Resource allocation in networks via coalitional games

Shams, F.

Chapter 4

Power trading coordination in smart grids

In traditional power distribution models, consumers acquire power from the central distribution unit, while "micro-grids" in a smart power grid can also trade power between themselves. In this paper, we investigate the problem of power trading coordination among such micro-grids. Each micro-grid has a surplus or a deficit quantity of power to transfer or to acquire, respectively. A coalitional game theory based algorithm is devised to form a set of coalitions. The coordination among micro-grids determines the amount of power to transfer over each transmission line in order to serve all micro-grids in demand by the supplier micro-grids and the central distribution unit with the purpose of minimizing the amount of dissipated power during generation and transfer. We propose two dynamic learning processes: one to form a coalition structure and one to provide the formed coalitions with the highest power saving. Numerical results show that dissipated power in the proposed cooperative smart grid is only 10% of that in traditional power distribution networks.

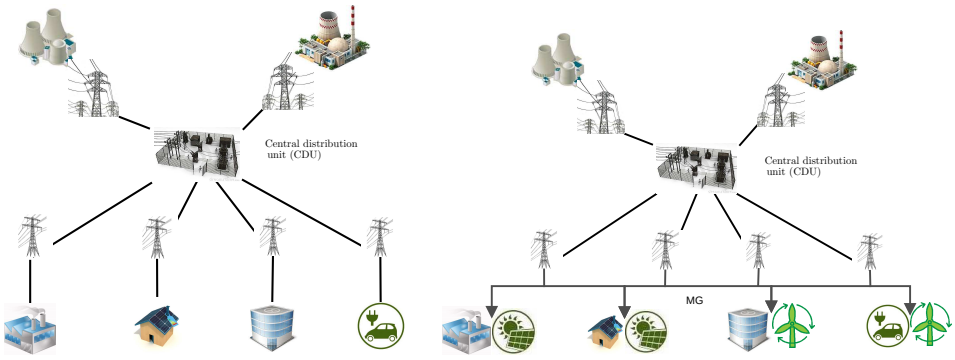
The remainder of the paper is structured as follows. After a brief motivation in the next section, Sect. 4.2 studies the circuit of transmission lines in a smart power grid. Sect. 4.3 consists of two subsections which contains the smart power network model, and models it as a coalitional game, respectively. The two subsections in Sect. 4.4 propose an innovative algorithm to form coalition and to maximize the performance in terms of power saving, respectively. Sect. 4.5 presents some simulation results and finally Sect. 4.6 concludes the paper.

4.1 Motivation

Interest in use of renewable electricity sources, the need for more efficient power distribution systems, lower energy delivery costs, and the improvement of reliability converge in the concept of the smart grid technology. The smart grid technology is an excellent candidate for exploiting information and communications technology (ICT) advantages at every line-powered device. The smart grid may be considered as the technological enabler of a variety of future applications that probably would not be available otherwise. The intelligence of smart grid relies upon the real-time communication either to or from the devices installed in homes and energy providers within the distribution and transmission grids.

While "macro-grids" were traditionally viewed as a technology used in remote area power supplies at a high-voltage, a "micro-grid" (MG) is introduced as a collective of geographically proximate, electrically connected loads and generators based on renewable energy technologies at a medium-voltage [290]. In general, an MG may or may not be connected to the wider electricity grid. Fig. 4.1 shows conceptual differences between the traditional grid, which is hierarchical, and a grid including MGs. In traditional electricity transmission, distribution networks act like the branches of the tree, interconnecting loads and the long-distance transmission network. The MG concept offers a path to autonomous, intelligent low-emissions electricity systems, by creating a localized smart grid that allows advanced and distributed control while being compatible with traditional electricity infrastructure. With such promise, MGs are of growing interest to grid operators, as a way of enhancing the performance of electricity systems [291]. However, realizing MGs is not without challenges, among which we will investigate the problem of coordination between MG electricity generators in order to minimize of the power dissipation over the transmission lines. The minimization of energy dissipation is and will be an important focus in electricity markets considering that power dissipated is accompanied with the cost of energy. Refs. [292–295] investigate the problem of power loss reduction in smart grids. Refs. [292, 293] focus on the optimization of various electricity parameters during peak load times.

Integration of renewable and distributed energy resources encompassing large scale at the transmission level, medium scale at the distribution level and small scale on commercial or residential building can present challenges for the controllability of



(a) Traditional power distribution (b) Smart power grid distribution

Fig. 4.1: *Traditional vs. smart power grid electricity distribution.*

these resources and for operation of the electricity system. Energy storage systems, both electrically and for thermally based, can alleviate such problems by decoupling the saving and delivery of energy. Smart grids can help through automation of control of generation and demand response to ensure balancing of supply and demand [296]. The potential economic impact of deploying power storage units and the possibility of having groups of storage and control units are studied in [297, 298].

Game theory [10] is a potential mathematical tool to model smart grid [298–300]. Non-cooperative game theory can model the distributed operations in smart power grids and cooperative/coalitional game theory can model the cooperation among nodes [301, 302]. To the best of our knowledge, in the existing literature there are only two game theory based algorithms with aim at minimization of power dissipated. Refs. [294, 295] propose coalitional game theory based algorithms which form a set of disjoint coalitions and enable MGs in each coalition to trade power among themselves with aim at minimization the overall power dissipated. The authors in [294] maximize the utility function of the grand coalition (coalition of all MGs), while the authors in [295] maximize Shapley values [48] of MGs. The results of [294] show that the proposed algorithm improves the performance reaching up to 31% (with 30 MGs) compared to the traditional transmission grids. Unfortunately, the performance of the proposed algorithm in [295] is evaluated with non realistic smart grid network area $10 \times 10\,\mathrm{km}^2$. It is obvious that in small areas the amount of power dissipated is

less than in a realistic network area $100\times100\,\mathrm{km}^2$. In addition, in Refs. [294, 295], the circuit of transmission lines is not realistic (We explain why in Sect. 4.2).

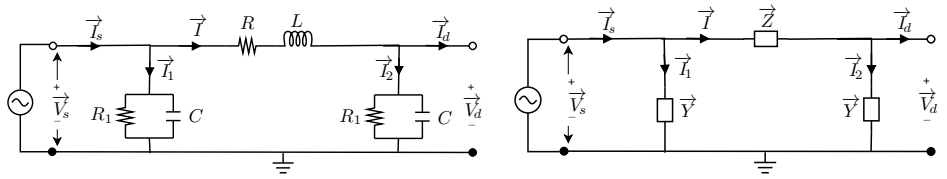
In this work, we investigate the problem of power trading coordination between MGs in smart power grids. Each MG has either a surplus or deficit quantity of power. Each MG with a surplus amount of power can transfer to the central distribution unit (CDU), and meanwhile it can serve MGs with deficit power. Each MG with a deficit amount of power can receive from the CDU and from MGs with surplus power. With the aim of power loss minimization during power generation and transfer, we will introduce a power trading coordination strategy among MGs. We formulate the problem using coalitional game theory in which MGs form a set of not necessarily disjoint and possibly singleton coalitions. MGs in each coalition can trade power between themselves and with the CDU. Each non-singleton coalition consists of one MG with surplus power which will serve the needed MG(s) with deficit power. On the other hand, the micro-grid in a singleton coalition only trades directly with the CDU.

