Aggregation and Control of a Heterogeneous Population of Solar Panels over the Grid Frequency

Andrea Peruffo, *Student, IEEE,* Emeline Guiu, Patrick Panciatici, *Fellow, IEEE,* Alessandro Abate, *Senior Member, IEEE*

***Abstract*— The increasing presence of solar energy sources has radically transformed the continental electrical network. Whilst solar panels affect the grid in a distributed manner and an individual device has negligible weight, the presence of a large population engaged in energy production can affect the overall network dynamics. Whilst this effect might in some scenarios be disruptive, solar panels can offer potential for frequency regulation, if their power output is to be controlled. In this work we present models for the aggregation of a heterogeneous population of solar panels connected to the electrical grid. The electrical network frequency is modelled centrally: the model accounts for primary frequency control and its inertia is a function of the penetration of PV generation. We study how features of the PV population affect the overall frequency signal. In particular, we investigate scenarios of power production (infeed) incidents: we simulate sudden generation losses, and study the corresponding frequency response at varying levels of renewable penetration. Simulations show good performance of the aggregated models when compared to the physical system, and quantitatively indicate that population heterogeneity is a desirable feature to avoid load shedding. We furthermore put forward a control architecture to regulate the power generation of the PV population: this intentionally simple architecture is aligned with primary control and is tailored to a decentralised implementation. We show that the control scheme can avoid load shedding, despite increasing penetration of solar generation, thus contributing to network resilience.**

***Index Terms*— photovoltaic panels, grid frequency dynamics, heterogeneity, population models, Markov models, decentralised control**

I. Introduction

ACADEMIA, industry and markets alike have recently shown an increased interest on renewable energy sources, which are seen as a core component towards environmental conservation and mitigation of rapid climate changes. Bolstered by initiatives such as the 2017 Paris Agreement [1] on climate, penetration of renewables, such as hydroelectric, wind, and solar energy, has considerably grown [2]–[4] in every part of the world.

A.Peruffo and A.Abate are with the Department of Computer Science, University of Oxford, OX1 3QD Oxford, UK, e-mail: andrea.peruffo@cs.ox.ac.uk; aabate@cs.ox.ac.uk.

E. Guiu and P. Panciatici are with RTE France, Paris, France, e-mail: emeline.guiu@rte-france.com; patrick.panciatici@rte-france.com.

Conventional power sources (e.g. coal, nuclear) entail sizeable production sites endowed with mechanical inertia due to the presence of rotating bodies [5]. Their models are well known, uncertainty on model parameters is limited and their estimation is by and large reliable. Our context is, instead, rather novel: we study a heterogeneous, large, distributed population of power generating devices with no inertia and high levels of uncertainty and of volatility. Nowadays, these populations significantly contribute to the electric grid and their use raises issues of network reliability.

In the field of renewables, much interest is placed on photovoltaics (we will interchangeably denote photovoltaic panels as PV- or solar panels) [6], which is now the third most important renewable energy source, after hydro and wind power, in terms of global installed capacity [4]. In view of the coupling between a solar panel and its inverter (namely the device that converts the continuous output of a solar panel to alternate current that is fed to the electricity grid), whenever we refer to solar panels, we implicitly consider the combination panel-inverter. A significant part of literature on PV focusses on the electrical features of panels in terms of photovoltaic cells [7], on the maximisation of power output (e.g. [8]), and on economic aspects of power production [9].

Despite this interest on PVs, not much works addresses the connection between populations of solar panels (as aggregation of individual devices) and the dynamics of the electric grid. Large PV farms have been modelled and studied aggregatively in [10], however there is no similar study of distributed solar power generation. At a household or building level, the assumption is that the produced power is limited and often consumed at the source, and hence that the net contribution to the grid is null. This assumption is evidently not tenable in view of the growing importance of populations of PV: as an example, in Germany PV generated power – distributed over 1.6 million PV setups – has provided for approximately 7.2% of the electricity demand in 2017, with peaks of 60% during weekends and holidays [3]. The growing relevance of PV panels justifies the modelling, analysis and control of this energy source. Another element that distinguishes it from conventional sources is that power production does not follow usual demand patterns: the power production of a panel, in a clear day, follows the irradiance of the sun, with a maximum at around midday; this produced power is usually injected

into the grid, due to limited consumption at those times. This unbalanced flow, compounded with the dependence of production on weather, if not handled correctly might lead to network issues (in the worst scenario, to blackouts) [11].

The models in this work (cf. Section III) are based on the following functional description of a solar panel (Section II). An panel-inverter device is equipped with a sensor, which is used to sample the network frequency, and with an internal counter. Two corresponding quantities are the key to model a panel-inverter behaviour: $\mathcal{I}_f$, the working interval of network frequencies, and τ_r, the time delay required for a safe connection to the network (the network frequency needs to remains inside $\mathcal{I}_f$ before the panel connects back to the grid). Each device, in principle, can have a different admissible frequency range and time delay. Based on this setup, this work amalgamates and extends previous contributions. [12] presents a framework for modelling and abstraction of a large population of photovoltaic panels. Further, [13] addresses the issue of evaluating the relationship between the panels' working interval $\mathcal{I}_f$ and the stability of the electrical network. Finally [14] adds a simple control scheme, originally developed in [15], [16] for refrigerating/cooling devices to continuous-time models of solar panels and thermostatically controlled loads (TCLs), and exploits their potential to support the network resilience after a generation loss incident.

In this work, we discuss a new and general analysis of the dynamics of a population of solar panels that is connected to the grid, particularly as a function of their heterogeneity [13]. As mentioned above, the disconnection-reconnection mechanism of a panel depends on the periodic sampling of the network frequency signal, which leads to a new discrete-time modelling framework. Models for the population are connected to the dynamics of the network frequency, resulting in an extended coupled model for the whole system [12]. The study shows that the new models are a closed representation of the system of interest. We further expand the study of the decentralised proportional control scheme from [14] to counterbalance a generation loss incident, within a network with reduced inertia due to the presence of a large population of renewables. We in particular investigate *load shedding* (also known as load reduction): this is a process that is activated to mitigate the lack of balance on the network, and to prevent potential subsequent blackouts. This procedure is essentially a forced stopping of electricity delivery, in order to avoid a complete shut-down of the electric grid. We show that this load shedding procedure can be avoided if the individual solar devices are equipped with a proportional control.

The remainder of the paper is organised as follows. In Section II we describe the behaviour of a single PV panel in frequency. Population models and the frequency dynamics are explained in Section III. Section IV introduces a grid model based on ENTSO-E [17], while the control architecture is designed in Section V. Experiments and case studies are in Section VI. Conclusions are drawn in Section VII.

II. Operation of a Solar Panel

We illustrate the functioning of a single solar-inverter device: we focus on how an inverter-panel system reacts as a function of the local electric grid. In this work we consider household devices, namely PV panels that are installed on the roof of private houses, as opposed to larger, less diverse industrial setups. Different manufacturers, weather conditions, ages, regulations, render this population naturally prone to a high level of heterogeneity.

TABLE I

Behaviour of single photovoltaic panel at time k. Key quantities are: panel state $q \in \{ON, OFF\}$; network frequency f; operating frequency band $\mathcal{I}_f$; clock/counter τ and re-connection delay τ_r.

Current state $q(k)$	Frequency	Delay	Next state $q(k+1)$
OFF	$f(k) \in \mathcal{I}_f$	$\tau(k) \geq \tau_r$	ON
ON	$f(k) \in \mathcal{I}_f$	–	ON
ON	$f(k) \notin \mathcal{I}_f$	–	OFF
OFF	$f(k) \in \mathcal{I}_f$	$\tau(k) < \tau_r$	OFF

A PV panel can be either active or disconnected: accordingly, we define two working states, ON and OFF. At time k, the switch amongst these two states depends on two values: the local network frequency $f(k)$ (whose nominal value is $f_0 = 50$ Hz in Europe) and an internal time delay τ_r. Whilst in practice τ_r is usually given in seconds, in this work we introduce a discretisation step h, and thus consider τ_r to be evaluated as a number of steps. Table I represents the switching behaviour of a PV panel, considering the contributions of the frequency signal and of the internal time delay. The device is connected to the electric grid and samples its frequency discretely in time. Safety regulations demand the panel to be active, i.e. in the ON mode, only when the network frequency lies within a given local frequency interval $\mathcal{I}_f$. If $f(k) \notin \mathcal{I}_f$ it must disconnect from the grid, i.e. switch to the OFF mode. We assume that the ON-to-OFF transition happens within a negligible time interval [18], [19], so at time $(k+1)$ the device is OFF. On the other hand, the OFF-to-ON switch cannot happen before an internal delay τ_r has passed, during which the frequency f must dwell within $\mathcal{I}_f$: this rule ensures the reconnection to a stable electric network, and avoids possible reconnection/disconnection scenarios (chattering) that may contribute to overall network instability. To guarantee that the frequency dynamics is stable, the inverter samples the grid frequency within the τ_r-long time interval: if the frequency is measured outside $\mathcal{I}_f$, the counter τ is reset. In practice, a PV panel senses the network frequency via a digital sensor with a defined sampling rate in the order of 200 ms, and τ_r is around 20 s. Notice that the network frequency sampled by different devices will differ, depending on several factors, such as natural frequency fluctuations, measurement noise, and dissimilar component quality. This motivates the use of a probabilistic model to describe this system. An alternative approach would require adding a measurement noise to individual devices: this, however, would incur in computational complexity associated to the modelling single devices. Instead, the proposed probabilistic model describes the aggregation of a population of heterogeneous devices in a simple and computationally effective manner.

The power output of an aggregation of panels adds uncertainty to the intrinsic diversity of the devices. In previ-

ous work [12], [13] we have focussed on a constant power production. This is a reasonable assumption over small time scales (seconds) and during a clear sunny day. However, this assumption might not hold during a cloudy day: as observed in [20], individual solar plants can significantly vary their outputs over the course of seconds. Weather should be encompassed in the model to have a better description a real system: we plan to add weather behaviour at a later stage, considering a longer timescale (hours or days). Additionally, in Section V we devise a decentralised control scheme over the power output of the panels: we show that solar panels can be equipped with an algorithm that tunes the output according to frequency deviations. In this work (Section VI on Case Studies), in order to model the unpredictability of the many factors (weather conditions, light occlusions) that influence the electric power production, we consider an additive Gaussian noise.

III. A Markov Model of a Population of PV Panels

In Section II we consider the behaviour of a single panel, which depends on its sampling of the frequency signal. This feature suggests using a discrete-time model. We employ a (discrete-time) Markov chain model to analyse the behaviour of a population of devices. We start the presentation by illustrating a model assuming constant power production as well as population homogeneity (i.e. all devices are characterised by the same parameters). We also assume that panels have no delays in (dis-)connection. We later introduce heterogeneity in the disconnection and reconnection thresholds of the PV devices; finally, we equip the models with a delayed-time structure. The preceding architectures (for delay-free homogeneous and heterogeneous populations) are introduced as key steps for the derivation of the final heterogeneous, delayed model, which is on of the main contributions of this work.

