Cyber-Attacks in the Automatic Generation Control

Maria Vrakopoulou, Peyman Mohajerin Esfahani, Kostas Margellos, John Lygeros and Göran Andersson

Abstract Power systems are traditionally monitored and controlled by an IT infrastructure, referred to as Supervisory Control and Data Acquisition (SCADA) system. The cyber-physical interaction of power systems (physical) and SCADA systems (cyber) rises security issues, since the links between those systems are vulnerable to cyber-attacks that can potentially lead to catastrophic economical and societal effects. In this chapter we focus on a specific cyber-physical link, the Automatic Generation Control (AGC), which is an automatic frequency control loop closed over the SCADA system. We provide an impact analysis in case of a cyber-attack on the AGC signal. We first carry out a feasibility analysis based on reachability and optimal control theory, that provides an information regarding the existence of an attack pattern that can disturb the power system. We then deal with the problem of synthesizing an attack signal and treat it as a nonlinear control synthesis problem. Third, performance of our methodologies are illustrated by means of dynamic simulations on IEEE-118 bus network.

1 Introduction

A well-functioning society relies heavily on the proper operation of the electric power system. Large power outages may be difficult and time- consuming to restore

Maria Vrakopoulou and Göran Andersson
Power System Laboratory, ETH Zurich, Physikstrasse 3, Zurich 8092, Switzerland, e-mail: vrakopoulou,andersson@eeh.ee.ethz.ch

Peyman Mohajerin Esfahani and John Lygeros
Automatic Control Laboratory, ETH Zurich, Physikstrasse 3, Zurich 8092 , Switzerland, e-mail: mohajerin,lygeros@control.ee.ethz.ch

Kostas Margellos
Department of Industrial Engineering and Operations Research, UC Berkeley, Hearst Avenue 2594, Berkeley CA 94720, US, e-mail: kostas.margellos@berkeley.edu

and may also have devastating economic and humanitarian consequences. The importance of electric power delivery is illustrated, for instance, by the economic and social impacts of the 2003 northeast American blackout during which 50 million people were affected [1]. Therefore, in large electric power systems, an Information Technology (IT) infrastructure, referred to as Supervisory Control and Data Acquisition (SCADA) system, provides system-wide supervision and control [2]. The SCADA system measures data through remote devices installed throughout the grid and gathers the information at a control center through communication channels, where from, after computer processing, control commands are sent back to the power system. The dependence of the power system (physical) on the IT infrastructure (cyber) constitutes a cyber-physical interaction that despite the fact that it is designed to lead to a more efficient operation of the power system, it renders it more susceptible to operational errors and external attacks.

The power system is typically divided in control areas, each of them monitored and controlled by a separate SCADA system. After gathering the measurements in the control center, state estimation is conducted so as to determine the most probable state of the system given that the measurements might be inaccurate or incomplete. Based on the estimated state, the SCADA system alerts the operator if control actions should be taken. The various power system points that are controlled by the SCADA system are the status of switches, generator voltage setpoints, generator active power setpoints, turns ratio of load tap changing transformers and other configuration settings. These control actions aim on a more efficient and secure operation of the power system and are typically manually driven. One of the few control loops that are closed over the SCADA system without human operator intervention is the Automatic Generation Control (AGC). This is a continuous[1] time control and involves the adjustment of the generator active power setpoints. All the aforementioned inputs and outputs of the control center (or the power system, respectively) constitute vulnerable points of the cyber-physical system that could be possibly manipulated as part of a cyber-attack to deteriorate the performance of the system.

Numerous analyses have described the systems potential vulnerabilities to cyber-attacks, while actual incidents have confirmed these vulnerabilities and underscored the importance of reducing them. The authors of [3] proposed a framework in order to clarify the interaction between the power system and the IT infrastructure and identify the vulnerabilities and the malfunctions of both that could lead to an abnormal operation of the power network. In [4–6], the vulnerabilities of a cyber-attack on the state estimation system were assessed. From another perspective the authors of [7] attempted to quantify the impact of a cyber-attack in a power market environment, while in [8–11] real examples of cyber-attacks were reported.

In this context, the VIKING research project [12] proposed a novel concept to address the challenges introduced by the interaction between the SCADA system and the power transmission and distribution systems. Tools and methodologies were developed to identify the vulnerabilities of these safety critical infrastructures [13],

[1] Practically is applied every 2-4 seconds.

to determine the impact that possible failures or attacks might have [14, 15] and to develop strategies to mitigate these effects [16].

Motivated by the research carried out within the VIKING project, in this chapter we enrich our earlier work [17, 18] with additional analysis tools and more informative case studies to investigate the impact of a cyber-attack on the AGC signal in a power system. This is a communication signal coming out of the SCADA system, serving as a command to the generators participating in the AGC loop to adjust their power set-point. Investigating how the attacker can gain access to this signal and specifying the exact intrusion path is outside the scope of this paper. For more details towards this direction we refer to [19], [13].

We focus on the AGC control loop, since its automatic nature renders it more susceptible to external attacks. AGC actions are usually determined for each control area at the control center. The main objective is to regulate frequency to its nominal value and maintain the power exchange between the control areas at the scheduled level. To achieve this, measurements of the system frequency and the tie line power flows are sent to the dispatch center and then a feedback signal that regulates the generated power is sent back to the generators, participating in the AGC, through the SCADA system.

In this chapter we assume that an attacker has gained access in one of the areas of the power system. We first provide a feasibility analysis and show whether there exists an attack signal that could irreversibly disturb the system. Our methodology employs tools from reachability theory and optimal control for nonlinear systems [20,21]. We next focus on the problem of synthesizing an attack signal; we treat it as a controller synthesis problem where the objective is to drive the system outside the safety margins. Different alternatives are provided ranging from open loop approaches, based on Markov Chain Monte Carlo (MCMC) optimization [22, 23], to closed loop schemes based on feedback linearization and gain scheduling [24, 25]. Due to the complexity (large-scale, nonlinear) of the models that describe power systems, for our analysis and synthesis investigation we use a two-machine frequency model where each machine represents a different control area. To evaluate the performance of our methodology, we apply the attack signal that is constructed based on the aforementioned abstraction to the detailed power system model. It should be noted that the proposed techniques are applied to nonlinear models with quite general structure and as such they have the potential to be applied to other cyber-physical problems and are not limited to the case of the AGC attack.

In Section 2 the physical description and the mathematical model of the two-machine power system is provided. Section 3 provides a feasibility analysis that provides intuition regarding how plausible it is for the system to be disturbed by an attack signal. Section 4 provides different attack signal synthesis alternatives and illustrates their efficacy on a detailed simulation environment. Finally, Section 5 provides some concluding remarks.

2 Power system modeling

The fact that power systems are generally exposed to disturbances originating from the uncertainty and variability of the loads, unpredictable line outages etc., has deemed necessary the integration of many control systems. These control systems aim to keep the power system within acceptable operating limits maintaining the security and the quality of supply in satisfactory levels. Due to the various time constants of the processes, the system is controlled in an hierarchical way. Some quantities are rapidly controlled locally and other, with a relatively slower response, via the SCADA system. The nonlinear nature of the power flow equations and the various control schemes that power systems are equipped with result in a very complex model characterized by large-scale nonlinear continuous and discrete dynamics. Such complex models cannot be used efficiently in the development of advanced control strategies thus we have to rely on different levels of abstraction that simplify certain elements of the initial model and/or take advantage of the possible de-coupling between control loops.

One example of a possible control loop de-coupling involves the two main control loops of power system. These are responsible for the regulation of the voltage magnitudes and the frequency of the system so as not to exceed certain specified limits. The time constants of the local Automatic Voltage Regulator (AVR) are quite smaller than the ones of the frequency control loops and hence in load-frequency studies one can use a quasi-state model that considers only the steady state point of the voltage control loop ignoring its fast dynamics.

In this chapter we investigate the impact of a cyber-attack on the AGC in one control area. To facilitate the needs of this study, we divide a network into two independent control areas, and consider the case where an attacker has gained access to the AGC signal of one of them being able to inject an undesirable input. Since we are dealing with load-frequency studies, we consider only frequency dynamics. For this purpose, a simplified nonlinear frequency model that includes governor and AGC dynamics is developed. We represent each control area by a single generator to apply and illustrate better the control design that will be presented in Section 4.

In the following subsections, we first describe the basic principles of the model abstraction we employ and the frequency dynamics and then present the two-area power system model.

2.1 Frequency dynamics in one Area

In this subsection the frequency dynamics of a single control area as driven by different frequency control levels are described. The analysis is mainly based on [26, 27]. As mentioned above, to simplify the dynamics of each area, we condense them into one single generating unit by considering aggregated quantities based on the center of inertia of the area. For that purpose we consider the following lumped quantities:

$$
\begin{aligned}
f &= \frac{\sum H_i f_i}{\sum H_i} && \text{Centre of inertia frequency}(Hz),\\
S_B &= \sum S_{B_i} && \text{Total rating (MVA)},\\
H &= \frac{\sum H_i S_{B_i}}{\sum S_{B_i}} && \text{Total inertia constant (sec)},\\
P_m &= \sum P_{m_i} && \text{Total mechanical power (MW)},\\
P_e &= \sum P_{e_i} && \text{Total electrical power (MW)},\\
\frac{1}{S} &= \sum \frac{1}{S_i} && \text{Equivalent droop constant (MW/Hz)},\\
\frac{1}{D_l} &= \sum \frac{1}{D_{l_i}} && \text{Equivalent damping coefficient (MW/Hz)},
\end{aligned}
$$

where $i \in G$ and G is the set of the generators that belong to control area i.

The aggregated principal frequency dynamics of each area can be described by

$$\Delta \dot{f} = \frac{f_0}{2HS_B}(\Delta P_m - \Delta P_e), \tag{1}$$

where operator Δ returns the deviation of its arguments from their reference values. The frequency of the rotor is denoted by f whereas P_m and P_e represent the generated (mechanical) power and consumed (electrical) power, respectively. A brief description of additional dynamics due to the term of the generated power (ΔP_m) and the consumed power (ΔP_e) follows.