For achieving the best power distribution over transmission lines, we propose two dynamic learning algorithms: 1) coalition formation dynamic learning, and 2) power loss minimization dynamic learning. The dynamic learning process 1) is nested in 2) one. In other words, the dynamic learning process 2) iteratively executes learning process 1) until it achieves a fixed point at which the performance can no longer improve. Coalition formation dynamic learning achieves a coalition formation structure and then the complementary power loss minimization dynamic learning leads the MGs to the maximum performance in terms of power saving. We show the stability (the convergence to a fixed-point) of both dynamic processes using the Kakutani fixed point theorem [303]. As our evaluation shows, our approach enables MGs to come to a power trading coordination among themselves that yields a significant power saving compared to the traditional power trading.

4.2 Transmission line model

Fig. 4.2(a) shows the single-phase equivalent[1] circuit of a medium-length (up to 200km) transmission line [290]. On an electricity transmission line from the sending-

[1]This circuit is suitable for analysing its symmetrical three-phase operation.



(a) Transmission circuit (b) π-equivalent circuit

Fig. 4.2: *Single-phase circuit of a medium-length transmission line.*

end s to the receiving-end d, the variables of interest are the (complex-valued) voltages and currents per phase $\overrightarrow{V_s}$, $\overrightarrow{V_d}$, $\overrightarrow{I_s}$, and $\overrightarrow{I_d}$ at the end-line terminals. The variables V_s and I_s denote the amplitudes of $\overrightarrow{V_s}$ and $\overrightarrow{I_s}$, respectively, and likewise for V_d and I_d. As the transmission line model in Fig. 4.2(a) has many parameters, it is convenient to replace each line with its π-equivalent shown in Fig. 4.2(b) where the impedance $\overrightarrow{Z} = R + i\omega L\ \Omega$ and the admittance $\overrightarrow{Y} = 1/R_1 + i\omega C\ \Omega^{-1}$ with i as the imaginary unit and $\omega = 2\pi f$ as the (sinusoidal) angular frequency in electrical radians at the frequency f.

Suppose the sending-end wants to supply a power amount P_s [W] and transfers it toward the receiving-end. A fraction of P_s will be dissipated during generation and transfer and the receiving-end d can load a power amount $P_d < P_s$. The real power loss is calculated by the following formula:

$$D_{sd} = P_s - P_d = V_s.I_s.\cos\theta_s - V_d.I_d.\cos\theta_d \quad [\text{W}] \tag{4.1}$$

where the parameter θ_s is the "power factor angle", in the range of $(-90°, +90°)$, and that is the phase difference between $\overrightarrow{V_s}$ and $\overrightarrow{I_s}$, and likewise for θ_d. So, the received power at the receiving-end d from the transmission line $s \to d$ can be formulated as $P_d = P_s - D_{sd}$.

In a transmission line, the amplitude of the voltage at sending-end V_s is known and we suppose $\overrightarrow{V_s}$ to be the reference vector, i.e, $\overrightarrow{V_s} = V_s\angle 0°$. We suppose also, as usual, that the power factor at the sending-end, $\cos\theta_s$, is known with the assumption that the current waveform comes delayed after the voltage waveform, i.e., $\overrightarrow{I_s} = I_s\angle -\theta_s$. For supplying and transferring P_s amount of power, the sending-end regulates the amplitude of the current as $I_s = \frac{P_s}{V_s.\cos\theta_s}$ and synchronizes its phase at $-\theta_s$. Knowing the parameters of the transmission line in the nominal π-circuit in Fig. 4.2(b), the

voltage and the current at the receiving-end are calculated by the following equations:

$$\begin{aligned}\overrightarrow{V_d} &= \left(1+\overrightarrow{Y}.\overrightarrow{Z}\right).V_s - \overrightarrow{Z}.\overrightarrow{I_s} \\ \overrightarrow{I_d} &= -\overrightarrow{Y}.\left(2+\overrightarrow{Y}.\overrightarrow{Z}\right).V_s + \left(1+\overrightarrow{Y}.\overrightarrow{Z}\right).\overrightarrow{I_s}\end{aligned} \tag{4.2}$$

Now, we can calculate the receiving-end's parameters V_d, I_d, and θ_d using (4.2) and then the amount of power loss is calculated as in (4.1). It is simple to show that the power loss equation (4.1) is a concave function of I_s when $L = C = 1/R_1 = 0$ (short-length transmission line model [290] as considered in Refs. [294, 295]), otherwise its shape strictly depends upon the values of the parameters.

4.3 System model and Problem formulation

4.3.1 System model

We study a smart power grid consisting of a single CDU which is connected to one or more main power plants at a high-voltage. The CDU is connected to K MGs denoted by the set $\mathcal{K}$. At any given time frame [299], each MG $k \in \mathcal{K}$ has a Q_k residual power load which is defined as the difference between the generated power and the overall demand. A positive quantity $Q_k > 0$ determines the surplus power that the MG can transfer to other MGs or to the CDU, whereas a negative quantity $Q_k < 0$ determines the deficit power the MG needs to acquire from other MGs or from the CDU. When $Q_k = 0$, the MG k meets its demanded power and it will not interact with any other MG or the CDU. We divide MGs into three groups: "suppliers" ($Q_k > 0$), "demanders" ($Q_k < 0$), and "inactives" ($Q_k = 0$) denoted by $\mathcal{K}^+$, $\mathcal{K}^-$, and $\mathcal{K}^0$, respectively, such that $\mathcal{K} = \mathcal{K}^+ \cup \mathcal{K}^- \cup \mathcal{K}^0$. In the rest of this paper, we assume $\mathcal{K}^0 = \emptyset$, $|\mathcal{K}^+| \geq 1$, and $|\mathcal{K}^-| \geq 1$.

Our goal is to develop an efficient power transfer policy between active MGs and the CDU themselves to minimize the overall power dissipation in the network. We will propose an algorithm which assigns to each supplier $s \in \mathcal{K}^+$ a subset of demanders in $\mathcal{K}^-$ and determines different fractions of Q_s to be transferred to each assigned demander and to the CDU. Doing so, traded power may be transferred on short distance lines. As values of each transmission line components (resistor, inductor, and capacitor) are an inverse function of the distance, the amount of dissipated power will be much lower than that in a traditional distribution. However, a MG can decide to

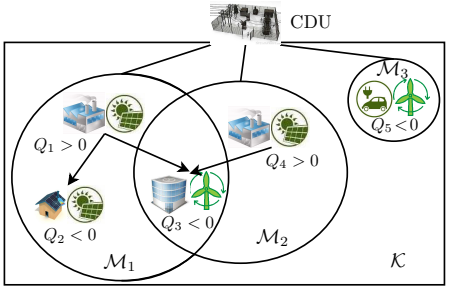


Fig. 4.3: *A sample of formed coalitions for cooperative power distribution.*

act as a non-cooperative MG which trades only with the CDU. Each MG in demand can be assigned to more suppliers to acquire the whole or a fraction of its own needed power. Two suppliers will not belong to the same coalition, since they do not trade power between themselves.

4.3.2 Problem formulation

To study the cooperative behavior of the MGs, we use a coalitional game theory framework [301]. A coalitional game is defined as $\mathcal{G} = (\mathcal{K}, \nu)$ where $\mathcal{K}$ is the player's set (the active MGs), and $\nu : 2^{\mathcal{K}} \longrightarrow \mathbb{R}$ is the characteristic function of each coalition (subset of $\mathcal{K}$) that assigns a real number representing the benefit earned by the coalition. We will propose an algorithm which forms a set of not necessarily disjoint and possibly singleton coalitions denoted by $\boldsymbol{\mathcal{M}} = \{\mathcal{M}_1, \ldots, \mathcal{M}_M\}$. The (unique) MG in a singleton coalition will trade only with the CDU. A non-singleton formed coalition consists of *only* one supplier and one or more demanders, i.e., the unique supplier MG provides the whole or a fraction of the overall power demanded by the assigned demander(s). Each demander in $\mathcal{K}^-$ may belong to more coalitions, i.e., the whole or a fraction of its needed power can be provided by more suppliers. Then, all MGs can trade with the CDU for the rest, if any. Fig. 4.3 shows a sample wherein two cooperative coalitions and one non-cooperative singleton coalition are formed. For instance, MG1 can decide to transfer 10% of its (surplus) quantity to MG2 and 30% to MG3 and the rest to the CDU. MG4 can decide to transfer 50% of its quantity to MG3 and the rest to the CDU. The rest of deficit quantities of MG3 and MG4 will be provided by the CDU. MG5 will be completely served by the CDU.

