# Large Scale Control of Deferrable Domestic Loads in Smart Grids

Ilenia Tinnirello, Giovanni Neglia, Laura Giarré, Giuseppe Di Bella, Alain Jean-Marie, Mariano Ippolito

# Large Scale Control of
# Deferrable Domestic Loads in Smart Grids

Ilenia Tinnirello*, Giovanni Neglia+, Laura Giarré*, Giuseppe Di Bella*, Alain Jean-Marie+, Mariano Ippolito*

*DEIM, Università di Palermo, Italy. Email: {firstname.lastname}@unipa.it.
+Inria, Maestro-team project, Sophia Antipolis, France. Email: {firstname.lastname}@inria.fr.

*Abstract*—In this paper, we investigate a realistic and low-cost deployment of large scale direct control of inelastic home appliances whose energy demand cannot be shaped, but simply deferred. The idea is to exploit 1) some simple actuators to be placed on the electric plugs for connecting or disconnecting appliances with heterogeneous control interfaces, including non-smart appliances, and 2) the Internet connections of customers for transporting the activation requests from the actuators to a centralized controller. Our solution requires no interaction with home users: in particular, it does not require them to express their energy demand in advance. A queuing theory model is derived to quantify how many users should adopt this solution in order to control a significant aggregated power load without significantly impairing their quality of service.

## I. INTRODUCTION

Load control in modern power grids is becoming increasingly essential for maintaining a balance between energy supply and demand. Traditionally, demand was much more variable and less controllable than supply, so energy balance was achieved by adapting dynamically generation levels to match consumption. Now, the increasing penetration of unpredictable renewable energy has radically changed the scenario. The possibility of controlling power demand is becoming more appealing both for the energy utilities (who can better plan production as well as control the grid reliability) and the end customers (who can actively participate in the energy market). Two envisaged main approaches are: Demand-Response and Direct Load Control (DLC). The former refers to the possibility of the end users changing their normal consumption patterns in response to a dynamic price signal, the latter to the possibility of the energy utility (or third-party entities) switching some specific users appliances on and off during peak demand periods.

In this paper, we deal with DLC mechanisms for deferrable domestic loads, i.e. loads of residential users whose power demand can be postponed but not interrupted. Examples of deferrable domestic loads are dishwashers and laundry machines. Indeed, most of the current DLC programs work on large industrial loads or domestic thermostatic loads. Large industrial loads have a significant and well characterized power demand and can be disconnected (according to prearranged agreements between the energy utility and the customer) with very simple decisions and control networks (often based on voice dispatch) [1]. Domestic thermostatic loads, such as air conditioners and heating systems, allow a fine-tuning regulation of power demand [2], [3]. However, a larger number of customers has to be involved, thus making the control

algorithms more complex [4], [5]. Only recently, deferrable loads have been considered for DLC, by assuming that each appliance is able to provide its energy profile, as well as an operation deadline, to the controller [6].

In this work, we deal with a large number of domestic deferrable loads whose location affects a substantial power demand in a given geographic area. This choice of load allows the energy utility to potentially reach all the residential customers with limited investments. Indeed, rather than considering appliances with smart models supporting advanced control interfaces [6], it is sufficient to deploy simple controlled switches between the electric plugs and the loads and to rely on a stochastic characterization of the appliances' profile. Differently from previous works, this study assumes that the appliances can be disconnected only once (at their activation request), while the control period is longer than the few minutes characterizing the on/off cycle of the thermostatic loads in [7], but shorter than the several hours characterizing the complete charging cycle of electric vehicles in [8].

The main goal of this work is to evaluate if DLC applied to deferrable domestic loads may allow the energy utility to control a significant percentage of power at the geographical scale and time scale of interest, while maintaining an acceptable quality of service for the customers. Since loads can only be deferred, the customers experience the effect of the direct control in terms of a possible *service delay*. Obviously, the more significant the change required to the power demand profile is (either in terms of instantaneous power reduction or duration of the control time interval), the longer the user delay will be. At the same time, the larger the number of appliances under control is, the smaller the probability that the plug of a given customer will be disconnected. Therefore, we aim to answer the fundamental question of *how many customers should be controlled in order to have a desired power reduction with an acceptable delay experienced by each customer*.

To this purpose, we apply queuing theory to study the relation between the load *availability* and the *temporal constraints* on the acceptable service delay. Although queuing theory has been already used for scheduling the energy suppliers working on renewable sources [9] or characterizing the aggregated power demand of a number of users [8], our achievements are more general since we study the service delays experienced when the number of electric loads is limited by admission control. This framework also offers some theoretical challenges, since it requires studying the transient period of a queue with
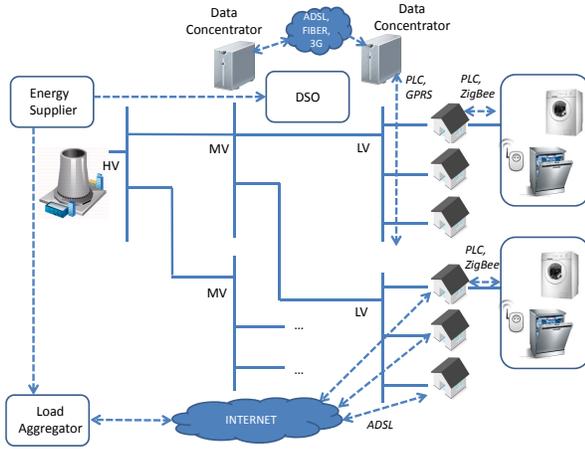
Fig. 1. Reference scenario: communication infrastructure and control system for DLC managed by a DSO and/or a load aggregator.

a dynamic number of servers.

The paper is organized as follows. In Section II we present our reference scenario and the architectural choices that allow DLC on deferrable domestic loads to be implemented with limited investments. In Section III we present the model used for characterizing the aggregated power demand of a large number of deferrable appliances, while in Section IV we model the effect of the admission control on the number of active appliances and quantify the service delay distribution. Section V presents some numerical results enlightening the tradeoffs between power reduction, service delay and number of controlled users. Finally, some concluding remarks are drawn in Section VI.

## II. SYSTEM DESCRIPTION

We consider the problem of power shaving for an energy supplier which wants to reduce its customers consumption during a given time interval. To achieve this goal, the energy supplier may interact with the distribution system operators (DSO), and/or with novel intermediate figures called load aggregators [10] by means of a communication infrastructure transporting the control messages. Both possibilities are illustrated in Fig. 1, where dashed lines indicate the flows of control messages.

The solution based on the DSO can take advantage of the advanced metering infrastructure (AMI), mainly based on PLCs (Power Line Communications) and GPRS (General Packet Radio Services) technologies, for reaching a large-scale number of domestic users in the low-voltage (LV) distribution grid [11], [12]. Conversely, the load aggregator can exploit standard Internet connectivity to interact with appliances at customers premises. The presence of the load aggregator allows to have an intermediary for interacting with large scale numbers of customers [13]. The aggregator can collect information on user profiles and willingness to respond to control actions for dispatching the controlled loads as a response of the energy supplier commands. Since the two approaches can coexist, we will talk generically about the load controller to denote the entity which drives the appliances.

The energy supplier specifies an high-level command for the load controller in terms of *maximum tolerable probability to*

*exceed a given power demand in the controlled area* [14], [15]. The load controller is responsible to meet this requirement by deciding about the activation of deferrable loads in the controlled area. To this purpose, each household can be equipped with a gateway which transmits and receivers the admission control messages and interacts with the domestic appliances by means of local area technologies, such as ZigBee and WiFi.