2.1.1 Generated power

The frequency of the system can be controlled adjusting properly the generated power. Every change in the setpoint of the generated power is first filtered by the turbine dynamics where it is converted in mechanical power. For simplicity we ignore here the turbine dynamics and hence for the setpoint of the generated power we can refer directly to ΔP_m. This setpoint depends on the output of the frequency control loops and on manual interventions. This can be expressed by

$$\Delta P_m = \Delta P_{m,p} + \Delta P_{m,AGC} + \Delta P_{m,set}, \tag{2}$$

where $\Delta P_{m,p}$ represents the change in the produced power due to the primary frequency control (governor) action, $\Delta P_{m,AGC}$ the change due to the secondary frequency control (AGC) action and $\Delta P_{m,set}$ a scheduled step change. In the following, modeling details for the primary frequency control and the AGC loop are presented.

- **Primary frequency control**

Primary frequency control refers to control actions that are done locally at every plant. The governor adjusts the setpoint of the produced power to bring the frequency close to its nominal value. The response should be in a scale of a couple of seconds. According to a simplified model of a governor with speed droop characteristic the rotor measured frequency is compared with the nominal one and the error signal is amplified to produce the control signal ΔP_p. Specifically the control law is given by $\Delta P_p = -\frac{1}{S}\Delta f$, where the proportional gain $1/S$ is referred to as droop or speed regulator.

However, every generator is set to have a specific reserve amount of power that is able to offer according to its availability, the optimal performance of the system, but also market rules. Hence, there are upper and down limits at the produced power deviation that primary and secondary frequency controllers can impose.

Therefore the final change of the produced power due to the primary control action is

$$\Delta P_{m,p} = \begin{cases} \Delta P_p^{min} \text{ if} & \Delta P_p \leq \Delta P_p^{min}, \\ \Delta P_p \text{ if} & \Delta P_p^{min} < \Delta P_p < \Delta P_p^{max}, \\ \Delta P_p^{max} \text{ if} & \Delta P_p \geq \Delta P_p^{max}. \end{cases} \tag{3}$$

- **Automatic Generation Control**

As already discussed, the main objectives of the AGC are to regulate frequency to the specified nominal value and maintain the interchanged power between the controlled areas to the scheduled values by adjusting the generated power of specific generators in the area. It consist the secondary frequency control loop acting in a scale of a couple of minutes.

AGC actions are usually determined for each control area at the control center via the SCADA system. Measured system frequency and tie line flows are sent to this center, where computer processing takes place, and finally a feedback signal that regulates the generated power is sent back to the generators.

However, as mentioned above, there is saturation at the imposed control signal. Therefore the final change of the produced power due to AGC signal in area i is

$$\Delta P_{m,AGC_i} = \begin{cases} \Delta P_{AGC_i}^{min} \text{ if} & \Delta P_{AGC_i} \leq \Delta P_{AGC_i}^{min}, \\ \Delta P_{AGC_i} \text{ if} & \Delta P_{AGC_i}^{min} < \Delta P_{AGC_i} < \Delta P_{AGC_i}^{max}, \\ \Delta P_{AGC_i}^{max} \text{ if} & \Delta P_{AGC_i} \geq \Delta P_{AGC_i}^{max}, \end{cases} \tag{4}$$

where ΔP_{AGC_i} is the AGC control signal before saturation, and $\Delta P_{m,AGC_i}$ is the AGC control signal that finally affects the produced power.

The secondary control of area i is typically a proportional-integral (PI) controller. To avoid wind up in case of saturation, an anti-wind up circuit is also used [28]. The overall block diagram for the AGC of a single area is shown in Fig. 1.

The error signal Δe_i, considering an interconnected system of N-areas each of them equipped with its own AGC controller, is:

$$\Delta e_i = \sum_{j \in \Omega_i} \Delta P_{ij} + B_i \Delta f_i, \tag{5}$$

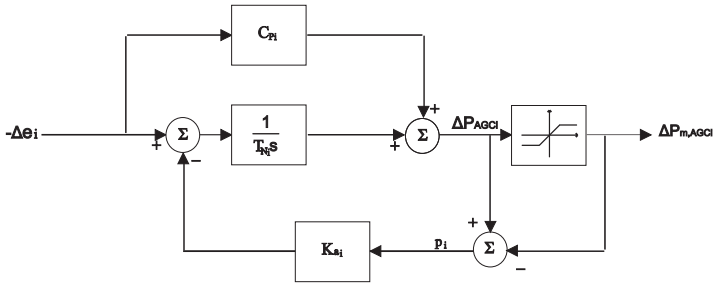

Fig. 1 PI controller with anti-wind up for the AGC loop.

where $\Delta f_i = f_i - f_0$, $\Delta P_{ij} = P_{ij} - P_{0_{ij}}$. Quantity P_{ij} is the power transmitted from area i to area j, $P_{0_{ij}}$ the scheduled transmitted power from area i to area j, Ω_i the set of indices corresponding to the areas connected to area i, f_i the frequency of area i and f_0 the nominal frequency of the system (same for all areas in steady state). Parameter B_i is the so called frequency bias factor and its value is given by $B_i = \frac{1}{S_i}$ (based on the non interactive control), where S_i corresponds to the equivalent total droop of area i.

The output of the AGC controller of area i is

$$\Delta P_{AGC_i} = -(C_{P_i} + \frac{1}{sT_{N_i}})(\sum_{j\in\Omega_i} \Delta P_{ij} + \frac{1}{S_i}\Delta f_i) - \frac{K_{a_i}}{T_{N_i}s} p_i, \tag{6}$$

or in the time domain

$$\Delta \dot{P}_{AGC_i} = -C_{p_i}(\frac{\Delta \dot{f}_i}{S_i} + \sum_{j\in\Omega_i} \Delta \dot{P}_{ij}) - \frac{1}{T_{N_i}}(\frac{\Delta f_i}{S_i} + \sum_{j\in\Omega_i} \Delta P_{ij}) - \frac{K_{a_i}}{T_{N_i}} p_i, \tag{7}$$

where C_{p_i} is the proportional factor of the AGC controller, T_{N_i} the integration time constant of AGC controller and

$$p_i = \begin{cases} 0 & \text{if} \quad \Delta P^{min}_{AGC_i} < \Delta P_{AGC_i} < \Delta P^{max}_{AGC_i}, \\ \Delta P_{AGC_i} - \Delta P_{m,AGC_i} & \text{else}, \end{cases} \tag{8}$$

where p_i and K_{a_i} refer to the anti-wind up circuit. The power of the tie lines depends on the state of the system and will be specified later in the modeling of the two-area power system.

2.1.2 Consumed power

A change in the consumed power can be due to a change in the actual load or due to frequency dependency of the load. For instance, the amount of power that motor loads consume differs with frequency since their speed changes also. Moreover, the fact that kinetic energy can be stored in rotating masses of large motors causes an

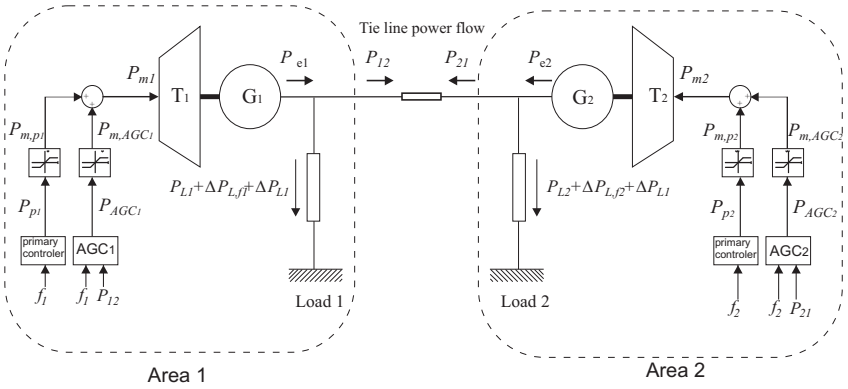

Fig. 2 Two-Area Power System. One generator model for each area equipped with primary control and AGC.

additional contribution depending on $\dot{f}$. Considering an area i in an interconnected system the load that the whole area has to compensate depends also on the power that is transmitted through the tie lines to other areas. Thus a change in the consumed power is expressed through

$$\Delta P_{e_i} = \Delta P_{L_i} + \Delta P_{L,f_i} + \sum_{j \in \Omega_i} (\Delta P_{ij}), \tag{9}$$

where ΔP_{L_i} is the actual deviation of the load, and $\Delta P_{L,f_i}$ is the deviation due to the frequency dependence of the load that is given by

$$\Delta P_{L,f_i} = \frac{1}{D_{l_i}} \Delta f_i + 2 \frac{W_{0_i}}{f_0} \Delta \dot{f}_i. \tag{10}$$

where the first term directly depends on the frequency, whereas the second one represents the fact that kinetic energy can be stored in the rotating masses of large motors.

2.2 Two-area power system model

Consider now the system of Fig. 2 that consists of only two control areas, each one equipped with its own AGC, connected by a tie line of reactance X. Each area is represented by an equivalent generating unit equipped also with an equivalent primary frequency control.

Based on the discussion of the previous subsection the frequency dynamics for area i that is connected with area j composing the two-are system are given by

$$\Delta \dot{f}_i = \frac{f_0}{2(H_i S_{B_i} + W_{0_i})} \Big(\Delta P_{m_i} - \Delta P_{L_i} - \frac{1}{D_{l_i}} \Delta f_i - \Delta P_{ij}), \tag{11}$$

All considered quantities are defined in the previous subsection except the power flow on the tie-line. The power flow from area i to area j is described by (12), where P_{ij} is positive when area i sends active power to area j. Also, since the active power losses on the line are neglected $P_{ji} = -P_{ij}$.

$$P_{ij} = \frac{V_i V_j}{X} \sin(\delta_i - \delta_j) = P_T \sin(\delta_i - \delta_j), \tag{12}$$

where $P_T = \frac{V_i V_j}{X}$ and X is the reactance of the tie line, V_i, V_j the voltage magnitude at the ends of the line, δ_i, δ_j the voltage angles at the ends of the line. Assuming the steady state point of the voltage controllers, we consider constant voltage magnitudes at the ends of the line during load deviation.