A coalition structure will be formed only if (i) surplus power quantities of all suppliers are completely loaded, and (ii) deficit power quantities of all demanders are completely served, as the following equations state:

$$\begin{cases} \sum_{d \in \mathcal{K}^-} P_{sd} + P_{s0} = Q_s \quad \forall s \in \mathcal{K}^+ & \text{(4.3a)} \\ \sum_{s \in \mathcal{K}^+} (P_{sd} - D_{sd}) \; + \; (P_{0d} - D_{0d}) = -Q_d \quad \forall d \in \mathcal{K}^- & \text{(4.3b)} \end{cases}$$

where the subscript 0 denotes the CDU index. The parameter P_{ij} denotes the amount of loaded power by the sending-end i as $P_{ij} = V_{ij}.I_{ij}.\cos\theta_{ij}$ to transfer over the transmission line $i \to j$. We suppose that the power network is *meshed*, i.e. each sending-end can regulate a different voltage, current, and power factor angle over each connected transmission line. The parameter D_{ij} is the amount of power dissipated, calculated as in (4.1) while the sending-end terminal i loads P_{ij} over the transmission line $i \to j$. Condition (4.3a) guarantees that the power quantity of each supplier $s \in \mathcal{K}^+$ is completely loaded over the connected transmission lines and condition (4.3b) guarantees that each demander $d \in \mathcal{K}^-$ receives the whole deficit quantity (demanded) power.

We divide the MGs in each formed coalition $\mathcal{M}_m$ into two subsets of suppliers and demanders denoted by $\mathcal{M}_m^+$ and $\mathcal{M}_m^-$, respectively. For a singleton coalition $\mathcal{M}_m$, either $\mathcal{M}_m^+$ or $\mathcal{M}_m^-$ is empty. In each non-singleton coalition $\mathcal{M}_m$, $\mathcal{M}_m^+$ is singleton since the supplier MG is unique. For each coalition $\mathcal{M}_m$ with $\mathcal{M}_m^+ = \{s\}$, we define the characteristic function as the inverse of total dissipated power over the distribution lines incurred by the power generation and transfer, as:

$$\nu(\mathcal{M}_m) = \left(\sum_{d \in \mathcal{M}_m^-} D_{sd} + D_{s0} + \sum_{d \in \mathcal{M}_m^-} D_{0d} \right)^{-1} \tag{4.4}$$

wherein the power minus accounts for the maximization problem, term D_{sd} refers to the power dissipated incurred by the generation and transfer of P_{sd} from the supplier s to each $d \in \mathcal{M}_m^-$, the term D_{s0} refers to the power dissipated incurred by the transfer of the power amount P_{s0} from the supplier s to the CDU, and D_{0d} refers to the power dissipated incurred by the power transfer from the CDU to the demanders in $\mathcal{M}_m^-$. For a singleton coalition which consists of one non-cooperative supplier MG, terms D_{sd} and D_{0d} are equal to zero and term D_{s0} is calculated with setting $P_{s0} = Q_s$.

On the other hand, for a singleton coalition which consists of one demander MG, terms D_{sd} and D_{s0} are equal to zero and the term D_{0d} is calculated with setting $P_{0d} = -Q_d + D_{0d}$, i.e., the demanded power $-Q_d$ will be served only by the CDU. We assign $\nu(\mathcal{M}_m) = -\infty$ when the power distribution conditions (4.3) do not hold. We assume the unit of currency as the price of unit amount of power. It is worthwhile to note that in the proposed framework it could also be considered coefficient minus instead of power minus in (4.4), since the power loss is neither a concave nor convex function.

For the grand coalition, the characteristic function is the inverse of the total amount of power loss in the entire network as the following formula:

$$\nu(\mathcal{K}) = \left(\sum_{s \in \mathcal{K}^+} \sum_{d \in \mathcal{K}^-} D_{sd} + \sum_{s \in \mathcal{K}^+} D_{s0} + \sum_{d \in \mathcal{K}^-} D_{0d} \right)^{-1} \tag{4.5}$$

We assign $\nu(\mathcal{K}) = -\infty$ when the power distribution conditions (4.3) do not hold.

A central question in a coalitional game is how to divide the earnings among the members of the formed coalition. The payoff of each MG (player in the game) can show the power of influence of the MG and it is obvious that a higher payoff is an incentive to cooperate more efficiently. The Shapley value [48] assigns a unique outcome to each MG. Let us denote ϕ_k as the Shapley value of MG $k \in \mathcal{K}$ in the game. For each MG $k \in \mathcal{K}$:

$$\phi_k = \sum_{\substack{\forall \mathcal{M}_m \subseteq \mathcal{K} \\ \text{s.t. } k \in \mathcal{M}_m}} \frac{(|\mathcal{M}_m| - 1)! \cdot (K - |\mathcal{M}_m|)!}{K!} \left(\nu(\mathcal{M}_m) - \nu(\mathcal{M}_m \backslash \{k\}) \right) \tag{4.6}$$

The expression $\nu(\mathcal{M}_m) - \nu(\mathcal{M}_m \backslash \{k\})$ is the marginal payoff of the MG k to the coalition $\mathcal{M}_m$. The Shapley value can be interpreted as the marginal contribution an MG makes, averaged across all permutations of MGs that may occur.

4.4 Best response algorithm

In this section we provide an answer to the question: *How do the MGs form the best coalition structure with aim at minimization of the amount of power dissipated?* Dynamic learning models provide a framework for analyzing the way players set their proper strategies. To address this question, we propose an "adaptive learning

algorithm" [304] in which the learners use the same learning algorithm. Typically, these types of algorithms are constructed to iteratively play a game with an opponent, and, by playing this game, to converge a solution. Forming a coalition structure by MGs is equivalent to distribute $Q_k \ \forall\, k \in \mathcal{K}$ over the transmission lines satisfying conditions (4.3). Obviously the solution is not unique. In the next subsections, we first propose a dynamic learning process to form a coalition structure where each formed coalition earns a positive bounded payoff and each MG earns a bounded Shapley value. We then introduce another complementary dynamic process which iteratively executes the coalition formation dynamic learning process in order to choose the best coalition structure with minimum power dissipated.

4.4.1 Coalition formation dynamic learning process

To realize a coalition formation structure, we use dynamic learning. Our goal is to determine the best amount of power loaded at the sending-end sides of the unidirectional transmission lines denoted by $\mathcal{L} = \{l_{sd}\} \cup \{l_{s0}\} \cup \{l_{0d}\}$, $\forall\, s \in \mathcal{K}^+$ and $\forall\, d \in \mathcal{K}^-$ where the symbol l_{ij} denotes the transmission line from the sending-end i to the receiving-end j. We denote $L = |\mathcal{L}| = |\mathcal{K}^+| \cdot |\mathcal{K}^-| + |\mathcal{K}^+| + |\mathcal{K}^-|$ as the number of individuals. We denote the parameters of the sending-end i over the transmission line l_{ij} (to the receiving-end j) by I_{ij}, V_{ij}, P_{ij}, and θ_{ij}. We will propose an algorithm to determine the best value of P_{ij}. The values of V_{ij} and θ_{ij} are known at the sending-end i and it will regulate the electricity current as $I_{ij} = \frac{P_{ij}}{V_{ij} \cos \theta_{ij}}$ to load P_{ij} toward the receiving-end j.

