**A Review of Graph Neural Networks and Their Applications in Power Systems**

* represented as graph-structured data with high-dimensional features and in‐ terdependency among nodes. The complexity of graph-struc‐ tured data has brought significant challenges to the existing deep neural networks defined in Euclidean domains
* the smart grid has evolved into a typical dynamic, non-linear, and large-scale control system. The multi-directional information makes it hard to find optimal solutions that coordinate all participants such as distribution systems operators, produc‐ ers, demand response aggregators, and consumers
* The integration of flexible sources, e. g., electric vehicles, poses revolutionary changes to radial distribution networks such as relay protection, bidirectional power flow, and voltage regu‐ lation [2]. Moreover, the deregulation of electricity markets makes it difficult to find a strategy that is beneficial to both customers and producers. In these cases, traditional modelbased methods are hard to fully meet the control and analy‐ sis requirements of power systems because of their uncertain‐ ty and complexity
* traditional model-based methods for scenario generations of RES are not able to ac‐ curately capture the probability distribution characteristics and fluctuations of power curves
* the existing DNNs in Euclidean domains such as CNNs and RNNs are not suitable for processing the graph-structured data [8], since they stack the features of nodes by a specific order and ignore the topo‐ logical information
* The classi‐ cal GNNs mainly include graph convolutional networks (GCNs), graph recurrent neural networks (GRNNs), graph at‐ tention networks (GATs), graph generative networks (GGNs), spatial-temporal graph neural networks (STGNNs), and hybrid forms of GNNs such as graph reinforcement learning (GRL) and graph transfer learning (GTL) [9], which have shown outstanding performance for the graph-struc‐ tured data
* Recently, spatial-based GCNs have developed rapidly, since spectral-based GCNs usually deal with the whole graphs si‐ multaneously, and they are difficult to extend to large-scale graphs. In addition to GCNs, many other GNNs generalized from traditional DNNs were developed over recent years. These models mainly include GRNNs, GATs, GGNs, STGNNs, and hybrid forms of GNNs
* Normally, the graph-structured data are represented as G = (VE), where E is a set of edges and V is a set of nodes [22]. Specifically, vi is the i th node and eij is the edge from the j th node to the i th node. For the i th node, its neighborhood can be denoted as N ( vi) = {ui /el V |( vi,ui) /el E}. The graph structured data generally have a nodal feature matrix Xnode of n x f scales, and a feature matrix Xedge of m x c scales for edg‐ es. The adjacency matrix A is a matrix of n x n scales where aij is equal to 0 if eij /el E and aij is equal to 1 if eij /el E
* For the spatial-temporal graph, it is an attributed graph where the features of nodes change with time [23]. The spa‐ tial-temporal graph can be represented as G(t) = (V, E, X(t) ).

A picture containing text, receipt, screenshot

Description automatically generated

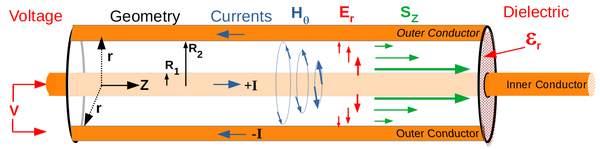
* For the directed graph, it has an asymmetric adjacency ma‐ trix, since these edges are directed from one node to another. Relatively, the edges of the undirected graph are all undi‐ rected, i. e., the adjacency matrix is symmetric
* Spectral-based GCNs. Unlike the images in Euclidean domains, the graph-structured data do not have the character‐ istics of translation invariance, which makes it hard to direct‐ ly define the convolutional operation in graph domains. In 2014, a spectral network was proposed in [20]. It transforms the samples into the Fourier domains to perform convolution‐ al operations through Fourier transform, and then the sam‐ ples are transformed back to the graph domains through in‐ verse Fourier transform
* Spatial-based GCNs. Analogous to the convolutional operation in the Euclidean domains, spatial-based GCNs di‐ rectly define the convolutional operation on the graph do‐ mains by operating on spatially close neighbors. The key challenges of these spatial-based GCNs are to define convo‐ lutional operations with the different number of neighbor‐ hoods and to keep the local invariance
* The main differences between spectral-based GCNs and spa‐ tial-based GCNs are as follows. Firstly, spectral-based GCNs either need to deal with the whole graphs simultaneously or perform eigenvector compu‐ tation, which leads to more computations induced by the for‐ ward and inverse graph Fourier transforms [33]. Relatively, spatial-based GCNs are extensible to large-scale graphs, since they directly define convolutional operations in graph domains. The computation of spatial-based GCNs can be per‐ formed in a batch of nodes in place of the whole graph-struc‐ tured data. Secondly, spectral-based GCNs which rely on the Fourier transform generalize unfavorably to various graphs. Any per‐ turbations in the graph-structured data will cause the eigenba‐ sis to change, because they assume that the graphs are fixed [34]. In contrast, spatial-based GCNs perform graph convolu‐ tional operations locally on each node, and the weights of networks can be easily shared across different locations. Thirdly, most of spectral-based GCNs are limited to han‐ dle undirected graphs [12], while spatial-based GCNs are more flexible to handle multisource graphs such as directed graphs [31], heterogeneous graphs [35], edge inputs [36], and signed graphs [37], since these graphs can be easily in‐ corporated into the aggregation function. In general, spectral-based GCNs perform convolutional op‐ erations in spectral domains through the complex Fourier transform, while spatial-based GCNs directly define convolu‐ tional operations in graph domains. Therefore, spatial-based GCNs show stronger generalization and flexibility compared with most of spectral-based GCNs
* The **GRNN**s are designed for the problems defined in graph domains, e.g., classifications in graph-level and nodelevel, which require outputting sequences.
* In the above-mentioned GCNs, the neighborhood of nodes is aggregated with equal or predefined weights. Neverthe‐ less, the impacts of neighbors may vary greatly [41]. There‐ fore, they should be learned in the process of training, in‐ stead of being predetermined. Activated by attention mecha‐ nisms**, GATs** introduce the attention mechanism into graph domains by revising the graph convolutional operation
* The purpose of GGNs is to generate some new graphstructured data by learning a series of given historical sam‐ ples. Similar to generative networks in Euclidean domains, the existing GGNs mainly include the graph automatic en‐ coders (GAEs), variational graph auto-encoders (VGAEs), and graph generative adversarial networks (GGANs) [45]. The GAEs consist of an encoder and a decoder [46]. First‐ ly, the features X and the adjacency matrix A of the nodes are fed to the encoder to obtain the embedding matrix ZGAE of the graph-structured data
* **Optimal Power Flow**: In addition to being used for power flow calculation, GNNs can also be further applied to the optimal power flow of distribution networks. For example, the spectrum-based GCNs are designed to optimize the reactive power of distri‐ bution networks in [6]. Specifically, the adjacency matrix is used to represent the topology information between the nodes in distribution networks, so as to mine the correlation of nodes.
* Then, the deep graph convolutional layer is used to capture the complex non-linear relationship between the state of the power equipment and the power loads. Simula‐ tion results show that the performance of this model is better than those of traditional data-driven methods such as CNNs, MLPs, and case-based reasoning. Similarly, the GNNs are designed to approximate the optimal power flow solution
* GNNs are local information and scalable processing ar‐ chitectures that mine the network structure of the input data. It is trained by taking a given network state as input and us‐ ing the output results to approximate the optimal solution of interior-point optimizer. Simulation results show that local solutions adequately exploit the latent grid structure and out‐ perform other comparable methods
* The ex‐ isting GNNs are difficult to account for the constraints of op‐ timal power flow such as voltage and current constraints, which may cause power grids to operate in an unsafe state. How to consider the constraints in GNNs can be further stud‐ ied.
* The existing methods assume that the topology of power systems is invariant, i. e., the adjacency matrix is fixed. However, the reconfiguration of the distribution net‐ work is also a way for regulating power flow. The distribu‐ tion network can be regarded as a kind of spatial-temporal graph, and the adjacency matrix and features of nodes change with times. In this case, none of the existing GCNs can be directly applied to optimal power flow calculations. Furthermore, spatial-temporal graph convolutional networks have shown outstanding performance in the field of comput‐ er vision for spatial-temporal graphs. Therefore, they may be extended to the optimal power flow of distribution networks
* The existing methods are only suitable to solve static op‐ timal power flow, which is difficult to be directly used in ac‐ tual engineering due to the fluctuation of power loads. GCNs may be extended to the dynamic optimal power flow
* **Optimal Load Shedding**: Load shedding is very important for operations of distribu‐ tion networks particularly under contingency events such as line failure. Since traditional methods need to solve complex optimization models, the computation time is very long, which cannot meet the real-time requirements of distribution networks. To solve these problems, the GCNs are used for load-shedding operations in [92]. Specifically, the features of nodes include voltage amplitude, voltage angle, active pow‐ er, and reactive power. The adjacency matrix represents the topology information of distribution networks. As shown in Table V, this model significantly outperforms the MLPs and linear regressions (LRs) in different testing systems. Further‐ more, the existing methods do not account for the impact of graph pooling layers and dropout layers on accuracy and computation time, which can be further discussed