We set $\phi_{ij} = \delta_i - \delta_j$ and then define the variables according to the deviation from their initial (scheduled) value, here highlighted by the '0' subscript (i.e $\phi_{ij} = \Delta\phi_{ij} + \phi_{0_{ij}}$). Since $\dot{\delta}_i = 2\pi\Delta f_i$, then $\Delta\dot{\phi}_{ij} = 2\pi(\Delta f_i - \Delta f_j)$ and P_{ij} and its derivative results in

$$\begin{aligned} \Delta P_{ij} &= P_T \sin(\Delta\phi_{ij} + \phi_{0_{ij}}) + P_{0_{12}}, \\ \Delta \dot{P}_{ij} &= 2\pi P_T (\Delta f_i - \Delta f_j) \cos(\Delta\phi_{ij} + \phi_{0_{ij}}). \end{aligned} \tag{13}$$

Based also on the discussion in the previous subsection we get the following equations for the dynamics of the two-area system for $(i, j) \in (1,2), (2,1)$:

$$\begin{aligned}
\Delta \dot{f}_i &= \frac{f_0}{2(H_i S_{B_i} + W_{0_i})} \Big(\Delta P_{m_i} - \Delta P_{L_i} - \frac{1}{D_{l_i}} \Delta f_i - P_T \sin(\Delta\phi_{ij} + \phi_{0_{ij}}) + P_{0_{ij}} \Big), \\
\Delta \dot{\phi}_{12} &= 2\pi(\Delta f_1 - \Delta f_2), \\
\Delta \dot{P}_{AGC_i} &= \Big(\frac{1}{D_{l_i}} \frac{C_{p_i} f_0}{2S_i(H_i S_{B_i} + W_{0_i})} - \frac{1}{S_i} \frac{1}{T_{N_i}} \Big) \Delta f_i - \frac{C_{p_i} f_0}{2S_i(H_1 S_{B_i} + W_{0_i})} \Delta P_{m_i} \\
&\quad - \Big(\frac{1}{T_{N_i}} - \frac{C_{p_i} f_0}{2S_i(H_i S_{B_i} + W_{0_i})} \Big) (P_T \sin(\Delta\phi_{ij} + \phi_{0_{ij}}) - P_{0_{ij}}) \\
&\quad - 2\pi C_{p_i} P_T (\Delta f_i - \Delta f_j) \cos(\Delta\phi_{12} + \phi_{12}) + \frac{C_{p_i} f_0}{2S_1(H_i S_{B_i} + W_{0_i})} \Delta P_{L_i} - \frac{K_{a_i}}{T_{N_i}} p_i, \\
\Delta P_{m,p_i} &= \begin{cases} \Delta P_{p_i}^{min} & \text{if} \quad \Delta P_{p_i} \le \Delta P_{p_i}^{min}, \\ \Delta P_{p_i} & \text{if} \quad \Delta P_{p_i}^{min} < \Delta P_{p_i} < \Delta P_{p_i}^{max}, \\ \Delta P_{p_i}^{max} & \text{if} \quad \Delta P_{p_i} \ge \Delta P_{p_i}^{max}, \end{cases} \\
\Delta P_{m,AGC_i} &= \begin{cases} \Delta P_{AGC_i}^{min} & \text{if} \quad \Delta P_{AGC_i} \le \Delta P_{AGC_i}^{min}, \\ \Delta P_{AGC_i} & \text{if} \quad \Delta P_{AGC_i}^{min} < \Delta P_{AGC_i} < \Delta P_{AGC_i}^{max}, \\ \Delta P_{AGC_i}^{max} & \text{if} \quad \Delta P_{AGC_i} \ge \Delta P_{AGC_i}^{max}, \end{cases}
\end{aligned} \tag{14}$$

$$\begin{aligned}
\Delta P_{p_i} &= -\tfrac{1}{S_i}\Delta f_i, \\
\Delta \dot{P}_{AGC_i} &= -C_{p_i}(\tfrac{\Delta \dot{f}_i}{S_i} + \Delta \dot{P}_{ij}) - \tfrac{1}{T_{N_i}}(\tfrac{\Delta f_i}{S_i} + \Delta P_{ij}) - \tfrac{K_{a_i}}{T_{N_i}} p_i, \\
p_i &= \begin{cases} 0 & \text{if} \quad \Delta P^{min}_{AGC_i} < \Delta P_{AGC_i} < \Delta P^{max}_{AGC_i}, \\ \Delta P_{AGC_i} - \Delta P_{m,AGC_i} & \text{else.} \end{cases}
\end{aligned} \tag{15}$$

where $\Delta \phi_{ij} = -\Delta \phi_{ji}$.

For the analysis of the following sections we consider the model in (15) and assume that an attacker has disabled the AGC signal in the second control area and applies an arbitrary input $u \in U \subseteq \mathbb{R}$. Under this assumption and using a compact notation (15) is transformed in a continuous time, non nonlinear control system of the form

$$\dot{x} = f(x,w) + g(x,w)u, \tag{16}$$

where $x = [x_1,\ x_2,\ x_3,\ x_4]^T = [\Delta f_1,\ \Delta f_2,\ \Delta \phi_{12},\ \Delta P_{AGC_1}]^T \in \mathbb{R}^4$, $u \in U \subseteq \mathbb{R}$ is the attack input, and w is a vector containing all constants parameter in (11).

Moreover, Let $\mathscr{U}_{[t,t']}$ denote the sets of Lebesgue measurable functions from the interval $[t,t']$ to U. Following [20], if U is compact, f is Lipschitz in x and continuous in u, and $T \geq 0$ is an arbitrary time horizon, then this system with initial condition $x(t) = x \in \mathbb{R}^4$ admits a unique solution $x(\cdot) : [t,T] \to \mathbb{R}^4 \quad$ for all $\quad t \in [0,T]$, $x \in \mathbb{R}^4$, $u(\cdot) \in \mathscr{U}_{[t,T]}$. For $\tau \in [t,T]$ we will use $\sigma(\tau,t,x,u(\cdot)) = x(\tau)$ to denote this solution.

3 Feasibility of AGC attack

3.1 Safety Considerations

The AGC scheme outlined in the previous section is vital to the satisfactory performance of the power system, since it tries to keep the system frequency to its nominal value because too large deviations could damage the power system devices. This action may in the end jeopardize the stability of the entire system and in the worst case lead to a system blackout. In normal operation the frequency deviation of each area should not exceed 1.5Hz.

The amount of power that a line can transfer is also limited to maintain reliability and stability in the system. The limiting value for the permissible power transfer is influenced, according to the line length, by three factors: the thermal limit, the voltage drop and the stability limits. In the case-study of the two-area system, the amount of power that can be transferred is considered to be limited only by the steady state stability limit. This limit is a percentage of the maximal power P_T. We consider a minimum allowable steady state margin of 30% [29] which implies that $\Delta P_{12} \in [-70\%P_T, +70\%P_T]$. Since P_T is assumed constant, the aforementioned limits are translated into a bound $x_3 \in [-44^\circ, 44^\circ]$ in the phase angle difference.

In summary we consider the system to be safe when the state trajectories of (16) lie inside the following safe set of the state space:

$$x_1 \in [-1.5,+1.5],\ \ x_2 \in [-1.5,+1.5],\ \ x_3 \in [-44^\circ,44^\circ] \tag{17}$$

We consider the model in (16) and investigate whether there exists a policy $u(\cdot)$ for the attacker, that can drive the system trajectories $\sigma(\cdot,t,x,u(\cdot))$ outside the safety margins in (17), and/or lead to unstable swinging in the power exchanged between the two control areas by exceeding the limits of x_3 for a sufficiently large amount of time. It should be noted that power swinging results in large power oscillations in the tie-line which are undesirable and can lead to triggering out-of-step protection relays that trip generating units in order to avoid potential damaging and mechanical vibrations [29].

3.2 Violating the safety margins

We first examine if the attacker, selecting a suitable policy, can lead the system trajectory outside the safe region defined in (17). Define $K_1 \subset \mathbb{R}^4$ by

$$K_1 := \{x \in \mathbb{R}^4 \mid |x_1| \leq 1.5, |x_2| \leq 1.5, |x_3| \leq 44^\circ\}, \tag{18}$$

and let $l_1(\cdot):\ \mathbb{R}^4 \to \mathbb{R}$ be the signed distance to the set K_1, defined by $l_1(x) = \min\{x_1+1.5, 1.5-x_1, x_2+1.5, 1.5-x_2, x_3+44^\circ, 44^\circ-x_3\}$, for any $x \in \mathbb{R}^4$. Clearly, $K_1 = \{x \in \mathbb{R}^4 \mid l_1(x) \geq 0\}$. Note that the last state x_4, which corresponds to the AGC signal in the first area is restricted indirectly due to the line saturation.

