Small-Scale Communities Are Sufficient for Cost- and Data-Efficient Peer-to-Peer Energy Sharing

Downloaded from: https://research.chalmers.se, 2024-04-18 10:44 UTC

Citation for the original published paper (version of record):

Duvignau, R., Heinisch, V., Göransson, L. et al (2020). Small-Scale Communities Are Sufficient for Cost- and Data-Efficient Peer-to-Peer Energy Sharing. e-Energy 2020 - Proceedings of the 11th ACM International Conference on Future Energy Systems: 35-46.
http://dx.doi.org/10.1145/3396851.3397741

N.B. When citing this work, cite the original published paper.

research.chalmers.se offers the possibility of retrieving research publications produced at Chalmers University of Technology. It covers all kind of research output: articles, dissertations, conference papers, reports etc. since 2004. research.chalmers.se is administrated and maintained by Chalmers Library

(article starts on next page)

Small-Scale Communities Are Sufficient for Cost- and Data-Efficient Peer-to-Peer Energy Sharing

Romaric Duvignau, Verena Heinisch, Lisa Göransson, Vincenzo Gulisano, Marina Papatriantafilou
{duvignau,verena.heinisch,lisa.goransson,vinmas,ptrianta}@chalmers.se
Chalmers University of Technology, Gothenburg, Sweden

ABSTRACT

Due to ever lower cost, investments in renewable electricity generation and storage have become more attractive to electricity consumers in recent years. At the same time, electricity generation and storage have become something to share or trade locally in energy communities or microgrid systems. In this context, peer-to-peer (P2P) sharing has gained attention, since it offers a way to optimize the cost-benefits from distributed resources, making them financially more attractive. However, it is not yet clear in which situations consumers do have interests to team up and how much cost is saved through cooperation in practical instances. While introducing realistic continuous decisions, through detailed analysis based on large-scale measured household data, we show that the financial benefit of cooperation does not require an accurate forecasting. Furthermore, we provide strong evidence, based on analysis of the same data, that even P2P networks with only 2-5 participants can reach a high fraction (96% in our study) of the potential gain, i.e., of the ideal offline (i.e., non-continuous) achievable gain. Maintaining such small communities results in much lower associated costs and better privacy, as each participant only needs to share its data with 1-4 other peers. These findings shed new light and motivate requirements for distributed, continuous and dynamic P2P matching algorithms for energy trading and sharing.

CCS CONCEPTS

• **Networks** → **Peer-to-peer networks**; • **Hardware** → *Energy generation and storage*; *Renewable energy*; *Energy distribution*; *Power networks*; *Smart grid*; Batteries; • **Applied computing**; • **Computing methodologies** → *Distributed algorithms*;

KEYWORDS

prosumer communities, P2P energy trading, distributed matching

ACM Reference Format:
Romaric Duvignau, Verena Heinisch, Lisa Göransson, Vincenzo Gulisano, Marina Papatriantafilou. 2020. Small-Scale Communities Are Sufficient for Cost- and Data-Efficient Peer-to-Peer Energy Sharing. In *The Eleventh ACM International Conference on Future Energy Systems (e-Energy'20), June 22–26, 2020, Virtual Event, Australia.* ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3396851.3397741

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
e-Energy'20, June 22–26, 2020, Virtual Event, Australia
© 2020 Association for Computing Machinery.
ACM ISBN 978-1-4503-8009-6/20/06…$15.00
https://doi.org/10.1145/3396851.3397741

1 INTRODUCTION

As politics drive large initiatives to mitigate global warming, high penetration of local renewable electricity generation is expected, thanks to its affordable prices and modularity. Residential households are shifting from pure consumers into prosumers [39], both producing and consuming energy on their own property. The development towards decentralized electricity systems and the establishment of an increasing number of sharing economies (such as sharing one's home, car, ride, etc) has resulted in *peer-to-peer* (P2P) energy sharing, or alternatively named "trading/exchanges" (see e.g. [2, 21, 24, 38, 43, 51]). P2P energy sharing allows end-users to cooperate distributively, potentially bypassing the centralized grid system, and share their energy resources (distributed small-scale generation and energy storage) to achieve greater self-sufficiency and less overall financial cost. This allows consumers to make the most of their local energy resources. The financial advantage for prosumers of sharing electricity locally is a consequence of the difference in price paid when buying from the grid (at wholesale market prices topped off by electricity taxes and grid fees) and selling it to the grid (often just slightly higher than wholesale market levels). Hence, cooperative households have interest to use their local electricity instead of the one from the grid whenever it is available [16]. Local balancing between electricity generation and consumption can also reduce the need for grid investments and the amount of information to share across the electricity system.

P2P energy sharing has been in the focus of studies within microgrids [34, 40], in the form of real-time trading processes among peers [41, 51], which induced high maintenance cost (due to the very fine-grained time aspect) and was very disruptive to the current power grid. Later instances of P2P energy trading assume longer time scale, with households forming sharing communities, with a predefined and pre-agreed cost-sharing mechanism [2, 15] among the participating households. Virtual currency and blockchain-based approaches have been proposed as concrete payment schemes in such P2P trading markets [4, 24, 28].

Motivations and Challenges: The current state-of-the-art does not tell, however, if households always have interest to team up (we show here that there is rather low expected benefit to cooperate within a community made only of prosumer households) and how to form P2P energy sharing communities. There is a strong need to understand how to form such decentralized communities in a context where energy sharing is accessible to both traditional consumer and prosumer households: which households should be included, in which way and what are the associated costs (e.g. P2P connections to maintain over time) and benefits obtained versus quantity of data to share and transfer over the network. The difficulty of forming such communities comes from the need for the peers to cooperate

in a challenging environment where distributed decisions must be taken with limited amount of knowledge of the future.

This work is about adaptive digital power systems in which renewable energy plays a dominant role. Key aspects are (i) the new predominant role of information technology (IT) and (ii) the possible shift from traditional centralized decision systems to distributed ones. Due to the distributed nature of having an increased number of active market players and forthcoming energy technology, there is need for improved understanding of:

- the granularity in terms of time-intervals and amount of data the different actors can benefit from sharing and
- the continuous, evolving nature of the decisions made by different groups of prosumers controlled by internal factors (e.g. the effects of previously made decisions).

Research Questions: In particular, this work, through detailed real-world data analysis, aims to understand when cooperation is beneficial: for which community sizes, for which production levels characterized by the sizes of photovoltaic (PV)-battery systems, and for which distribution of generation power among participants, against non-cooperative con/pro-sumers. The work also studies the energy cost optimization problem as a distributed continuous decision-making one, and paves the way to design and use online algorithms for building dynamically P2P energy trading communities, ready to be implemented in practical instances and for appropriate time-granularity. The objective is threefold:

(1) to understand which configurations of cooperative consumer and prosumer communities lead to noticeable cost savings;
(2) to identify ranges of sizes for energy production, where cooperation within energy sharing communities becomes interesting;
(3) to identify from which community sizes the gain starts to become important (community size determines the quantity of information to share on the network).

Contributions: We present a detailed study to check practical feasibility, efficiency and network costs of P2P energy sharing in concrete scenarios, based on large-scale real measured data. The starting point of the study is to extend previous energy trading community analysis [2, 15, 38, 43] towards realistic P2P instances with the following distinctions from previous state-of-the-art models:

- **Forecast Range:** we explore different prediction power while replacing *perfect foresight* by *limited prediction*: at each given time, only an estimated fraction of the future produced data is available. This turns the original optimization problem, where a single optimal decision is computed in a batch-fashion, into an online decision-making problem where continuous actions are taken, which influence the future outcomes.
- **Community Compositions:** we vary the sizes for PV and battery systems (from Very Small to Very Large local generation and storage capacities) and form communities constituted of both prosumers and non-prosumers (cf. Section 5).
- **Gain-sharing Mechanisms:** investment costs in renewable energy generation can be partially offset by sharing energy with peered households. Different strategies to split the cooperative gain (average financial advantage of cooperating, see Section 3.1) are investigated (cf. Section 5.5).

We clearly show the potential financial advantage of cooperation, when there are (i) enough local production and (ii) a non-empty set of non-producers in the community. Moreover, we show that most of the economical gain is reachable by organizing the P2P energy sharing network into a set of independent small communities, which can reduce to a large extent the amount of data to share compared to large communities (e.g. 100 or more households); our results emphasize that by grouping households into small communities of 2-5 peers, based on their consumption/production affinity (in terms of financial savings), we can reach a high gain, while keeping a low degree of data sharing over the network.

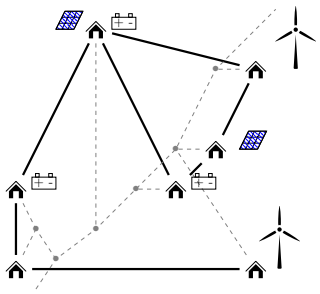

Figure 1: From fixed grid communication topology (dashed lines) towards P2P energy sharing (plain thick lines).