In our dynamic learning process, each transmission line $l_{ij} \in \mathcal{L}$ is an individual learner which, during learning process learns about the best amount of the power P_{ij} to be loaded by the sending-end i to transfer over the transmission line l_{ij}. By doing so: 1) each supplier $s \in \mathcal{K}^+$ will be aware of the amount of power to load over the transmission lines toward all demanders and the CDU, i.e., over l_{s0} and $\{l_{sd}\}$ $\forall\, d \in \mathcal{K}^-$, and 2) the CDU will be aware of the amount of power to load over the transmission lines toward all demanders, i.e., over $\{l_{0d}\}$ $\forall\, d \in \mathcal{K}^-$. If one supplier s loads a power $P_{sd} > 0$ over the transmission line l_{sd}, the MGs s and d will belong to the same coalition. If one supplier s loads $P_{s0} = Q_s$ over l_{s0}, the supplier s will form a singleton coalition. One demander $d \in \mathcal{K}^-$ will form a singleton coalition if it receives the whole $-Q_d$ from the CDU, i.e., $P_{0d} - D_{0d} = -Q_d$.

For each P_{ij}, let us denote the maximum power constraint and the best power amount by $\bar{P}_{ij}$ and P^*_{ij}, respectively. We set $\bar{P}_{sd} = \min\{Q_s, -\beta Q_d\}$, $\bar{P}_{s0} = Q_s$, and $\bar{P}_{0d} = -\beta Q_d$ with $\beta > 1$ for $\forall\, s \in \mathcal{K}^+$ and $\forall\, d \in \mathcal{K}^-$. About $\bar{P}_{s0}$, it is obvious that loaded power by a supplier cannot be greater than its residual power load. With respect to $\bar{P}_{0d}$, whenever the CDU wants to completely serve one demander d, it must load a power amount $P_{0d} > -Q_d$ to overcome power loss over the transmission line $0 \to d$ and so to guarantee received power $-Q_d$ at the proper demander side. Now, the setting of $\bar{P}_{sd}$ is obvious. As is apparent, choosing an appropriate value for β strictly depends upon the parameters of the transmission lines and the value of Q_ks. In our simulation in Sect. 4.5, we set $\beta = 2$, i.e., the CDU should load at most $-2Q_d$ for completely serving the proper demander (the maximum power loss over the line $0 \to d$ is $-Q_d$). Instead of using the parameter β, the appropriate solution is to answer this question: How much must be the load power P_s at the sending-end s, if it wants to guarantee a power amount P_d at the receiving-end d? Multiplication the both sides of $\overrightarrow{V_d}$ in (4.2) to the correspondence sides of $\overrightarrow{I_d}$ approaches a quadratic equation of I_s with complex coefficients for whose roots, in general case, we could not find a presentable algebraic expression. For the special case of $C = L = 1/R_1 = 0$, we refer the reader to [294, 295].

During the learning process, in each (discrete) time step t every learner $l_{ij} \in \mathcal{L}$ individually and distributively updates its temporary power value P^t_{ij} in a myopic manner, i.e., supposing that all other learners $\mathcal{L}\backslash\{l_{ij}\}$ are inactive and their power values are fixed. In the following, we propose a method to distributively update the power values P^t_{ij} in order to lead it to the best value and we will show that the updating process has a fixed point, i.e., for a given t large enough $P^{t+1}_{ij} = P^t_{ij}$ $\forall\, l_{ij} \in \mathcal{L}$, and then we set $P^*_{ij} = P^{t+1}_{ij}$.

The power value of learners l_{sd} appear in both conditions (4.3), whereas that of l_{s0} only in (4.3a) and that of l_{0d} only in (4.3b). From conditions (4.3), the following

equations are derived:

$$\begin{cases} P_{ij} = C_{ij}\,; & \forall\, l_{ij} \in \mathcal{L} \\ C_{s0} = Q_s - \sum_{d \in \mathcal{K}^-} P_{sd} & \forall\, l_{s0} \in \mathcal{L}\,; \qquad //\forall\, s \in \mathcal{K}^+ \\ C_{0d} = -Q_d - \sum_{s \in \mathcal{K}^+} P_{sd} + D_{0d} + \sum_{s \in \mathcal{K}^+} D_{sd} & \forall\, l_{0d} \in \mathcal{L}\,; \qquad //\forall\, d \in \mathcal{K}^- \\ C_{sd} = 0.5 \left(Q_s - Q_d - \sum_{d \neq d' \in \mathcal{K}^-} P_{sd'} - \sum_{s \neq s' \in \mathcal{K}^+} P_{s'd} - P_{s0} - P_{0d} + D_{0d} + \sum_{s' \in \mathcal{K}^+} D_{s'd} \right) \\ \qquad\qquad \forall\, l_{sd} \in \mathcal{L}. \qquad //\forall\, s \in \mathcal{K}^+,\ \forall\, d \in \mathcal{K}^- \end{cases} \tag{4.7}$$

According to (4.7), the rationality of each learner $l_{ij} \in \mathcal{L}$ is focusing on updating the proper value P_{ij} in order to achieve the value of C_{ij} *exactly*, i.e., $P_{sd} = C_{sd}$, $P_{s0} = C_{s0}$, and $P_{0d} = C_{0d}$. At each time step t, the learner l_{ij} will update the proper power value P_{ij}^t in a myopic manner, i.e., supposing that C_{ij}^t is a constant. To this end, we define the following learning utility function for each learner $l_{ij} \in \mathcal{L}$:

$$\begin{cases} u_{ij} = -\sqrt{|P_{ij} - C_{ij}|} + \alpha \cdot \mathrm{u}\,(P_{ij} - C_{ij})\,; & 0 \leq C_{ij} \leq \bar{P}_{ij} \\ u_{ij} = -\alpha & \text{otherwise.} \end{cases} \tag{4.8}$$

where $\mathrm{u}\,(\cdot)$ is the step function, with $\mathrm{u}\,(y) = 1$ if $y \geq 0$ and $\mathrm{u}\,(y) = 0$ otherwise (see Fig. 4.4), and α is a constant. When the value C_{ij} is out of the range of $[0, \bar{P}_{ij}]$, the proper learner will gain the minimum possible payoff, $-\alpha$. For a $C_{ij} \in [0, \bar{P}_{ij}]$, if $P_{ij} = C_{ij}$, the learner l_{ij} earns the highest possible payoff $u_{ij} = \alpha$. Whereas when $P_{ij} \neq C_{ij}$, the learner l_{ij} gets a payoff lower than α. The factor α is a sufficiently large positive constant that ensures u_{ij} to be positive when $P_{ij} > C_{ij}$. This is an expedient to let the learners distinguish the value of P_{ij} that is either smaller or larger than C_{ij} only by knowing their own payoffs. In practice, it is obvious that the payoff of each learner is bounded from below by the finite negative number $-\alpha$. Setting an appropriate value to α in order to guarantee a positive u_{ij} when $P_{ij} \geq C_{ij}$ is strictly dependent upon the values of Q_k. Note that the proposed framework could also consider any bounded utility function which increases as its argument moves from $\pm\infty$ to 0. This means that, for any $P_{ij} \neq C_{ij}$, each learner l_{ij} has an incentive to move towards the point zero where $P_{ij} = C_{ij}$ and the learner gains the highest possible payoff. When all learners gain α, then all conditions (4.3) hold, and obviously $P_{ij} = P_{ij}^*$.