**Power System Essentials (PDF)**

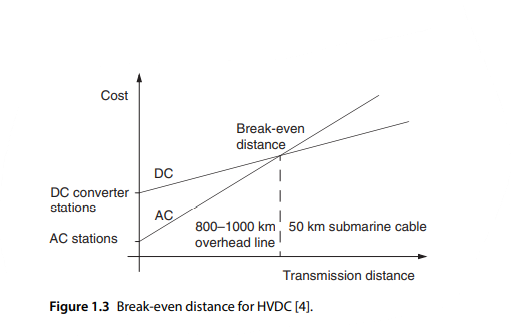
* the models for the power system components that are used in this book have a limited validity; they are only valid in the steady-state situation and for the analysis of low-frequency phenomena
* Steady-state analysis covers a variety of topics such as planning, design, economic optimization, load flow/power flow computations, fault calculations, state estimation, protection, stability, and control
* Milliseconds for dynamic analysis (kHz): Understanding the dynamic behavior of electric networks and their components is important in predicting whether the system, or a part of the system, remains in a stable state after a disturbance. The ability of a power system to maintain stability depends heavily on the controls in the system to dampen the electromechanical oscillations of the synchronous generators.
* Although the power system itself remains unchanged when different time scales are considered, components in the power system should be modeled in accordance with the appropriate time frame
* For steady-state computations at power frequency, the wavelength of the sinusoidal voltages and currents is 6000 km (in the case of 50 Hz): lambda = v / f = 3 \* 10^5 / 50 Hz = 6000 KM -> the transmission line is, so to speak, of “electrically small” dimensions compared to the wavelength of the voltage. The Maxwell equations can therefore be approximated by a quasi-static approach, and the transmission line can accurately be modeled by lumped elements
* When the effects of a lightning stroke have to be analyzed, frequencies of 1 MHz and higher occur and the typical wavelength of the voltage and current waves is 300 m or less. In this case the transmission line is far from being “electrically small,” and it is not allowed to use the lumped-element representation anymore. The distributed nature of the transmission line has to be taken into account, and we have to calculate with traveling waves.
* Despite the fact that we mainly use lumped-element models in our book, it is important to realize that the energy is mainly stored in the electromagnetic fields surrounding the conductors rather than in the conductors themselves
* . The Poynting vector, being the outer product of the electric field intensity vector and the magnetic field intensity vector, indicates the direction and intensity of the electromagnetic power flow: S = E x M

**Diagram, engineering drawing

Description automatically generated**

****

* Due to the finite conductivity of the conductor material and the finite permeability of the transformer core material, a small electric field component is present inside the conductor and a small magnetic field component results in the transformer core: E = J / σ: J the current density vector [A/m2] σ the conductivity [S/m]. H = B / μ: B the magnetic flux density vector [T= A H/m2] μ the permeability [H/m] This leads to small Poynting vectors pointing toward the conductor and the transformer core: the losses in the transmission line and the transformer are fed from the electromagnetic field, as is the power consumed by the load.
* Most of the power systems are 50 or 60 Hz three-phase AC systems.The voltage levels used are quite diverse
* The incremental costs of DC transmission over a certain distance are less than the incremental costs of AC, because in a DC system two conductors are needed whereas three-phase AC requires three conductors. On the other hand, the power-electronic converters for the conversion of AC to DC at one side, and from DC to AC at the other side, of the DC transmission line are more expensive than the AC transmission terminals. If the transmission distance is sufficiently long, the savings on the conductors overcome the cost of the converters and DC transmission is, from a capital investment point of view, an alternative to AC -> DC > AC for long transmission distances, AC > DC for relatively short distances

****

* a sinusoidal excitation of a linear system results in a sinusoidal response. Therefore, all the voltages and currents in the power system are sinusoidal and have the same frequency so that the components in the system can be designed for this wave shape.
* So summing up, the RMS or effective value of a sinusoidal alternating voltage (or current) is equal to the value of the equivalent direct voltage (or current) that dissipates the same amount of power in a given resistor during one time period of the alternating voltage (or current)
* It is the RMS value of the sinusoidal voltage and current that is read by the common type of voltmeters and ammeters
* When we speak of a voltage of 230 V, the RMS value of the voltage amounts to 230 V, while the peak value of the sinusoidal voltage is √ 2 ⋅ 230 = 325V.
* We see that when the applied voltage remains the same, a higher system frequency (ω =2πf ) results in a lower effective value of the magnetic flux (|Φ|) so that we can use a smaller cross-sectional area for the iron core when we keep the magnetic flux density constant. When there is freedom to choose the system frequency, a higher frequency can be very advantageous, especially in the case that weight and volume play a role, for example, on board of airplanes and ships.
* The transmission and distribution systems are three-phase systems. In this book we restrict ourselves to balanced three-phase power systems. In the case of a balanced three-phase system, the sinusoidal voltages are of equal magnitude in all three phases and shifted in phase by 120∘
* why do we apply a three-phase system and not a four- or five-phase system? This is because each phase requires its own conductor, and the balanced three-phase system is the system with the smallest number of phase conductors capable of delivering constant instantaneous power.
* When a compass needle is positioned in the middle of the three-phase coil system, the needle keeps pace with the rotating field, which is a crude equivalent of the synchronous motor. When a copper cylinder is placed in the center of the three-phase coil system, the rotating field drags the cylinder around with it, and we have a primitive equivalent of the induction motor.
* The voltage ratings in a power system can be divided into three levels: • The generation level: 10–25 kV. The power is generated at a relatively low voltage level to keep the high-voltage insulation of the generator armature within limits. • The transmission level: 110–420 kV and higher (in the former Soviet Union even the 1200 kV level is in operation). • The distribution level: 10–72.5 kV
* Power system calculations in the steady-state situation are considerably simplified by introducing the phasor. A phasor is an arrow in the complex plane that has a one-to-one relation with a sinusoidal signal
* General sinusoidal voltage and current expressions:
  + v(t) = √ 2 \* |V| cos(ωt)
  + i(t) = √ 2 \* |I| cos(ωt - φ)
* The phasor represents a sinusoidal signal: the length of the phasor equals the effective or RMS value of the signal, and the angle of the phasor matches the phase shift of the signal with respect to a reference (an ideal cosine (or sine)).We see that the frequency component is absent when we use the phasor notation
* The voltages produced by the synchronous generators in the power system are 50 Hz (or 60 Hz) sinusoidal voltages. As the power system is supposed to be a linear system (see Section 1.3.1) in steady state, the voltage out of the AC outlet at home is also a sinusoidal voltage with a frequency of 50 Hz; only the amplitude of the voltage differs and a phase shift may have occurred. Therefore, in steady-state calculations, the frequency gives no extra information and can be omitted so that we can do our calculations with the phasors, fixed in the complex plane. No relevant information is lost as the information with respect to the phase angle and the amplitude is still available in this phasor.
  + The phasor of the sinusoidal signal v(t)=141.4 cos(ωt + π/6) is written as V =100∠30∘

**Diagram, engineering drawing

Description automatically generated**

**Table

Description automatically generated with medium confidence**

Text, letter

Description automatically generated

**A picture containing application

Description automatically generated**

**Text

Description automatically generated**

* transfer of energy toward the load. The amplitude of this oscillating power is called imaginary or reactive power Q and is defined as Q = |V||I|sin(φ) .Q the reactive/imaginary power [var; reactive volt-amperes]
* Therefore, of the instantaneous power p, consumed by an element at any instant, a part (viR) is utilized for permanent consumption, such as conversion into heat. This part always has a positive value, that is, it cannot be returned to the rest of the circuit. The remainder (viX) is used to establish either a magnetic or electric field, that is, it is taken up and returned to the circuit with the rhythm of double the power frequency.
* We adopt the sign convention recommended by the International Electrotechnical Commission (IEC). A capacitor supplies reactive power, whereas an inductor consumes reactive power. Or in other words, the reactive power absorbed by an inductive load has a positive sign, and the reactive power absorbed by a capacitive load a negative sign. In the case that α>β (see Figure 1.30), the current lags the voltage. Therefore, the load is inductive and, in line with the IEC convention, consumes reactive power. To obtain the proper sign for the reactive power, it is necessary to calculate VI\* .