Paper's organization: In the following section, we provide a general view of P2P energy sharing and the specific problem we are addressing. In Section 3 we provide necessary definitions and we model the optimization of the billing cost in a P2P energy community. In Section 4, we describe how the different parameters are instantiated and how the optimization problem is addressed. Section 5 provides an extensive analysis of benefits and cost-saving impacts of having small-scale P2P energy sharing communities. Finally, related work is presented in Section 6, followed by conclusions and interesting directions of future research in Section 7.

2 PROBLEM DESCRIPTION

2.1 P2P Energy Sharing

Let us illustrate the paradigm change from non-cooperating consumers to cooperating energy sharing communities by Figure 1. In the *traditional grid infrastructure* (shown by dashed lines), prosumers (illustrated as equipped with PV panels, wind turbines and battery storage) communicate only with the grid to take local decisions about when to store energy, when to consume stored energy, when to buy and sell energy at the grid, depending on local parameters (e.g. amount of stored energy, current energy price and possibly some level of forecasting). In the *cooperation infrastructure alternative*, households communicate with peered households (via e.g. short-range wireless connection, Internet or Power Line Communication (PLC) network of the grid) to jointly take decisions on their energy trading. More precisely, they have the ability to "exchange" electricity in the sense that it is possible for a household to buy some amount of energy at zero cost, in exchange of the same amount of energy being sold to the grid by another participating household, also at zero cost. Such exchange, at the end of a predefined billing period (e.g. one year), can be satisfied by distributing among participating parties an amount of money corresponding to the "saving" obtained through cooperation. Here, the *saving* comes

from the difference in price between the energy that should have been bought at retail price and the one that should have been sold at net price. Formally, once one billing period elapses, *saving* is the difference between the bill payed by all parties together working individually and the one payed by all parties working in a cooperative fashion (taking profit of the "free" energy exchange mechanism). The individual billing process, once cooperation is activated, has to decide which part of the total bill each household should pay and is equivalent to deciding how to share the cooperative gain among participants. For instance, two households have an annual bill of respectively 1000€ (prosumer household A) and 2000€ (non-prosumer household B); if they cooperate, their combined annual bill becomes 2700€, i.e. 300€ has been saved through cooperating; one possibility is to share it by charging 850€ to A and 1850€ to B (i.e. split evenly the gain between the two).

2.2 Parameters of the Problem

In the problem we are considering, we want for each household to optimize its yearly electricity bill based on the following inputs and parameters that influence the cost-benefits of cooperation within P2P communities (with example values that stem from the data that the particular study analysed, shown in the tables below):

Input	Range
Electricity demand as hourly consumption	0-17.39 kWh
Electricity generation per installed PV capacity	0-0.86 kWh/kWp
Electricity hourly wholesale price	0-0.28 €/kWh

Parameter	Range
Number of participants in the full network	$2-100$
Individual capacities for solar generation	0-20.16 kWp
Individual capacities for battery systems	0-50.41 kWh
Quality and quantity of predicting inputs	0-48h
Number of communities formed by consumers	1-50

We consider different degrees of cooperation, from fully non-cooperative (no communication except with the centralized grid) to fully cooperative (all households form a single trading community). In between those extreme cases, several smaller communities can be managed (from 2 to 50 households) independently of each other and of the remaining households (that will work individually). The inputs must be forecast for online decision making of each household's interaction with the grid; forecasted local data needs to be shared in each community to be able to take coordinated decisions.

3 MODEL

3.1 Definitions

We consider a set of households $\mathcal{H}$ among which a subset is equipped with PV-panels and battery systems [1], so that for each household $h \in \mathcal{H}$, we have $\mathbf{PV}_h$ denoting its PV-capacity and $\mathbf{B}_h$ its battery capacity (potentially $\mathbf{PV}_h = \mathbf{B}_h = 0$). We consider discrete interval steps (hourly) during a 365-days year; hence we denote by $\mathcal{T} = \{0, 1, \ldots, 8760\}$ the set of possible time steps and by $t \in \mathcal{T}$ the number of elapsed hours from the start of the year. We use $\mathbf{cons}(h, t)$ to denote the energy consumption in kWh for household h during the t-th hour. We further provide some definitions needed to describe our model:

Average load: for a household $h \in \mathcal{H}$, it is its average hourly consumption, i.e., $\mathbf{avg_load}(h) = \sum_{t \in \mathcal{T}} \mathbf{cons}(h, t)/|\mathcal{T}|$.

Array-To-Load Ratio(ALR) [2]: for an $h \in \mathcal{H}$, ALR is the relative size of h's PV capacity $\mathbf{PV}_h$ (expressed in kWp) divided by its average load $\mathbf{avg_load}(h)$ (in kWh), i.e., $\mathbf{ALR}_h = \mathbf{PV}_h/\mathbf{avg_load}(h)$. An ALR of x for h means that h produces on average x times the electricity that it consumes and is likely to often store or sell generated electricity (when $x > 1$) or buy electricity (when $x < 1$).

Battery-To-Demand Ratio (BDR): it works in a similar fashion as ALR, but for the battery capacity, $\mathbf{BDR}_h = \mathbf{B}_h/\mathbf{avg_load}(h)$ where $\mathbf{B}_h$ is expressed in kWh. BDR can be understood as how many hours of average load can be stored by a given household.

Production level: a pair (ALR,BDR) for a given household h. In the following, we will usually deduce the values of $\mathbf{PV}_h$ and $\mathbf{B}_h$ from a given production level, i.e., a production level of (3,5) for a household h with an average load of 1.5 kWh entails $\mathbf{PV}_h = 3 \times 1.5 = 4.5$ kWp and $\mathbf{B}_h = 5 \times 1.5 = 7.5$ kWh.

Pure-consumer: a household with production level (0,0). When it is clear from the context, we use "consumer" for pure-consumer.

Prosumer: a household that is not a pure-consumer, i.e., it has at least some degree of energy production and/or storage.

Energy bill: the yearly cost for electricity for a given household. For prosumers, it depends on consumption policy (e.g. to use stored energy or buy from the grid) but for pure-consumers, it depends only on how much they consume along the year.

Community: we use "community" as a generic term to designate a set of households with a common will of optimizing their common energy bill. An ***n*-community**: a cooperative set of n households, i.e., a community of size n, and an ***m*/*n*-community** is an ***n*-community** containing exactly m prosumers.

Households pool: a set of available households that can participate in forming communities, and an n-pool is a pool of size n. The households in question may be prosumers or pure-consumers.

Configuration: it describes the production level for each household of a given pool.

Cooperative gain: (also called *saving*) for a households pool is the difference between the sum of individual bills in a given configuration (i.e., with no communication), and the single bill for the same configuration, working all together as a single community (i.e., full sharing of local data). If the pool is divided into several independent communities, the difference applies between the sum of individual bills and the sum of the bills of each community.

Perfect foresight: describes the situation where households have full knowledge of the future and can therefore always compute the optimal choice in order to minimize their energy bill. It is used here as an optimal baseline (as in the literature), but we also restrain here the knowledge available at each time step.

Limited foresight: reduces the availability of future data, in terms of duration (data for only a limited amount of future hours is available) and precision (data that is predicted may differ from the truth).

Note, with perfect foresight, the cooperative gain is always positive as each household can turn to deal with the grid if this turns out to be the best choice. However, with limited foresight, the cooperative gain can become negative as there is no guarantee that decisions taken online are the best among all possible choices.

[1] how those quantities have been set in our study is further described in Section 5.

[2] ALR/BDR were introduced in [47] and our formulation is similar to the one in [37].

Let us further for $h \in \mathcal{H}$ and $t \in \mathcal{T}$ define the following variables (all non-negative, except from the cost, that can become negative for net sellers):

- $\mathbf{bat}(h,t)$: the battery level in kWh of household h after t hours (or equivalently at time step t);
- $\mathbf{sun}(t)$: the solar profile in kWh/kWp during the t-th hour;
- $\mathbf{sun}(h,t) = \mathbf{sun}(t) \cdot \mathbf{PV}_h$, the electricity generated from solar panels for household h and time t;
- $\boldsymbol{el_{in}}(h,t)$: the electricity bought in kWh from the system for household h during the t-th hour;
- $\boldsymbol{el_{out}}(h,t)$: the electricity sold in kWh to the system for household h during the t-th hour;
- $\mathbf{price}(t)$: the electricity price per kWh for hour t, and it is assumed to be the same for all involved households;
- $\mathbf{cost}(h,t)$: the energy cost (or bill) for household $h \in \mathcal{H}$ for time step t, and $\mathbf{cost}(h,i,j) = \sum_{i \le t \le j} \mathbf{cost}(h,t)$ the cost during time span $[i,j]$ with $i,j \in \mathcal{T}$ and $i \le j$, and finally $\mathbf{bill}(h) = \mathbf{cost}(h,0,|\mathcal{T}|)$ the yearly bill of h.

The battery level at time $t \ge 1$ is assumed to behave as:

$$\mathbf{bat}(h,t) = \mathbf{bat}(h,t-1) + \mathbf{sun}(h,t) - \mathbf{cons}(h,t) + \boldsymbol{el_{in}}(h,t) - \boldsymbol{el_{out}}(h,t),$$

with the additional assumption that $\mathbf{bat}(h,0) = 0$.