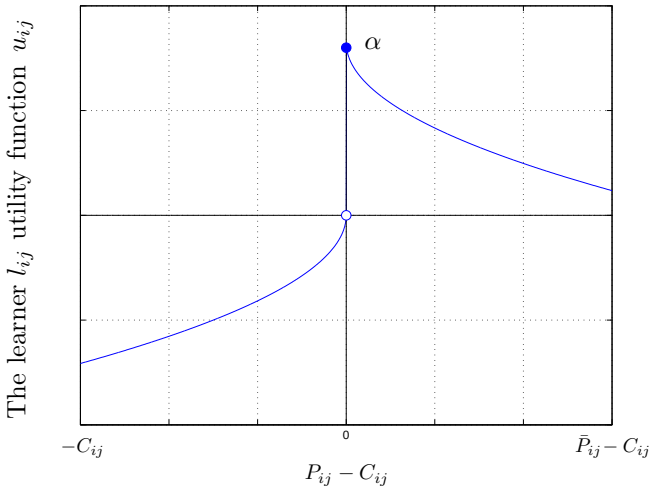


Fig. 4.4: *Learner utility as a function of* $P_{ij} - C_{ij}$ *with* $0 \leq C_{ij} \leq \bar{P}_{ij}$ *and* $\alpha \gg 0$.

The pseudocode in Tab. 4.1 shows how each learner l_{ij} takes its myopic decision during time step t. In this algorithm, sign$(\cdot)$ is our sign function, with sign$(y) = 1$ if $y \geq 0$ and sign$(y) = -1$ otherwise. The parameter $\tilde{u}_{ij}$ is the "trial" value of the current payoff of the learner l_{ij} when the tentative power $\tilde{P}_{ij}$ is adopted. As each learning time step, the power step $\Delta\tilde{P}_{ij}$ is the particular outcome (value) of a random variable uniformly distributed between 0 and $\overline{\Delta P}_{ij}$, with $\overline{\Delta P}_{ij} \ll \bar{P}_{ij}$. Optimal values for $\overline{\Delta P}_{ij}$ can be found in order to minimize the computational load of the algorithm, based on experimental results. If $u_{ij}^t < 0$, then $P_{ij}^t < C_{ij}$, and the best strategy for the learner l_{ij} is to increase its power so as to increase its payoff. Consequently, the tentative power is a random number in the interval $[P_{ij}^t, \bar{P}_{ij}]$. The learner l_{ij} accepts this value if and only if the utility u_{ij}^t increases, otherwise it ends up to keep its previous value. If $0 \leq u_{ij}^t < \alpha$, the learner l_{ij}'s best strategy is on the contrary to decrease P_{ij}^t, and thus the tentative (random) power level belongs to the interval $[0, P_{ij}^t]$. As is apparent, the convergence speed of the algorithm depends also on the choice of the maximum update step $\overline{\Delta P}_{ij}$.

The process starts at time step $t = 0$ with $P_{ij}^{t=0} = 0$ for all learners in $\mathcal{L}$. Thus, at $t = 0$, we have $C_{s0}^{t=0} = Q_s = \bar{P}_{s0} \;\; \forall s \in \mathcal{K}^+$ and $C_{0d}^{t=0} = -Q_d < \bar{P}_{0d} \; \forall d \in \mathcal{K}^-$, i.e.,

```
function P_ij^{t+1} = Powerupdate(l_ij, t)
   if (u_ij^t = α) or (u_ij^t = −α), then P_ij^{t+1} = P_ij^t, exit;
   P̃_ij = P_ij^t; //saving the current power
   repeat
      P̂_ij = P̃_ij; //saving tentative power
      compute ũ_ij; //tentative payoff
      ΔP̃_ij = unif [0, ΔP̄_ij]; //random power step
      P̃_ij = P̃_ij − sign(u_ij^t) · ΔP̃_ij; //tentative power
   until (ũ_ij > u_ij^t) or (P̃_ij > P̄_ij) or (P̃_ij < 0)
   if (ũ_ij > u_ij^t), then P_ij^{t+1} = P̂_ij; //accept
                       else P_ij^{t+1} = P_ij^t; //discard
end function.
```

Tab. 4.1: *Myopic decision of the learner l_{ij} at time step t.*

$u_{s0}^{t=0}$ and $u_{0d}^{t=0}$ are in the range of $(-\alpha, 0)$. Depending on Q_s, $-Q_d$, and β, the values of $C_{sd}^{t=0}$ can be either in or out of the range of $[0, \bar{P}_{sd}]$. At $t = 1$, each learner l_{s0} and l_{0d} may increase the proper power amount which results that some of $C_{sd}^{t=1} < C_{sd}^{t=0}$. After some time steps, some of the learners l_{sd} will have $0 \leq C_{sd}^{t} \leq \bar{P}_{sd}$ and they can update the proper powers.

Each learner l_{ij} individually decides to adjust its power value P_{ij}^t in a myopic manner while supposing that the value of C_{ij}^t is fixed and all other learners are inactive. The value of C_{ij}^t depends upon the power values of other learners and therefore the decision of adjusting the power value by each learner influences other C values in conflicting and incompatible ways which prevent learners from gaining the expected payoff. At each time step t, adjusting P_{ij}^t changes all C_{s0}^t and C_{0d}^t $\forall s \in \mathcal{K}^+$ and $\forall d \in \mathcal{K}^-$. To reduce the number of occurrences of this event, we modify our algorithm by requesting each learner l_{ij} *not* to update its power value at every time step with a probability $\lambda_{ij} \in [0, 1]$. At each time step t, every learner l_{ij} picks a random number ξ_{ij}^t uniformly distributed in $[0, 1]$. If $\xi_{ij}^t > \lambda_{ij}$, then the learner applies the algorithm and possibly update P_{ij}^{t+1}, otherwise $P_{ij}^{t+1} = P_{ij}^t$. Note that each learner l_{ij} is aware of the value of λ_{ij}. The algorithm is executed in a central computing unit [9], e.g. the CDU, and it can transmit the best amount of power load over each transmission line to MGs.

We show now that our proposed algorithm reaches a fixed point at which $P_{ij}^t = C_{ij}^t$ $\forall\, l_{ij} \in \mathcal{L}$. We model the evolution of the algorithm as the output of a Markov chain