A. A Homogeneous Markov Model Without Delays

We define the quantity P_{PV}, expressing the weighted power production, as

$$P_{PV} = \frac{1}{N}\sum_{i=1}^{N} P_{PV,i},$$

where N is the number of panels in the population, and $P_{PV,i}$ is the i-th panel power output. Consider the normalised power production $R(k)$, at time k, as

$$R(k) = \frac{1}{NP_{PV}}\sum_{i=1}^{N} q_i(k)P_{PV,i},$$

where $q_i(k) \in \{0,1\}$ denotes whether the i-th device at time k is in the OFF or ON state, respectively. In view of the population homogeneity, all the devices behave in a synchronous way and $R(k) \in \{0,1\}$, $\forall k$. We consider $R(k)$ as a Bernoulli random variable, and introduce a new variable

$$x(k) = \mathbb{E}[R(k)] = \mathbb{P}[R(k) = 1],$$

defined as the expected value of $R(k)$ at time k, and which additionally denotes the probability of being in the ON state at that time. By the law of total probability,

$$\begin{aligned} x(k+1) &= \mathbb{P}[R(k+1) = 1] = \\ &= \mathbb{P}[R(k+1) = 1|R(k) = 1] \cdot \mathbb{P}[R(k) = 1] + \\ &+ \mathbb{P}[R(k+1) = 1|R(k) = 0] \cdot \mathbb{P}[R(k) = 0]. \end{aligned}$$

Let us introduce $a(k) = \mathbb{P}[R(k+1) = 0|R(k) = 1]$, so that $\mathbb{P}[R(k+1) = 1|R(k) = 1] = (1 - a(k))$, and similarly $b(k) = \mathbb{P}[R(k+1) = 1|R(k) = 0]$. The previous equation can be rewritten as

$$\begin{aligned} x(k+1) &= \mathbb{P}[R(k+1) = 1] = \\ &= (1 - a(k))x(k) + b(k)(1 - x(k)). \end{aligned} \tag{1}$$

This relation describes how the probability of being ON gets updated at time k. In the adopted framework, the transition probability ON-to-OFF (quantity $a(k)$) and OFF-to-ON ($b(k)$) are governed by the value of the network frequency $f(k)$, namely whether or not $f(k) \in \mathcal{I}_f$. Since at the moment we assume population homogeneity, these values are binary: for instance, when $f(k) \in \mathcal{I}_f$, $a(k) = 0$ and $b(k) = 1$.

Whilst quantities $a(k)$ and $b(k)$ are deterministic in a homogeneous population, this assumption is restrictive in the case of a heterogeneous population. In the latter case, we assume different panels to be subjected to different disconnection and reconnection thresholds, in view of different ages, manufacturing processes, weather conditions, etc. This results in two modelling choices. We either make the educated assumption that the different thresholds in a population are distributed according to a Gaussian. Alternatively, we consider a χ^2 distribution, which has different features: it is asymmetric and has a semi-infinite support – in contrast with the support of the Gaussian distribution. Evidently, other choices are possible: in fact, our study can encapsulate the use of any probability distribution that might be empirically fit population data. In conclusion, the extension of the model to a heterogeneous population requires describing its heterogeneity with a probability distribution over $a(k)$ and $b(k)$.

B. A Heterogeneous Markov Model Without Delays

In this subsection we introduce heterogeneity on the frequency intervals $\mathcal{I}_f$ of different panels. As a consequence, each panel presents different frequency thresholds, which are constant in time: each panel reacts to the network frequency distinctively. Each panel can therefore disconnect or reconnect at different values of the frequency. We assume that these thresholds over the population are distributed continuously according to known probability distributions. These distributions model the transition from $R(k) = 1$ to $R(k+1) = 0$, and vice versa, and can in practice be regressed from population data.

We introduce $a(k)$ and $b(k)$ as the ON-to-OFF and OFF-to-ON transition probability, respectively (cfr. Fig. 2). Notice that in this case $a(\cdot)$ and $b(\cdot)$ are not binary: we obtain them by integrating probability distribution functions, as seen shortly. The current value of network frequency $f(k)$ is used as one of the extrema of such integral: in this way we obtain the

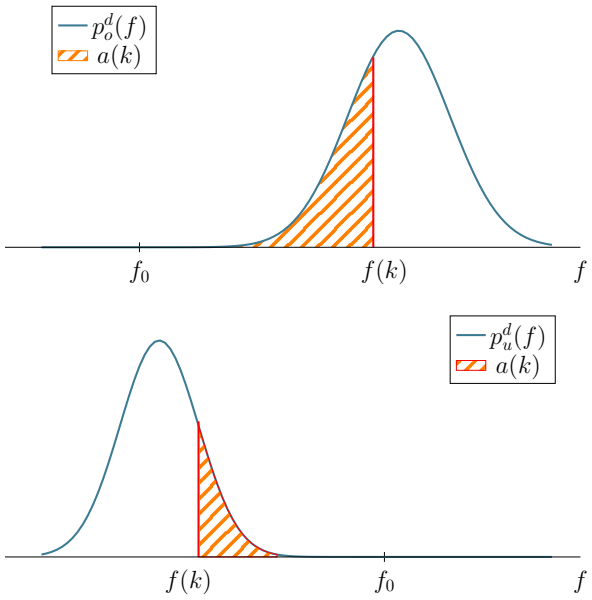


Fig. 1. Pictorial representation of $a(k)$ in over-frequency, i.e. $f(k) > f_0$ (top) and in under-frequency, i.e. $f(k) < f_0$ (bottom). The value of $f(k)$ is indicated as a red vertical line, which defines the upper or lower integration extrema in over- and under-frequency, respectively. In general the distributions $p_u^{(\cdot)}$ and $p_o^{(\cdot)}$ are not symmetric with respect to f_0, nor belong to the same distribution family.

portion of panels that are enabled to transition to the other state. Formally,

$$a(k) = \begin{cases} \int_{-\infty}^{f(k)} p_o^d(u)du & \text{if } f(k) > f_0 \\ \int_{f(k)}^{+\infty} p_u^d(u)du & \text{otherwise,} \end{cases}$$

$$b(k) = \begin{cases} \int_{f(k)}^{\infty} p_o^r(u)du & \text{if } f(k) > f_0 \\ \int_{-\infty}^{f(k)} p_u^r(u)du & \text{otherwise,} \end{cases}$$

where p_i^j denote the probability distribution functions mentioned above (recall that the superscripts d and r indicate the disconnection and reconnection, respectively), and where the subscripts o and u further denote over- and under-frequency, respectively. Note that intuitively $a(k)$ is computed as the part of the integral that is closer to f_0, whereas $b(k)$ is the part that is further away from f_0. The evaluation of $a(k)$ is shown in Fig. 1. As an alternative, $a(k)$ and $b(k)$ can be expressed via corresponding cumulative distributions.

Equation (1) now defines a two-state Markov chain with time varying transition probabilities, as depicted in Fig. 2. Note that $a(\cdot)$ and $b(\cdot)$ are function of the current frequency $f(k)$: to ease notations, we will denote them as $a(k)$ and $b(k)$ instead of $a(f(k))$ and $b(f(k))$.

C. A Heterogeneous Markov Model With Delays

We now introduce a framework to encompass delays in the population model (cf. Fig. 3): as observed in practice, panels

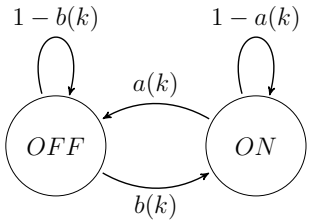


Fig. 2. A time-varying Markov chain for the aggregated dynamics.

cannot reactivate instantaneously but must wait a delay. We generalise the discussion of Section II by allowing for possibly random delays. Recall that each device is equipped with an internal counter for the OFF-to-ON transition. We assume to know the delay probability distribution function, and introduce transition values τ_i: value τ_i represents the probability of switching to state ON, given that the panel has been *waiting* for i time instants (which means that $f(\cdot) \in \mathcal{I}_f$ for i time instants). We utilise n states, defined as w_i, $i = 1, \dots, n$, representing the i-th time step in which $f(\cdot) \in \mathcal{I}_f$ with the device still OFF.

In view of the discussions in Section II, we focus on the case $n \gg 1$. In the i-th delay state, three outgoing transitions are present: one towards the ON state, a second towards state $i+1$, and one back to the OFF state (cf. Fig. 4). The probability associated with the latter transition is $1 - b(k)$, which is the $f(k) \notin \mathcal{I}_f$ probability. The first outgoing probability is $\tau_i b(k)$: τ_i is the delay probability value to transition from state w_i to state ON, multiplied by $b(k)$.

Recall that, to allow for a reconnection, the condition $f(k) \in \mathcal{I}_f$ must be satisfied at all times k when the counter $\tau(k) \leq \tau_r$. Since at those times the frequency must dwell within $\mathcal{I}_f$, every transition towards the ON state is multiplied by $b(k)$. In particular, transitions between waiting states can be thought of as a Binomial variable where the probability at each time step is given by $b(k)$. Notice that we have tacitly assumed that the probability distributions of frequency thresholds and time delays are independent. In practice, there can be correlation between these two quantities, however this does not invalidate our model: such a case requires computing integrals of joint probability distributions.

We further assume that delays are distributed according to a geometric distribution: this is often used to characterise arrival processes or waiting-time random variables. This distribution has the property that $\forall i, \tau_i \geq \tau_{i+1}$ and $\sum_i \tau_i = 1$.

The dynamics of the Markov chain in Fig. 3 can be summarised as

$$\begin{cases} x(k+1) = (1 - a(k))x(k) + b(k) \sum_{i=1}^{n} \tau_i w_i(k) \\ w_1(k+1) = b(k) \left[1 - x(k) - \sum_{i=1}^{n} w_i(k)\right] \\ w_i(k+1) = b(k)(1 - \tau_{i-1}) w_{i-1}(k) \\ w_n(k+1) = b(k) \left[(1 - \tau_{n-1}) w_{n-1}(k) + (1 - \tau_n) w_n(k)\right], \end{cases} \tag{2}$$

where $x(k)$ represents the probability of being in the ON state at time k. We alternatively interpret this value as the *portion of panels* that are ON at time k: this modified point of view allows us to translate thinking from a model of a *single* panel to an *aggregated* model for the population. Similarly, $w_i(k)$

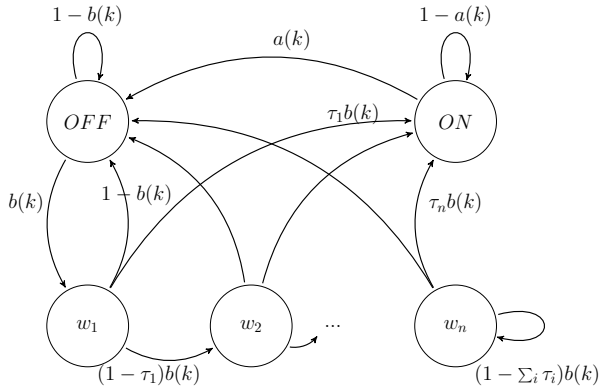


Fig. 3. A Markov model for the aggregated dynamics. This represents the heterogeneous population in presence of a delayed reconnection.

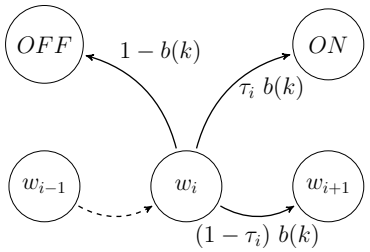


Fig. 4. Outgoing transitions of the single waiting state (refer to Fig. 3): towards OFF if the frequency trips outside $\mathcal{I}_f$; when the frequency is within $\mathcal{I}_f$ towards ON if the counter expires, or towards the next waiting state otherwise.

indicates the portion of panels waiting to turn ON for i time steps at time k and $a(k), b(k)$ are the *cdf*'s of the distributions of frequency thresholds in the population of panels.