The energy bill $\mathbf{cost}(h,t)$ for hour $t \in \mathcal{T}$ is then:

$$\mathbf{cost}(h,t) = \boldsymbol{el_{in}}(h,t) \cdot (\mathbf{price}(t) \cdot \mathbf{tax} + \mathbf{el_tax}) - \boldsymbol{el_{out}}(h,t) \cdot (\mathbf{price}(t) + \mathbf{el_net}).$$

3.2 Optimization Model

For household h, the optimization problem is then formulated as:

Objective function: minimize $\mathbf{bill}(h) = \sum_{0 \le t \le |\mathcal{T}|} \mathbf{cost}(h,t)$.

Constraints: $\forall t \in \{0,1,\ldots,8760\},\ 0 \le \mathbf{bat}(h,t) \le \mathbf{B}_h$.

Optimization variables: $\{\boldsymbol{el_{in}}(h,t), \boldsymbol{el_{out}}(h,t) \mid 0 \le t \le \mathcal{T}\}$.

We assume no further constraints on the system. Hence the full battery can be charged/discharged within one hour with no battery degradation, i.e., the battery of any household $h \in \mathcal{H}$ can be fully recharged (up to its maximum level $\mathbf{B}_h$) and $\boldsymbol{el_{in}}(h,t)$ can still cover h's hourly consumption $\mathbf{cons}(h,t)$, and conversely, h's battery system can be fully discharged at a rate that permits $\boldsymbol{el_{out}}(h,t)$ to reach $\mathbf{B}_h + \mathbf{sun}(t) \cdot \mathbf{PV}_h$. Furthermore, we assume no energy transmission losses, and no constraints on connection capacity to the households. Since communication faults are not considered here, each community is equivalent to a single prosumer with aggregated PV and battery capacities. This simplification allowed us to study in depth the many parameters of the P2P network (size of the communities, energy production levels, matching of the participants and prediction models) while keeping natural and reasonable assumptions about the underlying cost model [15, 16].

The target for household h is to **minimize** its bill $\mathbf{cost}(h,i,j)$ for some time-span $[i,j]$ and in particular its yearly bill. It can be shown that in an optimal solution any scenario where a household is both importing and exporting energy during the same hour leads to an unprofitable situation for h (either leading to a higher cost for the same battery level at time $t+1$, or a higher battery level for identical cost), i.e., whenever $\boldsymbol{el_{in}}(h,t) > 0$ then $\boldsymbol{el_{out}}(h,t) = 0$ and vice-versa $\boldsymbol{el_{out}}(h,t) > 0$ entails $\boldsymbol{el_{in}}(h,t) = 0$. Hence, we can summarize h's decision for time step t, to be taken after $t-1$ hours have elapsed, to setting the value for its hourly electricity balance for the next hour or equivalently its battery level after hour t has elapsed. In an online setting, h's decision for time t, i.e., the values of $\boldsymbol{el_{in}}(h,t)$ and $\boldsymbol{el_{out}}(h,t)$, determines $\mathbf{bat}(h,t)$ which in turn can be used to compute h's next decision and so on, provided that the input values (solar profile, energy consumption and electricity prices) ranging over the optimized interval are provided.

4 STUDY METHODOLOGY AND TOOLS

4.1 Dataset

We have based our study on residential households, some of which equipped with solar panels and battery systems. As shown in previous studies [15, 16], the most determinant parameter to incentivize prosumers to more collaboration and more self-consumption of electricity is the difference in price between electricity bought from and sold to the grid (itself determined by electricity spot market prices, transfer fees, tax rates, etc).

Households Demand: Consumption profiles for the 100 Northern European households of the study originate from [37] and were pre-processed as in [15]. Each trace is an hourly-measured year of electricity consumption for a different household in 2012-2013. Out of the original profiles, 100 were selected for belonging to neighboring geographical areas. The profiles have a relatively wide range of consumption behaviors due to different types of heating equipment and appliances, as well as different occupant capacities (see Figure 2 for range of electricity bills).

Electricity Prices: The basis for the hourly variable electricity price (wholesale electricity price) used in this study corresponds a projection for year 2030 of a scenario representing a European system pathway with a high share of renewable power generation, shorter lifetimes for nuclear power plants, and no carbon capture and storage. Energy tax, grid fees, VAT, small reimbursement for selling to the grid and investment costs have been set as described in Section 3.2 based on current assumptions for residential PV battery systems for the year 2030 (for more details see [15, 16] and references therein).

Solar Production: The solar profile is based on the geographical location of the households as described in [36]. We assume here that they all have similar roof panel orientations such that their respective PV production can be computed directly from the common profile and the capacity of their respective installation.

4.2 Taxes and Investment Costs

We have the following tax levels [3] $\mathbf{tax} = 1.25$, $\mathbf{el_tax} = 0.0690$ €/kWh and $\mathbf{el_net} = 0.0058$ €/kWh. The annual costs for the PV installation (per kWp) and battery system (per kWh) are obtained by assuming the following raw costs for the different equipments with corresponding lifetime and an interest rate of 4%:

Equipment	Nominal Cost	Lifetime	Annual Cost
PV panels	1200 € per kWp	30	69.4 € per kWp
Inverter	120 € per kWp	15	10.79 € per kWp
Battery	150 € per kWh	12.5	15.48 € per kWh

The annual investment cost is obtained by multiplying 80.1891 (sum of the annual costs of PV panels and Inverter) by each household's PV capacity and 15.48 by its battery capacities, which in turns

[3] the value of **el_tax** is in line of a taxable transfer fee of 2.3 € cents/kWh in addition to a non-taxable rate of 4.03 € cents/kWh.

are obtained by multiplying a given production level (ALR,BDR) by the average household load. *Saving including investment costs* reduces the reduction in the yearly electricity bill by the annual investment cost, and can easily become negative when the original saving does not cover up for the investments.

4.3 Forecasting

Limited foresight: In order to minimize $\mathbf{cost}(h, i, j)$, values for the different inputs over the timespan $[i, j]$ are required. In an online setting, such values are not known ahead of time and need therefore to be forecasted. In this work, instead of using a specific recent technique for performing this form of **Model Predictive Control (MPC)** (many can be found in the literature), we have kept our modeling generic and have accounted different predictive accuracies to test the influence of different forecasting power.

The main parameters in our abstracted forecast are the **window size n** and the **accuracy level**:

- **n** denotes how many hours are forecasted at a time. With a window size of n hours and time step $t < |\mathcal{T}| - n$, we want to optimize $\mathbf{cost}(h, t, t + n)$ by forecasting the solar profile, electricity prices and consumption for household h, over the interval $[t, t + n]$.
- **accuracy level** represents how precise the forecasted values are, compared with the ground truth.

We set three different levels of accuracy namely:

(1) **Truth**: all values used are exactly the measured ones (note this is different from perfect foresight as the forecasted window limits the size of available data at each optimization step), hence

$$\mathbf{truth}(h, t) = \mathbf{value}(h, t),$$

with $\mathbf{value}(h, t)$ being either $\mathbf{sun}(h, t)$, $\mathbf{price}(t)$ or $\mathbf{cons}(h, t)$ for the different inputs of the optimization problem.

(2) **Average**: each value is forecasted to be the average of the last seven previous days at the same hour, i.e.,

$$\mathbf{avg}(h, t) = \frac{1}{7} \sum_{i=1}^{7} \mathbf{value}(h, t - i \cdot 24).$$

(3) **Linear**: each value is a linear mix of the truth and the previously defined average such that while forecasting interval $[t, t + n]$, the value used for hour $t + i$ for $1 \leq i \leq n$ is assumed to be

$$\mathbf{linear}(h, t + i) = \frac{n - i}{n} \cdot \mathbf{truth}(h, t + i) + \frac{i}{n} \cdot \mathbf{avg}(h, t + i).$$

Our model allows to abstract the forecasting power into simple categories: *best possible* (but with limited depth), a *basic* one (a simple average over the orevious week, also referred as "persistance" prediction) and *intermediate* (between the two aforementioned). The linear accuracy level offers a natural way to simulate degrading accuracy performance, with first hours close to the truth thanks to potentially elaborate prediction mechanism, while later hours being poorly predicted (not much better than taking the average). In our dataset, the *truth* level produces no error, the *average* entails an average absolute error of 0.59 kWh for electricity consumption (31% of average load), 0.054 kWp (42% of average profile) for solar generation and 0.0096 €/kWh for electricity prices (23% of average prices), and finally the *linear* presents an error of i/n times the average error for forecasting the i-th next hour within an n-hour window .

4.4 Myopic Decisions

Those are common heuristic online approaches presented here for comparison purposes with limited forecast. Instead of forecasting a fixed-length window of future inputs with different degrees of accuracy, the approach takes the decision solely based on the current state of the system. We present two online strategies here, depending if the households work individually or within a group; the latter is introduced and presented in detail in [2].

Greedy Individual: demand is satisfied first by own PV generation; any surplus is stored in the local battery (electricity is sold to the grid whenever the battery is full); any remaining demand is satisfied by the locally stored electricity and in last resort the grid is used to buy electricity when the battery is empty.