The problem of interest can be though of as a reachability problem where the objective is to compute the set of states at some initial time $t < T$ for which there exists a control policy that can drive (at least for some time instance) the system trajectories in K_1^c, i.e. outside the safe region (17). The desired set can be encoded by

$$\begin{aligned} Reach(t,K_1) = \{x \in \mathbb{R}^4 \mid \exists u(\cdot) \in \mathscr{U}_{[t,T]} \\ \exists \tau \in [t,T]\ \sigma(\tau,t,x,u(\cdot)) \notin K_1\}. \end{aligned} \tag{19}$$

It is shown in [20, 21] that $Reach(t,K_1)$ can be related to the zero sub-level set of

$$V(x,t) = \inf_{u(\cdot)\in\mathscr{U}_{[t,T]}} \min_{\tau\in[t,T]} l_1(\sigma(\tau,t,x,u(\cdot))). \tag{20}$$

In particular, $Reach(t,K_1) = \{x \in \mathbb{R}^4 \mid V(x,t) < 0\}$ and $V(x,t)$ is the unique, bounded and uniformly continuous viscosity solution to the Hamilton-Jacobi equation

$$\frac{\partial V}{\partial t}(x,t) + \min\{0, \inf_{u\in U} \frac{\partial V}{\partial x}(x,t) f(x,u)\} = 0, \tag{21}$$

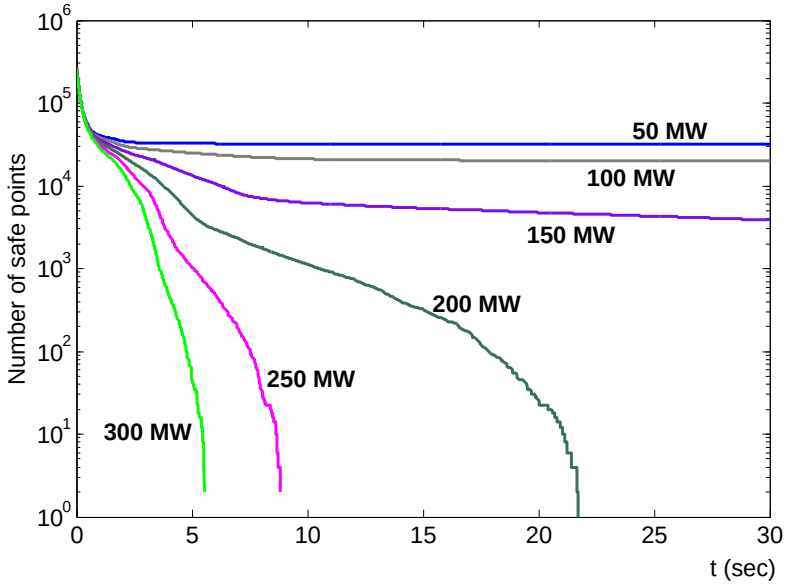

Fig. 3 Family of curves depicting how the volume (i.e. the number of safe points) of the safe set (complement of $Reach(t,K_1)$) changes over time for different bounds of the attack signal. For an attack authority greater than or equal to $200MW$, the volume of the safe set vanishes after a given time instance.

with terminal condition $V(x,T) = l_1(x)$.

Therefore, to compute $Reach(t,K_1)$ it suffices to solve the partial differential equation in (21). The latter can be achieved using standard numerical tools for such problems based on Level Set Methods [30].

For the analysis of this section, we performed a series of reachability computations for different bounds of the attack input. These bounds correspond to different limits on the AGC signal that the attacker has gained access to. Fig. 3 shows a family of curves that correspond to the different bounds of the attack signal. These curves quantify how the volume of the safe set changes in time. The safe set is defined as the complement of $Reach(t,K_1)$ since it includes all states from which the system trajectories can remain safe for the entire horizon. By inspecting this figure, since the volume of the safe set vanishes for an attack authority greater than or equal to $200MW$, we can conclude that the attacker would need a signal at least $200MW$ to disturb the system starting from the nominal operating point. It should be noted that the attacker does not need to have access to any power plant; the attack signal is just a communication signal coming out of the SCADA system, with amplitude in the order of $5V$. This signal serves then as a command to the generators participating in the AGC loop to generate power of certain magnitude, in this case at least $200MW$.

In Fig. 4, we show the result of a reachability calculation for the case where the attacker is able to inject an arbitrary signal up to 100MW, i.e. $|u| \leq 100$MW. The calculation is performed backwards in time, so that if the system trajectories start

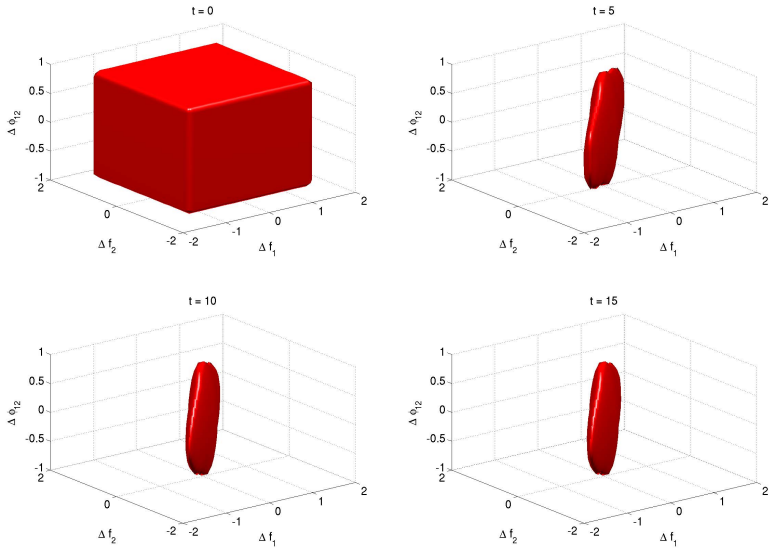

Fig. 4 Safe set for the case where $u \in [-100, +100]MW$ and $x_3 \in [-44^\circ, 44^\circ]$. After approximately 10 seconds, there are still some states, including the nominal point, from which system trajectories can start and remain in the safe region ("red" set in the upper-left panel).

from the "red" set of the last panel they can reach the "red" set of the first panel, irrespective of the pattern of the attack signal. As expected from Fig. 3, the safe set saturates after approximately 10 seconds, which means that there are still some states, including the nominal point, that system trajectories can start and remain in the safe region K_1 of (18). The complement of the "red" surfaces correspond to $Reach(t, K_1)$.

3.3 Power swinging between two areas

Next we consider the possibility of keeping the angle x_3 outside $[-44^\circ, 44^\circ]$ for a sufficiently large amount of time, thus leading to an unstable power swinging between the two areas. To this end we define the set $K_2 \subset \mathbb{R}^4$ by

$$K_2 := \{x \in \mathbb{R}^4 \mid |x_3| > 44^\circ\}, \tag{22}$$

and a function $l_2(\cdot) : \mathbb{R}^4 \to \mathbb{R}$ to be the signed distance to the set K_2, defined by $l_2(x) = \min\{-x_3 - 44^\circ, x_3 - 44^\circ\}$, for any $x \in \mathbb{R}^4$. Clearly, $K_2 = \{x \in \mathbb{R}^4 \mid l_2(x) \geq 0\}$.

We first perform a so called viability computation and determine the set of states for which there exists an attack policy such that the emanating trajectories remain in K_2 for the entire horizon. The desired set can be encoded by

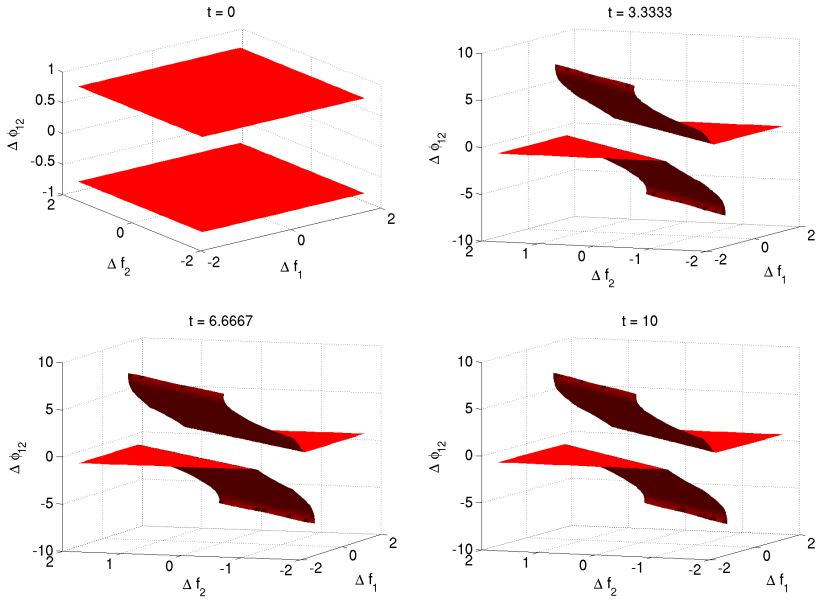

Fig. 5 Computation of the viable set $Viab(t,K_2)$ for an attack signal bounded in $[-350\ 350]$ MW. $Viab(t,K_2)$ is saturated in approximately 7 seconds; namely, there exists a non-empty set such that if the system starts from that set, the attacker can construct an input sequence to keep the angle above or below 44° for the specified time horizon.

$$Viab(t,K_2) = \{x \in \mathbb{R}^4 \mid \exists u(\cdot) \in \mathscr{U}_{[t,T]} \\ \forall \tau \in [t,T]\ \sigma(\tau,t,x,u(\cdot)) \in K_2\}. \tag{23}$$

It is shown in [20] that $Viab(t,K_2)$ can be related to the zero sub-level set of

$$\widetilde{V}(x,t) = \sup_{u(\cdot)\in\mathscr{U}_{[t,T]}} \min_{\tau\in[t,T]} l_2(\sigma(\tau,t,x,u(\cdot))). \tag{24}$$

In particular, $Viab(t,K_2) = \{x \in \mathbb{R}^4 \mid \widetilde{V}(x,t) \geq 0\}$ and $\widetilde{V}(x,t)$ is the unique, bounded and uniformly continuous viscosity solution to the Hamilton-Jacobi equation

$$\frac{\partial \widetilde{V}}{\partial t}(x,t) + \min\{0, \sup_{u\in U} \frac{\partial \widetilde{V}}{\partial x}(x,t) f(x,u)\} = 0, \tag{25}$$

with terminal condition $\widetilde{V}(x,T) = l_2(x)$.

