Sustainable Energy Reviews*, 77:193–214, September 2017. ISSN 13640321. doi:10.1016/j.rser.2017.03.107. URL https://linkinghub.elsevier.com/retrieve/pii/S1364032117304392.

Bernardo A. Huberman, Eytan Adar, and Leslie R. Fine. Valuating Privacy. *IEEE Security and Privacy Magazine*, 3(5):22–25, September 2005. ISSN 1540-7993. doi:10.1109/MSP.2005.137. URL http://ieeexplore.ieee.org/document/1514395/.

Martin Hupez, Jean-François Toubeau, Zacharie De Grève, and François Vallée. A New Cooperative Framework for a Fair and Cost-Optimal Allocation of Resources within a Low Voltage Electricity Community. *IEEE Transactions on Smart Grid*, 12(3): 2201–2211, May 2021. ISSN 1949-3053, 1949-3061. doi:10.1109/TSG.2020.3040086. URL http://arxiv.org/abs/2010.05544.

IBM. User's Manual for CPLEX, 2020. URL https://www.ibm.com/docs/en/icos/20.1.0?topic=cplex-users-manual.

IEA. *World Energy Outlook 2018*. Paris, 2018. URL https://doi.org/10.1787/weo-2018-en.

Information Security Stack Exchange. In PGP, why not just encrypt message with recipient's public key? Why the meta-encryption?, September 2012. URL https://security.stackexchange.com/questions/20134.

Rich H. Inman, Hugo T.C. Pedro, and Carlos F.M. Coimbra. Solar forecasting methods for renewable energy integration. *Progress in Energy and Combustion Science*, 39(6): 535–576, December 2013. ISSN 03601285. doi:10.1016/j.pecs.2013.06.002. URL https://linkinghub.elsevier.com/retrieve/pii/S0360128513000294.

Silviu Ionita. Multi Agent Holonic Based Architecture for Communication and Learning about Power Demand in Residential Areas. In *2009 International Conference on Machine Learning and Applications*, pages 644–649, Miami, FL, USA, December 2009. IEEE. ISBN 978-0-7695-3926-3. doi:10.1109/ICMLA.2009.87. URL http://ieeexplore.ieee.org/document/5381373/.

IPCC. Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate

Change [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)]. Technical report, IPCC, Geneva, Switzerland, 2014a. URL https://www.ipcc.ch/site/assets/uploads/2018/02/SYR_AR5_FINAL_full.pdf.

IPCC. 2014: Summary for Policymakers. In: Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)]. Technical report, IPCC, Geneva, Switzerland, 2014b. URL https://www.ipcc.ch/site/assets/uploads/2018/02/AR5_SYR_FINAL_SPM.pdf.

Paulin Jacquot, Olivier Beaude, Stéphane Gaubert, and Nadia Oudjane. Demand side management in the smart grid: An efficiency and fairness tradeoff. In *2017 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe)*, Torino, Italy, September 2017. IEEE. ISBN 978-1-5386-1953-7. doi:10.1109/ISGTEurope.2017.8260265. URL http://ieeexplore.ieee.org/document/8260265/.

Paulin Jacquot, Olivier Beaude, Stéphane Gaubert, and Nadia Oudjane. Analysis and Implementation of an Hourly Billing Mechanism for Demand Response Management. *IEEE Transactions on Smart Grid*, 2018. ISSN 1949-3053, 1949-3061. doi:10.1109/TSG.2018.2855041. URL https://ieeexplore.ieee.org/document/8410042/.

Paulin Jacquot, Olivier Beaude, Pascal Benchimol, Stéphane Gaubert, and Nadia Oudjane. A Privacy-preserving Disaggregation Algorithm for Non-intrusive Management of Flexible Energy. In *2019 IEEE 58th Conference on Decision and Control (CDC)*, pages 890–896, Nice, France, December 2019. IEEE. ISBN 978-1-72811-398-2. doi:10.1109/CDC40024.2019.9029991. URL https://ieeexplore.ieee.org/document/9029991/.