Greedy Coalition: it is composed of 4 phases:

(1) *Satisfy local demand*: First households satisfy their local demand from their own production, then either (i) for the remaining demand, they use the stored electricity from their local battery system, and any remaining demand is added to a *common demand* pool, or (ii) the surplus generated electricity is stored first in the local battery, and when the battery is full, it is added to a *common generation* pool.
(2) *Satisfy remote demand*: the common generation pool is used to satisfy any remaining demand in the group (common demand).
(3) *Charge remote batteries*: if the common generation pool is not empty, it is used to charge remote batteries in the group.
(4) *Grid*: The grid is then used in last resort to buy what is left in the common demand, or to sell electricity that could not have been stored (what is left in the common generation).

4.5 Distributed Approach and Data Sharing

Perfect Foresight: For a single household, the optimal solution (minimal yearly bill) has been computed using a single run of a linear programming solver (gplx through pyomo interface [12]). For a given community, the optimal solution is obtained by considering the full community as a single large household (summing their respective consumptions, PV and battery capacities) and taking decisions for all coordinated households. This model, used as a baseline in our work, presents the best possible obtainable cooperative gain as it uses perfect knowledge of the future.

Limited Foresight: Here the decisions at each given step are computed in an iterative fashion based on current forecasted data and the previous online decisions taken. For a single household, this represents $|\mathcal{T} - n|$ optimization problems to solve in order to compute its yearly bill for a forecasting window size of n hours. Community solving works similarly to the perfect foresight case and considers a set of households as forming a single virtual household. This implies that between each given time step, all necessary information (each household forecasted consumption values for the next window size) must be shared among all peers. This directly influences the communication overhead. In particular, amount of communication among a community grows quadratic to its size when data is flooded every hour and decisions are computed in a distributed fashion, and at least linearly when data is gathered at one peer, that needs, after computing the solution on shared data, to propagate the common decisions to all the other peers. One may note, however, that the *average* accuracy level of prediction does not need all the forecasted

window but only a single value to be shared among participating peers, as each peer can recompute the average prediction locally, based on previously shared data (keeping track in this case of the last 7 days of data for each cooperating household).

Myopic Decisions: Those approaches do not require any future data and rely only on the current state. Hence, only the current production/consumption values need to be shared among the peers in the greedy coalition.

4.6 Prosumers-Consumers Matchings

In our study, we explore in more detail the smallest of the "small-scale communitities" corresponding to dividing the P2P network into independent pairs. This organization provides the best solution in terms of data sharing costs and privacy preserving considerations as local data need only to be shared with a single other household. We will consider in this work several ways of achieving this division.

First, let us recall that a matching or independent edge set in a graph is a set of edges that do not share any endpoints. In our setting, our underlying graph is the complete graph whose vertex set is the households pool. Moreover, we will focus on the case where the pool is split into two sets, prosumers on one side and pure-consumers on the other and we will only form prosumer/consumer pairs. The problem of finding a matching set in a bipartite graph (also referred as "assignment problem") consists in finding a partner for each member of the first set (here, prosumers). Edges in the prosumer/consumer graph also have a natural weight in our context: the cooperative gain achieved by the given pair. Hereafter, matchings of particular interest:

Maximum-weight matching: corresponds to the matching of prosumers with consumers whose sum of weights is maximum over all matchings (corresponds to best average cooperative gain).

Greedy matching: the greedy matching selects first the available pair that entails the highest saving and continue until all prosumers are matched.

"Greedy-largest" matching: pairs first prosumers with the largest average load available consumers (starting with the prosumer with highest load) until all prosumers are matched.

Finding the maximum-weight matching in weighted bipartite graph is a classical problem that is solved by the "Hungarian algorithm" [25]. Let us remark that the distributed version is much more difficult [30, 46] and to be complete in our case, we need to add the extra difficulty that the weights are not provided but can only be estimated (potentially using past and forecasted data) and require extra communication between peers.

4.7 Gain-Sharing Mechanism

The way the cooperative gain is split among the participating peers may influence the target size of households wishing to invest in producing their electricity locally, and in the same time desire for pure-consumers to integrate P2P sharing communities. We explore 4 different splitting strategies for the cooperative gain:

- **No-split**: the cooperative gain is shared only among prosumers in the community.
- **1/2-split**: half of the gain is shared evenly among prosumers and half evenly among all participants.

Name	ALR	BDR	Bill	Inc.	PV	Bat.	# PV	Surf.
None	0	0	2034	2034	0	0	0	0
Very Small	0.5	1	1872	1975	0.93	1.85	3	5
Small	1.5	2.5	1610	1904	2.78	4.63	9	15
Medium	3	5	1284	1872	5.55	9.25	18	30
Large	4.5	10	950	1905	8.33	18.51	27	45
Very Large	6	15	658	1978	11.1	27.76	36	60

Table 1: Selected production levels with corresponding ALR/BDR, average yearly bill with and without investment costs (€/year), PV capacity (kWp), Battery capacity (kWh), and as indication the equivalent number of common 300W panels and covered roof surface (in m^2 for 1.63m^2 panels).

- **3/4-split**: similar to 1/2-split but 3/4 of the gain is redirected towards prosumers.
- **Even-split**: the cooperative gain is shared evenly among all participants in the community (irrespective of their consumption and generation).

5 SMALL-SCALE COMMUNITIES: A CASE STUDY OF 100 HOUSEHOLDS

We start our case study by setting different production levels and then investigate, in each of the following subsections, the impact of one or more of the parameter(s) on the cooperative gain.

Production Levels: The study aims to explore how the size of the community and the predictive power (among other parameters) influence the potential financial interest for residential households to share their energy resources (here, PV panels and battery systems). To investigate different configurations, we define in the first place 6 production levels considered *reasonable* for residential households that are further detailed in Table 1. The selected production levels range from "Very Small" to "Very Large" PV/Battery systems and have been chosen to cover a wide range of common households' installation sizes where one would invest in larger capacities on both electricity generation and storage in a similar manner (further details are provided in Appendix A). The selected values allow to simply reason in terms of average load, electricity generation and storage, to relate one level to another, and the selected levels correspond to similar options in recent literature [1, 17, 19, 37].

Yearly Bills: Figure 2 describes a grouping of the distributions of the yearly bills (excluding investment) over the full dataset, with each house individually optimizing its bill with perfect foresight, for

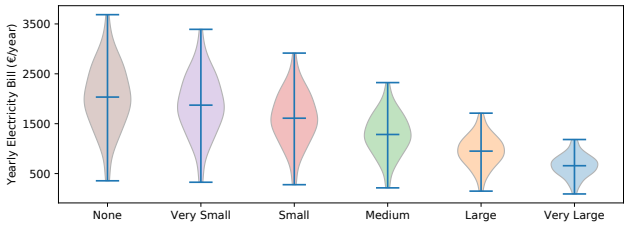

Figure 2: Distribution of the yearly electricity bills over the dataset for the different selected production levels.

different production levels (each with a distinct color, preserved in the rest of the study). In quantitative terms, the yearly bill varies by tenfold (ranging from 355 to 3685€ for pure-consumers, down to 91-1182€ when households have "Very Large" productions) and most household bills are around the average values (given in Table 1).

5.1 Which Types of Configurations Induce High Cooperative Gain?

Figure 3 shows the exploration of the relation between the average cooperative gain and the fraction of prosumers in the community. The figure has been generated using 300 experiments and averaging the results for each displayed configuration. Each curve corresponds to a different fraction of prosumers in the community whereas the x-axis ranges over the average ALR in the full community (including both prosumers and pure-consumers). The figure has two take-home messages. First, the fewer prosumers the higher the average gain; even though this might sound slightly counter-intuitive at first, it follows from the fact that cooperation is most useful to help, especially during high electricity price periods, households with (close to) no resources. Second, all configurations take advantage of a higher amount of resources in the community while cooperating, except from the 100% prosumers case, where the gain caps around an ALR of 3 and decreases afterwards (the latter is in line with previous studies for yearly optimization [15, 37]).

5.2 How Dependent on the Size of the Community Is the Cooperative Gain?

Let us follow by showing that small-scale communities provide already a high share of the cooperative gain in Figure 4. The figure presents the average saving per household for different production levels, Very Small (thinnest error bar) to Very Large (largest ones), and for different configurations [4] (for m/n-communities with $n \geq 10$, m has been selected to produce the best cooperative gain with a Very Large production level). As a comparison point, the lowest gain for 100 households is displayed: it corresponds to a 100% prosumers community of households with homogeneous production levels.

For each configuration of the form "x equipped households over y total households", we run 50 experiments selecting first y households randomly over 100; then x over y households are randomly selected and equipped with electricity production and storage. Each

[4] Results for all configurations are provided in Appendix B.

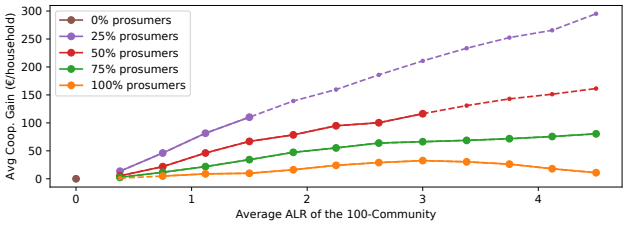

Figure 3: Average cooperative gain versus average ALR for different fraction of prosumers (dashed lines for average prosumers ALR outside [0.5,6], cf. Production Levels § 5).