Remark 1: In practice the reconnection interval is usually divided into two sub-intervals: during an early sub-interval no device is allowed to reconnect, whereas during a following sub-interval the reconnection happens stochastically. Our modelling handles both scenarios. The first one is represented by setting the τ_i to zero; the second one setting the values of τ_j according to the given probability distribution. As an example, imagine a minimum deterministic delay of 20 time steps with no reconnection, and a subsequent interval of 20 time steps when the devices might reconnect. The delay states needed to encompass this scenario is a total of 40. The vector T containing τ_i, $i \in [1, 40]$, is split as follows: $\tau_i = 0$, for $i = 1, \dots 20$, whereas $\tau_j > 0$, for $j = 21, \dots 40$. In the following, in order to ease the discussion we only consider the second condition. □

D. Simplification of the Heterogeneous Model with Delays

Towards a simplified and more explicit analysis of the dynamics of the model with delays, we aggregate the n waiting states into a single location (denoted by WAIT in Fig. 5), which thus represents the portion of devices that are waiting to turn ON. We associate the aggregated location to a new variable $y(k) = \sum_i w_i(k)$. To formulate the corresponding

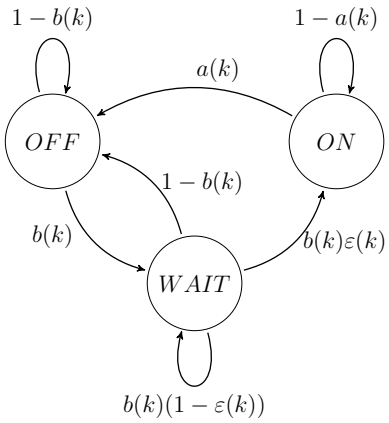


Fig. 5. Simplification of the Markov model in Fig. 3. States w_i are lumped into state *WAIT*.

dynamics, we rewrite the term $\sum_i \tau_i w_i(k)$ as a function of $y(k)$, and introduce a new term $\varepsilon(k) \in [0, 1],\ \forall k$, such that

$$\sum_{i=1}^{n} \tau_i w_i(k) = \varepsilon(k) \sum_{i=1}^{n} w_i(k),$$

namely

$$\varepsilon(k) := \frac{\sum_{i=1}^{n} \tau_i w_i(k)}{\sum_{i=1}^{n} w_i(k)}. \tag{3}$$

The new model has only three states, as depicted in Fig. 5, whose transition equations are

$$\begin{cases} x(k+1) = (1-a(k))x(k) + b(k)\varepsilon(k)y(k) \\ y(k+1) = b(k)(1-x(k)) - b(k)\varepsilon(k)y(k) \end{cases}, \tag{4}$$

in view of the definition of $\varepsilon(k)$ from Eq. (3).

The new model is smaller, and prone to explicit analysis. However, in general we do not know the exact value of $\varepsilon(k)$ and we need to estimate it. Knowing the values τ_i, we can estimate $\varepsilon(k)$ by estimating the states $w_i(k)$. The estimation of the states $w_i(k)$ can be attained knowing the state variable $f(k)$ at each step k (a model for the frequency dynamics will be introduced in the next Section). The state estimation algorithm is detailed in Appendix I.

Remark 2: The observability and estimation analysis are significant in practice, as they allow a precise monitoring of the state of the aggregated population of PV panels simply from a frequency measurement. □

IV. Model of the Electric Network

The model for the electric grid is derived from the ENTSO-E report [17]. It represents a geographically large network with several energy power sources and load characteristics. Its parameters represent an average over the whole network. It consists of a continuous-time model of the electric network in the form of a second-order transfer function, $G(s)$, which relates the photovoltaic power deviation, $\Delta P_{PV}(t)$ (its input) to the frequency change $\Delta f(t)$ (output). Fig. 6 shows the block diagram representation of $G(s)$, and which is shaped as:

$$G(s) = \frac{s+1}{T_L s^2 + (k_a + T_L)s + (k_a + k_{PU})},$$

where the quantity T_L represents the time to launch (related to the inertia of the system), whereas k_a and k_{PU} are gains of the load and of the primary control unit, respectively. The transfer function includes the primary control of the grid, which is designed as an internal feedback loop: this represents a realistic approximation of the network dynamics over a short timescale (seconds).

Remark 3: Power generation and consumption need to be matched in order to secure a high-quality operation of the electric network: this is attained by a control architecture known as *Load-Frequency Control* [17]. Control actions are performed in successive phases, each with different characteristics and goals: Primary Control stabilises the network frequency at a stationary value after a disturbance or incident, and acts within a time-frame of seconds without restoring the frequency to its nominal value. Secondary Control maintains a balance between generation and consumption in a time-frame from 30 seconds up to minutes after an incident. Tertiary Control replaces Secondary Control by re-scheduling power generation, and acts after several minutes. □

Note that the continuous-time feature of the grid model is in contrast with our discrete-time models above: we will perform a description and the analysis of the network in continuous-time, and introduce a time discretisation when connecting it to the aggregated model of the population of PV panels. The discrete-time version will be employed in the experiments of Section VI.

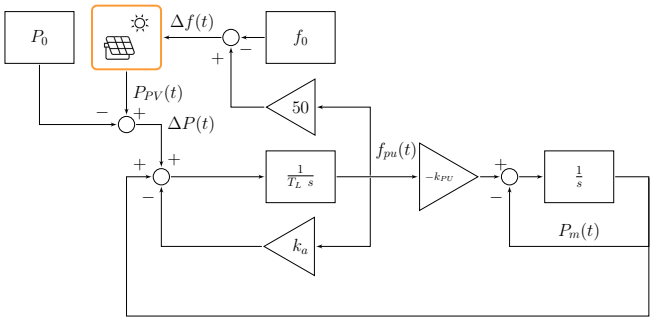


Fig. 6. Network block diagram representation. Quantities $\boldsymbol{f(t)}$ and $\boldsymbol{f_{pu}(t)}$ represent the frequency and the frequency per unit, respectively. $\boldsymbol{P_m(t)}$ denotes the power coming from traditional machines, whereas $\boldsymbol{P_{PV}(t)}$ is the solar power. The PV model is represented by the orange block.

Models of the national or European grid [5], [21], [22] are developed with the presence of synchronous machines and under inertia. In our context the total moment of inertia associated to synchronous machines is considered to be constant. This assumption is justified with traditional power sources connected to the network, however when dealing with renewable power sources the concept of inertia is more difficult to handle. For instance, the notion of *synthetic inertia* [23] is introduced when wind generation is under study (here potential energy is stored in their blades). However, evidently solar panels instead have no moving parts, which implies they come with no mechanical inertia at all: if solar irradiance suddenly subsides, a solar device immediately stops generating electricity. The opposite is also true: whenever the sun shines onto a PV panel, power is immediately generated.

As a result, when the penetration of solar power sources in a network is significant, their contribution must be accounted for. We include this contribution in the transfer function $G(s)$, which we regard as a function of the amount of the *conventional power* generated by synchronous machines.

We assume a linear relation between the time to launch T_L and the amount of conventional power generated in the network. Intuitively, the more conventional generators in the network, the more inertia governs the frequency evolution and the larger the the starting time:

$$T_L = \frac{CP}{k_T},$$

where CP represents the conventional power expressed in GW and k_T is a constant value. Note that whenever $CP = 0$ we obtain a time to launch that is equal to 0 seconds. We will not discuss the $CP = 0$ scenario, as it is out of the scope of this work: this situation leads to additional considerations about safety and management of such a network.

We substitute T_L in the expression of $G(s)$, obtaining

$$G(s, CP) = \frac{(s+1)}{\frac{CP}{k_T}s^2 + \left(k_a + \frac{CP}{k_T}\right)s + (k_a + k_{PU})}. \quad (5)$$

Notice that $S \geq CP$, where the equality represents a network without any renewable power.

Note that the considered network model does not account for local changes in frequency, nor for the topology of the network, aiming instead at modelling a large grid network. It is in practice considered by TSOs to be a reliable model for the continental electricity grid, as the frequency can be considered homogeneous across countries. So rather than being limited in its global perspective (in that it does not encompass local or regional variations), the main limitation of our network model is the homogeneous description of all energy sources in the network. As an example, in the network model we consider only one equivalent synchronous generator to represent all the synchronous power plants: this does not take into account the difference in dynamics between, say nuclear, coal or hydraulic power plants, which have different characteristics. Further, we have modelled a dependence of the inertia from the solar penetration in the network, whereas other renewables, e.g. wind power plants, might also have an impact. More detailed models exist (dynamic tools as Eurostag [24], among others), however they come with higher computational costs and would furthermore not improve the present study on the impact of PV panels dynamics on the network.

A. Influence of Solar Penetration on Network Dynamics: Root Locus Analysis

We study how the ratio CP/S influences the stability of the network. This is achieved via root locus analysis [25]. We assume to utilise a proportional control, as explained in Section V, under different ratios of CP/S. Fig. 7 shows the root locus in three different conditions: $CP/S = 1$, i.e. no renewable power in the network; $CP/S = 0.8$; and $CP/S = 0.5$, with $S = 220$ GW. Note that in every configuration, the function

$G(s, CP)$ is stable, i.e. its poles are in the negative real half plane, as shown in Fig. 7.

The effect of renewables on the network dynamics is quite straightforward: the more renewable power, the more oscillatory the system becomes, leading to larger overshoots with respect to the setting with only traditional power. On the other hand, we notice that also the real part increases in absolute value, meaning a faster convergence. Most importantly, wider oscillations could eventually cause issues in the electric grid setting: the grid operators should avoid wide variations of the frequency signal, as this brings the network to its operational limits and potentially causes load shedding.

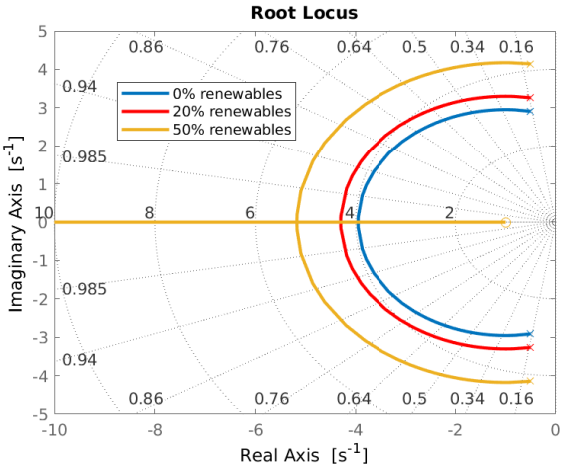


Fig. 7. Root locus with just traditional power sources (blue solid line), 20% (red dashed) and 50% (yellow dotted) renewable power.

B. Closed-Loop Dynamics of Solar Devices Connected to the Grid: Stability Analysis

We now introduce the population of solar panels, and connect it to the electric network in feedback, as per Fig. 6. As mentioned in the previous section, we continue the analysis after discretisation of the network transfer function. Let us focus on a network with a constant amount of conventional power, and refer to the transfer function as $G(s)$, rather than as $G(s, CP)$. Formally, we have that

$$\Delta f(t) = G(s)[P_{PV}(t) - P_{PV,0}].$$

Here $P_{PV}(t)$ encompasses the power generation of solar at time t, whereas $P_{PV,0}$ is the solar power injection to maintain the frequency $f(t)$ at the equilibrium f_0. Recall that $G(s)$ relates the power deviation to the frequency deviation $\Delta f(t) = f(t) - f_0$.

We now discretise $G(s)$ from Eq. (5) into $G(z)$ via the *step response invariance method* [26]. This approach guarantees perfect matching of the continuous-time signal with the discrete-time signal at sampling times. Formally,

$$G(z) = \frac{\beta_1 z + \beta_2}{z^2 + \alpha_1 z + \alpha_2}, \tag{6}$$

where α_1, α_2, β_1, β_2 are constants that depend on the chosen sampling time. In this work we select the same sampling time as that of the inverters (0.2 s). We can write the discrete-time equivalent of the previous equation as

$$\Delta f(k) = G(z)[P_{PV}(k) - P_{PV,0}], \tag{7}$$

where $\Delta f(k) = f(k) - f_0$, f_0 represents the nominal value of the network frequency, $f(k)$ is the value of the frequency at time k, while $P_{PV}(k)$ is the power output of the population of PV panels at time k and (as above) $P_{PV,0}$ the power produced when $f(k)$ is at the equilibrium f_0. We set

$$P_{PV}(k) = \bar{P} N x(k),$$

where $\bar{P}$ is assumed to be the constant power output of a single PV panel, and N is the total number of panels. As such, we obtain a proportional relation between the total power output and the portion of panels in the ON mode. This assumption simplifies the formal analysis on the feedback models in Equations (8) and (9).