**Diagram

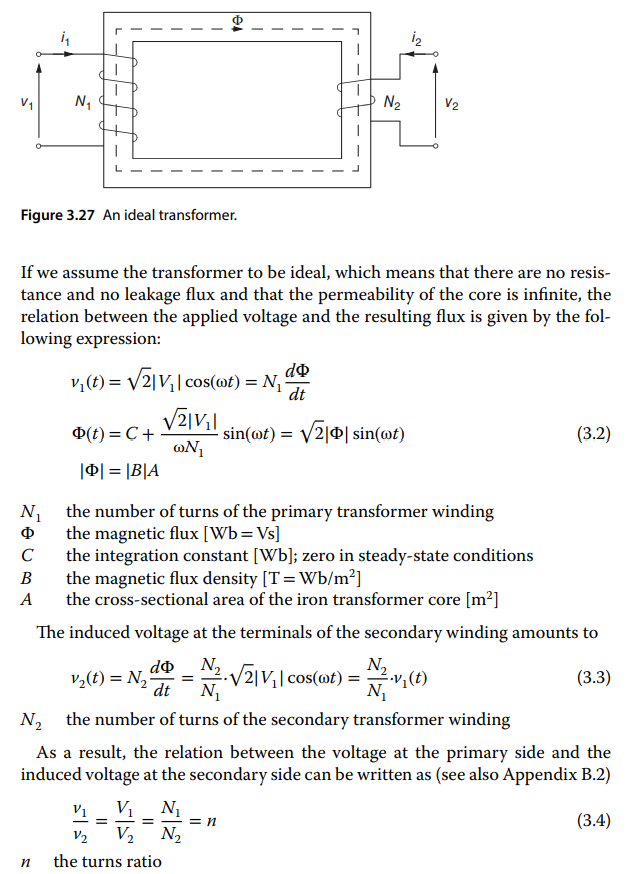
Description automatically generated**

* The term cos(φ) is called the power factor, being the cosine of the phase shift between the voltage and current, that is, the cosine of the phase angle between the voltage and current phasor. In fact, the power factor is that part of the apparent power that is related to the mean energy flow, like mechanical energy in the case of a machine or heat in the case of a resistor
* e. The power factor of the inductive load amounts to 0.8 and is to be increased to the value of 1. This can be done by connecting a capacitor in parallel to the load. The value of the capacitance is such that the capacitor supplies the amount of reactive power that is consumed by the inductive load.

**Diagram

Description automatically generated**

* cheap | more reliable but costlier | most reliable and costliest
* The simplest way to look at the power system is to consider it as a collection of nodes, which we call substations, and connecting power carriers, such as overhead lines and underground cables
* In normal operating conditions, a three-phase power system can be treated as a single-phase system when the loads, voltages, and currents are balanced (see also Section 1.3.3), and one can successfully use a single-line lumped-element representation of the three-phase power system for calculation. A fault brings the system to an abnormal condition. Short-circuit faults are especially of concern because they result in a switching action, which often results in transient overvoltages.
* Line-to-ground faults are faults in which an overhead transmission line touches the ground because of wind, ice loading, or a falling tree limb. A majority of transmission line faults are single line-to-ground faults
* Line-to-line faults are usually the result of galloping lines because of high winds (see also Section 3.9.1) or because of a line breaking and falling on a line below
* Double line-to-ground faults result from causes similar to that of the single line-to-ground faults but are very rare
* Three-phase faults, when all three lines touch each other or fall to the ground, occur in only a small percentage of the cases but are very severe faults for the system and its components. In the case of a symmetrical three-phase fault in a symmetrical system, one can still use a single-phase representation for the short-circuit and transient analysis.
* However, for the majority of the fault situations, the power system has become unsymmetrical. Symmetrical components and, especially, the sequence networks are an elegant way to analyze faults in unsymmetrical three-phase power systems because in many cases the unbalanced portion of the physical system can be isolated for a study, the rest of the system being considered to be in balance.
* A reliable protection is indispensable for a power system. When a fault or an abnormal system condition occurs (such as over-/undervoltage, over-/underfrequency, overcurrent, and so on) the related protective relay has to react in order to isolate the affected section while leaving the rest of the power system in service. The protection must be sensitive enough to operate when a fault occurs, but the protection should be stable enough not to operate when the system is operating at its maximum rated current
* There are also faults of a transient nature, a lightning stroke on or in the vicinity of a transmission line, for instance, and it is undesirable that these faults would lead to a loss of supply. Therefore, the protective relays are usually equipped with auto-reclosure functionality. Auto-reclosure implies that the protective relay, directly after having detected an abnormal situation leading to the opening of the contacts of the circuit breaker, commands the contacts of the circuit breaker to close again in order to check whether the abnormal situation is still there. In case of a fault of a transient nature, the normal situation is likely to be restored again so that there is and was no loss of supply. When the abnormal situation is still there, the protective relay commands the circuit breaker to open its contacts again so that either the fault is cleared or consecutive autoreclosure sequences can follow. In most cases, the so-called backup protection is installed in order to improve the reliability of the protection system
* When protective relays and circuit breakers are not economically justifiable in certain parts of the grid, fuses can be applied. A fuse combines the “basic functionality” of the current transformer, relay, and circuit breaker in one very simple overcurrent protection device.The fuse element is directly heated by the current passing through and is destroyed when the current exceeds a certain value, thus leading to an isolation of the faulted sections or components. After the fault is repaired/removed, the fuse needs to be replaced so that the isolated grid section can be energized again.
* A fuse is a weak link in a circuit and as such has one important advantage over circuit breakers. Because the element in the fuse has a much smaller cross section than the cable it protects, the fuse element will reach its melting before the cable. The larger the current, the quicker the fuse element melts. The fuse interrupts a very large current in a much shorter time than a circuit breaker does – so short in fact that the current will be cut off before it reaches its peak value, which in a 50 Hz system implies operation in less than 5 ms, and serious overheating and electromechanical forces in the system are avoided. This current-limiting action is an important characteristic that has application in many industrial low-voltage installations. The single-shot feature of a fuse requires that a blown fuse has to be replaced before service can be restored. This means a delay, the need to have a spare fuse and qualified maintenance personnel who must go and replace the fuse in the field. In a three-phase circuit, a single-phase-to-ground fault will cause one phase to blow and the other two phases stay connected.
* A high-voltage circuit breaker is an indispensable piece of equipment in the power system. The main task of a circuit breaker is to interrupt fault currents and to isolate faulted parts of the system. Besides short-circuit currents, a circuit breaker must also be able to interrupt a wide variety of other currents at system voltage such as capacitive currents, small inductive currents, and load currents. We require the following from a circuit breaker:
  + In closed position it is a good conductor
  + In open position it behaves as a good isolator between system parts.
  + It changes in a very short period of time from close to open.
  + It does not cause overvoltages during switching.
  + It is reliable in its operation.
* Overvoltages, stressing a power system, can generally be classified into two categories regarding their origin:
  + External overvoltages, generated by lightning strokes, which are the most common and severe atmospheric disturbances
  + Internal overvoltages, generated by changes in the operating conditions of the network, such as switching
* Transformers are essential components in the AC power system as they make it possible to convert electrical energy to different voltage levels with an efficiency of more than 99%. That enables us to generate power at a relatively low voltage level (10–25 kV, limited by the insulation of the generator), to transport it at high voltage levels (110–420 kV and higher) to reduce the losses during transportation, whereas domestic consumption can take place at a low and (more or less) safe voltage level (400 V and below). Transformers consist essentially of two coils on a common iron core.