The result of this calculation is shown in Fig. 5, where it was assumed that the attack signal is bounded in $[-350\ 350]$ MW due to the AGC saturation. It can be observed that the viability set $Viab(t,K_2)$ is saturated in approximately 7 seconds; namely, there exists a non-empty set such that if the system starts from that set, the attacker can construct an input sequence to keep the angle above or below 44° for

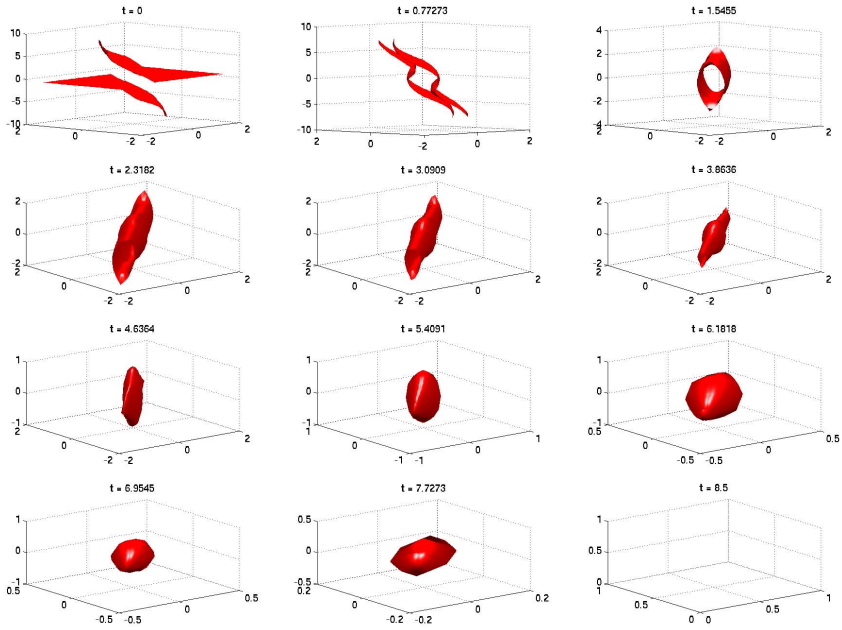

Fig. 6 Computation of the reachable set $Reach(t,K_3)$. Since the safe set (denoted by "red", complement of $Reach(t,K_3)$) is empty, for every initial condition, there exists at least one control policy for the attacker so as to reach the viability set $Viab(t,K_2)$ (outcome of the calculation in Fig. 5) in 8.5 seconds.

the specified time horizon. Notice that since the other states (except x_3) are free in this case, the constraint ($|x_3| > 44^\circ$) in the definition of K_2 divides the state space to two parts; one part between the two surfaces of Fig. 5, and one outside. The latter is the set where the attacker is trying to steer the system trajectories.

Having computed the set $Viab(t,K_2)$ it remains to verify whether the attacker is able to force the system to that set. If so, then once reaching the viability set, the attacker could change his control policy and keep the angle deviation in the unsafe region for sufficiently large amount of time. The latter can cause power swinging and its undesirable consequences. For this purpose, we define the set K_4 by

$$K_3 := \{x \in \mathbb{R}^4 \mid \widetilde{V}(x,0) > 0\}, \tag{26}$$

where $\widetilde{V}(x,0)$ is the value function that characterizes the set $Viab(0,K_2)$ obtained from the viability computation. We then compute the set $Reach(t,K_3)$ defined as in (19) with K_3 in place of K_1.

As shown in Fig.6, since the safe set (denoted by "red", complement of $Reach(t,K_3)$) is empty, for every initial condition, there exists at least one control policy for the attacker so as to reach the viability set $Viab(t,K_2)$ (outcome of the calculation in Fig. 5) in 8.5 seconds. Then, the attacker could switch policy and keep the state trajectory in K_2.

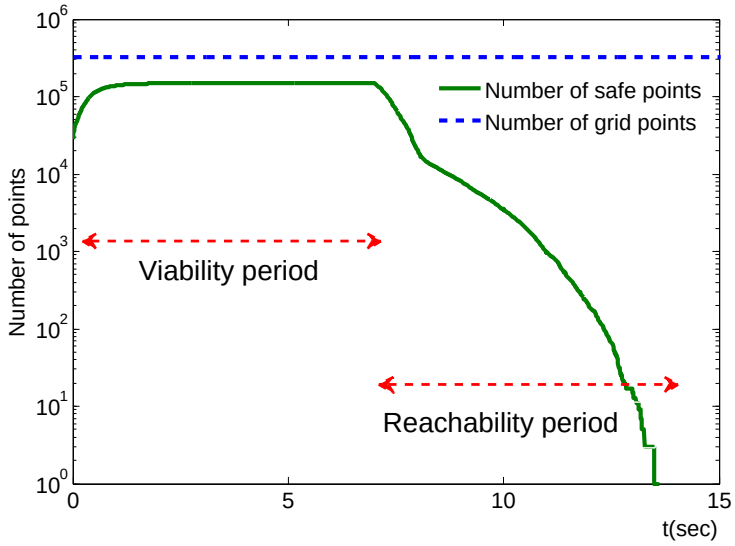

Fig. 7 The dashed "blue" line is the total number of grid points; the solid "red" line indicates how the volume of the safe set changes over time after the viability and reachability calculations.

Fig. 7 summarizes the previous analysis, which comprises of two stages; in the first stage we compute the viability set $Viab(t,K_2)$ of K_2, whereas in the second stage we compute the reachability set $Reach(t,K_3)$. One can see how the volume of the safe part of the state space changes over time. Following the definition of the reachable and the viability set, at the first stage the safe set coincides with the viable set, and in the second one it corresponds to the complement of the reachable set. At the 7th second the viability set is saturated, and the attacker could change policy so as to keep the angle increasing. That way, the power will start swinging and this in turn might lead to activation of the out-of-step protection relays.

4 Attack signal synthesis

Using the frequency model of Section 2 it was shown in Section 3 that if an attacker gains access to the AGC signal in one control area, then she can cause undesirable effects to the network. Using tools from optimal control and reachability theory the existence of such an attack policy was verified. Such a policy would be a minimum time to disruption signal, however, to construct it is a difficult task since it is based on the spatial derivatives of the value functions V, $\widetilde{V}$, whose computation is affected by discretization errors. To overcome this difficulty we present here different alternatives for a synthesis of an attack signal.

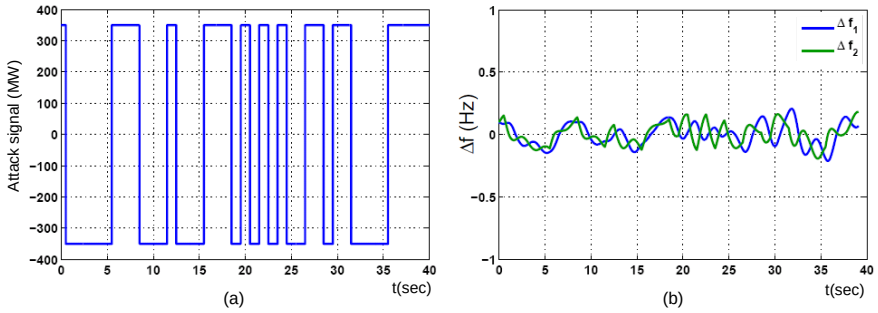

Fig. 8 (a) Naive bang-bang attack signal, (b) frequency deviation of the two areas.

4.1 Open loop

4.1.1 Naive attack signal

The optimal attack input generated based on the analysis of the previous section is shown to be a bang-bang signal. Motivated by this fact, we show here that constructing a bang-bang input sequence with arbitrarily selected switching instances is not sufficient for the attacker to disturb the system.

We select the naive, bang-bang attack signal to be a pulse sequence as the one shown in Fig. 8(a). By inspection of Fig. 8(b), such an attack signal leads only to minor deviations in the frequency of each area from its nominal value, and to affordable oscillations in the power exchanged between the two areas. Other random bang-bang signals have been tested as well, leading to similar performance. This reveals the need to resort to more sophisticated attack synthesis techniques.

4.1.2 MCMC based attack signal

It was shown that applying a naive, open loop bang-bang signal is not sufficient for the objectives of an attacker. Here we construct an open loop signal of similar type, but selecting this time the switching instances by means of an optimization problem. Specifically, we assume that the nominal parameter values are available to the attacker, i.e. $w = w_0$ for the model (16) considered in the proposed design.

We consider input sequences of the form

$$\mathbf{u}(t) = u_\kappa, \;\; \frac{T}{N}\kappa \le t < \frac{T}{N}(\kappa+1), \;\; \text{for} \;\; \kappa = 0,\ldots,N-1, \tag{27}$$

where $u_\kappa \in \{-350, 350\}$, T denotes the optimization horizon and T/N is the time discretization step. Identifying an optimal control policy for the attacker that leads the system trajectories outside the safety margins in (17) can be thought of as an

Algorithm 1 MCMC algorithm

1: Let θ_0 denote an initial choice for θ.
2: Define as N the total number of iterations.
3: **For** $i = 0, \ldots, N$
4: Fix $\alpha > 0$ and extract $\theta_{i+1} \sim p_\theta(\cdot|\theta_i)$

$$\rho = \min\big\{1, \frac{p_\theta(\theta_i|\theta_{i+1})}{p_\theta(\theta_{i+1}|\theta_i)} \frac{J(\theta_{i+1})^\alpha}{J(\theta_i)^\alpha}\big\},$$

$$\theta_{i+1} = \begin{cases} \theta_{i+1} & \text{with probability } \rho, \\ \theta_i & \text{with probability } 1-\rho, \end{cases}$$

5: **end**

optimization problem. We seek for a vector $\theta = (u_1, \ldots, u_N) \in \{-350, 350\}^N$ that maximizes the objective function $J = e^{\int_0^T x_2^2 dt}$, subject to the system dynamics (16). By maximizing the criterion J we implicitly maximize the deviation of the frequency from its nominal value and hence force the system trajectories outside the safe region in (17).