While smart appliances can be natively able to communicate with the gateway, the choice of working with deferrable loads allows to easily control also dumb appliances by means of programmable switches. Examples of these switches are the *smart plugs*, which can act as a *physical interface* between the appliances and the control application by exposing simple control primitives such as the activation of disactivation of the electric plug [7] or the setting of an operation timer [16].

### A. Use Cases

According to the employed admission control logic, to the actors (energy utilities, DSO, load aggregators) involved in the control and to the geographical scale of the controlled area, DLC can be used for different goals, such as improving the grid reliability or operation savings [17].

- *Medium Voltage/Low Voltage (MV/LV) line*: This control level can help to foster the integration of distributed generation from renewable sources in electrical MV networks. An adequate control of the demand, indeed, based on the variations of the power injected by not dispatchable supply, can attenuate some of the typical problems of voltage regulation in the MV lines. Furthermore, the possibility to operate actions of peak-shaving and/or load shifting in the hours of maximum load may allow to delay interventions.
- *High Voltage/Medium Voltage (HV/MV) station*: Similarly to the previous case, the control actions enforced by the energy utilities on the demand side can help to ease the regulation process of the voltage profile in the network, since the set-point in the automatic regulator of the station affects the voltage profile across the entire MV network. For advanced systems, with a management mode inspired by the smart-grid model, this second level of control may also participate in the control of the frequency, especially in emergency conditions in which the grid may operate in stand-alone configuration.
- *Power System Zone*: While the control actions at the previous levels are implemented by the electric utilities, the third level admission control can be performed at the community of customers level to respond to price signals, with the goal to make the power demand elastic in the electricity market. For this type of applications, the control of the energy consumption can be realized on a quite long time scale (from several hours to several days).

Regardless of the control level and involved actors, we assume that a centralized controller implements an admission control logic devised to enforce a probabilistic cap on the maximum power absorbed during a specific time interval, in the sense that the probability that the instantaneous power exceeds a desired limit is bounded (including a deterministic bound in case such a probability is set to zero). Such a

bound, as detailed in the Section IV under the assumption of homogeneous appliances, can be equivalently given in terms of maximum number of appliances that can be active at the same time.

## III. MODEL OF APPLIANCES

In this section, we characterize the power consumption originated by the aggregation of *one type only* of deferrable appliances, during activation time can be flexible for the users (e.g. dishwashers or laundry machines). Since each user consumption coincides with its appliance consumption, we indifferently refer to the total number of appliances or users.

Recent studies have characterized the percentage of users activating a specific residential appliance along different intervals of the day [18] and estimated their statistical properties. Moreover, it has been shown [19] that the aggregation of even a few thousands of customers makes residential demand uniform enough in a given period. As reported in Table I, in these studies the day is divided into equal duration intervals and the percentage of active users over a given population is averaged in each interval. Assuming that the user population $U$ is large and considering an observation time of one day, we can model the activation of a new appliance as a non-homogeneous Poisson process with arrival rate $\lambda(t)$. We also assume that the arrival rate $\lambda(t)$ is perfectly known.

The working interval of the reference appliance is usually deterministic or a deterministic function of the machine working program. Let $D$ be the time interval during which the appliance keeps working after its activation. We assume every appliance has the same deterministic operation time $D$ and its instantaneous power consumption is a random variable $X(t)$ with known time-invariant probability density function.

Let $u(t)$ be then the total number of users whose appliance is on at time $t$. Since, in absence of critical problems, the energy production is able to follow exactly the energy demand, we can model the $u(t)$ random process as the number of jobs in a $M(t)/D/\infty$ queue.[1] The instantaneous power consumption $P(t)$ at time $t$ can then be calculated as

$$P(t) = \sum_{i=1}^{u(t)} X_i(t). \tag{1}$$

We can easily characterize the probability distribution $p_i(t) = Pr\{u(t) = i\}$, that is the probability to find $i$ active users at time $t$. Since all the appliances activated before $t - D$ are deterministically switched off in $t$, we find $i$ users at time $t$ if exactly $i$ new appliances have been switched on during the last $D$ interval:

$$p_j(t) = \frac{[\bar{\lambda}_D^{t-D}D]^j}{j!} e^{-\bar{\lambda}_D^{t-D}D} \tag{2}$$

where $\bar{\lambda}_D^{t-D} = 1/D \int_{t-D}^{t} \lambda(x)dx$ is the average arrival rate in an interval of length $D$ starting at $t - D$.[2]

[1] Here "M(t)" (for "Markovian") denotes the Poisson time-varying appliance activation process, "D" indicate that the operation time is deterministic and $\infty$ the presence of an infinite number of servers, so that new requests can immediately be served. The reader can refer to [20] for a basic introduction to queuing theory.

[2] Note that for large values of $j$, the probabilities $p_j$ in Eq. 2 (as well as in Eq. 9) need to be evaluated approximating the Poisson distribution by the Gaussian one.
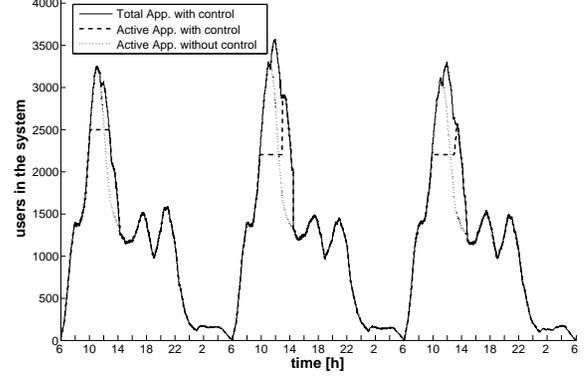


Fig. 2. Random realizations of the number of active appliances for a population of 10000 total users, with and without control.

We observe that there is no correlation between the number of active users in $t$ and $t + D$. Conversely, we can derive the probability distribution $p_j(t + \Delta t|u)$ of $j$ active users at $t + \Delta t$ for $\Delta t < D$ given that there are $u$ active users at $t$, by considering the joint probability to have $k$ departures and $j - u + k$ arrivals in $[t, t + \Delta t]$. The probability $a(t, \Delta t)$ that an appliance active in $t$ is switched off at time $t + \Delta t$ can be expressed as the probability that its arrival has occurred at time $t - D + \Delta t$:

$$a(t, \Delta t) = \frac{\int_{t-D}^{t-D+\Delta t} \lambda(x)dx}{\int_{t-D}^{t} \lambda(x)dx}. \tag{3}$$

Eq. 3 follows from the definition of conditional probability. Therefore, the probability $d_k(t, \Delta t|u)$ that $k$ appliances among the $u$ active at time $t$ are switched off during $[t, t + \Delta t]$ is expressed by the binomial:

$$d_k(t, \Delta t|u) = \binom{u}{k} a(t, \Delta t)^k (1 - a(t, \Delta t))^{u-k},$$

for $k = 0, 1, \ldots u$. Being $l = \max\{0, u - j\}$, the conditional probability to have $j$ active users can be expressed as:

$$p_j(t + \Delta t|u) = e^{-\bar{\lambda}_{\Delta t}^t \Delta t} \sum_{k=l}^{u} d_k(t, \Delta t|u) \frac{(\bar{\lambda}_{\Delta t}^t \Delta t)^{j-u+k}}{(j - u + k)!}. \tag{4}$$

Under the assumption that the load controller can monitor the actual number of active users at the current time instant $t$, Eq. 4 allows us to evaluate the distribution of the number of active users at $t + \Delta t$.