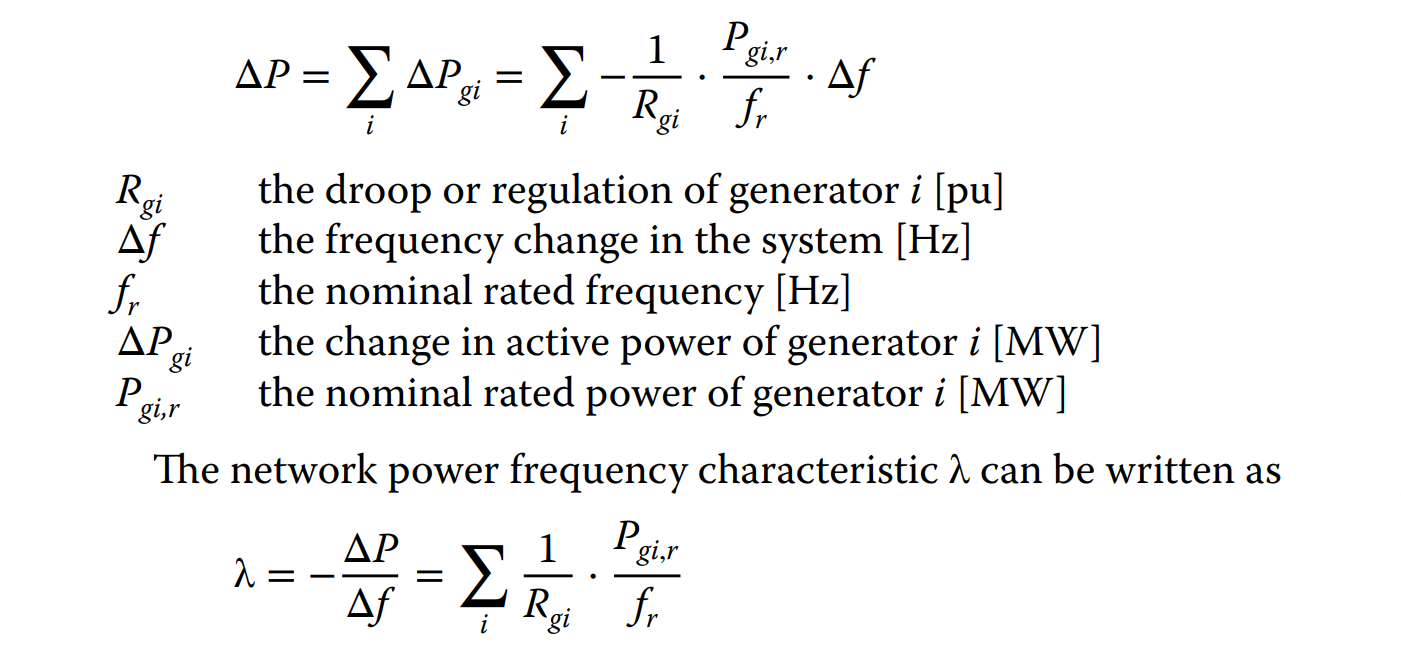
****

* wye-connected coils are applied for the higher voltage levels. Delta-connected coils are advantageous when a Dy transformer serves as a distribution transformer: the single-phase loads, which are connected to the secondary wye-connected coils, “spread out” over two coils at the primary
* side, so that possible unbalances between the phases at the secondary side are smoothed at the primary side of the transformer
* Consumer demand is not constant but varies from hour to hour each day, from day to day within a week, and from season to season
* The minimum load of the day, which is called the valley load, is about 40–60% of the system peak load and usually occurs between 4 and 5 a.m. The generation should be able to fulfill the fluctuating demand, and the transmission and distribution systems have to be able to facilitate the flow of energy
* Most power systems are vertically operated, which means that the grid consists of large power plants feeding bulk power into the high-voltage transmission network that in turn supplies the distribution substations. A distribution substation serves several feeder circuits, and a feeder circuit supplies numerous loads of all types. A light to medium industrial customer can be supplied from the distribution feeder circuit primary busbar directly, while a large industrial load complex usually is served directly from the bulk transmission system.
* The electricity generated in power plants travels to our homes and offices through the distribution grid. The voltage of the distribution grid is typically less than 20 kV. Voltage ratings of 12, 11, 10, and 7.2 kV are quite common
* The consumers are supplied with the requested amount of active and reactive power at constant frequency and with a constant voltage. Loads are switched on and off continuously, and because electricity cannot efficiently be stored in large quantities, the balance between the amount of generated and consumed electricity has to be maintained by control actions. The consumers are supplied with the requested amount of active and reactive power at constant frequency and with a constant voltage. Loads are switched on and off continuously, and because electricity cannot efficiently be stored in large quantities, the balance between the amount of generated and consumed electricity has to be maintained by control actions.
* The active power balance is controlled by the generators.There is also another option, and that is to reduce the active power consumption by disconnecting parts of the load (this is called load shedding), but this is merely an emergency measure and not common practice
* The reactive power balance can be controlled by rotating equipment (generators and motors) and by static components (capacitors and inductors). The synchronous generator is the most important component in the system for maintaining the active and the reactive power balance
* without the appropriate control actions, both the power system frequency and the voltage would be far from constant.
* increased active power consumption reduces the frequency and that, in the example shown, an increased reactive power consumption reduces the voltage
* In transmission systems where most of the control actions take place, the following approximations can be made:
  + Because the resistance of transmission links is much smaller than the reactance values (R ≪ X), the resistance of the transmission link can be neglected: Z = |Z|∠ρ = X∠(π∕2).
  + Because the difference between the voltage angles is rather small, we can replace sin(δ1 − δ2) by δ1 − δ2 and cos(δ1 − δ2) by the value 1