Jade. JAVA Agent DEvelopment Framework, 2021. URL https://jade.tilab.com/.

Binod Prasad Koirala, Ellen van Oost, and Henny van der Windt. Community energy storage: A responsible innovation towards a sustainable energy system? *Applied Energy*, 231:570–585, December 2018. ISSN 03062619. doi:10.1016/j.apenergy.2018.09.163. URL https://linkinghub.elsevier.com/retrieve/pii/S0306261918314740.

Sivanand Kumar Nunna and Suryanarayana Doolla. Multiagent-Based Distributed-Energy-Resource Management for Intelligent Microgrids. *IEEE Transactions on Industrial Electronics*, 60(4):1678–1687, April 2013. ISSN 0278-0046, 1557-9948. doi:10.1109/TIE.2012.2193857. URL http://ieeexplore.ieee.org/document/6179527/.

Antoine Lefort, Romain Bourdais, Guillaume Ansanay-Alex, and Hervé Guéguen. Hierarchical control method applied to energy management of a residential house. *Energy and Buildings*, 64:53–61, September 2013. ISSN 03787788. doi:10.1016/j.enbuild.2013.04.010. URL https://linkinghub.elsevier.com/retrieve/pii/S0378778813002454.

Thillainathan Logenthiran, Dipti Srinivasan, and Ashwin M. Khambadkone. Multi-agent system for energy resource scheduling of integrated microgrids in a distributed system. *Electric Power Systems Research*, 81(1):138–148, January 2011. ISSN 03787796. doi:10.1016/j.epsr.2010.07.019. URL https://linkinghub.elsevier.com/retrieve/pii/S0378779610001823.

Esther Mengelkamp, Johannes Gärttner, Kerstin Rock, Scott Kessler, Lawrence Orsini, and Christof Weinhardt. Designing microgrid energy markets. *Applied Energy*, 210: 870–880, January 2018. ISSN 03062619. doi:10.1016/j.apenergy.2017.06.054. URL https://linkinghub.elsevier.com/retrieve/pii/S030626191730805X.

Andrés Molina-Markham, Prashant Shenoy, Kevin Fu, Emmanuel Cecchet, and David Irwin. Private memoirs of a smart meter. In *Proceedings of the 2nd ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Building - BuildSys '10*, page 61, Zurich, Switzerland, 2010. ACM Press. ISBN 978-1-4503-0458-0. doi:10.1145/1878431.1878446. URL http://portal.acm.org/citation.cfm?doid=1878431.1878446.

MSI. Marché des Équipements Domestiques de Production d'Eau Chaude Sanitaire en France 2019. Technical report, April 2019.

Angelos I. Nousdilis, Georgios C. Christoforidis, and Grigoris K. Papagiannis. Active power management in low voltage networks with high photovoltaics penetration based on prosumers' self-consumption. *Applied Energy*, 229:614–624, November 2018. ISSN 03062619. doi:10.1016/j.apenergy.2018.08.032. URL https://linkinghub.elsevier.com/retrieve/pii/S0306261918311875.

Frauke Oldewurtel. *Stochastic model predictive control for energy efficient building climate control*. PhD thesis, ETH Zurich, 2011. URL https://www.research-collection.ethz.ch/handle/20.500.11850/153334.

Frauke Oldewurtel, Alessandra Parisio, Colin N. Jones, Dimitrios Gyalistras, Markus Gwerder, Vanessa Stauch, Beat Lehmann, and Manfred Morari. Use of model predictive control and weather forecasts for energy efficient building climate control. *Energy and Buildings*, 45:15–27, February 2012. ISSN 03787788. doi:10.1016/j.enbuild.2011.09.022. URL https://linkinghub.elsevier.com/retrieve/pii/S0378778811004105.

François Pacaud. *Decentralized Optimization Methods for Efficient Energy Management under Stochasticity*. PhD thesis, Paris-Est, October 2018. URL https://pastel.archives-ouvertes.fr/tel-02134163.