*Numerical Example.* For the laundry machine data in Table I, assuming $D$ equal to 90 min, the thin dotted curves in Fig. 2 shows a realization of the random process $u(t)$ with a total population of 10000 users over three days. In the figure we can clearly identify the peak hours and the effect of the time-varying activation rates. The maximum number of users into the system is about 3100, which corresponds to a load of 4.65 MW if we consider a power consumption equal to 1.5 KW for each laundry machine. At noon, being $\bar{\lambda}_{1.5}^{10:30} = 0.38 \cdot 10000/2h = 1900$ arrivals/h, the average number of users is $\bar{\lambda}_{1.5}^{10:30}D = 2850$.

## IV. LOAD MODEL UNDER CONTROL

Consider now the effect of load control on the number of active users, when a maximum number $c$ of appliances can

| Appliance/Hours | 6-8 | 8-10 | 10-12 | 12-14 | 14-16 | 16-18 | 18-20 | 20-22 | 22-24 | 24-6 |
|---|---|---|---|---|---|---|---|---|---|---|
| Dishwasher. | 3 | 9 | 9 | 3 | 13 | 0 | 16 | 38 | 13 | 3 |
| Laundry m. | 16 | 28 | 38 | 19 | 16 | 19 | 16 | 16 | 3 | 6 |

TABLE I

PERCENTAGE OF ACTIVE APPLIANCES OVER A POPULATION OF 30000 USERS ([18], [19])

be simultaneously active during a given time interval. The dashed and continuous curves in Fig.2 show the number of active appliances and the total number of appliances (active and disconnected appliances), respectively, when $c$ is set to 2500 (first day) or 2200 (second and third day). The figure also shows the effect of different resumption policies after load control: in the first and second day, all the disconnected appliances are immediately activated at the end of the control interval, thus leading to a potentially sudden increase of the power demand (second day); in the third day, the waiting users are admitted gradually by considering a maximum increment rate for the limit $c$. From the figure, it is evident that load control performs similarly to load-shifting: for example, in the second day the peak hour is shifted from noon to 2 p.m. Load shifting capacity is constrained by the number of users involved into the admission control and by the temporal delay that can be tolerated. To design the control mechanism, we need to quantify these two constraints.

*A. Power constraints*

In order to achieve a reduction of the power demand expected during a time interval $[T_s, T_e]$, we assume that the controller would like to enforce a maximum power consumption level $P_g$ for the set of appliances under control with probabilistic guarantees: the instantaneous power consumption ($P(t)$) can exceed $P_g$ with probability at most $\epsilon$, i.e., $\text{Prob}\{P(t) > P_g\} \leq \epsilon$. The problem that we consider is finding the maximum number $c$ of appliances active during the control period which guarantees that $\text{Prob}\{P(t) = \sum_{i=1}^{u(t)} X_i > P_g\} \leq \epsilon$, when $u(t) \leq c$. For simplicity, we develop the calculations for the case when $c$ is large enough for the aggregated power consumption to be approximated by a normal distribution. This case is also likely the most relevant from a practical point of view, given that we are interested in controlling a large number of appliances and we do not want to excessively limit the number of appliances active at a given time instant (otherwise users would experience too large delays). It is clear that $\text{Prob}\{P(t) = \sum_{i=1}^{u(t)} X_i > P_g\}$ is maximized for $u(t) = c$. In such a case, the average of the instantaneous power in Eq. 1, is given by $\text{E}[P(t)] = c\text{E}[X]$ and $\text{Var}(P(t)) = c\text{Var}(X)$. Under the normal approximation, $\text{Prob}\{P(t) > P_g\}$ is lower than $\epsilon$ if and only if $(P_g - c\text{E}[X])/\sqrt{c\text{Var}(X)} \geq z_{1-\epsilon}$, where $z_{1-\epsilon}$ is the $\epsilon$ percentile of the standard normal distribution. The power consumption profile is then satisfied if: $c\text{E}[X] + z_{1-\epsilon}\sqrt{c\text{Var}(X)} \leq P_g$. Then $c = \lfloor n^* \rfloor$, where $n^*$ is the smallest solution of the quadratic equation

$$\left(n^*\text{E}[X] - P_g\right)^2 = z_{1-\epsilon}^2 n^* \text{Var}(X). \qquad (5)$$

*B. Active and disconnected appliances*

We assume that the load controller can only work on the new activation requests, because disconnecting appliances whose working cycle is in progress can be inefficient and
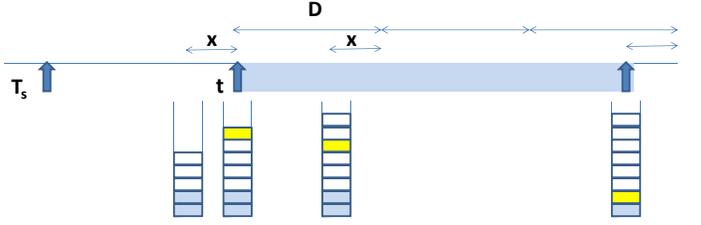


Fig. 3. A possible arrival and queue scenario (the yellow is the new arrival)

uncomfortable for the users. In order to guarantee that less than $c$ appliances are active at time $T_s$, the load controller anticipates the action at $T_{sc} = T_s - D$, because the appliances active at time $T_{sc}$ (whose number can be larger than $c$) will have necessarily terminated their operation at time $T_s$.

During the control period $[T_{sc}, T_e]$, each appliance activation is conditioned to an admission control: the appliance can be switched on only if the number of appliances already active is lower than $c$. When the new appliance cannot be activated, the controller disconnects the relative plug until some power resources become available. We assume that there is no limit to the maximum number of disconnected appliances and that they can be orderly reconnected into the system (when possible) according to the arrival time of their activation request. Under these assumptions, we can model the controlled load process as an $M(t)/D/c$ process.

Recall that $u(t)$ is the number of appliances at time $t$ that are working or are waiting for some available power in order to start. The probability distribution $p_j(T_{sc})$ of $u(T_{sc})$ can be evaluated on the basis of the previous analysis without admission control. Consider now a generic time duration $\Delta t < D$. The number of users $u(T_{sc}+\Delta t)$ in the system can be obtained from $u(T_{sc})$ considering the difference between the number of new activation requests and the number of appliances that have been switched off in the interval $[t, t + \Delta t]$. While the new requests depend on the average arrival rate $\bar{\lambda}_{\Delta t}^{T_{sc}}$, the number of appliances switched off depends on the number of arrivals in the previous interval $[T_{sc} - D, T_{sc} - D + \Delta t]$, being the total number of arrivals in the interval $[T_{sc} - D, T_{sc}]$ equal to $u(T_{sc})$. In other words, $p_j(T_{sc} + \Delta t | u(T_{sc}))$ can be expressed using Eq. 4. By weighting each conditional probability with the probability of the conditioning event to have $u(T_{sc})$ users in $T_{sc}$, we can find the probability to have $j$ users in the system at time $T_{sc} + \Delta t$, given by

$$p_j(T_{sc} + \Delta t) = \sum_{u=0}^{\infty} p_j(T_{sc} + \Delta t | u) p_u(T_{sc}). \qquad (6)$$