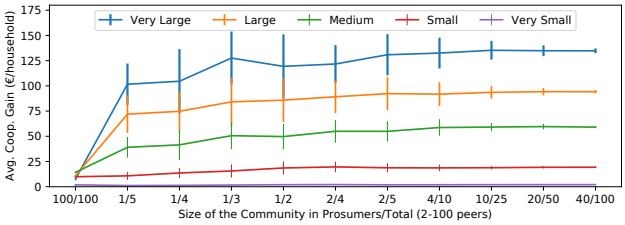

Figure 4: Small-Scale Effect: Avg. cooperative gain (€/year) in small-scale communities (2-5 peers) against best configurations in 10-100 communities for different production levels.

error bar in the figure, centered at the average saving for the corresponding configuration, spans the standard deviation of the 50 experiments. For Very Large production levels, a 1/2-community reaches 88% of the best configuration, a 1/3-community 94%, a 2/4-community 90% and a 2/5-community 97%. Moreover, small communities with low-share of prosumers such as 1 prosumer for 3 or 4 pure-consumers, can achieve 75-77% of the best average achievable cooperative gain. We note this also holds for smaller production levels.

The gain is explained by the ability of small-scale communities to produce similar levels of self-consumption as the one of large communities such as 100 households as depicted on Figure 5 for a Medium production level. In fact, they reproduce very similar increases in self-consumption as a large community with the same ratio of prosumers; let us note, such an increase is also highest when about 50% of the household pool are prosumers.

5.3 What Influences Cooperative Gain in a Households Pool Divided in 2-Communities?

The precedent analysis suggests that one can obtain the cooperative gain of large communities of up to 100-households (in the best 50%-prosumers configuration) with only pairs of households at a much lower cost. We further explore in this section the relation between forming pairs of cooperative households and the consequence in cooperative gain. For this purpose, we pick the Very Large production

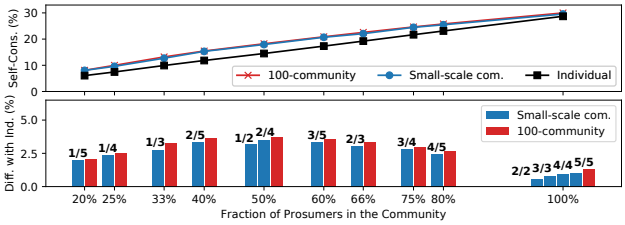

Figure 5: Electricity local self-consumption (upper graph) within small-scale communities (2-5 peers) compared with a 100-community and a household pool without cooperation (individual); difference in self-consumption between cooperative and individual (lower graph).

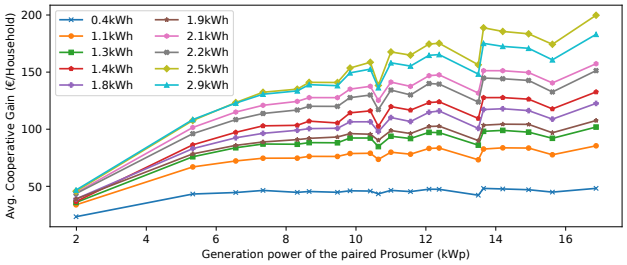

Figure 6: Avg. cooperative gain for different pairs of households: 10 pure-consumers illustrated by a different color/symbol each are matched with 20 prosumers.

level as it offers the biggest gain in absolute value and as similar outcomes can be presented for smaller production levels.

As shown on Figure 4, a pair of households with a prosumer matched with a pure-consumer offers on average 120€ of cooperative gain, which is about 89% of the best possible gain when the 100 households are working together. However, the gain is directly impacted by the chosen pair and the 2-households cooperation is the one displaying the largest variation with a standard deviation about 17 times larger than the 100-community. Hence, the way households are paired together plays a great role on how much gain the full community is able to reach from cooperating.

Figure 6 presents the cooperative gain for 10 different pure-consumers with varying average consumption (from average load of 0.4kWh up to 2.9kWh), paired with 20 different prosumers at Very Large production level and also progressive average load (from 1.99kWp to 16.9kWp). The figure illustrates several phenomena:

(1) the main factor for the amount of cooperation is the respective consumption and production level in the paired couple; in general, large consumers and large producers will entail more gain than small consumers/producers (up to more than 4-folds); however, after a certain threshold, matching a consumer with larger producers does not imply much more cooperative gain;
(2) some households are particularly poor cooperative prosumers (two down-peaks on the curves at 10.64 and 13.5kWp): regardless of the paired consumer, they will provide less gain than similar producers;
(3) some households are particularly good (e.g. the ones with average loads 1.4kWh and 2.5kWh) or bad (e.g. 1.9kWh and 2.2kWh) consumers, in the sense that, independently of which prosumer they are paired with, they will always entail better/worst cooperative gain than households with similar average load.

Finding "poor/good" prosumers and consumers in the pool is explained by differences in the individual load patterns of each household, and such differences have a larger impact on a cooperative pair if compared with a larger group. For instance, a typical affinity is to have daily peaks at different time, potentially allowing both peaks to be covered by local electricity production and storage.

This observation motivated us to further explore how to match a set of prosumers and consumers into pairs (that can be generalized to other small scale such as 3/4/5-matchings) in order to achieve a better overall average gain for the full set.

Individually, any household will have a tendency to prefer to be matched [5] with the available household that has the largest average load, once neglecting preferences as shown in Figure 6.

We have explored the consequence of pairing a fixed set of households in different manners and present the results in Figure 7. The household pool is constituted here of either 20 or 100 households, among which the same subset of 10 households are equipped with a Large production level, the remaining ones are pure-consumers, and we consider only pairing prosumers with non-prosumers. The 20-pool corresponds to selecting for each prosumer the consumer with closest load among the 90 of them in the 100-pool, hence it features a very similar distribution of loads between prosumers and consumers. Figure 7 presents the average cooperative gain achieved when the households are paired with each other randomly [6], the worst and best matchings [7], the greedy-largest and the greedy matching (cf. Section 4.6). We added and took as baseline the saving obtained when the 20-pool forms a single community (optimal solution in that case), and when the 100-pool forms a single community. Note that even though the 100-community provides less per household saving, the total absolute saving is about twice what is achievable when only 20 of them cooperate.

All matchings in the 20-pool produce cooperative gain within 80-96% of the optimal solution and a random matching already achieves 88% of it. Moreover, the greedy choice lies within 0.05% behind the best possible matching of the households hence provides a natural choice for forming rapidly 2-communities. If there exist more consumers than prosumers (100-pool case), then the choice of matching (which implies a selection on whose households do participate in energy sharing at the end) has a much greater impact. In that situation, the greedy matching is again a very good candidate lying less than 1% behind the best matching; on the contrary, choosing a matching poorly provide a 58% decrease in gain and on average a 29% reduction is observed.

At last, we note that pairs may not entail the greater total benefits that can be achieved by involving more pure-consumers in the sharing process (instead of using only 40 over the 100 available households). This advocates in this particular 20/100-community example to use larger peer neighborhoods, such as matching 3 to 4 consumers to a single prosumer to form 1/4- or 1/5-communities.

[5] This is true if e.g., the gain offered by the pair would be shared between the pair and not with the rest of the pool, see Section 5.5.
[6] averaged by enumerating all possible matchings for the 20-pool and after generating 10^7 random matchings in the 100-pool.
[7] computed exactly by the Hungarian algorithm.

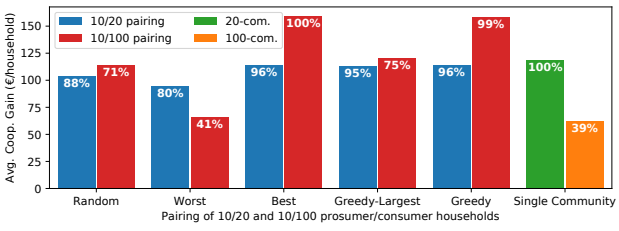

Figure 7: Average cooperative gain for different topologies.

5.4 How Much Prediction Power Is Needed for Achieving High Cooperative Gain?

All results presented so far have been focusing on what is achievable in the best possible scenario, i.e., with perfect foresight of the future. As appropriate with online, continuous decision making, we have also investigated if and how this changes with lower degree of forecasting power, by solving the iterative online optimization problem (c.f. Section 4.5) for each household equipped with Medium production level. Results are presented in Figure 8: under the three levels of predictor's accuracy, and with a variable number of forecasted hours (or prediction window size), the fraction of the optimal saving is displayed (average over all 100 households). Already about 57% of the gain is achieved using a very basic prediction (average over past 7 days using a window size of about 8 hours) or with greedy decisions, whereas more accurate predictions reach about 90% with the same window size and over 96% with 16 hours of predictions. The linear accuracy, whose prediction is almost as good as the truth for the first hours but then degrades to the same level as the average prediction, performs almost as good as the truth predictor. Hence, the experiment shows that we only need to accurately predict the first 4 (or resp. 8) hours to achieve about 90% (resp. 96%) of the optimal reachable saving of a medium-range production.