Anil Pahwa, Scott A. DeLoach, Bala Natarajan, Sanjoy Das, Ahmad R. Malekpour, S M Shafiul Alam, and Denise M. Case. Goal-Based Holonic Multiagent System for Operation of Power Distribution Systems. *IEEE Transactions on Smart Grid*, 6(5):2510–2518, September 2015. ISSN 1949-3053, 1949-3061. doi:10.1109/TSG.2015.2404334. URL http://ieeexplore.ieee.org/document/7057666/.

Kaveh Paridari, Alessandra Parisio, Henrik Sandberg, and Karl Henrik Johansson. Demand response for aggregated residential consumers with energy storage sharing. In *2015 54th IEEE Conference on Decision and Control (CDC)*, pages 2024–2030, Osaka, December 2015. IEEE. ISBN 978-1-4799-7886-1. doi:10.1109/CDC.2015.7402504. URL http://ieeexplore.ieee.org/document/7402504/.

Kaveh Paridari, Alessandra Parisio, Henrik Sandberg, and Karl Henrik Johansson. Robust Scheduling of Smart Appliances in Active Apartments With User Behavior Uncertainty. *IEEE Transactions on Automation Science and Engineering*, 13(1):247–259, January 2016. ISSN 1545-5955, 1558-3783. doi:10.1109/TASE.2015.2497300. URL http://ieeexplore.ieee.org/document/7346517/.

Alessandra Parisio, Christian Wiezorek, Timo Kyntäjä, Joonas Elo, and Karl Henrik Johansson. An MPC-based Energy Management System for multiple residential microgrids. In *2015 IEEE International Conference on Automation Science and Engi-*

neering (CASE), pages 7–14, Gothenburg, August 2015. IEEE. ISBN 978-1-4673-8183-3. doi:10.1109/CoASE.2015.7294033. URL http://ieeexplore.ieee.org/document/7294033/.

Alessandra Parisio, Christian Wiezorek, Timo Kyntäjä, Joonas Elo, Kai Strunz, and Karl Henrik Johansson. Cooperative MPC-Based Energy Management for Networked Microgrids. *IEEE Transactions on Smart Grid*, 8(6):3066–3074, November 2017. ISSN 1949-3061. doi:10.1109/TSG.2017.2726941. URL https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=8004502.

Andrew J. Paverd and Andrew C. Martin. Modelling and automatically analysing privacy properties for honest-but-curious adversaries. 2014.

Peter Pflaum, Mazen Alamir, and Mohamed Yacine Lamoudi. Comparison of a primal and a dual decomposition for distributed MPC in smart districts. In *2014 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pages 55–60, Venice, Italy, November 2014. IEEE. ISBN 978-1-4799-4934-2. doi:10.1109/SmartGridComm.2014.7007622. URL http://ieeexplore.ieee.org/document/7007622/.

Peter Pflaum, Mazen Alamir, and Mohamed Yacine Lamoudi. Probabilistic Energy Management Strategy for EV Charging Stations Using Randomized Algorithms. *IEEE Transactions on Control Systems Technology*, 26(3):1099–1106, May 2018. ISSN 1063-6536, 1558-0865. doi:10.1109/TCST.2017.2695160. URL https://ieeexplore.ieee.org/document/7915748/.

Pyomo Forum - Google groups. model succeeds when equality constraint is replaced by two equivalent inequality constraints, January 2018. URL https://groups.google.com/g/pyomo-forum/c/YbUtkWzYXj0?pli=1.

Sylvain Quoilin, Konstantinos Kavvadias, Arnaud Mercier, Irene Pappone, and Andreas Zucker. Quantifying self-consumption linked to solar home battery systems: Statistical analysis and economic assessment. *Applied Energy*, 182:58–67, November 2016. ISSN 0306-2619. doi:10.1016/j.apenergy.2016.08.077. URL http://www.sciencedirect.com/science/article/pii/S0306261916311643.