Consider now a generic instant $t > T_{sc} + D$. The number of appliances $u(t)$ is now depending on $u(t-D)$, because the appliances disconnected at $t - D$ are still into the system at time $t$. Indeed, $u(t)$ is given by the sum of $u(t-D) - c$ (if positive) and the new arrivals, i.e. we can find $j$ appliances in

$t$ if we have $j + \min\{0, c - u(t - D)\}$ new arrivals, with a probability given by:

$$p_j(t) = \sum_{i=0}^{j+c} p_i(t - D) \frac{[\bar{\lambda}_D^{t-D} D]^{j+\min\{0, c-i\}}}{(j + \min\{0, c-i\})!} e^{-\bar{\lambda}_D^{t-D} D}. \quad (7)$$

Since Eq. 6 yields $p_i(t)$ for $t \in [T_{sc}, T_{sc} + D]$, for $t \in [T_{sc} + D, T_e]$ it is possible to apply Eq. 7 in $k = \lfloor (t - T_{sc})/D \rfloor$ consecutive time intervals starting from the distribution $p_i(T_{sc} + (t - T_{sc})\%D)$. Therefore, the behavior of the user population can be characterized during the whole control period $[T_{sc}, T_e]$. The probability distributions $q_j(t)$ to have $j$ disconnected appliances is obviously $p_{j+c}(t)$ for $j > 0$ and $\sum_{i=0}^{c} p_i(t)$ for $j = 0$.

## C. Delay analysis

During the control period, some appliances cannot be activated exactly when the user makes the activation request. In this case, they experience a *service delay* until some power resources become available. Consider a generic appliance whose activation request is originated at time $t \in [T_{sc}, T_e]$. Let $E[W(t)]$ be its average service delay. Since the arrival rate of activation requests at time $t$ is given by $\lambda(t)$, the average delay experienced by a random user when the load control is applied can be expressed as

$$\bar{W}_{T_e - T_{sc}}^{T_{sc}} = \frac{\int_{T_{sc}}^{T_e} \lambda(t) E[W(t)] dt}{\int_{T_{sc}}^{T_e} \lambda(t) dt}. \quad (8)$$

The proof of Eq. 8 is in the appendix.

To derive $E[W(t)]$ we evaluate the cumulative distribution of the delay $W(t)$ experienced by a user arriving at time $t$ following the methods proposed in [21], [22]. We generalize the approach considering non-homogeneous arrival rates and an unknown initial state $u(T_{sc})$.

Let $t$ be the arrival instant of an appliance and $u(t) = kc+i$ with $i \in [1, c]$. We assume that the $u(t)$ appliances are ranked according to their arrival order. The appliance arrived at $t$ has to wait that $(k-1)c+i$ appliances complete their service before being reconnected (i.e. before being in the first $c$ positions). Since only $c$ appliances can complete their work in an interval equal to $D$, the new appliance is going to be reconnected in the interval $[t + (k-1)D, t + kD]$. Consider for example the scenario depicted in Fig. 3. The new arrival at time $t$ (the yellow one in the figure) is in position 7 in a system in which $c = 2$. Since it has to wait that 5 users complete their service, it will be activated after $t+2D$ and before $t+3D$. Specifically, its delay will be lower than $3D - x$ with $x \in [0, D]$, if the target appliance occupies one of the first $c$ positions of the queue by $t + 3D - x$. If we go backward in the past, this condition implies that the target appliance has to be in one of the first $2c$ positions by $t + 2D - x$, and in one of the first $3c$ positions by $t + D - x$. To satisfy this last condition, the number of users in the queue at $t - x$ plus the number of new arrivals before $t$ has to be strictly lower than $3c$ (note that the number of active appliances at $t - x$ are switched off by $t - x + D$ and therefore are not in the system anymore). We can generalize the previous considerations and evaluate the

cumulative probability $P\{W(t) \leq kD - x\}$ as the probability that the number of users $N_t(t+D-x)$ arrived before $t$ that are still into the system at $t+D-x$ is strictly lower than $kc$. For a given value of the queue $p$ in $t-x$, $N_t(t+D-x)$ is lower than $kc$ if the number of new arrivals is strictly lower than $kc - p$. Therefore $P\{W(t) \leq kD - x\}$ is 1 and for $kD - x \geq T_e - t$ is given by

$$\sum_{p=0}^{\infty} q_p(t - x) \sum_{j=0}^{kc-p-1} \frac{(\bar{\lambda}_x^{t-x} x)^j}{j!} e^{\bar{\lambda}_x^{t-x} x} \quad (9)$$

for $kD - x < T_e - t$. Obviously, for $t - x < T_{sc}$ (i.e. before the starting of the admission control), $q_0(t - x)$ is equal to 1 (there is no disconnected user in the system).

## V. NUMERICAL RESULTS

In order to answer our initial question about the scale of the energy demand controller, we quantify the service delay experienced by the users for different control parameters $(P_g, [T_s, T_e])$ and for different user populations.

We assume here that the controller manages a group of $U$ users served by the same primary substation. These users accept a delay in the appliances activation. In particular, we focus on the control of a given appliance type, which in our experiments is the laundry machine, whose uncontrolled activation rate generates the profile summarized in Table I. For simplicity we assume a constant power absorption per appliance equal to 1.5kW.

The energy supplier has historical knowledge about the power demand in similar days and can then characterize statistically the power demand in absence of any control. We assume that the energy supplier asks a power reduction $K$ to be applied starting from 10.00 a.m. (during the peak hours) to a reference level $P$ that can be the average power demand or the a given percentile of the power demand. The maximum aggregated power $P_g$ is then $P - K$ and the maximum number $c$ of appliances simultaneously active in $[T_s, T_e]$ can be determined.[3]

Consider first the case of the maximum size of user population, i.e. all the primary substation users (typically, tens of thousands) join the energy controller controller. Fig. 4 shows the average service delay experienced when the power reduction $K$ varies in the range 100 KW-5MW for a total population of 30000 users. To guarantee the desired power reduction $K$ at 10 a.m., the limitation on the number of active appliances is started at 8.30 a.m., i.e. at $T_s - D$. On the basis of historical consumption data, the reference level $P$ is evaluated as the average power demand at 10 a.m. in the left plot of the figure, and as the 90% percentile of the power demand in the right plot. The first case corresponds to a smaller value of $c$. From the figure, we can see that in the longest considered control period (namely, 180 min), asking for a power reduction of 0.5 MW from 10.00 a.m. to 13.00 a.m. leads to a service delay lower than 20 min. For higher power reduction values, the service delay can be still acceptable if the control time