Let us denote $\Delta x(k) = x(k) - x_0$, where x_0 represents the portion of active panels when $f(k)$ is at the equilibrium f_0. Transforming the frequency relation in Equation (7) into state-space form, and embedding it within the dynamics of the Markov chain with n waiting states, as per Equation (2), results in

$$\begin{cases} \Delta f(k+1) = \alpha_1 \Delta f(k) + \Delta\alpha_2 f(k-1) + \\ \qquad\qquad + \gamma_1 \Delta x(k) + \gamma_2 \Delta x(k-1) \\ x(k+1) = (1 - a(k))x(k) + b(k) \sum_{i=1}^{n} \tau_i w_i(k) \\ w_1(k+1) = b(k) \left[1 - x(k) - \sum_{i=1}^{n} w_i(k)\right] \\ w_i(k+1) = b(k)(1 - \tau_{i-1}) w_{i-1}(k) \\ w_n(k+1) = b(k)[(1 - \tau_{n-1}) w_{n-1}(k) + \\ \qquad\qquad + (1 - \tau_n) w_n(k)], \end{cases} \tag{8}$$

whereas the embedding within the Markov chain model with three states, as per Equation (4), results in

$$\begin{cases} \Delta f(k+1) = \alpha_1 \Delta f(k) + \alpha_2 \Delta f(k-1) + \\ \qquad\qquad + \gamma_1 \Delta x(k) + \gamma_2 \Delta x(k-1) \\ x(k+1) = (1 - a(k))x(k) + b(k)\varepsilon(k) y(k) \\ y(k+1) = b(k)(1 - x(k) - \varepsilon(k) y(k)). \end{cases} \tag{9}$$

Notice that we have introduced constants $\gamma_i = \bar{P} \cdot N \cdot \beta_i$, $i = 1, 2$. The models described by Equations (8) and (9) are utilised in the case studies in Section VI. The feedback connection of the Markov chains and the process $G(z)$ can be shown to be locally asymptotically stable around the nominal frequency. The proof can be found in Appendix II.

V. Decentralised Control of Solar Power Production

We have so far considered the power output of connected solar panels to be constant in time. However, in practice the operational mode of a solar panel is crucially determined by the so called Maximum Power Point Tracking (MPPT). This mechanism is embedded in the panel and employed to maximise its power output, and therefore its efficiency, which depends on several factors, such as solar radiation and external temperature [27], [28]. Within our modelling framework, the MPPT algorithm acts solely when a panel is active, namely in state ON, whereas it remains idle when the panel disconnects. Based on this practical setup, in this work we assume to

be able to control the power output of each device level by means of a proportional gain controller, which is introduced in Equation (10). We argue that in practice this control scheme can be implemented within the inverter alongside the MPPT method with little effort. The overall algorithm works as follows: it first finds the MPP; it then computes the desired power output via a proportional control (described below); finally it selects a new working point based on the ratio between the desired power and the maximum power.

Building on the described control algorithm for the single panel, we propose a decentralised scheme to control the entire population of PV panels, which can be interleaved with the current primary network control. We assume that each device, during normal operation, injects a power level P_{eq} into the grid that is strictly less than P_{MAX}, which is defined as the maximum available power to the panel. The quantity P_{eq} can be tuned according to specifications and requirements by the Transmission System Operator [29]. We utilise the following proportional control law:

$$P_{PV}(k) = P_{eq} + k_p \cdot \Delta f(k), \tag{10}$$

where k_p is a constant gain to be determined, and $\Delta f = f(k) - f_0$ represents the frequency deviation from the nominal value. We compute k_p via disturbance rejection at the steady state f_{ss}, representing the value of frequency at steady state, considering a step disturbance $d(t) = A > 0,\ t \geq 0$, so that

$$f_{ss} = A \cdot \frac{G(0)}{1 + k_p G(0)}, \tag{11}$$

where $G(0)$ is the steady-state gain of $G(s)$. Considering a maximum acceptable steady-state disturbance value f_{ss}^{max}, we can characterise a working region for the gain as

$$k_p \geq G(0)^{-1} \left(\frac{A \cdot G(0)}{f_{ss}^{max}} - 1 \right). \tag{12}$$

This results in a minimum value for k_p, which is then compared to the root locus analysis in Section IV (cf. Fig. 7). The root locus shows the coordinates of the poles of the interconnected system on the complex plane. The imaginary component results in oscillations of the response, which are in general undesirable: with the goal to avoid oscillations of the frequency response, we can select k_p resulting in real poles for the closed loop.

Lack of power balance between generation and load might result in an oscillatory frequency signal, which suggests the introduction of a (small) deadband for the controller design. Accordingly, regulation aims at maintaining the network frequency within a (small) interval around f_0, contrary to perfect tracking as usual in control theory. As such, a control action is performed solely when the frequency is outside this predefined interval, whereas if the frequency is close enough to the nominal value f_0 no control action is needed. In the case of underfrequency, we denote the interval over which the control act as $[f_u^{min}, f_u^{max}]$; similarly in case of overfrequency, $[f_o^{min}, f_o^{max}]$. Let us remark that these intervals are not related to the working thresholds of the devices. The control design in Equation (10) with deadband is depicted in Fig. 8.

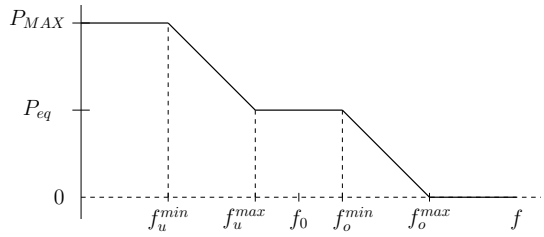


Fig. 8. Power output vs. network system frequency. The power profile shows a deadband $[f_u^{max}, f_o^{min}]$ around f_0 and a proportional gain in underfrequency interval $[f_u^{min}, f_u^{max}]$ and overfrequency interval $[f_o^{min}, f_o^{max}]$. P_{eq} represents the nominal power at f_0.

VI. Case Studies

In this section we test the quality of the developed models for the aggregation of a population of heterogeneous PV panels. Further, we utilise the control scheme, in order to illustrate the response of the grid frequency to a sudden generation or load loss, under several scenarios. Specifically,

1) Section VI-A shows the quality of our aggregated models, which are simulated and tested against a population of N individual devices.
2) Section VI-B studies the relationship between the distribution of frequency thresholds and the event of load shedding after an incident: we show that as the solar penetration grows, the chance of load shedding increases. However, heterogeneity can be exploited to enhance the network reliability and its resilience to incidents, much as we have argued earlier for stability analysis.
3) In Section VI-C we add a controlled power output to the previous scenarios: we show that the proposed control architecture avoids load shedding and provides a better frequency response.

In line with the ENTSO-E requirements [17], we consider a network with demands of $S = 220$ GW and 440 GW, respectively. This scenario fits the framework of disturbance rejection considered in Section V and accommodates the study of the following incidents. The Operation Security Network Code [30] divides incidents into three categories: *normal, exceptional*, and *out-of-range*. Following normal incidents, the function of the system must be maintained without violations of technical limits. Normal incidents are classified as a loss of up to 2 GW of load, and as a loss of up to 3 GW of power generation. In this work, we are interested in the consequences of such normal incidents under varying levels of penetration of solar energy, and under different distributions of the frequency intervals $\mathcal{I}_f$ over the population. In our simulations we utilise the maximum value of the infeed loss incident – 3 GW – and test the response of the network under different circumstances. On the contrary, as already noted in [13], [31], the load loss incident has practically no load shedding risk during normal operations: therefore, this test is omitted. Further, we are interested in testing the worst-case setup for the disconnection of devices: namely, we stress the network and attempt to attain the lowest value that $f(k)$ can reach in the electricity grid. To ensure this, the time-frame of power loss tests is set to 20 seconds, which prevents any panel from reconnecting –

the reconnection of devices can only increase $f(k)$, adding production to the power imbalance.

In order to simulate a power generation loss, we inject a negative step in the frequency loop corresponding to the incident. We utilise the network model with modified inertia from Section IV. As anticipated in earlier Sections, the unpredictable solar power output of single panels is encompassed via additive zero-mean Gaussian noise η_{PV}. Furthermore, the natural noisy nature of the electric frequency is described similarly with an additive Gaussian signal η_f. We add η_f to the dynamics of $\Delta f(k+1)$ in (8)-(9), and η_{PV} to the dynamics of $P_{PV}(k+1)$ derived from (8)-(9) and to Equation (10).

As we shall see, the injected noise has a small amplitude compared to the frequency and power signals, and will be negligible in the case of the 440 GW network.

The network model parameters are chosen according to the guidelines in [17][1]: the self-regulation of the load k_a is assumed to be 0.01 s, the average network primary control gain equal to $k_{PU} = 15000$ MW/Hz, and the system start time $T_L = CP/k_T$, where $k_T = 15 \cdot 10^3$. Power and frequency values are normalised relative to S and to 50 Hz. Power production of a single panel P_{MAX} is set to 3 kW. The variance of η_{PV} is set to 1% of P_{MAX}. The variance of η_f is computed from network frequency data [32] using a Maximum Likelihood Estimation technique: the obtained value is equal to 0.025 Hz^2. Time delays are modelled in accordance with [18], [19]: the minimum reconnection delay is set to 20 seconds, whereas the maximum to 40 seconds. Whilst these two quantities are handled deterministically, the population delays within these quantities are modelled via a geometric distribution, as discussed previously. According to requirements in [17], we set a frequency limit value of 49.2 Hz: in real setups, if the frequency trips below the limit of 49.2 Hz, an automatic procedure of load shedding is activated. In our simulations, we check if the frequency trips below this critical value, and stop the simulation.

Simulations are implemented in MATLAB. The grid frequency is sampled at a rate of 0.2 s, consistently with requirements introduced in [33]. The discussion is focused on the consequences of an incident after a few seconds, and the simulation time is set to $20\,s$. After this time interval, we assume the primary and secondary network controls have kicked in and shall stabilise the signal $f(\cdot)$ around its nominal value f_0 (cf. discussion in Remark 3).

A. Experimental Evaluation of the Aggregated Models

In order to show the precision of the aggregated models, we set up rounds of simulations comparing the two alternative models (the n-waiting-state in Fig. 3 and the three-state in Fig. 5) with the ground truth obtained from an explicit simulation of the entire population of PV panels within the power network (which we denote as the *explicit model*).

For the explicit model, each of the N panels has been given four different frequency thresholds (disconnection and reconnection in over- and under-frequency) and a time delay (as a number of time steps the device needs to wait before turning to state ON) between 20 and 40 seconds. These parameters have been generated according to given probability distributions for the population, which are then used in the computations for the aggregated models. We have set up the distributions of frequency thresholds as Gaussian, and that for the time delay as geometric, and have set $N = 10^6$.