**Text

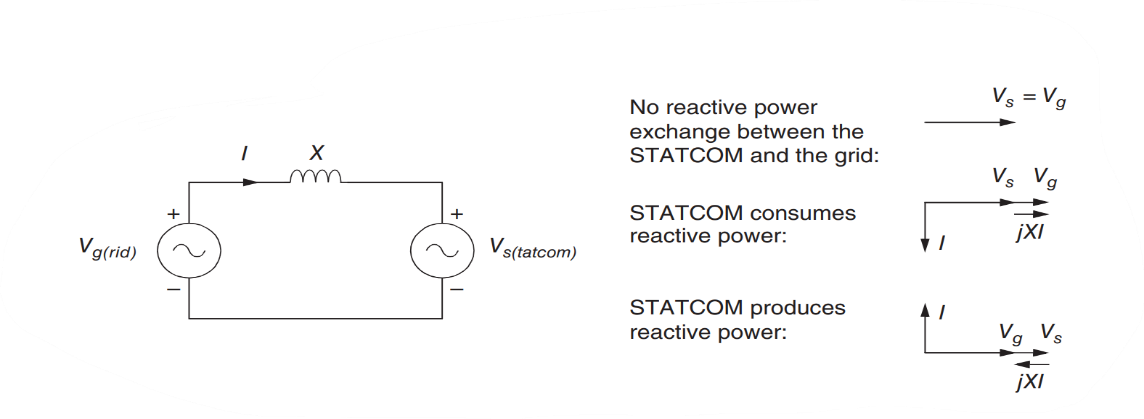
Description automatically generated**

* This “decoupling” between the active power and voltage angles and the reactive power and voltage magnitudes is of great importance for power system computations
* There is an important difference between the coupling of the active power and the frequency and the coupling of the reactive power and the voltage
* The frequency is a common parameter throughout the network: an increase of the active power consumption at a certain load point results in a reduction of the system frequency. An increase of the reactive power consumption at a certain load point is noticed only locally, as the voltage drops at the particular load point and at some nearby nodes.
* A change in the balance between the generation of active power and the consumption of active power changes the kinetic energy of the rotating mass of the generators and alters the system frequency
* 3 different cases for the speed governor:
  + stand-alone generator
    - The load dictates the amount of active power to be supplied by the generator; the speed governor determines the frequency. In Figure 5.5 (c), a stand-alone generator connected to a load experiences a sudden change of the active power balance: the load suddenly requires more active power P2 =P1 + ΔP. To cope with this, the kinetic energy of the rotating mass of the generator, and thus the frequency, drops. To restore the active power balance, the speed governor increases the mechanical power of the prime mover with ΔP, in accordance with the speed governor characteristic. The new frequency, for which the active power balance is fulfilled, is lower than the original frequency: f 2 =f 1 − Δf . See Example 5.4 (p. 192) for the illustration of this behavior
  + A generator connected to an infinite bus
    - The infinite bus dictates the frequency; the speed governor determines the amount of active power to be supplied by the generator. In Figure 5.5 (c), a generator connected to an infinite bus experiences a frequency drop f 2 =f 1 − Δf . As a result, the speed governor increases the prime mover power according to the speed governor characteristic with an amount of ΔP: P2 =P1 + ΔP
  + Two generators in parallel
    - In this case two generators supply the load; the frequency is set by both speed governors. The active power generation is shared by the two generators. When suddenly more active power ΔP is required by the load, the frequency drops with Δf (see Figure 5.6). To restore the active power balance, the speed governors increase the prime mover power according to the respective speed governor characteristics: ΔP = ΔPg1 + ΔPg2 (with ΔPg1 =Pg1,2 −Pg1,1 and ΔPg2 =Pg2,2 −Pg2,1).
* The case of two generators running in parallel, as described earlier, is in fact the most simple case of a multi-generator system. In a large-scale power system, with a large number of generators connected, it is the network power frequencycharacteristic that relates the difference between scheduled and actual system frequency to the amount of generation required to correct the power imbalance for that system: λ=−ΔP/ Δf (5.18)
  + λ the network power frequency characteristic [MW/Hz]
  + ΔP the amount of generation required to correct the power imbalance [MW]
  + Δf the difference between scheduled and actual system frequency [Hz]

****

* The area control error (ACE) is an indication of the surplus or the lacking amount of power in a particular control area. A negative ACE shows that the control area generates too little power to exchange the scheduled amount. A positive ACE means that the control area produces excess power and exchanges more than the scheduled amount. The ACE must be brought back to zero in order to establish the scheduled power exchange at the nominal frequency
* The frequency is sometimes called the system frequency: in an interconnected power system, the frequency has the same value everywhere in the system
* Voltage Control and Reactive Power**Diagram, schematic

  Description automatically generated**
* The frequency is sometimes called the system frequency: in an interconnected power system, the frequency has the same value everywhere in the system
* A similar “system voltage” does not exist: the voltage amplitude depends strongly on the local situation in the system. As a consequence, voltages in the power system can only be controlled locally: either at generator buses by adjusting the generator voltage control or at fixed points in the system where tap-changing transformers, capacitor banks, or other reactive power consumers/producers are connected
* The automatic voltage regulator (AVR) is the basis of the generator reactive power control. A simplified block diagram of an AVR is shown in Figure 5.9. The main task of the AVR is to keep the value of the voltage at the synchronous generator terminals at a specified level.
* The principle of operation of the AVR is rather straightforward.When the terminal voltage of the generator decreases (increases), the voltage regulator magnifies (reduces) the excitation, which results in a higher (lower) internal EMF and terminal voltage. The same result can be obtained by increasing (decreasing) the reference voltage that is offered to the voltage regulator (see Figure 5.9). When we consider a generator with a fixed terminal voltage, the effect that the reactive power output rises when the internal EMF increases is easily understood from the phasor diagram
* there is a tight relation between reactive power exchange and voltage level (Examples 5.2 (p. 186) and 5.3 (p. 189)): reactive power consumption (by an inductive component) at a network node results in a lower node voltage, but reactive power injection (by a capacitive component) gives a higher node voltage
* **Controlling active power flows**
* Apart from the phase shifter, which is described hereafter, there are other options to control the active power flows in the system. When HVDC links are part of the grid (see also Section 1.3.1), the active power flowing through the DC links can be influenced; this is an additional advantage of HVDC links. Another piece of equipment that is able to control active power flows is the unified power flow controller (UPFC);

****

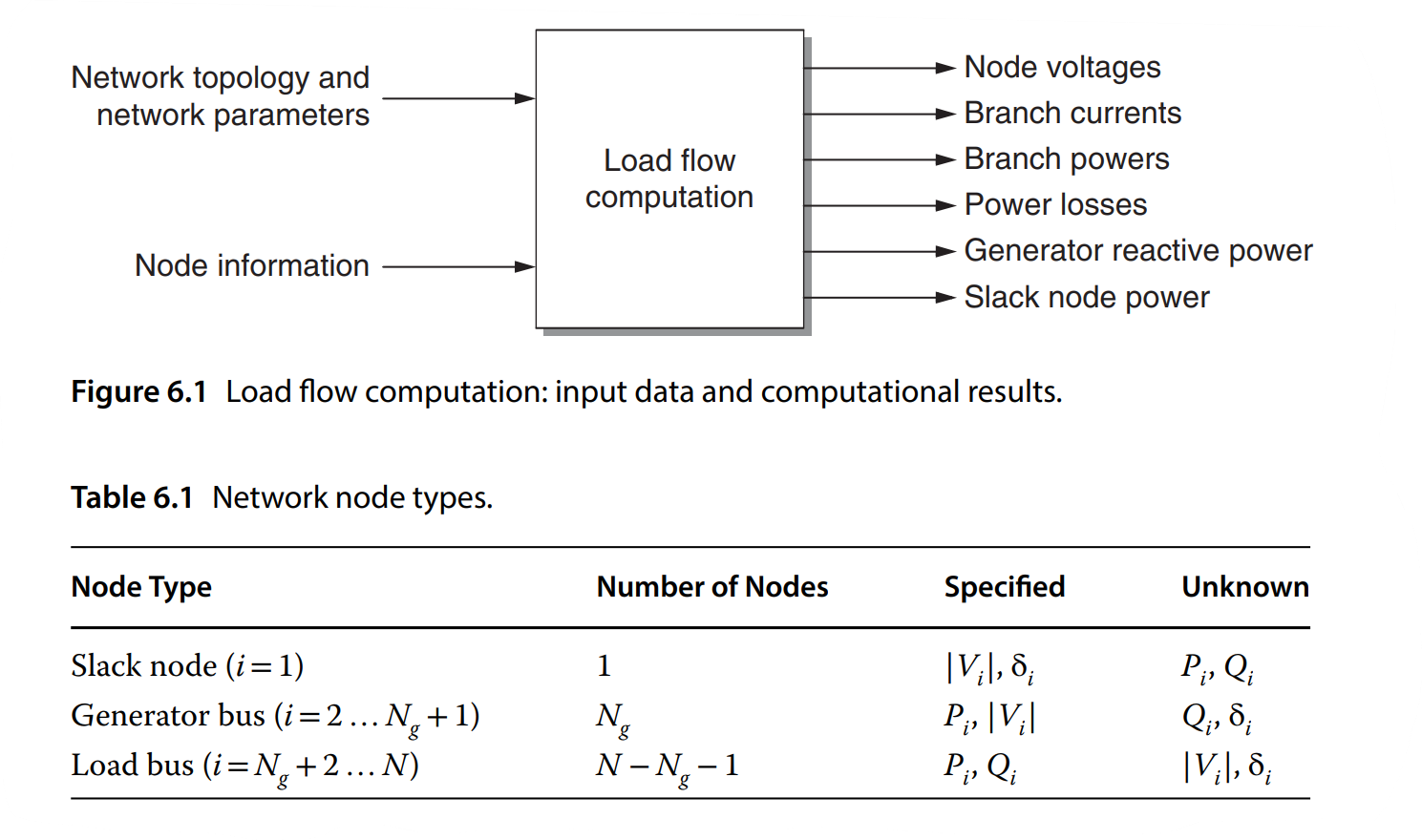
* **Controlling Reactive Power Flows**
* Reactive power flows in the system can be influenced by making use of series compensation: the effective series reactance of a transmission line is reduced in order to reduce the voltage drop across the line and the reactive power “loss” in the transmission line
* Static synchronous series compensator (SSSC) The static synchronous series compensator (SSSC) is depicted in Figure 5.27 and looks like the STATCOM (see Section 5.4.3). The SSSC is a device that can change the effective impedance of a transmission line by injecting a voltage ΔV that is in antiphase with the voltage drop across the transmission line. The output voltage ΔV is controlled such that it lags the current I by 90∘ and thus
* **Unified Power Flow Controller**
* The unified power flow controller (UPFC) is in fact an SSSC (see Section 5.5.2) that, additionally, has an active DC source. The UPFC is equipped with two converters (VSCs) operated from a common DC link as shown in Figure 5.28. The basic operating task of the first converter (VSC 1) is to supply and absorb the active power demanded from the second converter (VSC 2) at the DC link. The second converter (VSC 2) injects a voltage with a variable amplitude (0–ΔV max) and a variable phase angle (0–360∘) into the transmission line. Therefore, the injected voltage ΔV can take any position inside the circle
* The UPFC has a large number of system applications, such as voltage regulation, series compensation, and phase shifting:
  + Voltage regulation is achieved when the voltage ΔV is in phase with the voltage V1 , as shown in Figure 5.29 (b)
  + In Figure 5.29 (c) a combination of voltage regulation and series compensation is demonstrated. The voltage ΔV consists of two components: one that is in phase with the voltage V1 (voltage regulation) and another component that lags the line current I by 90∘ (series compensation).
  + In Figure 5.29 (d) a combination of voltage regulation and phase-angle regulation is pictured. The voltage ΔV consists of two components: one that is in phase with the voltage V1 (voltage regulation) and another component that shifts the resulting voltage by α degrees (phase shifting).
* **Diagram