---

[3]In this simple case where we consider appliances with constant power demand, we can calculate $c$ simply dividing $P_g$ by the power demand of a single appliance.
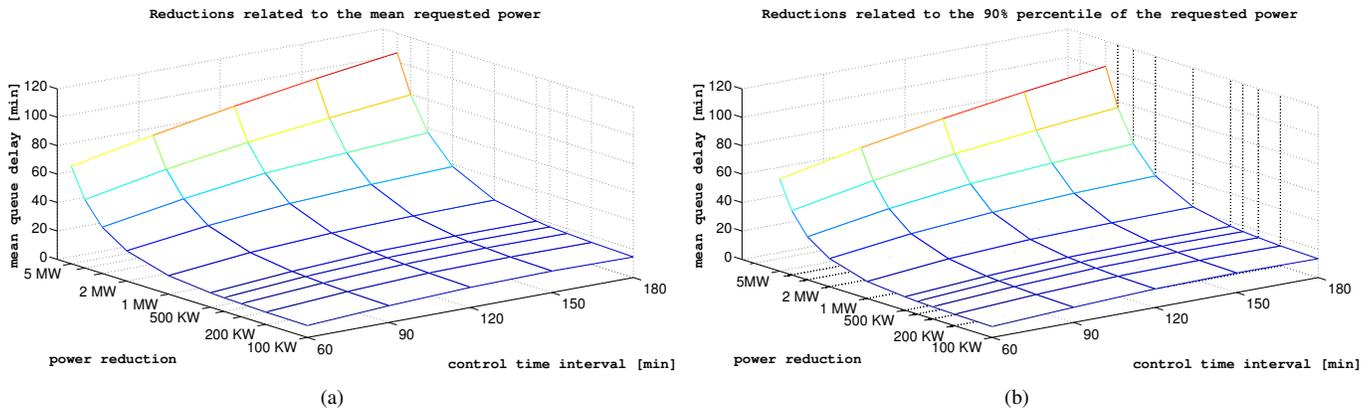
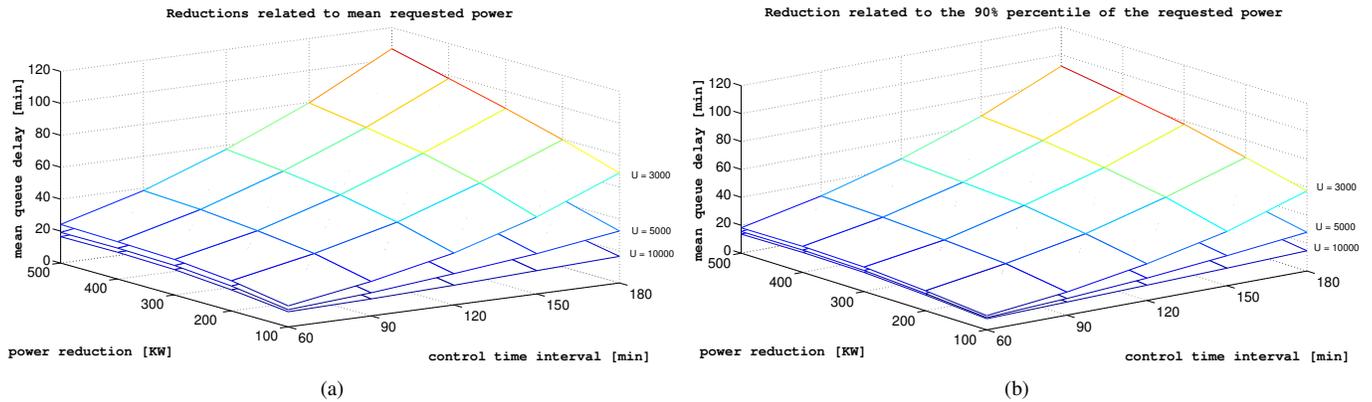Fig. 4. Average service delay for a population of 30000 laundry machines.



Fig. 5. Average service delay for different user populations: 3000, 5000 and 10000 laundry machines.
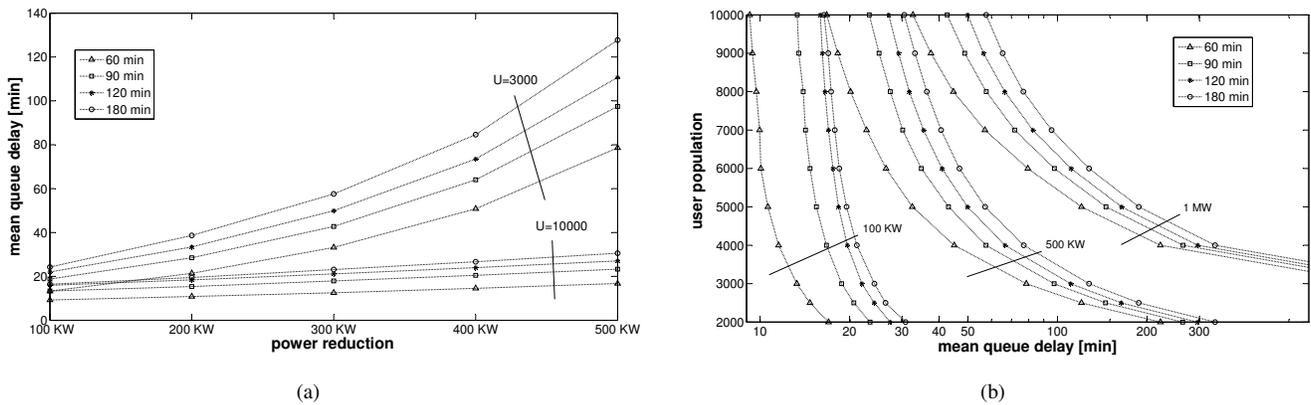


Fig. 6. Design curves with mean service delays vs. power reduction (a) and user population vs. mean service delays (b) under different control periods.

interval is limited. By comparing Fig. 4-a and 4-b, we can also observe that there is not a significant difference between applying the power reduction to the 90% percentile of the power demand or to the average value. This is due to the fact that we are considering a large population of users which corresponds to a small variability of the aggregated power demand.

Fig. 5 shows the curves (analogous to the ones plotted in Fig. 4) obtained for different user populations. As the number of users increases, for a given power reduction, the system is obviously able to provide a lower service delay. For example, for a power reduction of 100 KW and a control interval of 3

hours, the average service delay is lower than 20 min when $U = 10000$ and about 70 min when $U = 3000$. For small user populations, the service delay experienced when the power reduction refers to the 90% percentile of the power demand is significantly smaller than when it refers to the average value (e.g. about 10 min of reduction for $U = 3000$, $K = 100KW$ and $T_e - T_s$=3 hours).

The projection of the curves in Fig. 5-a is plotted in Fig. 6-a, for the cases $U = 10000$ and $U = 3000$, while a wide range of population sizes vs. the mean queue delay are considered in Fig. 6-b. Both the figures clearly show the tradeoff between user population, power reduction and

| $\Delta\lambda(t)$ | 1% | 2% | 5% | 10% | 15% | 20% |
|---|---|---|---|---|---|---|
| $\Delta E[W]$ | 1.3% | 2.6% | 6.1% | 11.2% | 16.2% | 21.3% |
| $\Delta D$ | 5 min | 10 min | 15 min | 20 min | 25 min | 30 min |
| $\Delta E[W]$ | 0.1% | 0.5% | 0.7% | 1.8% | 2.9% | 4.2% |

TABLE II

PERCENTAGE OF DELAY VARIATIONS IN CASE OF DATA UNCERTAINTY FOR $U = 3000$, $K = 500KW$, AND $T_e - T_s = 180$ MIN.

control period. For example, in Fig. 6-b, an average delay lower than 50 min can be guaranteed with $U = 3000$ for a power reduction of 500 KW for 60 min, or with $U = 8000$ for a power reduction of 1 MW for 180 min. By considering the maximum delay that users can tolerate and the desired power reduction, it is then possible to have an estimate of the percentage of residential users in a given area that should adopt our solution.