This is a nonlinear optimization problem over a discrete domain; to solve it we use the Markov Chain Monte Carlo (MCMC) method. This is a randomized optimization technique, which explores the search space via a Markov chain. It is an iterative process and at each iteration $i = 0, \ldots, N$ (step $3-5$ of Algorithm 1) we extract a candidate solution from a proposal distribution of our choice, which approximates the maximizer of the problem. Within each iteration, the algorithm involves a sophisticated accept-reject mechanism (step 4 of Algorithm 1), which quantifies the trade-off between the objective value of the sampled solution and its probability of occurrence.

At each iteration i, the extracted variable θ_i is accepted with a probability ρ, as this is defined in the algorithm above. Otherwise, it is rejected and the previous state of the chain is replicated. At the end of the algorithm, the extracted points are concentrated at different regions, and based on the peakedness, the optimal value for θ is determined. The probability density $p_\theta(\cdot|\theta_i)$ denotes the proposal distribution, which at the first step is chosen to be a uniform distribution, so as to search evenly the decision space. At a next step the entire process is repeated, this time sampling from gaussian distributions centered at the accepted samples of the first run. That way, a local search is performed and a more accurate maximizer is identified. To investigate the performance of the attack policy constructed according to the MCMC algorithm we considered two different case studies. In the first set-up the obtained solution, which is based on a model with $w = w_0$, is applied to a perfect model that has the nominal set of parameter values, i.e. $w = w_0$. The second study involves the application of the obtained solution on a model with $w \neq w_0$. For the MCMC algorithm we selected $T = 40$ sec and $N = 40$. We performed in total 82306 iterations until the accepted states of the chain were 50000. The ratio between accepted and total states of the chain is 0.61.

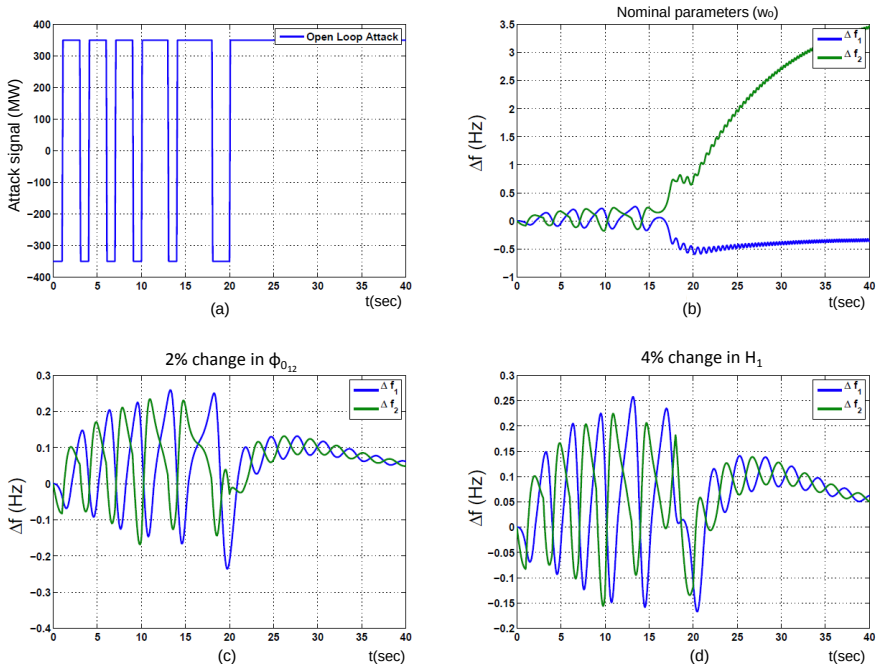

Fig. 9 (a) Open loop policy generated by the MCMC algorithm. (b) Frequency trajectories for perfect model with $w = w_0$. (c) Frequency trajectories for imperfect model with $w \neq w_0$ (2% mismatch in $\phi_{0_{12}}$). (d) Frequency trajectories for imperfect model with $w \neq w_0$ (2% mismatch in H_1).

Perfect Model: We consider a set-up with $w = w_0$.Fig. 9(a) depicts an open loop policy for the attacker, obtained via the MCMC optimization method. Fig 9(b) shows the frequency response in the two areas. Clearly, the impact of the attack signal is extremely severe. The swings of the transferred power on the tie line will result in triggering the out-of-step protection relays. If the system is not equipped with such a protection scheme, the generators of the second area would start to trip by the time that the frequency of that area would exceed the safety margins. The latter may lead to cascading failures and even to a wide-area blackout.

Imperfect model: In Fig. 9(c) and 9(d), we assume that the attacker does not have perfect information of the system. Specifically we consider the case where the angle $\phi_{0_{12}}$ and the inertia H_1 in the first area that the attacker considers in her design are 2% and 4%, respectively, higher than true parameter values. It is clear that the open loop strategy is extremely sensitive to such a model mismatch and hence the open loop policy does not serve practically as an efficient solution.

4.2 Closed loop

The poor performance of the naive attack signal and the sensitivity to parameter uncertainty of the MCMC based signal motivate the synthesis of a feedback attack policy. Two alternatives are proposed, one based on feedback linearization and one based on gain scheduling. For simplicity we assume perfect state information for the attacker. We refer to [18] for the case of partial state information, where a nonlinear observer is constructed.

4.2.1 Feedback linearization based attack signal

We consider an attack scheme that is based on feedback linearization and the MCMC algorithm presented in the previous subsection. Feedback linearization is based on applying a nonlinear coordinate transformation and a nonlinear feedback to transform a nonlinear input affine system as the one in (16) to a system that is linear in the new coordinates.

The feedback linearization procedure is based on the notion of relative degree γ. Specifically, for the nonlinear system (16) with output $y = l(x)$, for some $l(\cdot): \mathbb{R}^4 \to \mathbb{R}$, is said to have relative relative degree γ with $1 \leq \gamma \leq n$, in a region $D \subset \mathbb{R}^4$ if $L_g L_f^{i-1} l(x) = 0$ for $i = 1, 2, \cdots, \gamma - 1$, and $L_g L_f^{\gamma-1} l(x) \neq 0$ for all $x \in D$. Note that $L_f^1 l(x) = \frac{\partial l}{\partial x} f$ is called the Lie derivative of l with respect to f, whereas higher order Lie derivatives are defined recursively [24]. It should be also noted that for the relative degree to be well-defined $L_f^i l(x)$ needs to be differentiable and hence $f(\cdot, w)$ should be smooth. Due to the saturation of primary and secondary loop control this might not be the case; however, we assume that that none of the dynamic saturations is activated and $f(\cdot, w)$ is sufficiently smooth. The saturations will be explicitly taken into account in the design of the attack signal.

It can be easily seen that by choosing $y = l(x) = x_3$, (16) has relative degree $\gamma = 2$. It is then shown in [24] that, for every $x_0 \in D$, there exists a nonlinear transformation $T(\cdot,\cdot): \mathbb{R}^4 \times \mathbb{R}^3 \to \mathbb{R}^4$ such that $[\eta,\ \xi]^T = T(x, w)$, $\eta, \xi \in \mathbb{R}^2$, and a nonlinear feedback $v = \alpha(x, w) + \beta(x, w)u$ with $\alpha(x, w) = L_f^{\rho} l(x)$, $\beta(x, w) = L_g L_f^{\gamma-1} l(x)$, that results in a dynamical subsystem that is linear in the ξ coordinates.

Upon using the linearizing transformation T and the associated functions α and β, (16) is transformed to

$$\begin{aligned} \dot{\eta} &= f_0(\eta, \xi), \\ \dot{\xi} &= A_c \xi + B_c v, \\ y &= C_c \xi, \end{aligned} \tag{28}$$

where $A_c \in \mathbb{R}^{2\times 2}$, $B_c \in \mathbb{R}^2$ and $C_c \in \mathbb{R}^{1\times 4}$ are canonical controllability matrices [25] and $f_0(\cdot,\cdot): \mathbb{R}^2 \times \mathbb{R}^2$ is a nonlinear function referred to as zero dynamics. This form decomposes the system into a linear subsystem in the ξ coordinates and an internal nonlinear subsystem in the η coordinates. Here our main goal is to push the system trajectories to the unsafe region in contrast to the usual stabilization idea. Hence,

unstable behavior of the internal dynamics would be a benefit for our objectives, i.e. destabilize the system.

In the linearized subsystem we can apply the state feedback $v = K\xi$, which results in the feedback law

$$u(x,w,K) = \frac{K\xi - \alpha(x,w)}{\beta(x,w)} = \frac{K[0\ \ I]T(x,w) - \alpha(x,w)}{\beta(x,w)}, \tag{29}$$

in the original coordinates. The feedback gain $K \in \mathbb{R}^{1\times\gamma}$ is a constant vector. To consider the saturation limits of AGC, $|u(x,w,K)| \leq U_0 = 350$ MW, we pass the control law through a saturation operator as

$$\begin{aligned} \bar{u}(x,w,K) &= \text{sat}(u(x,w,K),U_0) \\ &= \begin{cases} u(x,w,K) & \text{if } |u(x,w,K)| \leq U_0, \\ U_0\ \text{sign}(u(x,w,K)) & \text{if } |u(x,w,K)| > U_0, \end{cases} \end{aligned} \tag{30}$$

where the sign$(\cdot)$ operator returns the sign of its argument.

The attack signal would be then given by (30), where the gain K remains to be defined. To determine K we seek to maximize $J = \max(\|x_1\|_\infty, \|x_2\|_\infty)$ subject to the dynamics in (16) when the feedback law in (30) is applied. That way the gain K chosen by an attacker would result in the maximum possible deviation of the frequencies from their nominal values and lead the system trajectories outside (17). To solve this maximization problems we use the MCMC algorithm described in the previous subsection. Note that J is different from the objective function used in the previous subsection since it resulted in this case to a better performance.