We have further explored if this also applies to cooperating communities. Figure 9 presents the average cooperative gain for 10 pairs (prosumer,pure-consumer) under the three prediction accuracies with 16 hours of forecast and the greedy coalition, and when each pair results from the greedy-largest matching of the 40-community. Truth and linear predictions achieve very close to perfect foresight results (on average 97.8% for truth, and 92.4% for linear) whereas the average predictor performs poorly with average 16% of the achievable gain with some pairs showing negative cooperative gain; the greedy coalition lies in between with 38% of the achievable gain on average. The cooperative gain has been computed here by comparing the bill of the pair to the sum of the individual perfect foresight bills, hence comparing cooperation with *realistic* predictions against the optimal individual solutions. Overall, the predicting power influences communities in similar manner as it would influence individual optimizing their own bill, and a high cooperative gain can be achieved in both cases (compared to the individual case with perfect foresight).

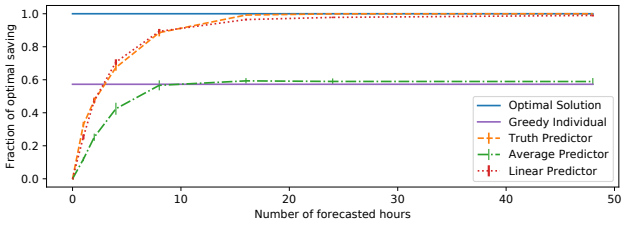

Figure 8: Average fraction of optimal saving achieved with different level of prediction versus number of forecasted hours (over 100 households).

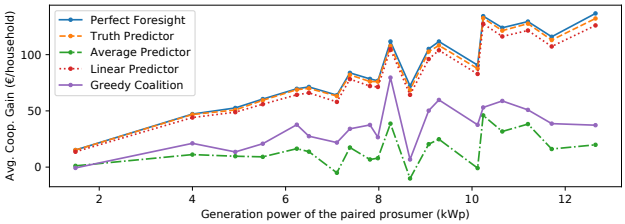

Figure 9: Average cooperative gain for the greedy-largest matching of 20 pure-consumers with 20 prosumers; the gain is calculated based on optimal (perfect forecast) individual bills.

5.5 How Does the Gain-Sharing Impact Motivations for Joining P2P Communities?

Figure 10 (up) presents the maximum average obtainable cooperative gain for equipped households (i.e., with the best equipped households distribution, around 40-60% depending on the configuration of the pool) whereas Figure 10 (down) presents the same but for unequipped households; both figures show the impact of the different splitting strategies. With "no split" strategy, there is a strong incentive to build the largest possible residential installation thanks to the additional financial benefit and there is no incentive to join the community for unequipped households. All other strategies offer a good trade-off between both incentives with the "even split" showing the greatest incentive for unequipped households while pushing prosumers to build larger installations than if P2P energy sharing is not present (the *working individually* curve) but not Very Large ones (ALR 6 and beyond). In practice, once the share among prosumers and pure-consumers has been decided, one may design strategies to distribute the gain, via e.g. allocating portions of the cooperative gain proportionally to each household's own share of the consumption of the group (or share of the sum of the individual bills) and allocating part of the gain to the utility company or service provider.

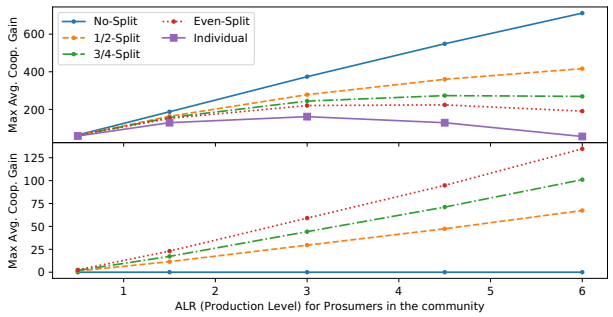

Figure 10: Gain from investing in renewable energy versus ALR (upper graph, in € per prosumer) and gain from joining a sharing community versus the ALR of prosumers in the joined community (lower graph, in € per pure-consumer), with different gain-splitting strategies.

5.6 Summary of the Results

(1) Small-scale communities can surprisingly obtain up to 88-97% of the same benefits (in terms of both cooperative gain or self-consumption) of any larger community with a similar fraction of prosumers versus consumers, leading to large reduction in the amount of data to share over the network.
(2) The cooperative gain in small communities is dependent on the choice of trading partners and there exist prosumers and consumers that are better or worse than other similar peers for saving financial cost through cooperation. In particular, matching prosumers with pure-consumers in the right way, such as greedy-matching them by their potential to save money, can lead to up to 59% improvement on the cooperative benefit.
(3) Using only 8h (resp. 16h) of predictions with moderate forecasting accuracy, we can obtain up to 90% (resp. 96%) of the optimal saving on electricity bill, both when considering individual prosumers and for cooperating communities, advocating that the problem can be solved in a continuous fashion.
(4) Splitting the cooperative gain among the peers is a trade-off. Distributing too large a fraction towards prosumers creates incentives to over-invest in the future and reduces the advantage of consumers to join the sharing process.

6 RELATED WORK

The concept of P2P energy sharing (alternatively named P2P energy trading) has emerged in the last couple of years [43] extending previous works on real-time trading of distributed energy resources [27]. It belongs to a wider research trend towards efficient management of distributed energy resources, smart homes and smart cities (see among others [14, 18, 20, 31]) such as virtual power plants [35], transactive energy [41], collective PV farms [26], as well as newly introduced notion of cloud energy sharing [29, 51] where a central component overviews the local P2P communities.

The initial moves towards energy sharing has focused on energy storage sharing [21, 42, 48] using game-theoretic approaches. Then integration of distributed electricity generation has shifted the focus towards how to implement the local energy market: trough e.g. bidding systems [49, 50], some sort of cost-sharing mechanisms [2] or financial incentives [23]. Recent developments in this area have seen secured blockchain-based trading [4, 33] and tests of P2P energy trading in microgrids [3, 38, 40]. Also motivations to join a P2P energy community has been explored in [11] where price is shown to be the main motivation and where the desire of greater energetic independence appears in second place.

Our work departs from those aforementioned works and similar previous prosumers community analysis [15, 16] by being more specific. We take into account here the online aspect of the decision problem as in [2, 10] but also the uncertainty due to forecasts (no prior knowledge), the impact of the P2P community size and the induced amount of data sharing. Indeed, no previous works to our knowledge have shown how beneficial small-scale P2P communities can be for energy sharing and on its direct impact on quantity of data to share and the proportion of the cooperative gain reached. Also, previous works often assume that only prosumers may take part in the P2P community (whereas we show here that cooperation is not very interesting in that situation) and provide non-sufficient evidence based on real-world data.

Our results advocate the benefits of having many small communities of a few households, optimizing their energy bill in an autonomous fashion. We have also shown that households do show preferences to be matched with some others and depending how the total money gained from cooperating is distributed among the peers, prosumers might also be inclined to have preferences in choosing their sharing partner. Also, as both preferences may change over time, communication is needed outside of the "autonomous" communities for the full network to converge on a P2P topology to use for decreasing the cost of sharing the data. This links the current work to distributed algorithms able to match nodes using preferences [8, 22]. The adaptation of such algorithms, and in particular greedy matchings [32], to the use-case presented here (i.e., with weights that require prediction and communication to estimate) while considering the stability of the network under arrival/departure of peers [5, 9], would provide mechanisms to build those sharing communities in a distributed and efficient manner.

7 CONCLUSION

We explored how useful small-scale P2P communities are for optimizing sharing energy resources among prosumers and consumers. We show, based on analysis of real-world data and considering communities of up to 100 households, the financial gains and increase in local self-consumption of electricity can be obtained with small-scale communities of 2 to 5 peers. This entails great benefits in terms of number of peers with which data needs to be shared, thus improving households' privacy while reducing the network's bandwidth consumption. We also show that the cooperative gain does not need high forecasting power (predictions of 8 to 16 hours in the future suffice for good results) and is high as long as a sufficient amount of peers are non-prosumers. By using distributed energy resources at a greater efficiency locally, our results can make renewable electricity generation more attractive, while providing financial motivation for non-prosumers to join P2P energy sharing communities and reduce their electricity bill.

This work can be extended in several directions. While focusing on if, when and at which scale, P2P energy sharing is beneficial for residential households, it leaves algorithms to form small-scale communities open for development. In particular, the relation among the quantity of data to share, its privacy aspects and the financial gain ought to be further studied. A key dimension in future research is how to transition from batch-based to online analysis, since the latter is required for the analysis itself to scale according to the size of modern cyber-physical systems [6, 7, 13, 44, 45]. We have shown that, even with limited forecasting power, households can practically achieve an almost optimal financial cost saving using greedy matching, however, it is important to notice that finding, in a distributed and continuous fashion, the ideal partner for each household is a challenging task and requires dedicated algorithms.

ACKNOWLEDGMENT

The work was supported by the Chalmers Area of Advance Energy project ADAPT and by the EU Horizon 2020 Framework Programme under grant agreement 773717.