### A. System uncertainty

Our model requires some knowledge about the user population, in particular about the appliances' operation time, their power consumption and finally the activation rate.

Even if one of our model's assumptions is a homogeneous deterministic operation time $D$, we carried out some numerical Monte-Carlo simulations to evaluate how its variability affects the experienced mean queue delay. In particular we have considered the case when the operation time is uniformly distributed in the range $[D-\Delta D, D+\Delta D]$, where $D = 90$ min and $\Delta D$ varies from 5 to 30 min with steps of 5 minutes. As shown in Table II for the case $U = 3000$ and $K = 500KW$, the mean queue delay, expressed in terms of percentage variation compared to the case of constant operation time, is quite small: it ranges from less than 1% up to about 4% when $\Delta D = 30$ min. Fig. 7 shows a random realization of the service delays experienced in case of deterministic operation time and random operation time with $\Delta D$=30 min, which intuitively justifies the model robustness to the variability of the operation time[4]. This leads us to conclude that using a model with constant operation times provides provide good estimates of the results obtained also in case of heterogeneous operation times.

The uncertainty about the instantaneous power consumption $X$ is taken into account by the procedure described in Sec. IV and, in particular, by the term $\text{Var}(X)$ in Eq. (5). A too conservative choice for $\text{Var}(X)$ leads to a too small threshold $n^*$ and then to uselessly longer activation delays. On the contrary, underestimating $\text{Var}(X)$ leads to violate the probabilistic bound.

About the activation rate, energy utilities are currently able to forecast the aggregate consumption of a large population (millions of users) with very high accuracy (a few percents of error) one day-ahead. DLC solutions, like the one proposed in this paper, would not reach such deployment on a short time-horizon, but, at the same time, they need consumption forecasts on shorter timescales. An underestimate of the appliance activation rates, obviously, results in an increased service delay

---

[4]Note that, for an activation request arrived at $t$, the waiting time is the minimum between the queuing time of the finite servers model and $T_e - t$, because at $T_e$ all the appliances are reactivated; therefore, when equal to $T_e - t$ it does not depend on the operation times of the active appliances.
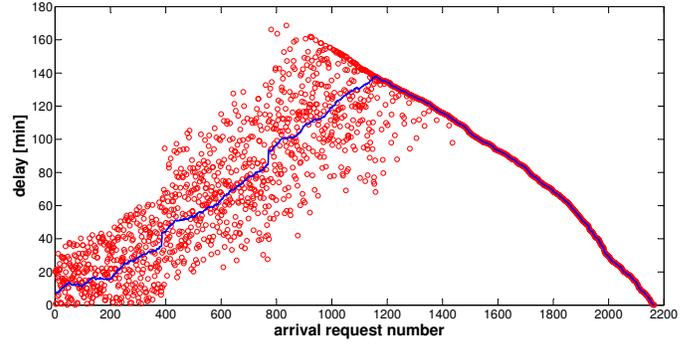


Fig. 7. Random realization of queuing delays of the activation requests arrived during the control interval for constant $D$ values (blue) and uniform $D$ values in $[60min, 120min]$ (red), $U = 3000$, $K = 400KW$.

that, according to the simulations summarized in Table II, is (in percentage) comparable to the error on the activation rate. We also considered additional experiments with a non-Poisson exogenous activation request process. The results suggest that our findings are quite robust to deviation from the Poisson assumption.

### VI. FINAL REMARKS

In this paper, we propose a large scale control mechanism for deferrable domestic loads, whose deployment requires minimal local communication overhead to allow a prompt user penetration, and to quantify the aggregation scale that should be reached in order to have appreciable power reduction with a minimal user discomfort.

Unlike other load control proposals published so far, which mainly work on thermostatic loads, the proposed scheme acts on deferrable loads (such as dishwashers or laundry machines), whose activation requests have to be accepted or postponed by a central server. A queuing theory model is derived for quantifying the user population to achieve a significant aggregated power load. Although the model has been derived for the case of homogeneous appliances, our numerical results are to be quite robust to predict the system performance in the case of heterogeneous operation times.

We are currently working on different model extensions for taking into account heterogeneous appliances, periodic sampling of instantaneous power consumption and alternative admission mechanisms, including distributed decisions. Moreover, in [23] we introduce the concept of energy-bucket to take into account power consumption correlation.

### REFERENCES

[1] D. Callaway and I. Hiskens, "Achieving controllability of electric loads," *Proc. of the IEEE*, vol. 99, no. 1, pp. 184–199, Jan 2011.

[2] L. Mathieu, S. Koch, and D. Callaway, "State estimation and control of electric loads to manage real-time energy imbalance," *Power Systems, IEEE Trans. on*, vol. 28, no. 1, p. 430 440, 2013.

[3] H. Hao, Y. Lin, A. Kowli, P. Barooah, and S. Meyn, "Ancillary service to the grid through control of fans in commercial building HVAC systems," *IEEE Trans. on Smart Grid*, vol. 5, no. 4, 2014.

[4] Y. Y. Hsu and C. C. Su, "Dispatch of direct load control using dynamic programming," *IEEE Trans. on Power Systems*, vol. 6, no. 3, p. 1056 1061, 1991.

[5] G. B. Sheble and K. H. Ng, "Direct load control-a profit-based load management using linear programming," *IEEE Trans. on Power Systems*, vol. 13, no. 2, pp. 668–694, 1998.

[6] G. O'Brien and R. Rajagopal, "Scheduling non-preemptive deferrable loads," *Power Systems, IEEE Transactions on*, vol. PP, no. 99, pp. 1–11, 2015.

[7] T. Keep, F. Sifuentes, D. Auslander, and D. Callaway, "Using load switches to control aggregated electricity demand for load following and regulation," in *Power and Energy Society General Meeting, 2011 IEEE*, July 2011, pp. 1–7.

[8] M. Alizadeh, A. Scaglione, J. Davies, and K. Kurani, "A scalable stochastic model for the electricity demand of electric and plug-in hybrid vehicles," *IEEE Trans. on Smart Grid*, vol. 5, no. 2, pp. 848–860, March 2014.

[9] M. Neely, A. Tehrani, and G. Dimakis, "Efficent Algorithms for Renewable Energy Allocation to Delay Tolerant Consumers," in *IEEE Int. Conf. on Smart Grid Communications*, 2010.

[10] N. Ruiz, I. Cobelo, and J. Oyarzabal, "A Direct Load Control Model for Virtual Power Plant Management," *IEEE Trans. on Power Systems*, vol. 24, no. 2, pp. 959–966, 2009.

[11] "An overview to new plc applications and new challenges - maxium integrated products." [Online]. Available: http://www.ieee-isplc.org/2011/Razazian_talk.pdf

[12] "Demand-response technology roadmap." [Online]. Available: http://innovation.edf.com/fichiers/fckeditor/Commun/Innovation/conference/Dherbecourt-Technology_Roadmap_DR_28th_May_2009_V1_1_va.pdf

[13] E. Ancillotti, R. Bruno, and M. Conti, "The role of communication systems in smart grids: Architectures, technical solutions and research challenges," *Computer Communications*, vol. 36, no. 1718, pp. 1665 – 1697, 2013.

[14] G. Di Bella, L. Giarré, M. Ippolito, A. Jean-Marie, G. Neglia, and I. Tinnirello, "Modeling Energy Demand Aggregator for Residential Users," in *IEEE CDC*, 2013, pp. 6280–6285.