In order to validate our modelling framework, let us test the power response – namely the evolution of $x(k)$ – injecting a controlled frequency signal into the feedback models in Eq. (8) and Eq. (9). Fig. 9 shows the power response of the explicit model, as well as of the $(n+2)$-state and 3-state models, together with the injected frequency signal. The frequency evaluates to 50 Hz for the first 10 seconds, 49.5 Hz between 11 and 40 seconds (causing a partial disconnection of the population), 48 Hz between 41 and 100 seconds (resulting in a total disconnection of the population), and increasing back to 50 Hz from 101 seconds on. First, we note the perfect overlapping between the $(n+2)$-state and the 3-state models, as they are analytically equivalent. We then show the small discrepancies between the explicit and the two abstract models in Fig. 10. Specifically, Fig. 10a zooms in on the difference between the $x(k)$ signals after the first disconnection period, which evaluates to around 10^{-4} (dimensionless), whereas Fig. 10b depicts the gap at the end of the reconnection process, evaluating at around 3.5. Recall that these gaps should be multiplied by the total power of the aggregation $N \cdot P_{MAX}$, resulting in deviations of 0.3 MW and 100 MW in a 220 GW network, respectively. These modelling gaps derive from our approximation: disconnection and reconnection thresholds are defined over a *continuous* domain, whereas the explicit model is characterised by a large-but-finite number of devices. Further, the reconnection of the abstract models shows an exponential rate – in view of the term $(1-\tau_n)w_n$ in the last waiting state in Eq. (2), whereas the explicit model displays a less smooth behaviour. In the explicit model all devices deterministically reconnect after 40 seconds, thus reaching $x(k) = 1$ after a finite time horizon (see Fig. 10b where at 141 seconds $x(k)$ becomes 1); the abstract models show instead an exponential convergence to $x(k) = 1$, thus resulting in a continuously decreasing gap between the power signals.

In terms of computational complexity, the explicit model is much heavier to simulate. At each time step, each device compares the current value of the network frequency with its frequency thresholds (and delay counter), and disconnects (or reconnects) accordingly. The Markov model, on the other hand, computes a single integral per time step, relieving the computational effort significantly.

B. Testing Distributions and Load Shedding Relation

Previously the quality of the aggregated models has been tested on a simulation benchmark, which has shown their precision. In this section we employ the aggregated models for the population of PV panels in several scenarios.

Notice that different distributions over $\mathcal{I}_f$ have different outcomes and bring to different conclusions about heterogeneity.

[1]G(z) includes a variation of load in response to frequency deviations, encoded as parameter k_a. The system start time is the time that a device needs to accelerate from zero to the working speed.

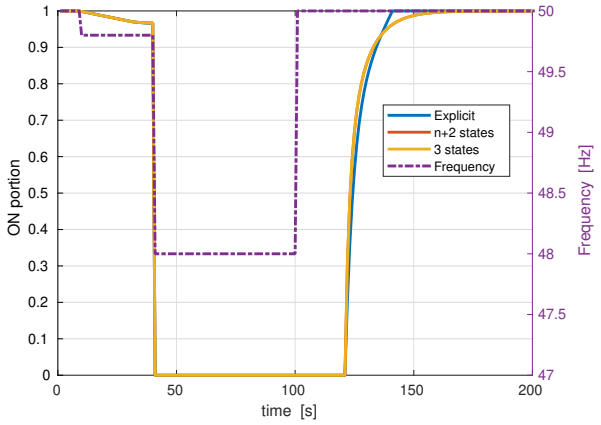


Fig. 9. Power response for the explicit model ($\boldsymbol{N}$ individually simulated devices, in blue), for the $(\boldsymbol{n+2})$-state model (red), and for the 3-state Markov model (yellow) to the frequency signal (purple, dashed).

TABLE II
TEST RESULTS FOR 3 GW POWER LOSS SCENARIO OVER 220 GW NETWORK LOAD.

% PV	Mean under	Mean over	Variance	Load Shedding
10	49.5	50.5	0.05−0.25	0.15−0.25
20	49.5	50.5	0.05−0.25	0.10−0.25
40	49.5	50.5	0.05−0.25	0.05−0.25
$\sim 10\%$	$49.8 - 49.0$	50.2	1−5	no

We then test two scenarios concerning population thresholds: a *Narrow Interval Scenario*, encompassing a limited working frequency interval and Gaussian distributions, and a *Composite Scenario* with several χ^2 distributions. As we will see, Gaussian distributions accurately model inverter measurement noise, whereas χ^2 define a minimum performance setting, in terms of a minimum working interval.

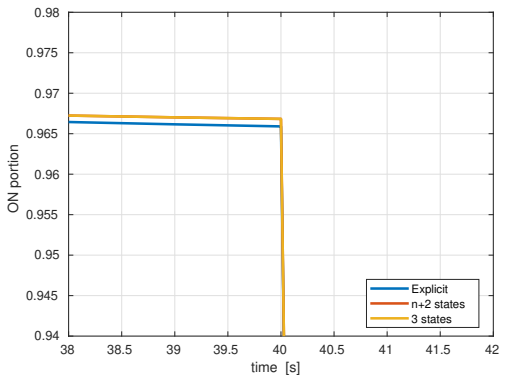
Tables II and III refer to network load scenarios of 220 and 440 GW, respectively. They show the results of tests as a function of the percentage of solar penetration in the network, and of the mean and variance of the considered distributions. In the simulations the amount of load shedding does not need to be computed: when the network frequency trips below 49.2 Hz, we assume that the procedure of load shedding is activated (as discussed in the Introduction) and we report it in the *Load Shedding* column of the Tables, along with the value of the variance. In other words, the *Load Shedding* column indicates which values of variance causes the load shedding procedure to activate. In the following we describe the outcomes of the two scenarios of thresholds distribution.

1. Narrow Interval Scenario: In the *Narrow Interval Scenario*, we consider three percentages of penetration of solar energy production in the network: 10%, 20%, 40% (first three rows of Tables II and III). We assume that the heterogeneity over the frequency intervals that model the reconnection/disconnection thresholds – in terms of ageing, manufacturing, performance deterioration, sensor noise, etc. – is well described by Gaussian distributions, which we consider over different values of mean and variance.

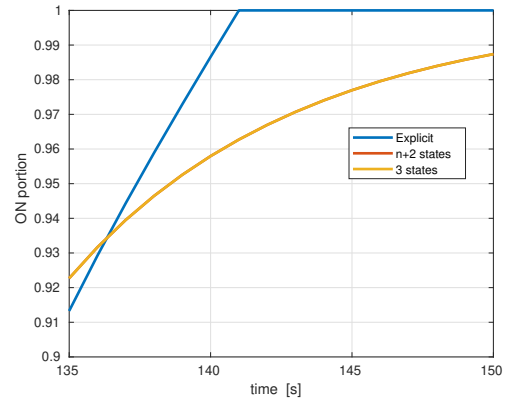
Thresholds average values are set to 49.5 Hz (underfrequency) and 50.5 Hz (overfrequency), resulting in a 500 mHz band around the nominal frequency of 50 Hz. We investigate five values of variance, with values ranging from 0.05 to 0.25 Hz^2 at intervals that are 0.05 apart. Simulations show, both in the 220 GW and 440 GW scenario, that larger values of the variance cause a higher number of panels to disconnect, resulting in load shedding. In the 440 GW network load with 10% solar penetration scenario, no value of variance was sufficient to trip the frequency below 49.2 Hz (cf. first row of Table III). Clearly, when the network load is higher, the reliability of the network itself is enhanced, having a higher time-to-launch and therefore a larger inertia. Fig. 11 shows the frequency drop in a 220 GW network load scenario, with 10% solar penetration (cf. first row of Table II). Different values of the variance (between $0.15 - 0.25$ Hz^2) result in different

TABLE III
TEST RESULTS FOR 3 GW POWER LOSS SCENARIO OVER 440 GW NETWORK LOAD. SAME SETUP AS IN TABLE II.

% PV	Mean under	Mean over	Variance	Load Shedding
10	49.5	50.5	0.05−0.25	no
20	49.5	50.5	0.05−0.25	0.15−0.25
40	49.5	50.5	0.05−0.25	0.15−0.25
$\sim 5\%$	$49.8 - 49.0$	50.2	1−5	no



(a)



(b)

Fig. 10. Zoom at the end of the first frequency drop (Fig. a) and at the end of the reconnection stage (Fig. b). The difference between the normalised power of the explicit and the abstract models is $\mathbf{1.3\ 10^{-4}}$ (dimensionless) after the first frequency drop, and $\mathbf{3.5}$ (dimensionless) at the end of the reconnection process.

TABLE IV
PV POPULATION WITH DISTINCT GROUPS OF PANELS, EACH WITH SPECIFIC UNDER-/OVER-FREQUENCY THRESHOLDS, AND ASSOCIATED ENERGY PRODUCTION (IN MEGAWATTS) AND RELATIVE SIZE (%).

Underfrequency threshold	49.8	49.7	49.5	49.0	47.5
Overfrequency threshold	50.2	50.2	50.2	50.2	50.2
Energy production [MW]	1000	2000	4500	4000	8500
Relative size [% of total]	5	10	22.5	20	42.5

numbers of panels to disconnect, potentially resulting in load shedding. Similar outcomes are shown in Fig. 12, where a network load of 440 GW and solar penetration of 20% are considered (cf. second row of Table III).

2. *Composite Scenario*: We slightly modify the settings used above to achieve a more faithful description of the heterogeneity of the grid. We employ data from European countries [31] as a reference point, as summarised in Table IV. We assume that the PV population is divided in five groups, each with different under- and over-frequency limits. A group characterises the smallest working interval that complies with regulations, and different groups comply with different regulations. A given panel must have a working frequency interval $\mathcal{I}_f$ that is wider than the characteristic interval of its group: as an example, panels installed with the 49.5 Hz underfrequency limit can have a threshold $f_{uf} \leq 49.5$. Similarly, panels are allowed to have $f_{of} \geq 50.2$ in the 50.2 Hz overfrequency case. The probability distribution that best describes this scenario is the χ^2 distribution. Note that, in contrast with the previous case study, the χ^2 distribution is not symmetric around its mean. The contribution in energy production [MW] from each group is on the third row of Table IV, whereas the fourth row shows the partition size in percent terms. The total amount of energy production from the entire population accrues to close to 10% of solar penetration in the 220 GW network, and to 5% in the 440 GW network: see the last row of Tables II and III, respectively.

Similarly to the previous case studies, we test five values of variance, ranging from 1 to 5 Hz^2 (with a step of 1), and how they possibly result to load shedding (cf. low end of values of variance for the 220 GW case on bottom row of Table II). Note that increasing variance leads to distributing the population thresholds away from the nominal frequency. Indeed, simulations show that increasing the variance enhances the reliability of the grid, as depicted in Fig. 13: whilst no scenario results in load shedding, a higher heterogeneity acts as a frequency response smoother.

Remark 4: In practical terms, it is interesting to emphasise that new European Union regulations [33] allow for a broader interval of frequency values for inverters than in the past: newly manufactured solar inverters should work within the interval $[47.5, 51.5]$ Hz. Based on the developed models, as demonstrated above and further discussed in [13], this larger frequency interval is likely going to contribute to the reduction of the risk of load shedding after normal incidents. On the other hand, old solar inverters do not abide by these new requirements and low-quality panels are likely to suffer from

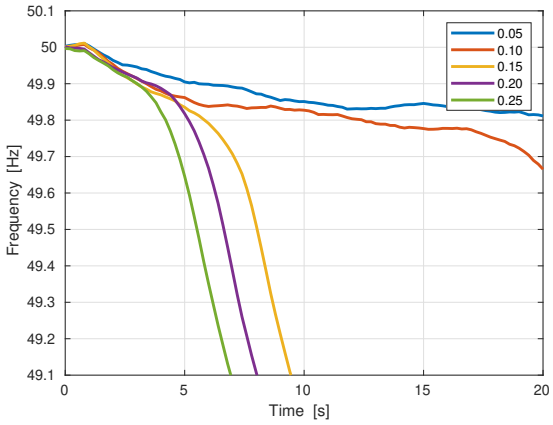


Fig. 11. Simulations for a generation loss of 3 GW on a 220 GW network load, with Gaussian distributions for the frequency thresholds under several values of variance. There is a $\mathbf{10\%}$ penetration of solar power generation.