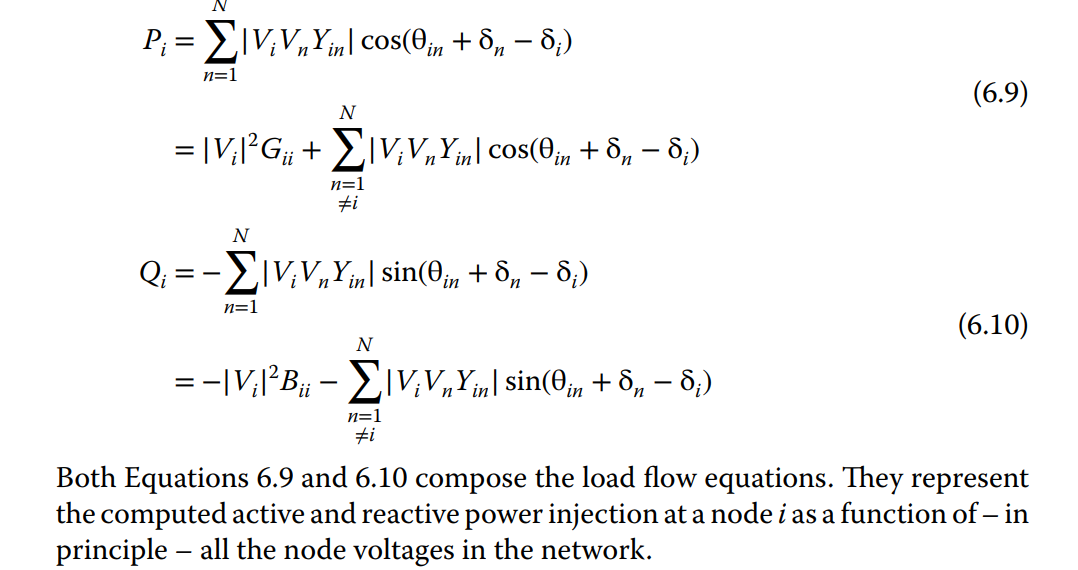
  Description automatically generated**
* In the power system, the power flow follows the laws of physics. In the past, the possibilities to control the power flows in the system were limited; they were mainly based on mechanical devices, such as transformer tap changers and turbine governors. These mechanisms automatically introduce a limitation to the speed of control. Nowadays, FACTS devices are available that enable a greater flexibility in the operation of AC power systems (FACTS is an acronym for flexible AC transmission systems) [2–6]. FACTS devices are large power-electronic controlled devices and can do control actions at a considerably higher speed. Some of these devices are operated in a shunt configuration for reactive power and voltage control, whereas others are put in series to control the power flow. This gives the system operator flexibility and a certain degree of freedom in operating the system, which is of great value in the present-day market environment

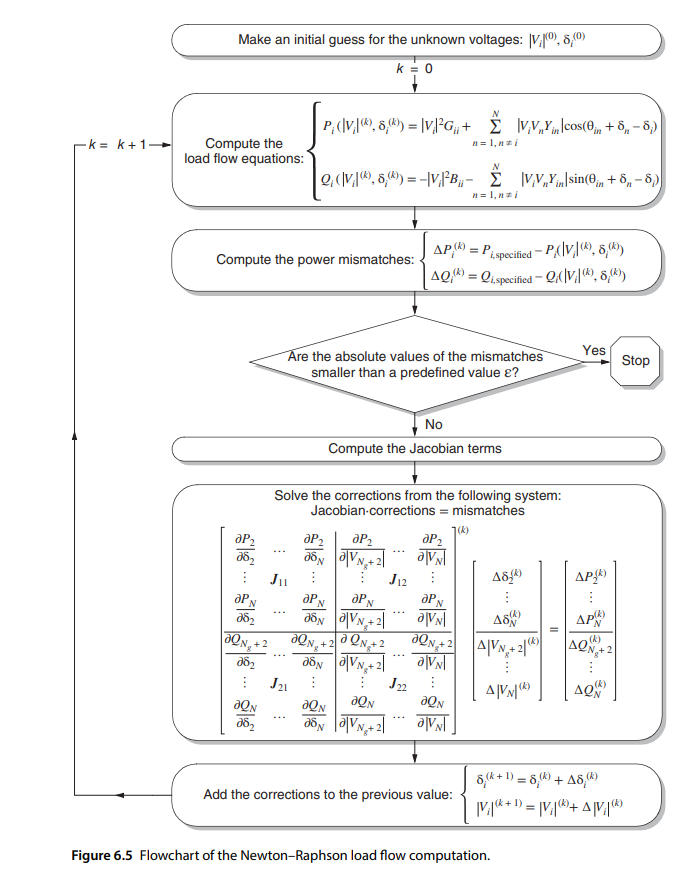
**Energy management Systems Chapter 6**

* In the control center, the transmission and distribution of electrical energy are monitored, coordinated, and controlled. The energy management system (EMS) is the interface between the operator and the actual power system
* An indispensable part of the EMS is the supervisory control and data acquisition (SCADA) system. The main functionalities of the SCADA system are to collect real-time measured data from the system and to present it to the computer screen of the operator and to control actual components in the network from the control center. The actual network status is stored in the “real-time telemetered database,” which is used as input for the other EMS functions
* The word load flow or power flow computation is self-explanatory: the flows of power in the network are computed. Later on we see that the name is rather misleading as first the node voltages in the network are computed before the power flows can be calculated. The load flow is the most important network computation, as it allows insight into the steady-state behavior of the power system
* When a transmission line is taken out of service temporarily, for planned maintenance, for example, the power originally flowing through the transmission line will find itself a new path to the loads.The operators want to be certain, in advance, that other transmission lines and/or cables in the vicinity are not overloaded after taking the particular line out of service. A load flow computation of the network configuration in which the transmission line is taken out of service gives insight into the new power flows and indicates possibly overloaded connections or components.
* The node voltages in the network should be kept within close limits during normal operation between 1.1 and 0.9 pu. A load flow computes the voltages in the network and visualizes the effect of tap-changing transformers, capacitor banks, and load shedding on the voltage profile in the system.
* The system operation should be robust and therefore the power system is operated n−1 secure. This means that a system component may fail without overloading other components or without violating the voltage limits. A list with transformers, transmission lines, cables, generators, and so on is available from which the components are taken out of service in a simulation one by one and each time a load flow is computed. When a load flow calculation shows an overloaded connection or transformer, preventive actions can be taken in the real network to prevent that particular situation. This analysis, based on a large number of load flow computations, is called the contingency analysis. The contingency analysis is performed in the control center of the utilities on a regular time basis.