REFERENCES

[1] Valentin Bertsch, Jutta Geldermann, and Tobias Lühn. 2017. What drives the profitability of household PV investments, self-consumption and self-sufficiency? *Applied Energy* 204 (2017), 1–15.

[2] Sid Chi-Kin Chau, Jiajia Xu, Wilson Bow, and Khaled Elbassioni. 2019. Peer-to-Peer Energy Sharing: Effective Cost-Sharing Mechanisms and Social Efficiency. In *Proceedings of the Tenth ACM International Conference on Future Energy Systems*. ACM, 215–225.

[3] Tianyi Chen and Shengrong Bu. 2019. Realistic Peer-to-Peer Energy Trading Model for Microgrids using Deep Reinforcement Learning. In *2019 IEEE PES Innovative Smart Grid Technologies Europe (ISGT-Europe)*. IEEE, 1–5.

[4] Can Dang, Jiangfeng Zhang, Chung-Ping Kwong, and Li Li. 2019. Demand Side Load Management for Big Industrial Energy Users under Blockchain-Based Peer-to-Peer Electricity Market. *IEEE Transactions on Smart Grid* (2019).

[5] Philippe Duchon and Romaric Duvignau. 2014. Local update algorithms for random graphs. In *Latin American Symposium on Theoretical Informatics*. Springer, 367–378.

[6] Romaric Duvignau, Vincenzo Gulisano, Marina Papatriantafilou, and Vladimir Savic. 2019. Streaming piecewise linear approximation for efficient data management in edge computing. In *Proceedings of the 34th ACM/SIGAPP Symposium on Applied Computing*. 593–596.

[7] Zhang Fu, Magnus Almgren, Olaf Landsiedel, and Marina Papatriantafilou. 2014. Online temporal-spatial analysis for detection of critical events in cyber-physical systems. In *2014 IEEE International Conference on Big Data (Big Data)*. IEEE, 129–134.

[8] Giorgos Georgiadis and Marina Papatriantafilou. 2012. Adaptive distributed b-matching in overlays with preferences. In *International Symposium on Experimental Algorithms*. Springer, 208–223.

[9] Giorgos Georgiadis and Marina Papatriantafilou. 2013. Overlays with preferences: Distributed, adaptive approximation algorithms for matching with preference lists. *Algorithms* 6, 4 (2013), 824–856.

[10] Giorgos Georgiadis and Marina Papatriantafilou. 2014. Dealing with storage without forecasts in smart grids: problem transformation and online scheduling algorithm. In *Proceedings of the 29th Annual ACM Symposium on Applied Computing*. ACM, 518–524.

[11] Ulf JJ Hahnel, Mario Herberz, Alejandro Pena-Bello, David Parra, and Tobias Brosch. 2019. Becoming prosumer: Revealing trading preferences and decision-making strategies in peer-to-peer energy communities. *Energy Policy* (2019), 111098.

[12] William E Hart, Carl D Laird, Jean-Paul Watson, David L Woodruff, Gabriel A Hackebeil, Bethany L Nicholson, and John D Siirola. 2017. *Pyomo-optimization modeling in python*. Vol. 67. Springer.

[13] Bastian Havers, Romaric Duvignau, Hannaneh Najdataei, Vincenzo Gulisano, Marina Papatriantafilou, and Ashok Chaitanya Koppisetty. 2020. DRIVEN: A framework for efficient Data Retrieval and clustering in Vehicular Networks. *Future Generation Computer Systems* 107 (2020), 1–17.

[14] Verena Heinisch, Lisa Göransson, Mikael Odenberger, and Filip Johansson. 2019. Interconnection of the electricity and heating sectors to support the energy transition in cities. *International Journal of Sustainable Energy Planning and Management* 24 (2019).

[15] Verena Heinisch, Mikael Odenberger, Lisa Göransson, and Filip Johnsson. 2019. Organizing prosumers into electricity trading communities: Costs to attain electricity transfer limitations and self-sufficiency goals. *International Journal of Energy Research* (2019).

[16] Verena Heinisch, Mikael Odenberger, Lisa Göransson, and Filip Johnsson. 2019. Prosumers in the Electricity System—Household vs. System Optimization of the Operation of Residential Photovoltaic Battery Systems. *Frontiers in Energy Research* 6 (2019), 145.

[17] Rodrigo Henriquez-Auba, Patricia Pauli, Dileep Kalathil, Duncan S Callaway, and Kameshwar Poolla. 2018. The Sharing Economy for Residential Solar Generation. In *2018 IEEE Conference on Decision and Control (CDC)*. IEEE, 7322–7329.

[18] Jean-Laurent Hippolyte, Shaun Howell, Baris Yuce, Monjur Mourshed, Hassan A Sleiman, Meritxell Vinyals, and Loïs Vanhée. 2016. Ontology-based demand-side flexibility management in smart grids using a multi-agent system. In *2016 IEEE International Smart Cities Conference (ISC2)*. IEEE, 1–7.

[19] Joern Hoppmann, Jonas Volland, Tobias S Schmidt, and Volker H Hoffmann. 2014. The economic viability of battery storage for residential solar photovoltaic systems–A review and a simulation model. *Renewable and Sustainable Energy Reviews* 39 (2014), 1101–1118.

[20] S Howell, Y Rezgui, JL Hippolyte, and M Mourshed. 2016. Semantic Interoperability for Holonic Energy Optimization of Connected Smart Homes and Distributed Energy Resources. In *eWork and eBusiness in Architecture, Engineering and Construction: ECPPM 2016: Proceedings of the 11th European Conference on Product and Process Modelling (ECPPM 2016), Limassol, Cyprus, 7-9 September 2016*. CRC Press, 259.

[21] Dileep Kalathil, Chenye Wu, Kameshwar Poolla, and Pravin Varaiya. 2017. The sharing economy for the electricity storage. *IEEE Transactions on Smart Grid* 10, 1 (2017), 556–567.

[22] Arif Khan, Alex Pothen, Md Mostofa Ali Patwary, Nadathur Rajagopalan Satish, Narayanan Sundaram, Fredrik Manne, Mahantesh Halappanavar, and Pradeep Dubey. 2016. Efficient approximation algorithms for weighted b-matching. *SIAM Journal on Scientific Computing* 38, 5 (2016), S593–S619.

[23] Ali Khodabakhsh, Jimmy Horn, Evdokia Nikolova, and Emmanouil Pountourakis. 2019. Prosumer Pricing, Incentives and Fairness. In *Proceedings of the Tenth ACM International Conference on Future Energy Systems*. ACM, 116–120.

[24] Mohsen Khorasany, Yateendra Mishra, and Gerard Ledwich. 2018. Market framework for local energy trading: a review of potential designs and market clearing approaches. *IET Generation, Transmission & Distribution* 12, 22 (2018), 5899–5908.

[25] Harold W Kuhn. 1955. The Hungarian method for the assignment problem. *Naval research logistics quarterly* 2, 1-2 (1955), 83–97.

[26] Stephen Lee, Prashant Shenoy, Krithi Ramamritham, and David Irwin. 2018. VSolar: Virtualizing community solar and storage for energy sharing. In *Proceedings of the Ninth International Conference on Future Energy Systems*. 178–182.

[27] Woongsup Lee, Lin Xiang, Robert Schober, and Vincent WS Wong. 2014. Direct electricity trading in smart grid: A coalitional game analysis. *IEEE Journal on Selected Areas in Communications* 32, 7 (2014), 1398–1411.

[28] Fabio Lilliu, Meritxell Vinyals, Roman Denysiuk, and Diego Reforgiato Recupero. 2019. A novel payment scheme for trading renewable energy in smart grid. In *Proc. of the 10-th ACM Int'l Conference on Future Energy Systems*. ACM, 111–115.

[29] Jingkun Liu, Ning Zhang, Chongqing Kang, Daniel S Kirschen, and Qing Xia. 2017. Decision-making models for the participants in cloud energy storage. *IEEE Transactions on Smart Grid* 9, 6 (2017), 5512–5521.

[30] Zvi Lotker, Boaz Patt-Shamir, and Seth Pettie. 2015. Improved distributed approximate matching. *Journal of the ACM (JACM)* 62, 5 (2015), 1–17.

[31] Xi Luo, Yanfeng Liu, Jiaping Liu, and Xiaojun Liu. 2020. Energy scheduling for a three-level integrated energy system based on energy hub models: A hierarchical Stackelberg game approach. *Sustainable Cities and Society* 52 (2020), 101814.

[32] Fredrik Manne, Md Naim, Håkon Lerring, and Mahantesh Halappanavar. 2016. On stable marriages and greedy matchings. In *2016 Proceedings of the Seventh SIAM Workshop on Combinatorial Scientific Computing*. SIAM, 92–101.

[33] Chathurika Prasadini Mediwaththe, Marnie Shaw, Saman Halgamuge, David Smith, and Paul Scott. 2019. An Incentive-compatible Energy Trading Framework for Neighborhood Area Networks with Shared Energy Storage. *IEEE Transactions on Sustainable Energy* (2019).