[15] G. Neglia, G. Di Bella, L. Giarré, and I. Tinnirello, "Unidirectional Probabilistic Direct Control for Deferrable Loads," in *IEEE INFOCOM Workshop on Communications and Control for Smart Energy Systems*, 2014.

[16] G. Galioto, N. Galioto, C. Giaconia, L. Giarré, G. Neglia, and I. Tinnirello, "Smart plugs: A low cost solution for programmable control of domestic loads," in *AEIT - From Research to Industry: The Need for a More Effective Technology Transfer*, Sept 2014, pp. 1–6.

[17] J. Chen, F. Lee, A. Breipohl, and R. Adapa, "Scheduling direct load control to minimize system operation cost," *IEEE Trans. on Power Systems*, vol. 10, no. 4, pp. 1994–2001, Nov 1995.

[18] R. Miceli, "Sustainable development and energy saving laboratory," DIEET - University of Palermo, Tech. Rep., August 2007.

[19] S. Pagliuca, I. Lampropoulos, M. Bonicolini, B. Rawn, M. Gibescu, and W. L. Kling, "Capacity assessment of residential demand response mechanisms," in *Universities' Power Engineering Conference (UPEC), Proceedings of 2011 46th International*, 2011, pp. 1–6.

[20] L. Kleinrock, *Queueing Systems. Volume 1: Theory.* Wiley, 1975.

[21] G. Franx, "A simple solution for the M/D/c waiting time distribution," *Operation Research Letters*, vol. 29, no. 5, pp. 221–229, 2001.

[22] ——, "The transient m/d/c queueing system," 2002.

[23] G. Neglia, G. Di Bella, L. Giarré, and I. Tinnirello, "Scalable and privacy-preserving admission control for smart grids," in *IEEE 54th Annual Conference on Decision and Control*, December 2015.

[24] E. Çinlar, *Introduction to Stochastic Processes.* New Jersey: Prentice-Hall, 1975.

## APPENDIX

We consider a non-homogeneous Poisson arrival process $\mathcal{N}(t), t \geq 0$ with a continuous intensity rate $\lambda(t)$ and expectation function $\Lambda(t) = \int_0^t \lambda(x) dx$. $\{\tau_1, \tau_2, \ldots \tau_n, \ldots\}$ denotes the sequence of arrivals for $t \geq 0$. Let $W_{\tau_i}$ be a stochastic process defined on the points of $\mathcal{N}(t)$.

We prove that:

$$\mathrm{E}\left[\frac{\sum_{i|\tau_i < T} W_{\tau_i}}{\mathcal{N}(T)}\right] = \frac{\int_0^T \lambda(t)\mathrm{E}\left[W_t\right]dt}{\int_0^T \lambda(t)dt} = \frac{\int_0^T \lambda(t)\mathrm{E}\left[W_t\right]dt}{\Lambda(T)},\tag{10}$$

where $\mathrm{E}\left[W_t\right]$ indicates the expected value of the random variable $W_t$ and it is then conditioned on the fact that there is an arrival in $t$. In particular, we are going to prove that:

$$\mathrm{E}\left[\sum_{i=1}^n W_{\tau_i}\Big|\mathcal{N}(T) = n\right] =$$

$$n\frac{\int_0^T \lambda(t)\mathrm{E}\left[W_t|\mathcal{N}(T) = n, \tau_j = t \text{ for some } j = 1, \ldots n\right]dt}{\Lambda(T)},\tag{11}$$

from which our result follows immediately by deconditioning. Let $f_{\mathcal{N}}\left(\tau_1 = s_1, \tau_2 = s_2, \ldots \tau_n = s_n|\mathcal{N}(T) = n\right)$ be the joint probability density function of the ordered arrival times of $\mathcal{N}(t)$ in the interval $[0, T]$. Let us define $\hat{\tau}_i = \Lambda(\tau_i)$. The point process $\{\hat{\tau}_i\}$ is a homogeneous Poisson process ($\hat{\mathcal{N}}(t)$) with rate 1 [24], then

$$f_{\hat{\mathcal{N}}}(\hat{\tau}_1 = u_1, \hat{\tau}_2 = u_2, \ldots \hat{\tau}_n = u_n|\hat{\mathcal{N}}(t) = n) = \frac{n!}{t^n}.$$

Clearly, the event $\mathcal{N}(T) = n$ corresponds to the event $\hat{\mathcal{N}}(\Lambda(T)) = n$. For a given number of arrivals $n$, let $(\lambda(s_i))_{ii}$ be the (diagonal) Jacobian of the transformation $u_i = \Lambda(s_i)$. We can express $f_{\mathcal{N}}(s_1, s_2, \ldots s_n|\mathcal{N}(T) = n)$ as:

$$f_{\hat{\mathcal{N}}}\left(\Lambda(s_1), \Lambda(s_2), \ldots \Lambda(s_n)|\hat{\mathcal{N}}(\Lambda(T)) = n\right)|(\lambda(s_i)_{ii})|\tag{12}$$

and combining the previous equations we obtain

$$f_{\mathcal{N}}\left(\tau_1 = s_1, \tau_2 = s_2, \ldots \tau_n = s_n|\mathcal{N}(T) = n\right) = \frac{n!}{\Lambda(T)^n}\prod_{i=1}^n \lambda(s_i).$$

Similarly we can calculate the joint probability density function $\tilde{f}_{\mathcal{N}}(s_1, \ldots s_n|\mathcal{N}(T) = n)$ of the arrival times as:

$$\frac{1}{n!}f_{\mathcal{N}}\left(\tau_1 = s_{(1)}, \ldots \tau_n = s_{(n)}|\mathcal{N}(T) = n\right) = \prod_{i=1}^n \frac{\lambda(s_i)}{\Lambda(T)},\tag{13}$$

where $s_{(1)}, \ldots s_{(n)}$ are the order statistics of $s_1, \ldots s_n$. Let $\mathbf{s}$ and $\mathbf{s}_{-i}$ denote the vectors $(s_1, \ldots s_i, \ldots s_n)$ and $(s_1, \ldots s_{i-1}, s_{i+1}, \ldots, s_n)$ respectively. We observe that both density functions $f_{\mathcal{N}}$ and $f_{\hat{\mathcal{N}}}$ are invariant under permutations of $\mathbf{s}$. We define $g_i(s_1, \ldots s_n) = \mathrm{E}\left[W_{s_i}|\tau_1 = s_{(1)}, \ldots \tau_n = s_{(n)}\right]$ and $g(s_1, \ldots s_n) = \sum_{i=1}^n g_i(s_1, \ldots s_n)$. Then, $g_i()$ is invariant under permutations of $\mathbf{s}_{-i}$, while $g()$ is invariant under permutations of $\mathbf{s}$. We are now ready to prove Eq. 11.