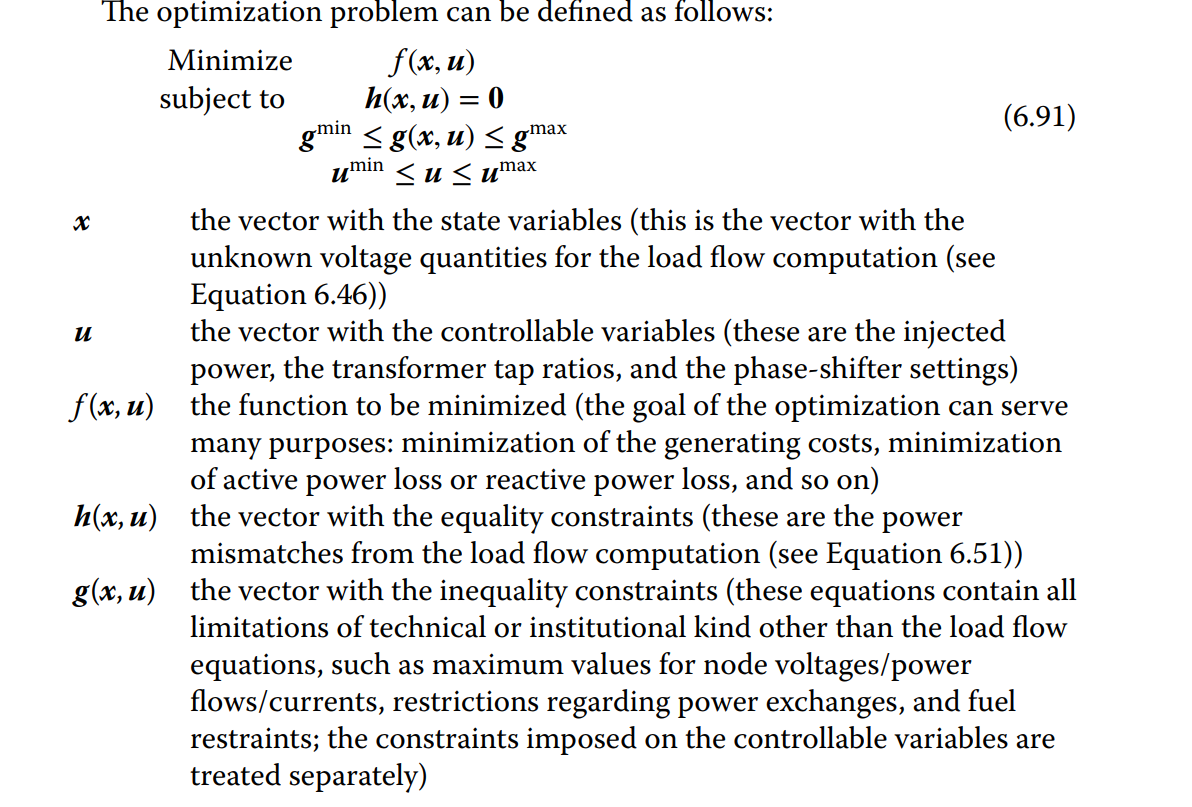
**Load Flow Equations**

* ****
* A network node is fully described (electrically) by four parameters:
  + The voltage phasor magnitude: |V|
  + The voltage phasor angle: δ
  + The injected active power: P
  + he injected reactive power: Q
* Three types of network nodes can be distinguished, and only two of the four parameters are known for each node:
  + The load bus Loads are modeled as constant power sinks instead of impedances. The reason for this is that the load flow is usually computed for the higher voltage levels (50 kV and above). The voltage deviations, which would result from changes in the load, are canceled out by changing the tap position of regulating transformers, and the system “sees” a load of constant power.
  + The generator bus The generator has two controls: the active power control and the voltage control. A wind generator, however, does not have these controls and is treated in the load flow as a load bus.
  + The slack node The load flow computation needs one network node to be addressed as slack node. The slack node serves as a reference for the other nodes: it is the only node of which the angle of the voltage phasor is specified. The actual value of this reference voltage angle is not of importance, because the other voltage angles are relative to this value. It is common practice to take as voltage angle δ =0. Another characteristic of the slack node is that the power injection at the node is not prescribed. This is a necessary requirement as becomes clear from the following. Let us consider a network of only load buses and generator buses. That would implicate that at all nodes, the active power injection is known (see Table 6.1) and, as a consequence, that the active power losses (|I| 2R) in the network are prescribed.This is not possible: the load flow needs to compute the unknown node voltages first before the line currents and the losses can be calculated. The difference (the slack) between the total active power input and total active power output plus the computed total |I| 2R losses is balanced by the slack node. Similarly for the reactive power, the difference (the slack) between the total reactive power input and total reactive power output plus the calculated total |I| 2X “losses” is balanced by the slack node. In the considered network with only load buses and generator buses, one of the generator buses must be assigned as slack node.
* **Text

  Description automatically generated**
* has the same value, as the net admittance connected between node j and node i is the same as the net admittance connected between node i and node j)
* Yij = |Yij|∠θij = |Yij|(cos θij + jsin θij) = Gij + jBij
* Gij the conductance Bij the susceptance
* ****
* The load flow computation is in fact the calculation of the voltage magnitude and angle at each bus of the power system under specified conditions of system operation. Other system quantities such as the current values, power values, and power losses can be calculated when the voltages are known. Speaking in mathematical terms, the load flow problem is nothing more than a system consisting of as many nonlinear equations as there are variables to be determined
* In other words, the state variables must be determined such that the power mismatches, being the difference between the specified and computed power injections, are equal to zero