with state space $\Omega = \left\{\omega = (\boldsymbol{u}) | \, \boldsymbol{u} \in [-\alpha, \alpha]^L\right\}$. At time step t, our system is in the state $\omega^t = (\boldsymbol{u}^t)$, where $\boldsymbol{u}^t$ is the set $\left\{u_{ij}^t\right\} \ \forall \, l_{ij} \in \mathcal{L}$. For all time steps t, since u_{ij} is continuous and bounded in $[-\alpha, \alpha]$, $\boldsymbol{u}^t$ is bounded in the L-dimensional space $[-\alpha, \alpha]^L$. So, $\boldsymbol{u}^t$ is compact and Lebesgue measurable. The evolution of the Markov chain is then dictated by the strategy of the learning process. The strategy of each learner l_{ij} is to find the best power amount P_{ij}^t that leads to an increase its own payoff u_{ij}^t. In practice, each learner l_{ij} autonomously decides whether to change its power value P_{ij}^t, making its payoff better off, or to keep the power at the same power level (when its payoff is equal to α or $-\alpha$). The transition from the state $\omega = (\boldsymbol{u})$ to a new state $\breve{\omega} = (\breve{\boldsymbol{u}})$ occurs if and only if the new state $\breve{\omega}$ "dominates" the state ω, i.e., compared to $\boldsymbol{u}$, no learner gets worse off in $\breve{\boldsymbol{u}}$. The CDU computes $C_{ij}^{t+1} \ \forall \, l_{ij}$ using the values of P_{ij}^{t+1} and then computes all utility payoffs $\boldsymbol{u}^{t+1}$. If the transition from the state ω^t to the state ω^{t+1} occurs, then the CDU communicates to all learners sending them the proper C_{ij}^{t+1}, otherwise announce them to keep the previous power amount, i.e., $P_{ij}^{t+1} = P_{ij}^t$. The Markov process asymptotically tends towards a stable power distribution state at which no learner has any incentive to change its power value. In other words, all learners get their maximum payoffs, $\boldsymbol{u} = \{\alpha\}^L$, and consequently $P_{ij}^t = C_{ij}^t \ \forall \, l_{ij} \in \mathcal{L}$, and then, obviously, no learner has any incentive to update its power value. Our algorithm guarantees tending the Markov chain towards a fixed point state with probability one when $t \to \infty$. Obviously, the stable state is not unique and according to the way the learners generate random numbers, the algorithm leads them to one of the possible solutions and then the power values are no longer updated.

Theorem 7 *The coalition formation learning process converges a stable state.*

Proof Denote by $\mathcal{P}$ the set of all possible $\boldsymbol{u}^t$'s. Obviously $\mathcal{P} = [-\alpha, \alpha]^L$ is not empty and it is Lebesgue measurable, compact and convex set. We construct a correspondence $\Gamma : \mathcal{P} \to \mathcal{P}$ with $\Gamma(\boldsymbol{u}) = \{\breve{\boldsymbol{u}}\}$ where $\{\breve{\boldsymbol{u}}\}$ is the set of all possible next states of the (current) state $\boldsymbol{u}$ in the proposed best-response algorithm, i.e., the states $\breve{\boldsymbol{u}}$s dominates $\boldsymbol{u}$. According to the best-response algorithm, the transition probability from $\boldsymbol{u}$ to one member of $\{\breve{\boldsymbol{u}}\}$ is one and zero to the others members. Show that a fixed point exists. Now we claim that there is upper hemi-continuity (uhc)[2] for a given $\boldsymbol{u}^t$. To this end, let $\boldsymbol{u}^{-t}$ (resp. $\breve{\boldsymbol{u}}^{-t}$), for t large enough, be some sequences in $\mathcal{P}$

[2]There is uhc for $f : X \to X$ when the graph of f is close i.e., for all sequences $\{x_n\}$ and $\{y_n\}$ such that $y_n \in f(x_n)$ for all n, $x_n \to x$, and $y_n \to y$, we have $y \in f(x)$ [10].

converging to the current $\boldsymbol{u}^t$ (resp. $\check{\boldsymbol{u}}^t$). Best response transition rule confirms that : $\check{\boldsymbol{u}}^{t-1} = \Gamma(\boldsymbol{u}^{t-1}) = \boldsymbol{u}^t$. Study such correspondence sequence in best-response process confirms that: $\check{\boldsymbol{u}}^{t-1} = \boldsymbol{u}^t = \Gamma(\ldots\Gamma(\Gamma(\boldsymbol{u}^{t=0})))$ after t times of Γ operation. This is somehow obvious to claim that $\check{\boldsymbol{u}}^t \in \Gamma(\boldsymbol{u})$. Then, by the arguments above, all the conditions for the Kakutani fixed point theorem [303] are satisfied, and there exists $\boldsymbol{u}^* \in \mathcal{P}$ such that $\boldsymbol{u}^* \in \Gamma(\boldsymbol{u}^*)$ which corresponds to the best value of $P_{ij}^* \ \forall\, l_{ij} \in \mathcal{L}$.■

Note that, during the learning process for achieving stable state, the payoffs of all coalitions are equal to $-\infty$ since the conditions (4.3) are not still satisfied. Once all learners achieve the coalition formation stable state, the conditions hold a possible coalition structure $\boldsymbol{\mathcal{M}}$ is formed, but the power dissipated is not necessarily the minimum one. In the next subsection, we propose another "complementary" Markov model which, in each of its time steps τ, executes coalition formation dynamic process and calls the achieved stable state $\omega = (\boldsymbol{u})$. The evolution during time τ leads to a network with the minimum amount of power dissipated.

4.4.2 Power loss minimization dynamic learning process

At the stable state of the coalition formation process, conditions (4.3) hold, a possible coalition structure $\boldsymbol{\mathcal{M}}$ is formed, and each coalition earns a limited payoff. The payoff of each coalition is the inverse of total power dissipated and so higher payoffs for coalitions means lower amount of power dissipated. Obviously, the coalitions' payoffs strictly depend upon the distribution of power over transmission lines, i.e., the achieved coalition formation stable state $\omega = (\boldsymbol{u})$. In general, the payoffs of coalitions and MGs is neither a convex nor a concave function of the loaded power over transmission lines. Hence, it is not possible to find an analytical solution for finding the absolute best payoffs of coalitions. With aim at minimization of the overall power dissipated, we propose a complementary dynamic learning process. In a similar fashion as the coalition formation dynamic learning, we model the evolution of these learning processes as a Markov model $\Pi = \left\{\pi = (\boldsymbol{u}, \rho) \mid \ \boldsymbol{u} = \{\alpha\}^L, \ \rho \subset \mathbb{R}^K\right\}$ wherein $\omega = (\boldsymbol{u})$ is the achieved coalition formation stable state. For the second input ρ, we will consider one of the characteristics of the network which is limited and Lebesgue measurable. In the simulation results, we will consider the following two values for ρ and we will compare the performance of them:

(i) We will set $\rho = \nu(\mathcal{K}) \in \mathbb{R}$. In fact, in this case we disregard the coalition

structure and MGs payoffs, but we aim at maximization of $\nu(\mathcal{K})$ which is, according to (4.5), limited and Lebesgue measurable.

(*ii*) We will set $\rho = \boldsymbol{\phi} \in \mathbb{R}^K$ where $\boldsymbol{\phi} = \{\phi_k\}$ denotes the MGs Shapley values. In this case, the structure of the coalitions and their payoffs are important.

In the simulation results we will show that maximizing Shapley values outperforms maximizing $\nu(\mathcal{K})$ in terms of power saving. At each time step τ, the central computing unit, first, executes the coalition formation dynamic learning. Once the coalition formation learning achieves a stable state, the complementary Markov model evolution is in the state $\pi^\tau = (\boldsymbol{u}^\tau, \rho^\tau)$, i.e., either $\pi^\tau = (\boldsymbol{u}^\tau, \nu(\mathcal{K})^\tau)$ if we choose (i), or $\pi^\tau = (\boldsymbol{u}^\tau, \boldsymbol{\phi}^\tau)$ if (ii), where $\boldsymbol{u}^\tau$ is the stable state of the coalition formation dynamic learning. At the next time step $\tau + 1$, the CDU re-executes the coalition formation dynamic learning picking different random numbers and achieves a new stable state $\boldsymbol{u}^{\tau+1}$ with a different power distribution and coalition structure. The transition from π^τ to $\pi^{\tau+1}$ occurs if and only if $\rho^{\tau+1}$ dominates ρ^τ, otherwise the Markov model Π stays at the same state π^τ. If we set $\rho = \nu(\mathcal{K})$ the transition occurs when the grand coalition earns a higher payoff, whereas with $\rho = \boldsymbol{\phi}$ it occurs when no MG will get worse off. At time step $\tau + 1$, if the coalition formation learning process happens to achieve the same $\boldsymbol{u}^\tau$, again the complementary Markov model stays at the same state π^τ. Since the values of ρ are limited, with a same discussions in Theorem 7, it is easy to show that the evolution of Π achieves a fixed point for ρ, i.e., for a given time step τ large enough, $\rho^{\tau+1} = \rho^\tau$ for a $\boldsymbol{u}^{\tau+1} \neq \boldsymbol{u}^\tau$. At the fixed point of the complementary dynamic learning process, in case (i) the grand coalition earns the highest possible payoff, and in case (ii) all MGs in $\mathcal{K}$ earn the highest possible Shapley value ϕ_k. This means that the amount of power dissipated is the lowest possible. Obviously, this algorithm can not guarantee the absolute minimum power loss, since the shape of the power loss function strictly depends upon the parameters of the transmission line. For this, the power loss minimization stable point is not necessarily unique and the proposed learning process leads micro-grids to one of the stable points. For the reader's convenience, the power loss minimization algorithm is summarized in Tab. 4.2 and for better readability, the coalition formation dynamic process part is colored blue.