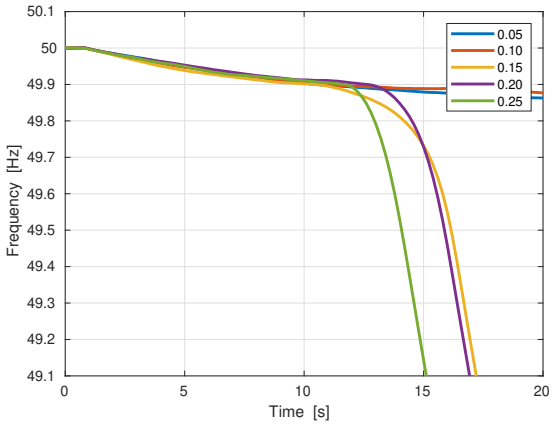


Fig. 12. Simulations for a power generation loss of 3 GW on a 440 GW network load, with a $\mathbf{20\%}$ penetration of solar power generation.

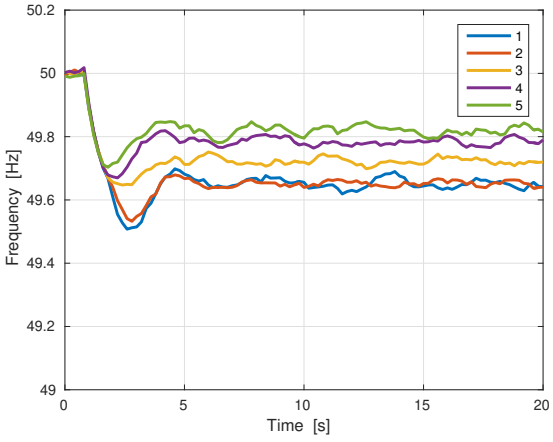


Fig. 13. Simulations for a power generation loss of 3 GW on a 220 GW network load with several values of variance of the $\boldsymbol{\chi^2}$ distribution for the thresholds. Solar penetration is around $\mathbf{10\%}$.

deteriorated frequency measurements, which might practically lead to smaller working intervals. □

C. Experiments with a Controlled Power Output

In this part we implement the control architecture discussed in Section V. In particular, we test the use of the proportional gain in those scenarios from above where the generation loss would result in the activation of the load shedding procedure. We aim at highlighting the potential contribution that common devices can bring to the resilience of the grid to a generation incident, namely the mitigation of load shedding scenarios, by implementing a simple control architecture.

Controllers utilise a frequency deadband of $[49.95, 50.05]$ Hz. We assume that every panel implements the same proportional control design, with $k_p = 4$. P_{eq} is set to 90% of the maximum available power output, as per Fig. 8 in Section V. As an example, in a 220 GW network with 10% solar penetration, the solar power output accounts for a total of 22 GW. P_{eq} is thus 19.8 GW, whereas the available control power for frequency regulation amounts to 2.2 GW. Fig. 14 depicts the frequency response after the infeed loss without any controlled device (solid lines), and with controlled PVs (dashed lines). Solar panels contribute to 10% of the 220 GW network. It can be seen that the controlled signals are able to sustain the network after the incident, and to prevent panels from disconnecting, which is the cause of load shedding. Similarly, Fig. 15 considers a 440 GW network load scenario with the solar penetration set to 20%. Finally, on a a 220 GW network load scenario Fig. 16 shows the frequency signals with uncontrolled and controlled power output, obtained with the highest and lowest values of variance tested. (The uncontrolled signals correspond to those in Fig. 13.) In both cases, the controlled output keeps the frequency very close to the nominal value (within tens of mHz). As evidenced in Fig. 13, no load shedding procedure is activated even in the absence of the control design. Nevertheless, a control architecture enhances the electricity grid, providing an immediate ancillary service and, in this scenario, avoids the disconnection of a portion of panels – the portion of the population with disconnection thresholds at 49.8 Hz and 49.7 Hz. Further, a frequency decrease of hundreds of mHz represents a stressful situation for the grid: with an active control the frequency decrease remains within tens of mHz, thus significantly improving the aftermath of an incident.

As our simulations clearly show, the implementation of a simple proportional control architecture is sufficient to avoid the activation of the load shedding procedure. Notice that the frequency signals carry a steady-state error due to the proportional nature of the control. As discussed in Remark 3, it is expected that secondary control starts up after the considered time horizon, thus it is not explicitly modelled in this study.

Remark 5: Our studies underline the relevance and the vulnerability associated to solar power generation. Solar panels are a useful resource of clean energy and of control power. Our initial modelling framework considers panels only as a ON/OFF devices (from the perspective of their power production) that might turn off under stress conditions in the

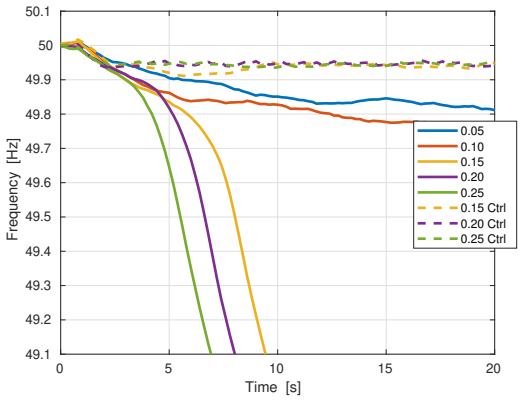


Fig. 14. Comparison between frequency response after a 3 GW loss of production incident, with $\mathbf{10\%}$ of renewable power on a 220 GW network load. The response without any control is depicted with continuous lines while in dashed lines the solar panels are controlled.

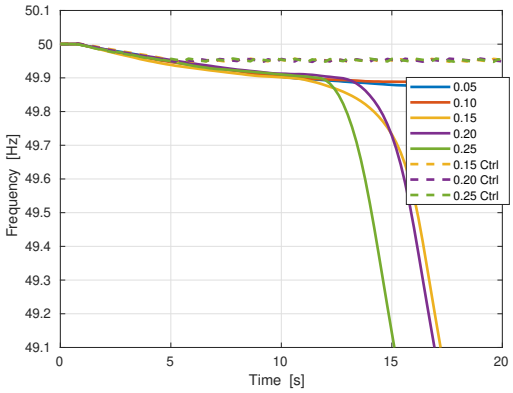


Fig. 15. Comparison between frequency response after a 3 GW loss of production incident, with $\mathbf{20\%}$ of renewable power on a 440 GW network load. The response without any control is depicted with continuous lines while in dashed lines the solar panels are controlled.

electricity grid. If they remain unable to actively participate to the grid regulation, they can have a destabilising impact on the overall network stability. On the other hand, it is easy to understand why panels cannot be the only source of regulation: our control approach is intended to show the potential of such a resource, however this has clear practical shortcomings. First, from the economic point of view, the constant 10% loss of production – by setting P_{eq} to 90% of the available power – is a non-negligible amount over an extended period of time. Further, renewables need a storage system to provide a long-term reliable regulation: weather unpredictability and occlusions are among the factors that need to be accounted for. Also, the proposed control strategy works over a short time interval: the maximum power output of panels changes during daytime. As such, the discussed control approach is useful in extreme situations, e.g. after an incident, rather than in daily regulation scenarios. □

VII. Discussion and Conclusions

In this work we have discussed a modelling framework for the aggregation of a heterogeneous population of solar panels

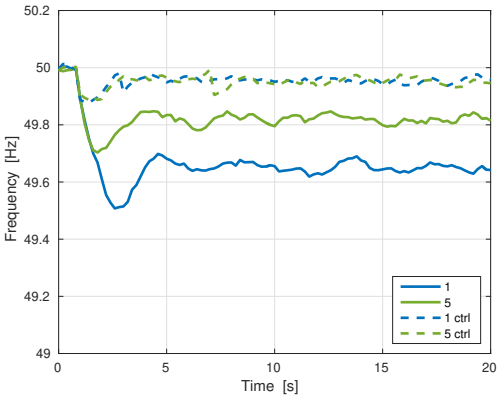


Fig. 16. Comparison between control-less (solid lines) and controlled (dashed) frequency response. The scenario represents a power generation loss of 3 GW on a 220 GW network load with χ^2 distribution for the thresholds. Solar penetration is around 10%.

connected to the electric grid via two dtMCs. The frequency dynamics are simulated with a discrete-time equation encompassing the primary control with modified inertia (function of the renewables penetration). Specifically, a linear dependency between the amount of conventional power and the time to launch of the network is used. As a result, we show that the introduction of solar power increases the oscillatory response of the grid after an incident. Wider overshoots are a dangerous inconvenience, as they can stress the physical network.

Further, we have tested the stability of the network in two scenarios of generation loss with different distributions of population thresholds. Experiments show a correlation between load shedding and the variance of the chosen distributions, or more precisely, to the number of panels that disconnect at values of frequency close to the nominal value. A Gaussian distribution has a symmetric shape around its average value: when its variance increases, the tails on both sides spread out. If frequency thresholds are distributed according to a Gaussian, increasing the variance of the thresholds distribution cause a higher number of panels (represented by the tails of the distribution) to have thresholds closer to f_0. Consequently, we observe more panels with a narrow working interval around the nominal frequency. As a result, the network is more likely to fail, because it becomes more susceptible to small frequency deviation: as highlighted by Tables II and III, this issue is more relevant when a larger population is connected to the grid. In the case of the (uni-directional) χ^2 distribution, increasing their variance leads to larger thresholds. As expected, the experiments show that, in this scenario, an increased heterogeneity guarantees a more reliable network.

Finally, we have introduced a proportional control scheme to control the aggregate power coming from solar panels. Panels individually regulate their power output to restore nominal frequency conditions after an incident. The control framework supports a decentralised implementation. The control design is simple, made to resemble the already existing primary frequency regulation. However, traditional primary network regulation does not act as fast as the panels do, being related to the inertia coefficient. Despite its sheer simplicity, the control implementation presented in this work clearly increases the resilience of the network in case of sudden losses for brief time instants. Simulations have proved that these devices are a useful resource for frequency regulation in extreme situations. We have verified that such a simple control could sustain the network and avoid load shedding.

This regulation approach can be implemented by flexible loads, such as in the case of cooler devices, as outlined in [14]–[16]. Extensions of this work might include modelling storage systems, such as batteries or electric vehicles.

We conclude with a practical recommendation: in order to guarantee the reliability of the electric network in case of incidents, it is important to have populations heterogeneous thresholds and larger working intervals. Household solar devices are widespread in the market and this population rarely participate in the retrofit programs proposed by the European Community, and we imagine them less frequently serviced than industrial installations.