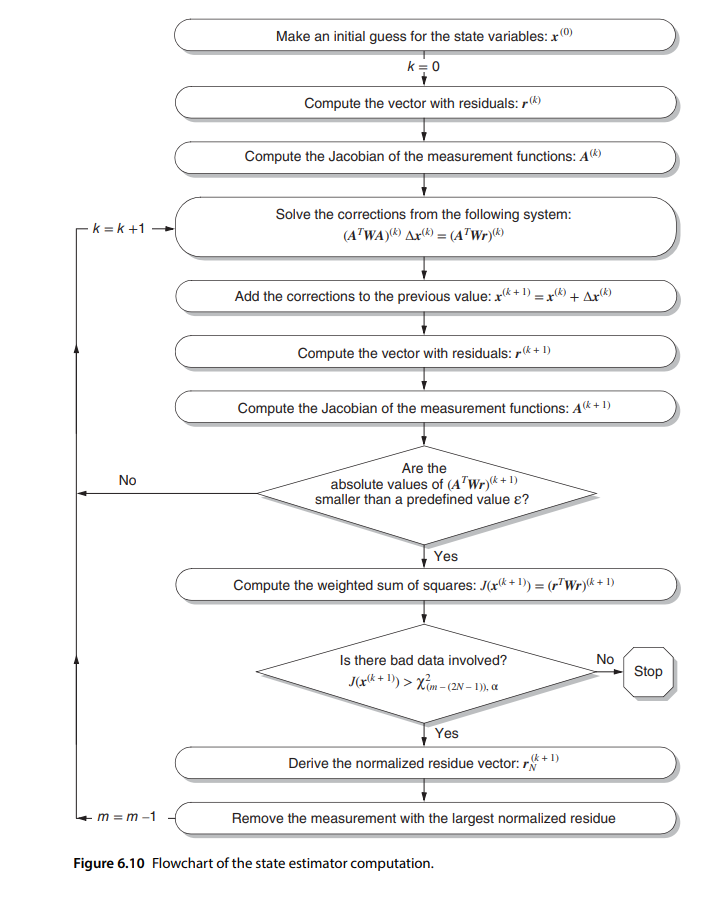
****

* there is a kind of “decoupling” between the active power and voltage angles on the one hand and the reactive power and voltage magnitudes on the other when some specific properties of power systems under normal operating conditions are taken into account:
  + The resistance of the overhead transmission lines is much smaller than its reactance.
  + The differences between the voltage angles are small.
* This “decoupling” can be recognized in the Jacobian matrix, as the elements of the off-diagonal submatrices J 12 and J 21 (the off-diagonal elements in Equation 6.62 (Example 6.6 (p. 234))) are rather small.
* Those equations seem to be decoupled as the real power mismatches are used only to calculate the voltage-angle corrections, whereas the reactive power mismatches are applied only to calculate the voltage-magnitude corrections. But the elements in the matrices J 11 and J 22 still depend on both the voltage magnitude and the voltage angle. In order to obtain an efficient computational procedure, further simplifications can be made in the matrices J 11 and J 22 when we take into account that during normal system operation [1]
  + The line susceptances are much larger than the line conductances: Gij sin(δj − δi ) ≪ Bij cos(δj − δi ).
  + The differences between the voltage angles are small: sin(δj − δi )= δj − δi and cos(δj − δi )=1
  + The reactive power injected into a node is much smaller than the reactive flow that would result if all lines connected to that bus were short-circuited to reference: Qi ≪ |Vi | 2Bii.
* In situations where a lot of load flow computations have to be made, as is, for instance, the case for reliability computations or for security analysis, a linear approximation of the load flow problem can be made to save computation time: this is called the **DC load flow**
* The DC load flow is principally different from the decoupled load flow. In the DC load flow, the nonlinear load flow equations are linearized to ease the calculation and to speed up the computation of the unknown voltages; this means that the actual model of the power system is altered, and this affects the final solution of the load flow. In the decoupled load flow, the nonlinear load flow equations are solved iteratively (the model of the power system remains unchanged), and approximations are made to the Jacobian matrix only; therefore only the speed of convergence is affected, but the final result remains the same. The DC load flow can also be applied to find a fairly good approximation of the unknown voltages that can be used as initial values in a Newton–Raphson/decoupled load flow calculation.
* Approximations:
  + The node voltage magnitudes are 1 pu.
  + The resistances of the transmission lines are neglected: Yij = Gij + jBij = |Yij|∠θij with Gij = 0 and θij = π∕2 rad
  + The differences between the voltage angles are small: sin(δj − δi )=δj − δi and cos(δj − δi ) = 1
* ->Now, a linear relation is obtained between the active power injections and the busbar voltage angles.
* **Optimal Power Flow**
* The load flow or power flow computation solves the node voltages in a given network under specified load conditions and a selected generation.
* Besides the generator power, also the injected power (e.g., reactive power from capacitor banks), the transformer tap ratios, and the phase-shifter settings can be chosen. It is this freedom of choice that makes it possible to optimize the system for certain criteria. **An optimal power flow is a load flow computation**, in which the calculation of the node voltages and the optimization of the controllable variables are carried out simultaneously.

****

* **State Estimator**
* The operator of a power system must be able to observe every part of the system at all times. The system view should be coherent; there should be no gaps, no incongruities, and no misinformation. The SCADA system, which collects real-time measurements from the power system, cannot fulfill all of these requirements. Some measurements may not be available, other measurements may be corrupted because of hardware failures or communication problems, redundant measurements of the same quantity are rarely the same, and even some measurements could be totally wrong. The combination with the state estimator overcomes the deficiencies of the SCADA system. In effect, the state estimator enhances and maintains the integrity of the real-time database, thus making it possible for the control center software to support the operator with a complete, consistent, and accurate system overview
* **Diagram

  Description automatically generated**
* The state of the power system is described by the bus voltage magnitudes and phase angles. The number of state variables amounts to 2N −1 (with N the number of nodes in the network), as one phase angle serves as a reference and is not included in the state vector:
  + **X = [**δ2,…, δN,|V|1,…, |V|N]
* **General Scheme of the State Estimator**
* The presence of bad data among the observations processed by a least-squares estimator is, as a rule, detrimental to the performance of the estimator and usually results in poor state estimates
* The statistical properties of the measurement errors facilitate the detection and identification of bad data. These properties are described in Section 6.4.3. Statistical theory shows that the weighted sum of squared residues has a chi-square distribution with m −(2N −1) degrees of freedom (as demonstrated in the following section). However, this is no longer the case if a measurement that is erroneous enough to violate the normality assumption of the measurement noise vector is present. Therefore, bad data can be detected by means of a chi-square test
* After each state estimation run, the weighted sum of squared residues is computed. This value is compared with a critical value from a chi-square distribution with m −(2N −1) degrees of freedom and a specified probability α, which is the probability that the sum of weighted squared residues exceeds the critical value. If the weighted sum of squared residues is larger than this critical value, one concludes that bad data are present, and an identification procedure can be invoked to find out which measurements are erroneous. Otherwise, the state estimates are accepted on the ground that there is not enough evidence to indicate the presence of bad data.
* Locating the bad data requires the individual examination of the estimation residues. A possible identification strategy could be to find the maximum residue and then to conclude that the corresponding measurement is the faulty one. However, this is not necessarily true for two reasons
  + The residues are, in general, correlated among themselves so that an error associated with a measurement can spread over other residues.
  + Meters (for different quantities) can have different accuracies and the variances of the corresponding measurements can be significantly different.
* As different types of meters can have different variances, a residual value that is an outlier for a specific measurement could be very well acceptable for another one. Therefore, a normalization of the residues is necessary to make up for that imbalance. A convenient and simple way to make a comparison of residues is to normalize them with respect to their standard deviations. After this normalization, the measurement that corresponds to the maximum normalized residue is, most likely, the bad measurement.

****

**WINNING THE L2RPN CHALLENGE: POWER GRID MANAGEMENT VIA SEMI-MARKOV AFTERSTATE ACTOR-CRITIC**

* While most approaches have focused on controlling the generation or the load of electricity (Venkat et al., 2008; Zhao et al., 2014; Huang et al., 2020), managing the power grid through the topology control (changing the connection of power lines and bus assignments in substations) would be the ultimate goal.
* By reconfiguring the topology of the power grid, it can reroute the flow of electricity, which enables the transmission of electricity from the producers to consumers efficiently and thus prevent surplus production
* One of the main challenges in RL for the real-world scale power grid management lies in its massive state and action space
* We address the problem by adopting a goal-conditioned hierarchical policy with the afterstate representation. First, we represent state-action pairs as afterstates (Sutton & Barto, 2018), the state after the agent has made its decision but before the environment has responded, to efficiently cover the large state-action space
* The afterstate representation can be much more succinct than the state-action pair representation when multiple state-action pairs are leading to an identical afterstate. For example, in the case of controlling the topology of the power grid, a pair of a current topology and an action of topology modification can be represented as a reconfigured topology, since the topology is deterministically reconfigured by the action
* Then the next state is determined by random external factors, such as the change of power demands in load. Second, we extend this idea to a hierarchical framework, where the high-level policy produces a desirable topology under the current situation, and the low-level policy takes care of figuring out an appropriate sequence of primitive topology changes. Combined together, our hierarchical policy architecture with afterstates facilitates effective exploration for good topology during training.
* The power grid is essentially a graph composed of nodes corresponding to substations that are connected to loads, generators, and power lines. The generator produces electricity, the load consumes electricity, and the power line transmits electricity between substations. The substation can be regarded as a router in the network, which determines where to transmit electricity. Grid2Op considers 2 conductors per substation, known as the double busbar system. This means that the elements connected to a substation, i.e. loads, generators, and power lines, can be assigned to one of the two busbars, and the power travels only over the elements on the same busbar. Thus, each substation can be regarded as being split into two nodes.
* The state of the power grid consists of various features such as a topology configuration (the connectivity of each power line and the bus assignment in each substation), as well as the amount of power provided by each generator, required by each load, transmitted in each line, and so on. The power supplied by generators and demanded by loads changes over time, and the power transmitted in lines also changes according to the current topology configuration together with supply and demand. In addition, each line has its own capacity to transmit electricity and can be automatically disconnected when there is an overflow of electricity

**Data(set)**

* Simbench
* Choose datasets with different bus numbers that provide valid ACOPF solutions
* Supervised: result from pandapower as target
* Unsupervised: GNN is to provide state features for every node (bus), which are adjusted/configured according to the result from pandapower when these state features are set.

**Pipeline:**

1. Uniformly sample p\_ref as in p\_L = ~ Uniform(0.9 \* p\_ref, 1.1 \* p\_ref) in net.load
2. Uniformly sample q\_ref as in q\_L = ~ Uniform(0.9 \*q\_ref, 1.1 \* q\_ref) in net.load
3. Set the values in net.load with results
4. Let Pandapower calculate the DCOPP solution
5. Replace voltage magnitudes with initial bus voltages and q\_mvars with reactive power demands from net.load
6. Let Pandapower calculate the interior point fixed ACOPF solution
7. Repeat to create 100 datasets from one grid example
8. Save each of X and Y in a Python list to later be partitioned as Train, Validation and Test Sets
9. Each “batch” in the trainingset will thus be a sampled entire grid with size = Number of buses in the grid