As in the case of the MCMC based attack design we investigated two cases according to whether the attacker has perfect model knowledge (i.e. $w = w_0$) or not. For the MCMC algorithm employed to determine the gain K we used $T = 40$ sec and $N = 40$. We performed in total 10000 iterations and the ratio between accepted and total states of the chain is 0.37.

Perfect Model: We considered a scenario with $w = w_0$. Fig. 10(a) shows the feedback policy of the attacker and Fig. 10(b) shows the frequency trajectory of each area and proves the severe impact that a suboptimal attack signal could have on the system.

Imperfect model: Similarly to the open loop simulations (Fig. 10(c), 10(d)), we assume that the attacker has an imperfect knowledge of the system, and the angle $\phi_{0_{12}}$ and the inertia H_1 in the first area that she considers in her design are 2% and 4%, respectively, higher than true parameter values. In contrast to the open loop strategy, the feedback policy is considerably robust to such a model mismatching and consequently it provides an effective and practical solution to construct an attack signal.

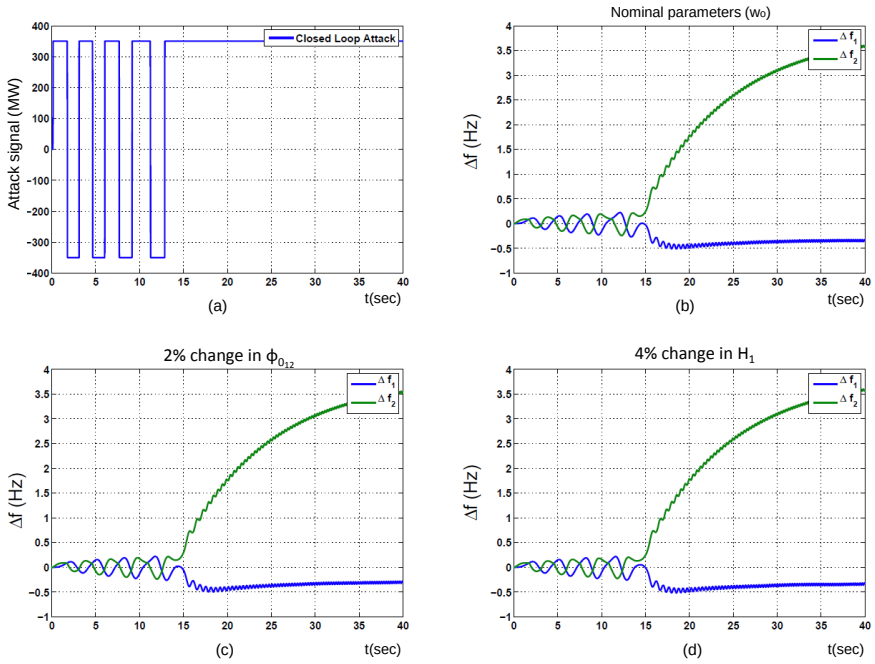

Fig. 10 (a) Closed loop policy generated using feedback linearization. (b) Frequency trajectories for perfect model with $w = w_0$. (c) Frequency trajectories for imperfect model with $w \neq w_0$ (2% mismatch in $\phi_{0_{12}}$). (d) Frequency trajectories for imperfect model with $w \neq w_0$ (4% mismatch in H_1).

4.2.2 Gain scheduling based attack signal

If the system dynamics in (16) were described by a linear dynamical system, a natural choice for the attacker would be to choose her signal among the class of linear feedback policies. The feedback gain would be then selected so that the eigenvalues of the linear system have positive real part, resulting in an unstable behavior.

In our case, however, (16) is nonlinear and in fact, due to the saturation limits of the primary frequency controller and the AGC, it involves multiple modes of operation. Motivated by control design techniques based on gain scheduling [25], we apply the following procedure [31]. We first linearize (16) around a nominal operating point at every mode of operation. In total 27 modes are distinguished due to the saturation limits of the primary frequency controller of each area and the AGC of the second area (the attacker has gained access of the AGC of the second area). We then have a family of linear systems. For each one we design a linear feedback so that the eigenvalues of the corresponding system have positive real parts.

The attack signal is then a switched linear feedback of the state, since the feedback gain changes according to the mode of operation. Moreover, to ensure that the

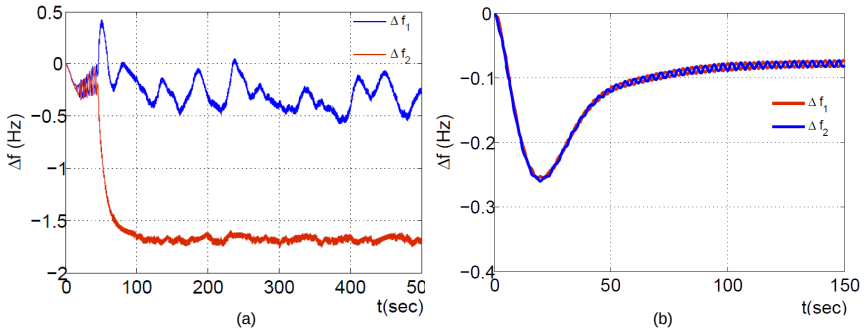

Fig. 11 Frequency response at each area as an effect of the gain scheduling based attack signal.

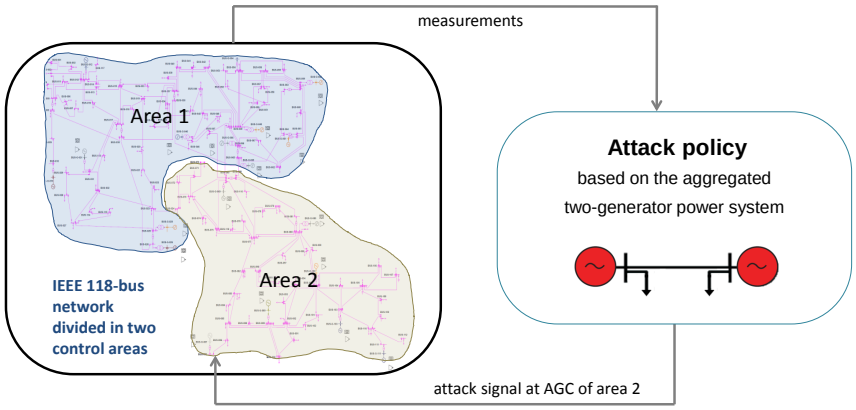

Fig. 12 Interaction between the power system and the cyber-attack policy. The abstracted model of Section 2 is used for synthesizing the attack signal that is applied to the two-area detailed model. The areas are determined arbitrarily.

attack signal satisfies the saturation limits of the AGC, we pass it through a saturation function as in (30).

Fig. 11(a) illustrates that by applying the attack signal generated by the aforementioned procedure, unacceptable deviations in the frequencies are obtained. It should be also noted that if the feedback gain is not updated according to the mode of operation then the effect of the corresponding attack signal is very different. Fig. 11(b) shows the negligible deviations in the frequency response for each area, to highlight the necessity of the gain scheduling scheme.

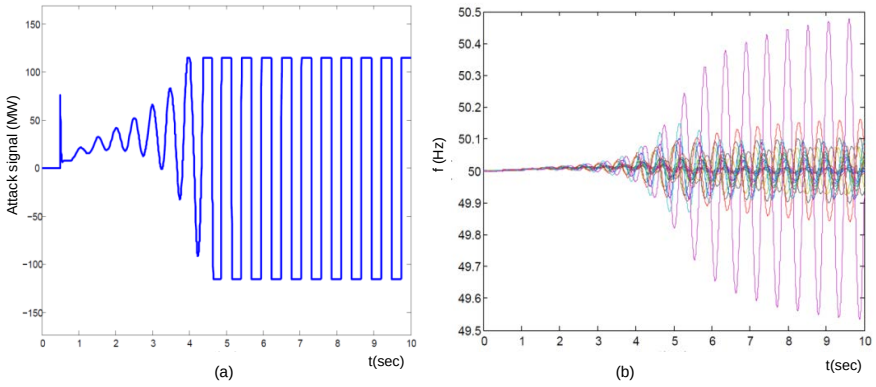

Fig. 13 (a)Attack signal, (b) Frequency response of the generating units.

5 Evaluation on a detailed simulation environment

In the previous sections, for the synthesis and the evaluation of the constructed attack signal, the two-generator power system model of Section 2 was employed. Here we investigate the performance of the constructed signals when they are applied to a detailed model of the network. For this study, we used the IEEE 118-bus network, and a detailed power system simulation environment implemented in MATLAB by [32]. All generators are represented by the classical model and are also equipped with primary frequency control, Automatic Voltage Regulator (AVR) and Power System Stabilizer (PSS). Moreover, we arbitrarily divide the network into two control areas, each one equipped with each own AGC loop. The data of the model are retrieved from a snapshot available at [33]. Since there were no dynamic data available, typical values provided by [34] are used for the simulations.

For the control synthesis, we use the abstracted two-generator model by aggregating each area into one generator based on the center of inertia as discussed in Section 2. In Fig. 12, we show the interaction between the power system and the cyber-attack policy. The model abstraction is only used for the attack signal synthesis and serves as feedback to the detailed simulation environment.

Fig. 13 shows the effect in the detailed system once the feedback linearization based attack signal is applied. In contrast to the results of the previous section, applying the attack signal to the detailed system does not lead to significant frequency deviations. However, as shown in Fig. 14, swings on the generator output and the power flows across the tie lines connecting the two areas are observed. The swinging behavior can be dangerous for the system since they may trigger out-of-step protection relays and cause a cascade of undesirable effects. It should be noted that the qualitatively different effect of the application of the attack signal to the detailed model compared with those obtained when it is applied to the model of Section 2, are due to the mismatch between the two models (abstraction error).