REFERENCES

[1] "The Paris Agreement," https://unfccc.int/process-and-meetings/the-paris-agreement/the-paris-agreement, Accessed: 09/12/2018.
[2] ENTSO-E, "Statistical factsheet," Tech. Rep., 2016.
[3] H. Wirth, "Recent Facts about Photovoltaics in Germany," Fraunhofer ISE, Tech. Rep., 2018.
[4] Solar Power Europe, "Global Market Outlook for Solar Power 2016-2020," Tech. Rep., 2016.
[5] G. Andersson, "Power System Analysis," *ETH Zurich*, 2011.
[6] S. Kouro, J. I. Leon, D. Vinnikov, and L. G. Franquelo, "Grid-connected Photovoltaic Systems: An Overview of Recent Research and Emerging PV Converter Technology," *IEEE Industrial Electronics Magazine*, vol. 9, no. 1, pp. 47–61, 2015.
[7] H. Tiam, F. Mancilla-David, and K. Ellis, "A Detailed Performance Model for Photovoltaic Systems," NREL, Tech. Rep., 2012.
[8] K. Tse, M. Ho, H.-H. Chung, and S. Hui, "A Novel Maximum Power Point Tracker for PV Panels using Switching Frequency Modulation," *IEEE transactions on power electronics*, vol. 17, no. 6, pp. 980–989, 2002.
[9] S. Rehman, M. A. Bader, and S. A. Al-Moallem, "Cost of Solar Energy Generated using PV Panels," *Renewable and Sustainable Energy Reviews*, vol. 11, no. 8, pp. 1843–1857, Mar 2007.
[10] Y. Ghiassi-Farrokhfal, F. Kazhamiaka, C. Rosenberg, and S. Keshav, "Optimal Design of Solar PV Farms with Storage," *IEEE Transactions on Sustainable Energy*, vol. 6, no. 4, pp. 1586–1593, Oct 2015.
[11] ENTSO-E, "Dispersed Generation Impact on CE Region, Dynamic study," Tech. Rep., 2014.
[12] A. Peruffo, E. Guiu, P. Panciatici, and A. Abate, "Aggregated Markov Models of a Heterogeneous Population of Photovoltaic Panels," *International Conference on Quantitative Evaluation of Systems*, pp. 72–87, 2017.
[13] ——, "Synchronous Frequency Grid Dynamics in the Presence of a Large-Scale Population of Photovoltaic Panels," in *2018 Power Systems Computation Conference (PSCC)*, June 2018, pp. 1–7.
[14] ——, "Impact of Solar Panels and Cooling Devices on Frequency Control after a Generation Loss Incident," in *Decision and Control (CDC) 2018 IEEE 57th Annual Conference Proceedings*. IEEE, 2018.
[15] S. H. Tindemans, V. Trovato, and G. Strbac, "Frequency Control Using Thermal Loads Under the Proposed ENTSO-E Demand Connection Code," in *2015 IEEE Eindhoven PowerTech*, June 2015, pp. 1–6.
[16] ——, "Decentralized Control of Thermostatic Loads for Flexible Demand Response," *IEEE Transactions on Control Systems Technology*, vol. 23, no. 5, pp. 1685–1700, Sept 2015.
[17] ENTSO-E, "Policy 1: Load-frequency Control and Performance," Tech. Rep., 2009.
[18] B. L. M. Jung, O. Wiss, "Analyses et conclusions – Tests en sous-frequence," RTE, Tech. Rep., 2016.
[19] ——, "Analyses et conclusions – Tests en sur-frequence," RTE, Tech. Rep., 2016.

[20] J. G. Kassakian, R. Schmalensee, G. Desgroseilliers, T. D. Heidel, K. Afridi, A. Farid, J. Grochow, W. Hogan, H. Jacoby, and J. Kirtley, "The Future of the Electric Grid," *Massachusetts Institute of Technology, Tech. Rep*, pp. 197–234, 2011.
[21] D. Jones, "Dynamic System Parameters for the National Grid," *IEEE Proceedings - Generation, Transmission and Distribution*, vol. 152, no. 1, pp. 53–60, Jan 2005.
[22] P. Kundur, N. J. Balu, and M. G. Lauby, *Power System Stability and Control*. McGraw-hill New York, 1994, vol. 7.
[23] J. Morren, S. W. De Haan, W. L. Kling, and J. Ferreira, "Wind Turbines Emulating Inertia and Supporting Primary Frequency Control," *IEEE Transactions on Power Systems*, vol. 21, no. 1, pp. 433–434, Feb 2006.
[24] Tractebel Engineering S.A., "Eurostag," http://www.eurostag.be, 1997–2018.
[25] R. C. Dorf and R. H. Bishop, "Modern Control Systems," 2011.
[26] K. Ogata *et al.*, *Discrete-time Control Systems*. Prentice Hall Englewood Cliffs, NJ, 1995, vol. 2.
[27] B. Bendib, H. Belmili, and F. Krim, "A survey of the most used MPPT methods: Conventional and Advanced Algorithms Applied for Photovoltaic Systems," *Renewable and Sustainable Energy Reviews*, vol. 45, pp. 637 – 648, May 2015.
[28] M. A. G. De Brito, L. Galotto, L. P. Sampaio, G. d. A. e Melo, and C. A. Canesin, "Evaluation of the Main MPPT Techniques for Photovoltaic Applications," *IEEE transactions on industrial electronics*, vol. 60, no. 3, pp. 1156–1167, 2013.
[29] European Commission, "Commission regulation (EU) 2017/1485 of 2 august 2017," Tech. Rep., 2017.
[30] ENTSO-E, "Network Code on Operational Security," Tech. Rep., 2013.
[31] ——, "Dispersed Generation Impact on CE Region, Dynamic study," Tech. Rep., 2014.
[32] "National Grid Status," https://www.gridwatch.templar.co.uk/, Accessed: 10/12/2018.
[33] European Commission, "Commission regulation (EU) 2016/631 of 14th april 2016," Tech. Rep., 2016.

APPENDIX I
OBSERVABILITY AND OBSERVER DESIGN

Whilst it is in principle possible to have real-time information about every photovoltaic panel in the network, this would be highly impractical. We thus consider a more realistic scenario, where the only accessible (i.e., observable) output is the frequency of the network. Note that if the value of $f(k)$ is known, also $a(k)$ and $b(k)$ can be computed, as we assume to know the population distributions.

Observability of the Models

Let us show that both the 3-state Markov Chain model and the n-state one (cf. Equations (8), (9)) are observable when the only output signal is $f(k)$. This means, in practical terms, that we can derive the value of the other variables of the system, reconstructing them solely from the value of $f(k)$.

Let us focus on a simplified version of the three-state Markov model, namely

$$\begin{cases} f(k+1) = \alpha_1 f(k) + \beta_1 x(k) \\ x(k+1) = (1-a(k))x(k) + b(k)\varepsilon(k)y(k) \\ y(k+1) = b(k)(1 - x(k) - \varepsilon(k)y(k)) \end{cases}.$$

This version, featuring first-order frequency dynamics, is simpler to analyse whilst retaining the significant features of the model in Equation (9). Observability analysis with the second-order frequency dynamics can be attained following a similar procedure. Limiting the access to the frequency, the output matrix is the constant quantity

$$C(k) \equiv C = [1\ 0\ 0],$$

and let us define $A(k)$, the Jacobian matrix of the linearised system at time k

$$A(k) = \begin{bmatrix} \alpha_1 & \beta_1 & 0 \\ 0 & \bar{a}(k) & b(k)\varepsilon(k) \\ 0 & -b(k) & -b(k)\varepsilon(k) \end{bmatrix},$$

where $\bar{a}(k) = 1 - a(k)$. The observability matrix results in

$$\mathcal{O}(k) = \begin{bmatrix} 1 & 0 & 0 \\ \alpha_1 & \beta_1 & 0 \\ \alpha_1^2 & \alpha_1\beta_1 + \beta_1\bar{a}(k+2) & \beta_1 b(k+2)\varepsilon(k+2) \end{bmatrix}.$$

The rank is full as long as $b(k)\varepsilon(k) \neq 0$, namely when either $b(k) \neq 0$ and $\varepsilon(k) \neq 0$. Let us now analyse what the conditions $b(k) = 0$ and $\varepsilon(k) = 0$ elicit.

1) $\varepsilon(k) = 0$ – In this case the system reduces to

$$\begin{cases} f(k+1) = \alpha_1 f(k) + \beta_1 x(k) \\ x(k+1) = \bar{a}(k)x(k) \\ y(k+1) = b(k)(1 - x(k)). \end{cases}$$

In this case $y(k)$ is clearly not observable, as $x(k+1)$ loses its dependency on $y(k)$ and at the same time $y(k)$ has no effect on the output. However, if we restrict our attention to the reduced system composed of the first two equations, this is observable: the observability matrix becomes

$$\mathcal{O}(k) = \begin{bmatrix} 1 & 0 \\ \alpha_1 & \beta_1 \end{bmatrix},$$

which is always full rank. Note that $y(k)$ depends only on $x(k)$: in the next section we will be able to compute its estimated value $\hat{y}(k)$ from $\hat{x}(k)$, the estimated value of $x(k)$.

2) $b(k) = 0$ – Similarly, the system reduces to

$$\begin{cases} f(k+1) = \alpha_1 f(k) + \beta_1 x(k) \\ x(k+1) = \bar{a}(k)x(k) \\ y(k+1) = 0 \end{cases},$$

and analogous conclusions can be drawn.

Regarding the observability of the (n+2)-state model, the analysis results analogous. Equation (8) suggests a chain-dependency of $w_i(k)$ from $w_{i-1}(k-1)$, which keeps the rank of the observability matrix full. This leads to a chain of substitutions that guarantees the observability in n steps, under the condition $b(k)\varepsilon(k) \neq 0$. If $b(k)\ \varepsilon(k) = 0$, then we can repeat a similar argument. Finally, the analysis is easily generalisable to the second-order frequency dynamics.

Observer Design

An observer can be built for both models in Equations (8) and (9). We start with the observer for the smaller model. Let us define $\hat{x}(k)$ and $\hat{y}(k)$ as the estimated values of $x(k)$ and of $y(k)$, respectively. Assume the unique output is the frequency of the network, namely $\hat{f}(k) = f(k)$. Our aim is the computation of $\hat{x}(k)$ and $\hat{y}(k)$. As above, considering the three-state model, we assume $\varepsilon(k)$ known as we carry on the analysis. Later, for the (n+2)-state model estimator computation, we build $\hat{\varepsilon}(k)$, the estimation of $\varepsilon(k)$.

Let us now focus on the three-state model and show how to build $\hat{x}(k)$ and $\hat{y}(k)$. After some algebra, from Equation (9) we obtain:

$$\begin{bmatrix} f(k-2) \\ \hat{x}(k-2) \\ \hat{y}(k-2) \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1/\beta_1 & -\alpha_1/\beta_1 \\ l_1 & l_2 & l_3 \end{bmatrix} \cdot \begin{bmatrix} f(k) \\ f(k-1) \\ f(k-2) \end{bmatrix},$$

where $[l_1,\ l_2,\ l_3]$ is equal to

$$\left[\frac{1}{\beta_1 b(k-2)\varepsilon(k-2)},\ -\frac{\bar{a}(k-2)+\alpha_1}{\beta_1},\ -\frac{\alpha_1\bar{a}(k-2)}{\beta_1}\right].$$

Our estimation is two steps behind the current time instant. We then set up a two-steps predictor following the dynamical equations of the system, after some algebra, as

$$\begin{bmatrix} \hat{f}(k) \\ \hat{x}(k) \\ \hat{y}(k) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ obs_{2,1} & obs_{2,2} & obs_{2,3} \\ obs_{3,1} & obs_{3,2} & obs_{3,3} \end{bmatrix} \begin{bmatrix} f(k) \\ f(k-1) \\ f(k-2) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ b(k-1) \end{bmatrix},$$

where

$$\begin{array}{lll} obs_{2,1} = \frac{\bar{a}(k-1)}{\beta_1}, & obs_{2,2} = -\frac{\bar{a}(k-1)\alpha_1}{\beta_1}, & obs_{2,3} = 0, \\ obs_{3,1} = -\frac{1+c_1}{\beta_1} - c_1, & obs_{3,2} = -\frac{\alpha_1}{\beta_1}(1+2c_1), & obs_{3,3} = 0, \end{array}$$

and where

$$c_1 = \frac{\varepsilon(k)\bar{a}(k-1)}{b(k-1)\varepsilon(k-1)}.$$

Regarding the (n+2)-state model, a similar situation is expected. In fact, for the dynamical equation we need an $(n+2)$-step predictor. The value of $\hat{x}(k)$ can be attained using two temporal values of $f(\cdot)$, as it is easy to note from Equation (8). The key part is how to connect the estimation of $x(k)$ to $w_1(k)$. We note that, again from Equation (8):

$$x(k+1) + \sum_i w_i(k+1) = \bar{a}(k)x(k) + b(k)(1-x(k)),$$

so that

$$\sum_i w_i(k) = \bar{a}(k-1)x(k-1) + \\ + b(k-1)(1-x(k-1)) - x(k).$$

We can substitute this into the estimator of $w_1(k)$, giving

$$\hat{w}_1(k) = b(k-1)[1-b(k-1)(1-\hat{x}(k-1)) - \bar{a}(k-1)\hat{x}(k-1)].$$

The observer and predictor can be built as follows:

$$\begin{aligned} &\hat{f}(k) = f(k), \\ &\hat{x}(k-i) = \frac{f(k-i+1) - \alpha_1 f(k-i)}{\beta_1},\ i = 1, 2, \\ &\hat{x}(k)|\hat{x}(k-1) = \bar{a}(k-1)\hat{x}(k-1) + b(k-1)\sum_i \tau_i \hat{w}_i(k-1), \\ &\hat{w}_1(k) = b(k)[1 - b(k-2)(1-\hat{x}(k-2)) + \\ &\qquad - \bar{a}(k-2)\hat{x}(k-2)], \\ &\hat{w}_i(k) = b(k-1)(1-\tau_{i-1})\hat{w}_{i-1}(k-1), \qquad i = 2, \dots n-1 \\ &\hat{w}_n(k) = b(k-1)[(1-\tau_{n-1})\hat{w}_{n-1}(k-1) + \\ &\qquad + (1-\tau_n)\hat{w}_n(k-1)]. \end{aligned}$$