Initialization:

//Initialization for $\Pi = \{\pi = (\boldsymbol{u}, \rho)\}$ *Markov model:*

set $\tau = 0$, $\rho^{\tau} = -\infty$, and a tolerance ε;

Power loss minimization process:

```
while 1
    //Coalition formation dynamic process:
    set t = 0, P_ij^t = 0;  ∀ l_ij ∈ L  //Initialization for Ω = {ω = (u)} Markov model
    compute C_ij^t;  ∀ l_ij ∈ L
    compute u^t;
    repeat
        execute  P_ij^{t+1} = Powerupdate(l_ij, t);  ∀ l_ij ∈ L      //parallel decisions
        t = t + 1;
        compute C_ij^t;  ∀ l_ij ∈ L
        compute u^t;
        if  ( u^t dominates u^{t-1} )  then  //ω^t = (u^t)
            ω^{t-1} ⟶ ω^t;      //the transition occurs
        else
            ω^t = ω^{t-1};      //discard:  P_ij^t = P_ij^{t-1} and then C_ij^t = C_ij^{t-1}, u^t = u^{t-1}
        end
    until min_{l_ij ∈ L} ((−ε ≤ u_ij^t < 0)  or  (u_ij^t ≥ α − ε));
    // ω^t = (u^t) is the stable state of the coalition formation dynamic learning process
     τ = τ + 1;
     u^τ = u^t;
    compute ρ^τ;    //either ρ = ν(K) or ρ = φ
    if  ( ρ^τ dominates ρ^{τ-1} )  then
        π^{τ-1} ⟶ π^τ;   //the transition to π^τ = (u^τ, ρ^τ) occurs
        if (| ρ^τ − ρ^{τ-1} | ≤ ε) then
            break;    //stable state
        end
    else
         π^τ = π^{τ-1};   //discard:  transition does not occur
    end
end
//π^τ is the stable state of the power loss minimization dynamic learning process.
```

Tab. 4.2: *Power loss minimization algorithm.*

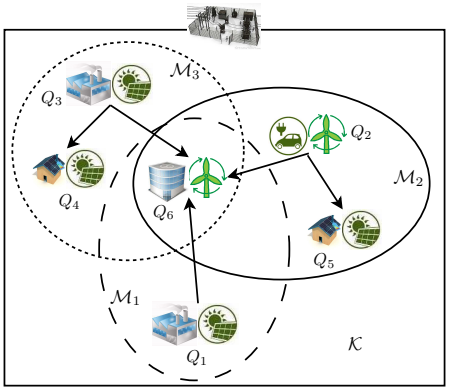


Fig. 4.5: *Snapshot of a formed coalition structure.*

4.5 Numerical results

In this section, we show the performance of the proposed algorithm. Throughout the simulations, we make use of the following transmission lines parameters: $\omega C = 4.518\ \mu\Omega^{-1}$/km, $\omega L = 367$ mΩ/km, $R = 37$ mΩ/km, and $R_1 = 1\ \Omega$ [290, p. 68]. Power transfer among MGs themselves is done at a medium voltage 100 kV, while between the central transmission unit and the MGs that is done at 345 kV [290, p. 68]. The MGs are uniform randomly distributed within an area with radius 100 km and the CDU located at the center. For each MG $k \in \mathcal{K}$, the amount of residual power Q_k is supposed to be a normal random distribution with zero mean and a standard deviation uniformly distributed in the range [100, 500] MW [291]. In the following set of evaluations, based on numerical optimizations, we consider the following parameters in the function Powerupdate: the power update step $\overline{\Delta P}_{sd} = \bar{P}_{sd}/5$, $\overline{\Delta P}_{s0} = \bar{P}_{s0}/10$, and $\overline{\Delta P}_{0d} = \bar{P}_{0d}/10$; the parameter that reduces the probability of conflicting decisions among learners $\lambda_{sd} = 0.7$, $\lambda_{s0} = 0.85$, and $\lambda_{0d} = 0.85$ for $\forall\ s \in \mathcal{K}^+$ and $\forall\ d \in \mathcal{K}^-$. All results are obtained by averaging over $2,000$ random realizations of a network with different positions and different values of Q_ks.

Fig. 4.5 reports a snapshot of the achieved coalition formation dynamic learning process stable state in a network consisting of $K = 6$ MGs randomly located in an area with residual quantities Q_ks equal to $\{+150, +350, +200, -150, -200, -450\}$ MW, respectively. After one execution of coalition formation dynamic learning process, the power distribution at the coalition formation stable state is as follows: $P_{16} =$

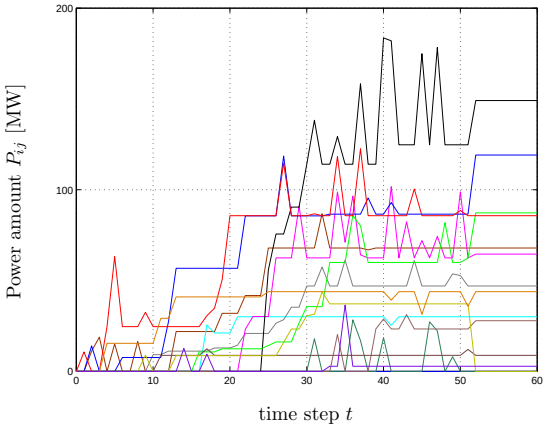


Fig. 4.6: *Learners behaviour during coalition formation process.*

$0.43Q_1$, $P_{25} = 0.38Q_2$, $P_{26} = 0.42Q_2$, $P_{34} = 0.53Q_3$, $P_{36} = 0.47Q_3$ and consequently the coalition structure is $\mathcal{M}_1 = \{1, 6\}$, $\mathcal{M}_2 = \{2, 5, 6\}$, $\mathcal{M}_3 = \{3, 4, 6\}$, and $\mathcal{K}$. The achieved coalitions payoffs are $\{107, 14.9, 29, 9.5\}$ nW^{-1}, respectively, and the Shapley values of the MGs are $\{4.426, -0.430, 0.038, 0.038, -0.430, 5.888\}$ nW^{-1}, respectively. The total amount of power dissipated, $1/\nu(\mathcal{K})$, is 59% of that of the traditional non-cooperative model which is calculated as a network wherein each MG trades only with CDU as in Fig. 4.1(a). Note that here we report the results at the stable point of one execution of coalition formation dynamic learning process only.