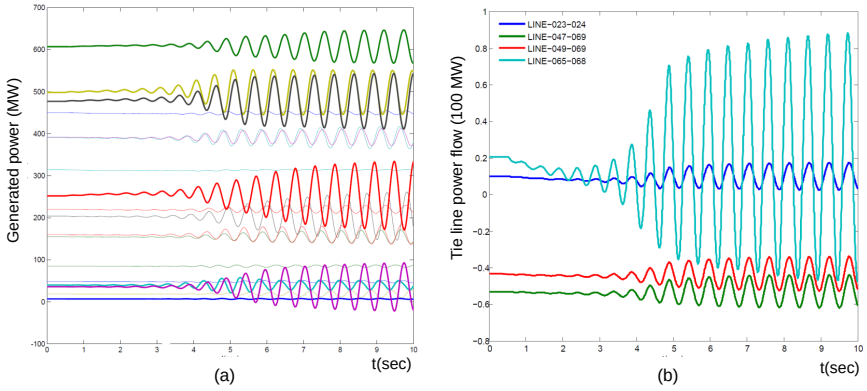

Fig. 14 (a) Swinging on the produced power of the generators and (b) power swinging between the two areas as an effect of the application of the attack signal.

6 Conclusion

In this chapter we investigated the impact that a cyber-attack on the AGC loop may have in the power system. We employed an abstraction of the detailed power system model and carried out an feasibility analysis based on reachability and optimal control theory. This analysis offered us intuition on whether there exist an attack pattern that can disturb the power system. We also investigated the problem of synthesizing an attack signal by using open and closed loop nonlinear control approaches. The efficacy of the proposed methods was investigated by means of simulations on the IEEE-118 bus network.

The fact that our results show that the power system can indeed be disturbed by an AGC attack, highlight the necessity of devising an attack detection scheme. A complimentary study towards this direction can be found in [35, 36] where a detection algorithm and an intuitive mitigation strategy is proposed.

In our analysis we considered complete information about the system state, i.e. frequencies, AGC signal, etc. For a more realistic implementation the case of partial information should be investigated. A preliminary discussion can be found in [18]. Moreover, to facilitate the use of the proposed control techniques more accurate model abstractions need to be developed. Another direction for future work, would be to investigate the impact of an attack, not necessarily at the AGC signal, but at a different interface point between the power network and the SCADA system.

Appendix

Nomenclature for the frequency model of Section 2

f	Frequency (rotor speed)
H	Inertia constant
S_B	Rated power
P_e	Electrical power
$\frac{1}{S}$	Droop constant
$\frac{1}{D_l}$	Load damping coefficient
$P_{m,p}$	Primary frequency control power after saturation
$P_{m,AGC}$	Secondary frequency control power after saturation
$P_{m,set}$	Power setpoint
P_p	Primary frequency control power prior to saturation
P_{AGC}	Secondary frequency control power prior to saturation
Δe_i	Error signal of the AGC controller
P_{ij}	Power transmitted from area i to area j
B_i	Frequency bias factor
C_p	Proportional factor of the AGC controller
T_N	Integration time constant of the AGC controller
p	Anti-wind up variable
K_a	Anti-wind up gain
P_L	Constant (not frequency dependent) load
$P_{L,f}$	Frequency dependent load
V	Bus voltage magnitude
δ	Bus voltage angle
ϕ_{ij}	Voltage angle difference between bus i and j
X	Line reactance
u	Attack signal
x	State vector
w	Parameter vector

The aforementioned quantities are scalar valued unless stated otherwise. The min or max superscript in any of these variables denotes their minimum and maximum limit, respectively. For the symbols of Sections 3 and 4 that do not correspond to physical quantities the reader is referred to the explanation provided in these sections.

Acknowledgements Research was supported by the European Commission under the project VIKING, FP7-ICT-SEC-2007-1. The authors would also like to thank the coordinator of the VIKING project Gunnar Björkman for his valuable input and insight in the current work. We would also like to thank Dr. Turhan Demiray for his support and for generously providing the Matlab-based detailed simulation environment.

References

1. G. Andersson, P. Donalek, R. Farmer, N. Hatziargyriou, I. Kamwa, P. Kundur, N. Martins, J. Paserba, P. Pourbeik, J. Sanchez-Gasca, R. Schulz, A. Stankovic, C. Taylor, and V. Vittal, "Causes of the 2003 Major Grid Blackouts in North America and Europe, and Recommended Means to Improve System Dynamic Performance," *IEEE Transactions on Power Systems*, vol. 20, no. 4, pp. 1922–1928, 2005.
2. M. Zima and M. Bockarjova, *Operation, Monitoring and Control Technology of Power Systems*. Lecture Notes, ETH Zurich, 2007.
3. D. Kirschen and F. Bouffard, "Keep the Lights On and the Information Flowing," *Power and Energy Magazine, IEEE*, vol. 7, no. 1, pp. 50–60.
4. A. Teixeira, S. Amin, H. Sandberg, K. Johansson, and S. Sastry, "Cyber security analysis of state estimators in electric power systems," in *Decision and Control (CDC), 2010 49th IEEE Conference on*, Dec 2010, pp. 5991–5998.
5. O. Vukovic, K. C. Sou, G. Dan, and H. Sandberg, "Network-aware mitigation of data integrity attacks on power system state estimation," *Selected Areas in Communications, IEEE Journal on*, vol. 30, no. 6, pp. 1108–1118, July 2012.
6. G. Hug and J. Giampapa, "Vulnerability assessment of ac state estimation with respect to data injection cyber-attacks," *IEEE Transactions on Smart Grid*, 2012.
7. M. Negrete-Pincetic, F. Yoshida, and G. Gross, "Towards Quantifying the Impacts of Cyber Attacks in the Competitive Electricity Market Environment," *IEEE Power Tech Conference*, 2009.
8. *Forbes, Congress Alarmed at Cyber-Vulnerability of Power Grid, available at http://www.forbes.com/2008/05/22/cyberwar-breach-government-tech-security_cx_ag_0521cyber.html*.
9. *CNN, Sources: Staged cyber attack reveals vulnerability in power grid, available at http://www.cnn.com/2007//US/09/26/power.at.risk/index.html*.
10. *Comptuterworld, DHS to review report on vulnerability in West Coast power grid, available at http://www.computerworld.com/s/article/9138017*.
11. J.-W. Wang and L.-L. Ronga, "Cascade-based attack vulnerability on the US power grid," *Elsevier, Safety science*, vol. 47, no. 10, pp. 1332–1336, 2009.
12. *VIKING Project, http://www.vikingproject.eu*.
13. "Vulnerability assessment of scada systems," *Deliverable D3.1, VIKING project*, 2011.
14. "Impact analysis of adverse events," *Deliverable D3.2, VIKING project*, 2011.
15. "Consequence and cost analysis of scada system vulnerabilities," *Deliverable D3.3, VIKING project*, 2011.
16. "Mitigation and protection strategies," *Deliverable D4.3, VIKING project*, 2012.
17. P. M. Esfahani, M. Vrakopoulou, K. Margellos, J. Lygeros, and G. Andersson, "Cyber attack in a two-area power system: Impact identification using reachability," in *American Control Conference*, 2010, pp. 962 – 967.
18. ——, "A robust policy for automatic generation control cyber attack in two area power network," in *49th IEEE Conference Decision and Control*, 2010, pp. 5973 – 5978.
19. "Report on case studies," *Deliverable D5.3, VIKING project*, 2012.
20. J. Lygeros, "On reachability and minimum cost optimal control," *Automatica*, vol. 40, no. 6, pp. 917–927, 1999.
21. I. Mitchell, A. M. Bayen, and C. Tomlin, "A time-dependent Hamilton-Jacobi formulation of reachable sets for continuous dynamic games," *IEEE transactions on Automatic Control*, vol. 50.
22. A. Lecchini-Visintini, J. Lygeros, and J. Maciejowski, "Stochastic optimization on continuous domains with finite-time guarantees by markov chain monte carlo methods," *Automatic Control, IEEE Transactions on*, vol. 55, no. 12, pp. 2858–2863, Dec 2010.
23. C. Robert and G. Casella, *Monto Carlo Statistical Methods*. Springer Verlag.
24. S. Sastry, *Nonlinear systems: analysis, stability and control*. New York: Springer-Verlag, 1999.

25. H. Khalil, *Nonlinear Systems*. NJ: Prentice-Hall, Upper Saddle River, Third Edition, 2002.
26. P. Kundur, *Power System Stability and Control*. McGraw-Hill, 1993.
27. G. Andersson, *Dynamics and Control of Electric Power Systems*. ETH Zürich, 2009.
28. G. F. Franklin, J. D. Powell, and A. Emami-Naeini, *Feedback Control of Dynamic Systems*. Prentice Hall, 2002.
29. P. Kundur, *Power System Stability and Control*. McGraw-Hill Inc., 1994.
30. I. Mitchell, "Application of level set methods to control and reachability problems in continuous and hybrid systems," *Stanford University, PhD thesis*, 2002.
31. A. Panagou, *Cyber-security issues in the Automatic Generation Control*. Semester thesis, Power System Laboratory, ETH Zurich, Switzerland, 2013.
32. T. Demiray, *Simulation of Power System Dynamics using Dynamic Phasor Models*. PhD thesis, Diss. ETH No.17607, ETH Zurich, Switzerland, 2008.
33. *Power Systems Test Case Archive*. College of Engineering, University of Washington, URL: http://www.ee.washington.edu/research/pstca/.
34. P. M. Anderson and A. A. Fouad, *Power System Control and Stability*. IEEE Computer Society Press, 2002.
35. P. M. Esfahani, M. Vrakopoulou, G. Andersson, and J. Lygeros, "A tractable nonlinear fault detection and isolation technique with application to the cyber-physical security of power systems," in *Decision and Control (CDC), 2012 IEEE 51st Annual Conference on*, Dec 2012, pp. 3433–3438.
36. P. M. Esfahani and J. Lygeros, "A tractable fault detection and isolation approach for nonlinear systems with probabilistic performance," Tech. Rep., Feb. 2013, [Online]. Available: http://control.ee.ethz.ch/index.cgi?page=publications&action=details&id=4344.