Abdorreza Rabiee, Mohammad Sadeghi, Jamshid Aghaeic, and Alireza Heidari. Optimal operation of microgrids through simultaneous scheduling of electrical vehicles and responsive loads considering wind and PV units uncertainties. *Renewable and Sustainable Energy Reviews*, 57:721–739, May 2016. ISSN 13640321. doi:10.1016/j.rser.2015.12.041. URL https://linkinghub.elsevier.com/retrieve/pii/S1364032115014240.

Ritchie and Max Roser. CO and GHG Emissions by sector, September 2020. URL https://ourworldindata.org/emissions-by-sector.

Mike B. Roberts, Anna Bruce, and Iain MacGill. Impact of shared battery energy storage systems on photovoltaic self-consumption and electricity bills in apartment buildings. *Applied Energy*, 245:78–95, July 2019. ISSN 03062619. doi:10.1016/j.apenergy.2019.04.001. URL https://linkinghub.elsevier.com/retrieve/pii/S0306261919306269.

Vinay Sajip. python-gnupg - A Python wrapper for GnuPG, March 2021. URL https://gnupg.readthedocs.io/en/latest/.

Bruce Schneier. *Applied cryptography*. Wiley, New York, 2nd ed edition, 1996. ISBN 9780471128458 9780471117094.

Gianluca Serale, Massimo Fiorentini, Alfonso Capozzoli, Daniele Bernardini, and Alberto Bemporad. Model Predictive Control (MPC) for Enhancing Building and HVAC System Energy Efficiency: Problem Formulation, Applications and Opportunities. *Energies*, 11 (3), March 2018. ISSN 1996-1073. doi:10.3390/en11030631. URL http://www.mdpi.com/1996-1073/11/3/631.

SolarPower Europe. Global Market Outlook for Solar Power 2018-2022. Technical report, institution, June 2018. URL https://www.solarpowereurope.org/global-market-outlook-2018-2022/.

Fabrizio Sossan, Anna Magdalena Kosek, Sergejus Martinenas, Mattia Marinelli, and Henrik Bindner. Scheduling of domestic water heater power demand for maximizing PV self-consumption using model predictive control. In *IEEE PES ISGT Europe 2013*, Lyngby, Denmark, October 2013. IEEE. ISBN 978-1-4799-2984-9. doi:10.1109/ISGTEurope.2013.6695317. URL http://ieeexplore.ieee.org/document/6695317/.

Kin Cheong Sou, James Weimer, Henrik Sandberg, and Karl Henrik Johansson. Scheduling smart home appliances using mixed integer linear programming. In *IEEE Conference on Decision and Control and European Control Conference*, pages 5144–5149, Orlando, FL, USA, December 2011. IEEE. ISBN 978-1-61284-801-3 978-1-61284-800-6 978-1-4673-0457-3 978-1-61284-799-3. doi:10.1109/CDC.2011.6161081. URL http://ieeexplore.ieee.org/document/6161081/.

Konstantinos Steriotis, Georgios Tsaousoglou, Nikolaos Efthymiopoulos, Prodromos Makris, and Emmanouel Varvarigos. Real-time pricing in environments with shared energy storage systems. *Energy Efficiency*, 12(5):1085–1104, June 2019. ISSN 1570-646X, 1570-6478. doi:10.1007/s12053-018-9723-8. URL http://link.springer.com/10.1007/s12053-018-9723-8.

Akin Tascikaraoglu, Nikolaos G. Paterakis, Ozan Erdinc, and Joao P. S. Catalao. Combining the Flexibility From Shared Energy Storage Systems and DLC-Based Demand Response of HVAC Units for Distribution System Operation Enhancement. *IEEE Transactions on Sustainable Energy*, 10(1):137–148, January 2019. ISSN 1949-3029, 1949-3037. doi:10.1109/TSTE.2018.2828337. URL https://ieeexplore.ieee.org/document/8340882/.