With this technique we prove that we are able to build an observer for the system, estimate the $w_i(k)$ and compute the $\hat{\varepsilon}(k)$ value at each time step as

$$\hat{\varepsilon}(k) = \frac{\sum_i \tau_i \hat{w}_i(k)}{\sum_i \hat{w}_i(k)}.$$

Note that the observer has a transient of $(n+2)$ steps, which is necessary to initially compute $\hat{w}_i,\ \forall i$. This value is then used in the estimation and prediction of the three-state model.

APPENDIX II
STABILITY ANALYSIS OF THE CLOSED-LOOP MODEL

We illustrate the stability analysis for the closed-loop models in Equations (8) and (9). In order to ease the notation and to make the discussion clearer, we utilise a reduced first-order transfer function, as done already in Appendix I. The analysis with the second-order transfer function introduced in the paper can be carried out similarly.

We start with the closed-loop three-state model: the analysis of the $(n+2)$-state model is analogous to what we present in the following. Whilst for simplicity we manipulate quantities $a(k)$ and $b(k)$, recall that they are functions of the frequency $f(k)$, namely we have $a(f(k))$ and $b(f(k))$. More specifically, they are cumulative distribution functions of the panels thresholds, where $f(k)$ enters as an extremum of the integral.

Let us re-write the second equation of Equation (9): we need the dependency on $\Delta x(k)$ to be explicit. Recall from Section IV that $\Delta x(k) = x(k) - x_0$, where x_0 is the portion of panels in state ON when $f(\cdot) = f_0$. In our setting, we assume $x_0 = 1$ (all panels ON), i.e. all panels are active when the frequency is at its nominal value. Notice that in this case the equilibrium point of this system is the origin, however if $x(0) \neq 0$, the following analysis must be modified to accommodate for this.

Substituting $x(k) = \Delta x(k) + x(0)$, we obtain an alternative formulation for the three-state model as

$$\begin{cases} \Delta f(k+1) = \alpha_1 \Delta f(k) + \beta_1 \Delta x(k) \\ \Delta x(k+1) = \bar{a}(k)\Delta x(k) + b(k)\varepsilon(k)y(k) + a(k)x(0) \\ y(k+1) = -b\Delta x(k) - b(k)\varepsilon(k)y(k) + b(k)(1-x(0)). \end{cases}$$

We compute the Jacobian as in the previous Appendix and evaluate it at the equilibrium point: this holds an eigenvalue in 1: the eigenvalue analysis is not sufficient in this case (the stability depends on the non-linear components of the vector field), and analysis via Lyapunov function is thus necessary to assess the stability of the equilibrium.

Recall that at $f(\cdot) = f_0$ the functions $\bar{a}(k) = 1$ $b(k) = 1$, $a(k) = 0$, $x(0) = 1$. Therefore we have $\varepsilon(k) = 0$, $y(k) = 0$, $\Delta x(k) = 0$, $\Delta f(k) = 0$.

Let us define the vector $\xi(k) = [\Delta f(k), \Delta x(k), y(k)]$. Notice that, thanks to the stability of the first-order transfer function, we ensure the stability of $f(k)$ when $\Delta x(k) \to 0$. We then select the following Lyapunov function:

$$V(\xi(k)) = (\Delta x(k) + y(k))^2,$$

which is positive everywhere but at the origin. In order to assess the stability of the interconnection, we compute

$$\begin{aligned} V(\xi(k+1)) - V(\xi(k)) &= [(\bar{a}(k) - b(k))^2 - 1]\Delta x(k) + \\ &+ 2[a(k)(\bar{a}(k) - b(k)) - y(k)]\Delta x(k) + a(k)^2 - y(k)^2. \end{aligned}$$

From the definitions of $a(\cdot)$ and $b(\cdot)$, we have

$$a(f_0) = 0, \quad \bar{a}(f_0) = 1, \quad b(f_0) = 1. \tag{13}$$

Substituting these values we obtain:

$$V(\xi(k+1)) - V(\xi(k)) = -(\Delta x(k) + y)^2 \leq 0.$$

The negativity of the Lyapunov function is guaranteed as long as the conditions in (13) hold. We argue that these conditions also hold in a non-trivial neighbourhood of f_0 (call it I_{f_0}): in practice, panels disconnect and reconnect at values $f_0 \pm \delta$, $\delta > 0$, which means that the network is stable as long as no panel disconnects (given $x_0 = 1$). As soon as $f(\cdot)$ exits I_{f_0}, the stability is not guaranteed. The transition from stability to instability depends on the threshold distributions for the population of PV panels, and in particular on their averages and variances. The stability of the interconnection depends heavily on the interval I_{f_0}: the larger is the interval, the more resilient is the system to oscillations. Let us focus on $a(k)$ for simplicity. Let us assume a distribution $a(k)$ that is the integral of truncated normal distribution. Let us focus on the condition $f(k) > f_0$ (the $f(k) \leq f_0$ side is handled symmetrically). Define $\lambda_o = [\lambda_1, \lambda_2]$ to be the support of $p^d_{t,o}(\cdot)$, and suppose that λ_0 supports a symmetric truncated normal $p^d_{t,o}$ (i.e. its average value is $\mu = (\lambda_1 + \lambda_2)/2$) for any choice of λ_1 and of λ_2. By definition of the truncated normal distribution $p^d_{t,o}$, the quantity $a(k)$ becomes

$$a(k) = \begin{cases} 0 & \text{if } f_0 \leq f(k) < \lambda_1 \\ \int_{\lambda_1}^{f(k)} p^d_{t,o}(u)du & \text{if } \lambda_1 \leq f(k) \leq \lambda_2 \\ 1 & \text{otherwise.} \end{cases}$$

In the case under consideration, the upper limit of I_{f_0} is λ_1. As such, the length of I_{f_0} increases as the support of $p^d_{t,o}$ is far from f_0, namely as λ_1 increases. With an increase of I_{f_0}, the stability of the interconnection is preserved for a larger set of network frequency values. On the other hand, also λ_2 plays an important role in the stability analysis. Increasing λ_2 enlarges the support of $p^{(\cdot)}_{t,o}$. This reduces the rate of increase of $a(k)$ when $f(k)$ increases: the distribution $p^{(\cdot)}_{t,o}$, over a larger support, has lower values. Note that the sign of $V(\xi(k-1)) - V(\xi(k))$ outside I_{f_0} depends heavily on the slope of $a(k)$.

Let us define F, a value of frequency so that $|F - f_0| >> 0$. When $f(\cdot) = F$, then $a(F) = 1$, $\bar{a}(F) = 0$, $b(F) = 0$. Let us analyse the Lyapunov function under these conditions:

$$V(\xi(k+1)) - V(\xi(k)) = -(\Delta x(k) + y(k))^2 + a(F)^2,$$

which is positive as long as $(\Delta x(k) + y(k))^2 < a(F)^2$. Note that $\Delta x(\cdot), y(\cdot) \in [0, 1]$, and are such that $\Delta x(k) + y(k) \leq 1$ $\forall k$, by definition. The slope of $a(\cdot)$ here plays a major role: if the growth rate of $a(k)$ is faster than the one of $(\Delta x(k) + y(k))$, the system becomes unstable.

Let us finally study the system for values of $f(k)$ so that $a(f),\ \bar{a}(f),\ b(f) \in (0, 1)$. The Lyapunov function becomes

$$\begin{aligned} &V(\xi(k+1)) - V(\xi(k)) = \\ &= -(\Delta x(k) + y(k))^2 + (a(k) + d(k)\Delta x(k))^2, \end{aligned}$$

where $d(k) = \bar{a}(k) - b(k)$. The analysis is similar to the one above: the Lyapunov function can become positive if $a(\cdot)$ and $d(\cdot)$ grow faster than $\Delta x(k)$ and $y(k)$.

In summary, under all the three conditions analysed, the increase of $a(\cdot)$ and $b(\cdot)$ is dictated by the variance of the disconnection and reconnection distributions, respectively. A small value of the variance renders the system less robust and more prone to instability. This situation occurs with a population having individual $\mathcal{I}_f$ with small variations. A higher value of variance corresponds to a higher degree of population heterogeneity, which we have shown to increase the stability region of the Lyapunov function. Heterogeneity, in this sense, can be exploited to enhance the network reliability and resilience against oscillations and incidents. On the other hand, as the analysis on the first condition suggests, instability can be mitigated by moving the average value away from f_0.

This stability analysis can be generalised to the second-order frequency dynamics using the same Lyapunov function. We consider the following for the (n+2)-state model

$$V(\xi(k)) = \left(\Delta x(k) + \sum_{i=1}^{n} w_i(k) \right)^2.$$

andrea.png

Andrea Peruffo (S'18) is a DPhil student in the Department of Computer Science at the University of Oxford. He received a Laurea in Information Engineering in October 2012 and an MS in July 2015 in Automation Engineering, both from from the University of Padova (IT). He worked as a research engineer at Inria Rennes in 2016 for the Lagadic group and visited the Erato MMSD Group at NII, Tokyo, in 2020. His research interests include hybrid systems, formal verification and control theory.

Emeline Guiu graduated from Ecole Centrale de Nantes, PhD degree with ONERA (France) in 2006. She joined RTE in 2006, She has a position of research engineer.

Patrick Panciatici (F'18) graduated from Supelec, joined EDF R&D in 1985 then he joined RTE (French Transmission System Operator) in 2003 and participated in the creation of the internal R&D department of RTE. He has more than 30 years' experience in the field of R&D for transmission systems. Presently, as a senior scientific advisor, he inspires, coordinates and supervises long term research activities in RTE. He is a member of CIGRE, a Fellow of IEEE and an Emerite member of SEE.

Alessandro Abate (S'02-M'08-SM'19) is Professor of Verification and Control in the Department of Computer Science at the University of Oxford, and a fellow of the Alan Turing Institute for Data Sciences in London. He received a Laurea in Electrical Engineering in October 2002 from the University of Padova (IT), an MS in May 2004 and a PhD in December 2007, both in Electrical Engineering and Computer Sciences, at UC Berkeley (USA). He has been an International Fellow in the CS Lab at SRI International in Menlo Park (USA), and a PostDoctoral Researcher at Stanford University (USA), in the Department of Aeronautics and Astronautics. He has also been an Assistant Professor at the Delft Centre for Systems and Control, TU Delft (NL).