[34] Esther Mengelkamp, Johannes Gärttner, Kerstin Rock, Scott Kessler, Lawrence Orsini, and Christof Weinhardt. 2018. Designing microgrid energy markets: A case study: The Brooklyn Microgrid. *Applied Energy* 210 (2018), 870–880.

[35] Ashot Mnatsakanyan and Scott W Kennedy. 2014. A novel demand response model with an application for a virtual power plant. *IEEE Transactions on Smart Grid* 6, 1 (2014), 230–237.

[36] Zack Norwood, Emil Nyholm, Todd Otanicar, and Filip Johnsson. 2014. A geospatial comparison of distributed solar heat and power in Europe and the US. *PloS one* 9, 12 (2014), e112442.

[37] Emil Nyholm, Joel Goop, Mikael Odenberger, and Filip Johnsson. 2016. Solar photovoltaic-battery systems in Swedish households–Self-consumption and self-sufficiency. *Applied energy* 183 (2016), 148–159.

[38] Amrit Paudel, Kalpesh Chaudhari, Chao Long, and Hoay Beng Gooi. 2018. Peer-to-Peer Energy Trading in a Prosumer-Based Community Microgrid: A Game-Theoretic Model. *IEEE Trans. on Industrial Electronics* 66, 8 (2018), 6087–6097.

[39] Ruggero Schleicher-Tappeser. 2012. How renewables will change electricity markets in the next five years. *Energy policy* 48 (2012), 64–75.

[40] Ashish Shrestha, Rajiv Bishwokarma, Anish Chapagain, Sandesh Banjara, Shanta Aryal, Bijen Mali, Rajiv Thapa, Diwakar Bista, Barry P Hayes, Antonis Papadakis, et al. 2019. Peer-to-Peer Energy Trading in Micro/Mini-Grids for Local Energy Communities: A Review and Case Study of Nepal. *IEEE Access* 7 (2019), 131911–131928.

[41] Chen Sijie and Liu Chen-Ching. 2017. From demand response to transactive energy: state of the art. *Journal of Modern Power Systems and Clean Energy* 5, 1 (2017), 10–19.

[42] Wayes Tushar, Bo Chai, Chau Yuen, Shisheng Huang, David B Smith, H Vincent Poor, and Zaiyue Yang. 2016. Energy storage sharing in smart grid: A modified auction-based approach. *IEEE Transactions on Smart Grid* 7, 3 (2016), 1462–1475.

[43] Wayes Tushar, Chau Yuen, Hamed Mohsenian-Rad, Tapan Kumar Saha, H. Vincent Poor, and Kristin L. Wood. 2018. Transforming Energy Networks via Peer-to-Peer Energy Trading: The Potential of Game-Theoretic Approaches. *IEEE Signal Processing Magazine* 35 (2018), 90–111.

[44] Joris van Rooij, Vincenzo Gulisano, and Marina Papatriantafilou. 2018. Locovolt: Distributed detection of broken meters in smart grids through stream processing. In *Proceedings of the 12th ACM International Conference on Distributed and Event-based Systems*. 171–182.

[45] Joris van Rooij, Johan Swetzén, Vincenzo Gulisano, Magnus Almgren, and Marina Papatriantafilou. 2018. echidna: Continuous data validation in advanced metering infrastructures. In *2018 IEEE International Energy Conference (ENERGYCON)*. IEEE, 1–6.

[46] Mirjam Wattenhofer and Roger Wattenhofer. 2004. Distributed weighted matching. In *International Symposium on Distributed Computing*. Springer, 335–348.
[47] Joakim Widén, Ewa Wäckelgård, and Peter D Lund. 2009. Options for improving the load matching capability of distributed photovoltaics: Methodology and application to high-latitude data. *Solar Energy* 83, 11 (2009), 1953–1966.
[48] Chenye Wu, Dileep Kalathil, Kameshwar Poolla, and Pravin Varaiya. 2016. Sharing electricity storage. In *2016 IEEE 55th Conference on Decision and Control (CDC)*. IEEE, 813–820.
[49] Chenghua Zhang, Jianzhong Wu, Meng Cheng, Yue Zhou, and Chao Long. 2016. A bidding system for peer-to-peer energy trading in a grid-connected microgrid. *Energy Procedia* 103 (2016), 147–152.
[50] Zhenyuan Zhang, Haoyue Tang, Qi Huang, and Wei-Jen Lee. 2019. Two-Stages Bidding Strategies for Residential Microgrids Based Peer-to-Peer Energy Trading. In *2019 IEEE/IAS 55th Industrial and Commercial Power Systems Technical Conference (I&CPS)*. IEEE, 1–9.
[51] Yanglin Zhou, Song Ci, Ni Lin, Hongjia Li, and Yang Yang. 2018. Distributed Energy Management of P2P Energy Sharing in Energy Internet Based on Cloud Energy Storage. In *Proceedings of the Ninth International Conference on Future Energy Systems*. ACM, 173–177.

APPENDIX

A SETTING REASONABLE PRODUCTION LEVELS

Figure 11 presents the saving on electricity bill, reduced by corresponding investment, for the six selected production levels illustrated by a different symbol each, among production levels with ALR ranging from 0 to 12 and BDR ranging from 0 to 17.5. The saving is computed assuming perfect foresight of all necessary data, hence represents the best achievable outcome when each house works individually. One may note, the selected levels do not correspond to the best choices for a given ALR or for a given BDR, as this is not necessary how customers pick their production level, but rather span the range of possibilities with both increase in generation and storage capacities. They all produce a yearly positive return on investment falling in three categories: 56-59€ for Very Small and Very Large production levels, 129-130€ for Small/Large levels and 162€ for Medium level. From ALR 12 and BDR 15 or from ALR 15 (level not shown on the figure), households are on average net sellers; from ALR 12 on, the investments are too high and always result in an average negative yearly outcome.

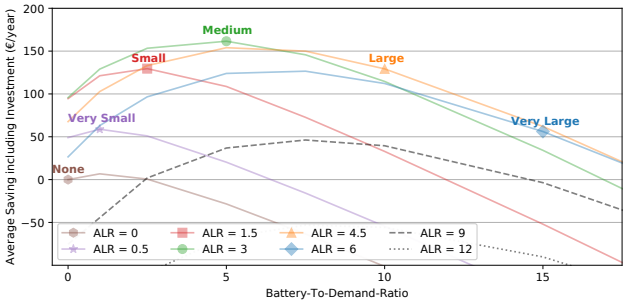

Figure 11: Average saving on yearly bill (including investments) per household for different production levels.

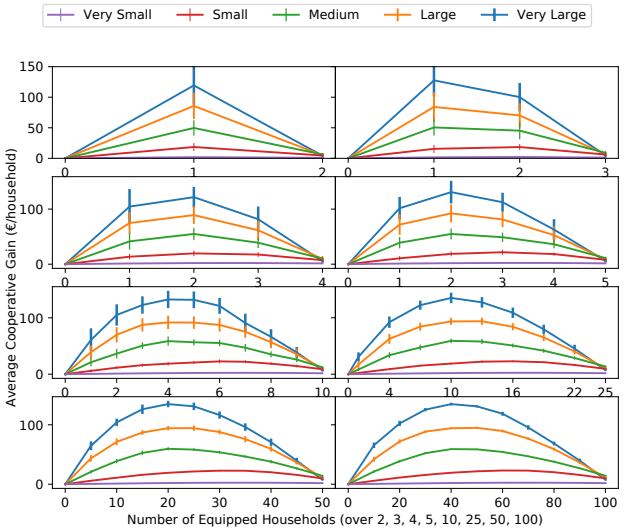

Figure 12: Average absolute cooperative gain (€/year).

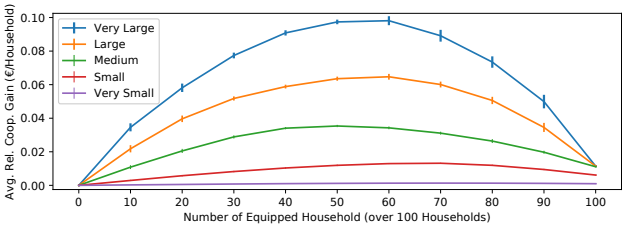

Figure 13: Relative avg. cooperative gain for 100 households.

B COOPERATIVE GAIN IN DIFFERENT CONFIGURATIONS

Capability for small-scale communities to achieve similar gains as larger ones for different configurations is shown in Figure 12. The figure presents average cooperative gain per household for different production levels, Very Small (thinnest error bar) to Very Large (largest ones), and for different community sizes of 2, 5, 10, 25, 50 and 100 households (same methodology as for Figure 4). For each pool size, we have experimented different proportions of prosumers (the remaining ones are pure-consumers). The lowest gain is always achieved by a full prosumers community for any community size and any production level, whereas the largest is around 40-60% of prosumers depending on the production level.

Figure 13 displays cooperative gain, in a similar fashion as Figure 12, for 100 households in relative terms (compared to average yearly bill). In relative value, the average cooperative gain per household can reach up to almost 10% of the average annual bill in the best configurations (with a share of 40-70% of prosumers, variable depending on the production level) but only up to 1% of annual bill for communities with only prosumers. Similar curve shapes are observed for other pool sizes.