Tamir Tassa and Dror J. Cohen. Anonymization of Centralized and Distributed Social Networks by Sequential Clustering. *IEEE Transactions on Knowledge and Data Engineering*, 25(2):311–324, February 2013. ISSN 1041-4347. doi:10.1109/TKDE.2011.232. URL http://ieeexplore.ieee.org/document/6081867/.

Dimitrios Thomas, Olivier Deblecker, and Christos S. Ioakimidis. Optimal operation of an energy management system for a grid-connected smart building considering photovoltaics' uncertainty and stochastic electric vehicles' driving schedule. *Applied Energy*, 210:1188–1206, January 2018. ISSN 03062619. doi:10.1016/j.apenergy.2017.07.035. URL https://linkinghub.elsevier.com/retrieve/pii/S0306261917309078.

Jean Thorey, Christophe Chaussin, and Vivien Mallet. Ensemble forecast of photovoltaic power with online CRPS learning. *International Journal of Forecasting*, 34(4):762–773, October 2018. ISSN 01692070. doi:10.1016/j.ijforecast.2018.05.007. URL https://linkinghub.elsevier.com/retrieve/pii/S016920701830089X.

UCI Machine Learning Repository. Individual household electric power consumption Data Set, August 2012. URL https://archive.ics.uci.edu/ml/datasets/individual+household+electric+power+consumption.

John S. Vardakas, Nizar Zorba, and Christos V. Verikoukis. A Survey on Demand Response Programs in Smart Grids: Pricing Methods and Optimization Algorithms. *IEEE Communications Surveys & Tutorials*, 17(1):152–178, 2015. ISSN 1553-877X, 2373-745X. doi:10.1109/COMST.2014.2341586. URL https://ieeexplore.ieee.org/document/6861959/.

VHK. Review study of ecodesign and energy labelling for water heaters and tanks. Technical report, October 2019. URL http://www.ecohotwater-review.eu.

Wikipedia. Pretty Good Privacy, October 2021. URL https://en.wikipedia.org/w/index.php?title=Pretty_Good_Privacy&oldid=1051147443. Page Version ID: 1051147443.

Ulf Wiström. Dymola - Dynamic Modeling Laboratory. 2013.

Michael J. Wooldridge. *An introduction to multiagent systems*. J. Wiley, New York, 2002. ISBN 978-0-471-49691-5.

World Bank Data Help Desk. How does the World Bank classify countries?, 2021. URL https://datahelpdesk.worldbank.org/knowledgebase/articles/378834-how-does-the-world-bank-classify-countries.

Matt Wytock, Nicholas Moehle, and Stephen Boyd. Dynamic energy management with scenario-based robust MPC. In *2017 American Control Conference (ACC)*, pages 2042–2047, Seattle, WA, USA, May 2017. IEEE. ISBN 978-1-5090-5992-8. doi:10.23919/ACC.2017.7963253. URL http://ieeexplore.ieee.org/document/7963253/.

Shangyu Xie, Han Wang, Yuan Hong, and My Thai. Privacy Preserving Distributed Energy Trading. In *2020 IEEE 40th International Conference on Distributed Computing Systems (ICDCS)*, pages 322–332, November 2020. doi:10.1109/ICDCS47774.2020.00078.

Andrew Chi-Chih Yao. How to generate and exchange secrets. In *27th Annual Symposium on Foundations of Computer Science (sfcs 1986)*, pages 162–167, Toronto, ON, Canada, October 1986. IEEE. ISBN 978-0-8186-0740-0. doi:10.1109/SFCS.1986.25. URL http://ieeexplore.ieee.org/document/4568207/.

Feng Ye, Zheyuan Cheng, Xianghui Cao, and Mo-Yuen Chow. A Random-Weighted Privacy-Preserving Distributed Algorithm for Energy Management in Microgrid with Energy Storage Devices. In *2020 2nd IEEE International Conference on Industrial Electronics for Sustainable Energy Systems (IESES)*, pages 249–254, Cagliari, Italy, September 2020. IEEE. ISBN 978-1-72814-017-9. doi:10.1109/IESES45645.2020.9210675. URL https://ieeexplore.ieee.org/document/9210675/.