Fig. 4.6 exhibits the behaviour of the learners during coalition formation dynamic process. At $t = 0$, all $P_{ij} = 0$ and only the learners of P_{s0} and P_{0d} $\forall\ s \in \mathcal{K}^+$ and $\forall\ s \in \mathcal{K}^+$ may increase the proper power, since they have a negative payoff not equal to $-\alpha$. As can be seen, after some time steps, some other learners earn a utility not equal to $-\alpha$, and so may update their power value. Despite conflicts between simultaneous and myopic decisions, Fig. 4.6 shows the convergence of P_{ij} the stable point after 52 time steps.

Fig. 4.7 reports the average number of time steps τ for achieving the stable point of the power loss minimization process in the network scenario with the above mentioned (fixed) residual quantities. The dashed line reports the grand coalition's payoff and other curves report the Shapley values of the MGs during the dynamic process as a function of τ. As can be seen, when we choose $\rho = \nu(\mathcal{K})$ as the input of the Markov

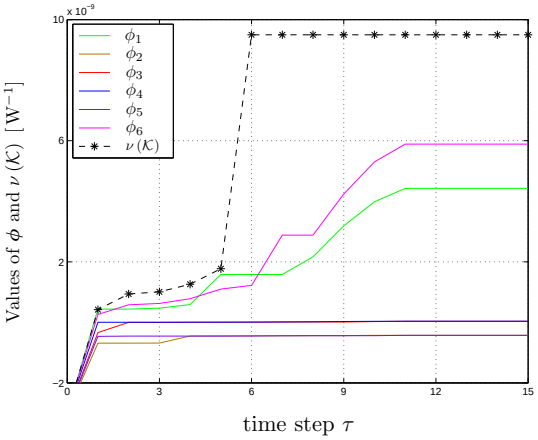


Fig. 4.7: *Shapley values and grand coalition's payoff during power loss minimization dynamic learning.*

state $\pi = (\boldsymbol{u}, \rho)$, the algorithm achieves stable state after 7 time steps, while this happens in $\tau = 12$ with $\rho = \boldsymbol{\phi}$. It is worthwhile to emphasize that, at each time step τ, first coalition formation dynamic learning process is executed (Fig. 4.6), and then the inputs of the state π^τ are updated.

To measure the power dissipated improvement, now we compare the performance of our proposed algorithm to the traditional non-cooperative power transmission. Fig. 4.9 reports the percentage of improvement in terms of average dissipated power amount expressed as the ratio of the whole power dissipated amount D^{NC} achieved by the traditional non-cooperative power transferring to the whole power dissipated amount D^C achieved by the proposed coalitional game theory based algorithm. The parameter $D^{NC} = \sum_{s\in\mathcal{K}^+} D_{s0} + \sum_{d\in\mathcal{K}^-} D_{0d}$ with $P_{s0} = Q_s$ and $P_{0d} = -Q_d$, and the parameter $D^C = 1/\nu(\mathcal{K})$. As can be seen, the growing rate of the curves slightly decreases after $K = 25$. The average of power loss improvement is 1.2% per MG with setting $\rho = \nu(\mathcal{K})$, while that is 1.5% per MG with $\rho = \boldsymbol{\phi}$. As a result, considering the Shapley values in which the coalition structure and all coalitions' payoffs are involved outperforms setting $\rho = \nu(\mathcal{K})$ in which only the grand coalition payoff is considered, at the cost of more computational complexity. The proposed algorithm in this paper outperforms the proposed coalition formation algorithm in [294].

Fig. 4.8 depicts the average of power loss fraction that is defined as the proportion of

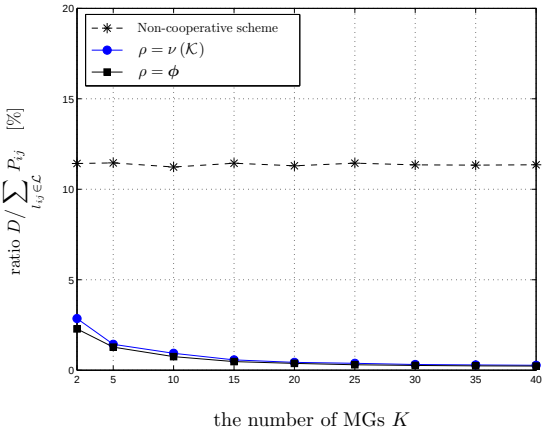


Fig. 4.8: *Average of loss fraction (power dissipated as a proportion of power loaded) as a function of K.*

the overall amount of dissipated power to the overall loaded power. The dashed line reports the traditional non-cooperative scheme with $D = \sum_{s \in \mathcal{K}^+} D_{s0} + \sum_{d \in \mathcal{K}^-} D_{0d}$ by setting $P_{s0} = Q_s$ and $P_{0d} = -Q_d$, and the solid line reports the proposed cooperative scheme with $D = 1/\nu(\mathcal{K})$. In other words, the total power dissipated in the traditional non-cooperative scheme is calculated as a network with K singleton coalitions. The average of the loss fraction in the non-cooperative scheme is 11% and that is not very sensitive to the number of MGs. However, in our cooperative scheme the average of the ratio decreases with the increasing number of MGs and its average is 0.83% with $\rho = \nu(\mathcal{K})$, and 0.69% with $\rho = \boldsymbol{\phi}$. Considering the range of the quantities (in $\pm[100,\ 500]$ MW), the power dissipated in non-cooperative scheme is significant and that reduces significantly in the cooperative scheme. As can be seen, with applying the proposed cooperative algorithm, the power dissipated will become 10% of the traditional distribution networks. As a result, setting $\rho = \boldsymbol{\phi}$ in which the coalition structure and all coalitions' payoffs are involved outperforms setting $\rho = \nu(\mathcal{K})$ in which only the grand coalition payoff is considered, at the cost of more computational complexity. Fig. 4.8 shows also that any value greater than 1.2 is a sufficient value for β in the power constraints $\bar{P}_{sd}$ and $\bar{P}_{0d}$, as we set $\beta = 2$.

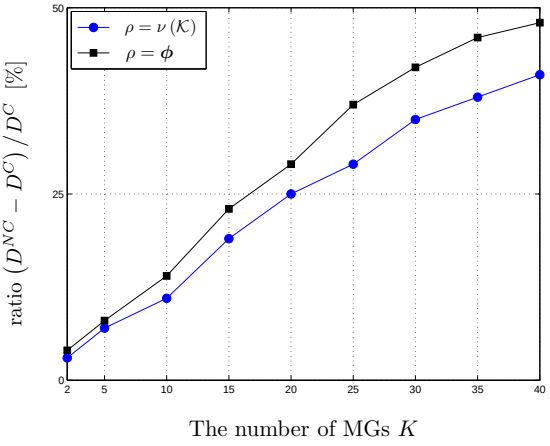


Fig. 4.9: *Percentage of power dissipated improvement as a function of K compared to the traditional power distribution.*

4.6 Conclusion

In this chapter, we have investigated the problem of power trading coordination in a smart power grid consisting of a set of micro-grids (MGs) each of which has a quantity of either surplus or deficit amount of power to sell or to acquire, respectively. The proposed coalitional game theory based approach allows micro-grids (MGs) to form a set of coalitions and trade power between themselves in order to achieve the minimum power dissipated during power generation and transfer. In this context, we propose a dynamic learning process to form a set of non necessarily disjoint and possibly singleton coalitions. In the proposed dynamic learning process each transmission line, as a learner, adjusts the best power load such that: *i*) all suppliers to be completely loaded, and *ii*) all demanders to be completely served. A complementary dynamic learning process is proposed to lead the coalition structure to the best performance in terms of power saving. Numerical results show power dissipated in the proposed cooperative smart grid is 10% of that in the traditional non-cooperative networks.