$$\mathrm{E}\left[\sum_{i=1}^n W_{\tau_i}\Big|\mathcal{N}(T) = n\right]$$

$$= \int_{0 \leq s_1 \cdots \leq s_n \leq T} \sum_{i=1}^n \mathrm{E}\left[W_{\tau_i}|\tau_1 = s_1, \ldots \tau_n = s_n\right] \cdot$$
$$f_{\mathcal{N}}\left(\mathbf{s}|\mathcal{N}(T) = n\right)d\mathbf{s}$$

$$= \frac{1}{n!}\int_{[0,T]^n} g(\mathbf{s})f_{\mathcal{N}}\left(\mathbf{s}|\mathcal{N}(T) = n\right)d\mathbf{s}\tag{14}$$

$$= \sum_i \int_{[0,T]^n} g_i(\mathbf{s})\tilde{f}_{\mathcal{N}}\left(\mathbf{s}|\mathcal{N}(T) = n\right)d\mathbf{s}$$

$$= \sum_i \int_{[0,T]}\left(\int_{[0,T]^{n-1}} g_i(\mathbf{s})\tilde{f}_{\mathcal{N}}\left(\mathbf{s}_{-i}|\mathcal{N}(T) = n-1\right)d\mathbf{s}_{-i}\right)\frac{\lambda(s_i)}{\Lambda(T)}ds_i\tag{15}$$

$$= \sum_i \int_{[0,T]}\left(\int_{0 \leq s_1 \cdots \leq s_{i-1} \leq s_{i+1} \leq \ldots s_n \leq T} g_i(\mathbf{s}) \cdot f_{\mathcal{N}}\left(\mathbf{s}_{-i}|\mathcal{N}(T) = n-1\right)d\mathbf{s}_{-i}\right)\frac{\lambda(s_i)}{\Lambda(T)}ds_i\tag{16}$$

$$= \sum_i \int_{[0,T]}\left(\int_{0 \leq s_1 \cdots \leq s_{i-1} \leq s_{i+1} \leq \ldots s_n \leq T}\mathrm{E}\left[W_{s_i}|\tau_1 = s_{(1)}, \ldots \tau_n = s_{(n)}\right] \cdot f_{\mathcal{N}}\left(\mathbf{s}_{-i}|\mathcal{N}(T) = n-1\right)d\mathbf{s}_{-i}\right)\frac{\lambda(s_i)}{\Lambda(T)}ds_i$$

$$= \sum_i \int_{[0,T]}\mathrm{E}\left[W_{s_i}|\mathcal{N}(T) = n, \tau_j = s_i \text{ for some } j\right] \cdot \frac{\lambda(s_i)}{\Lambda(T)}ds_i$$

$$= n\int_{[0,T]}\mathrm{E}\left[W_s|\mathcal{N}(T) = n, \tau_j = s \text{ for some } j\right] \cdot \frac{\lambda(s)}{\Lambda(T)}ds.$$

Eq. 14 follows from the invariance of $g(\mathbf{s})f_{\mathcal{N}}(\mathbf{s}|\mathcal{N}(T) = n)$ to permutations of $\mathbf{s}$. The equality in Eq. 15 relies on the factorization of $\tilde{f}()$ in 13. The invariance of $g_i(\mathbf{s})f_{\mathcal{N}}(\mathbf{s}_{-i}|\mathcal{N}(T) = n)$ to permutations of $\mathbf{s}_{-i}$ is used in Eq. 16.

**Ilenia Tinnirello** received her Ph.D. degree in telecommunications engineering from the University of Palermo in 2004. She is currently an associate professor at the University of Palermo. She has also been Visiting Researcher at the Seoul National University, Korea, in 2004, and at the Nanyang Technological University of Singapore in 2006. Her research activities have been focused on wireless networks, and in particular on the design and prototyping of protocols and architectures for emerging reconfigurable wireless networks. Recently, she is also working on the definition of novel services (smart grid, smart metering, indoor localization, etc.) enabled by the pervasive availability of ICT technologies. I. Tinnirello has been involved in several European research projects, among which the FP7 FLAVIA project, with the role of technical coordinator, and the H2020 WiSHFUL and Flex5Gware projects.

**Giuseppe Di Bella** is a researcher at Telecom Italia S.p.A. Currently, he works in the SWARM Joint Open Labs in the Innovation & Strategy Department, mainly engaged in research on distributed adaptive systems. He received a Ph.D. in telecommunication engineering from the University of Palermo in 2015 with a thesis on privacy-preserving approaches in smart grid scenarios. He graduated from the University of Palermo with B.S. and M.S. degrees in telecommunication engineering in 2009 and 2011, respectively.

**Giovanni Neglia** received the Master?s degree in electronic engineering and Ph.D. degree in telecommunications from the University of Palermo, Italy, in 2001 and 2005, respectively. He has been a researcher with the Maestro team, Inria, Sophia Antipolis, France, since September 2008. In 2005, he was a Research Scholar with the University of Massachusetts, Amherst, visiting the Computer Networks Research Group. Before joining Inria, he was a postdoctorate with the University of Palermo and an external scientific advisor with the Maestro team, Inria. His research is focused on modeling and performance evaluation of computer networks and more recently of smart grids.

**Alain Jean-Marie** obtained his PhD degree in Computer Science from the University of Paris XI at Orsay, France, in 1987. Since then, he has been researcher at Inria (French National Institute of Computer Science and Control) in Sophia-Antipolis and Montpellier, France. He has been Professor of Informatics at the University of Montpellier 2, from 1999 to 2003. His research interests range from applied probabilities and stochastic modeling to optimal control and game theory, with applications to scheduling, networking, data management, biological systems and natural resource management, statistical physics.

**Laura Giarré** received the Laurea degree in Electronic Engineering from the Universitá di Firenze, Italy and the Ph.D. degree in System Engineering from the Universitá di Bologna, Italy, in 1986 and 1992, respectively. . She has held visiting positions at the Department of Mechanical Engineering of the University of California at Santa Barbara (CA) and at the Laboratory for Information and Decision Systems and Computer Science Artificial Intelligence Laboratory at MIT, Cambridge, Boston (MA) (1999, 2006, 2015). Since 1993, she was Assistant Professor at Politecnico di Torino, Turin, Italy and since 1998 she has been Associate Professor at the Universitá di Palermo, Palermo, Italy. In 2013 she received the national qualification (ASN) for full professor. She served as Associate Editor of Systems and Control Letters, 1999 -2011 and IEEE Transaction on Automatic Control 2012-2015. In 2010 she was vice-chair of SING6 and in 2014 General Char of IEEE MED14. Her research interests include networked systems, according to a game theoretical approach with applications to WIFI networks, sensor networks, inventory networks, smart grid. She worked also on identification of LPV, nonlinear and networked systems and e-health solutions for interoperability. She has recently developed an interest in assistive technology for visually impaired persons, organizing and chairing three workshops (2012 Pisa, 2013 Boston, 2015 Milan EXPO) on this theme and she co-founded the startup In.sight to develop systems enhancing senses trough technology for which she is the CEO.

**Mariano Giuseppe Ippolito** was born in 1965. He received the Degree and the Ph.D. degree from the University of Palermo, Palermo, Italy, in 1990 and 1994, respectively, both in electrical engineering. He is currently an Associate Professor of power systems with the University of Palermo, Palermo, Italy, and a Tutor Professor of numerous Ph.D. students. Since 2001, he has been locally responsible for several research projects in the field of power distribution development, distributed generation, electrical storage systems, and smart grids. He is currently the Coordinator of the M.Sc. Program in Electrical Engineering and the Director of the Master's Degree Course (post grade) in Home and Building Automation. In his areas of interest, he has author or co-author of more than 170 technical papers published in international journals or presented at various national and international conferences. His main research interests include power systems analysis, optimal planning, and design and control of electrical distribution systems. In recent years, his research activities have included the integration of distributed generation and storage systems into electrical networks and smart grids.