Philip R. Zimmermann. *The Official PGP User's Guide*. MIT Press, Cambridge, MA, USA, 1995. ISBN 0262740176.

Acknowledgment

For all the icons in the figures: Icons made by monkik from www.flaticon.com

RÉSUMÉ

Cette thèse aborde le problème de la maximisation de l'autoconsommation de la production photovoltaïque résidentielle, par le contrôle optimal d'usages électriques domestiques. Dans la première partie, le problème d'optimisation est restreint au contrôle d'un unique chauffe-eau à effet Joule, et est réécrit comme un problème sans contraintes. Un algorithme d'optimisation efficace et rapide est proposé, et des expériences numériques montrent que ses performances dans un cadre déterministe sont meilleures que celles d'autres heuristiques. Afin d'étudier l'impact des incertitudes de production d'énergie photovoltaïque sur les performances des algorithmes de contrôle, ceux-ci sont évalués face à un grand nombre de scénarios. Une nouvelle méthode est présentée pour générer ces scénarios, de manière à ce que la répartition de l'ensemble représente l'étendue des réalisations possibles proportionnellement au niveau d'incertitude associé à la prévision initiale, et que chaque scénario individuel présente une variabilité intra-journalière réaliste. Des études numériques montrent qu'à un pas-de-temps de 30 minutes, l'impact d'une prévision de production photovoltaïque « parfaite » est négligeable en comparaison de l'impact du choix de l'algorithme d'optimisation. Dans la seconde partie de cette thèse, le problème est étendu au contrôle optimal d'une diversité d'usages répartis dans plusieurs foyers formant une communauté locale d'énergie, et partageant l'utilisation d'une batterie commune. Afin de modéliser aisément ces usages, la formulation retenue est celle d'un problème d'optimisation linéaire mixte. La contrainte de protection des données personnelles des participants au micro-réseau mène à proposer trois protocoles de communication préservant l'anonymat. Ces protocoles sont utilisés pour concaténer les contraintes des foyers individuels au sein d'un problème d'optimisation centralisé, dont le but est de minimiser le coût de fourniture de l'énergie à tout le micro-réseau, tout en préservant les données personnelles des participants.

MOTS CLÉS

Réseaux électriques intelligents, autoconsommation, prévision de production photovoltaïque, contrôle optimal de système énergétiques

ABSTRACT

This thesis addresses the problem of maximizing the self-consumption of residential photovoltaic power though the optimal control of household appliances. In the first part of the thesis, the optimization problem is restricted to the control of a unique electric water heater, and is formulated as an unconstrained problem. A novel and computationally efficient optimization algorithm is proposed, and is shown to perform better in a deterministic setting than other heuristics. In order to assess the impact of photovoltaic power production uncertainties on the control algorithms performances, the algorithms are evaluated under a large number of scenarios. A novel methodology is presented to generate these scenarios ensuring that each individual scenario presents a realistic intra-day variability, and that the set as a whole represents with proportionality the range of possible production outcomes. Numerical experiments show that at a 30-minute timestep, the impact of a ``perfect'' photovoltaic production forecast is negligible compared with the impact of the choice of the control algorithm. In the second part of the thesis, the problem is extended to the optimal control of a diversity appliances located in several households forming a local energy community and sharing the use of a community battery. A mixed integer linear programming formulation of the problem is adopted in order to ease the appliances modeling. The constraint of preserving the private data of the microgrid participants leads to proposing three privacy-preserving communication protocols. These protocols are used to communicate the individual households constraints for the microgrid energy management system to be able to concatenate them into a centralized microgrid optimization problem aiming to minimize the cost of providing energy to the whole community, while preserving privacy.

KEYWORDS

Smart grids, self-consumption, PV power production forecasts, optimal operation and control of power systems