Stochastic Information Management in Smart Grid

Hao Liang, Member, IEEE, Amit Kumar Tamang, Weihua Zhuang, Fellow, IEEE, and Xuemin (Sherman) Shen, Fellow, IEEE

Abstract—Rising concerns about the efficiency, reliability, economics, and sustainability in electricity production and distribution have been driving an evolution of the traditional electric power grid toward smart grid. A key enabler of the smart grid is the two-way communications throughout the power system, based on which an advanced information system can make optimal decisions on power system operation. Due to the expected deep penetration of renewable energy sources, energy storage devices, demand side management (DSM) tools, and electric vehicles (EVs) in the future smart grid, there exist significant technical challenges on power system planning and operation. Specifically, efficient stochastic information management schemes should be developed to address the randomness in renewable power generation, buffering effect of energy storage devices, consumer behavior patterns in the context of DSM, and high mobility of EVs. In this paper, we provide a comprehensive literature survey on the stochastic information management schemes for the smart grid. We start this survey with an introduction to the smart grid system architecture and the technical challenges in information management. Various component-level modeling techniques are presented to characterize the sources of randomness in the smart grid. Built upon the component-level models, we further explore the system-level stochastic information management schemes for smart grid planning and operation. Future research directions and open research issues are identified.

Index Terms—Demand side management, electric power system, vehicle information and communication systems, microgrid, renewable energy, smart grid, stochastic control, stochastic modeling, stochastic optimization.

I. INTRODUCTION

As named by the National Academy of Engineering (NAE) in the United States, electrification is “the most important engineering achievement of the 20th century”. Electricity (along with natural gas and refined petroleum products) is and will continue to be a major source of energy supply for residential, commercial, and industrial sectors in the foreseeable future. However, concerns have been raised about the efficiency, reliability, economics, and sustainability of the decades-old electric power grid. Penetration of renewable energy sources is increasing at a rapid rate, thanks to government incentives, falling installation costs, and rising fossil fuel prices. According to the International Energy Agency (IEA) forecast, electricity generation from renewable energy sources will be nearly tripled from 2010 to 2035, reaching 31% of the world’s total power generation. Hydro, wind, and solar are three of the major renewable energy sources, which will provide 50%, 25%, and 7.5% respectively of the total renewable power generation in 2035 [1]. On the other hand, to reduce the greenhouse gas (GHG) emissions in energy consumption, electricity customers have been participating and will continue to participate in the demand side management (DSM) programs which provide incentives via energy bill savings. For instance, 4.7 million smart meters have been installed in Ontario, Canada, as of February 2012 and 3.8 million Ontarians are on time-of-use rates [2]. In addition to the residential, commercial, and industrial sectors which are the main consumers of electricity, there is an inevitable trend of electrification in the transportation sector to further reduce the GHG emissions. According to the Electric Power Research Institute (EPRI), the electric vehicle (EV) penetration level in the United States can reach 35%, 51%, and 62% by 2020, 2030, and 2050, respectively [3]. Also, it is estimated that there will be at least 500,000 highway-capable EVs on Canadian roads by 2018, as well as a possibly larger number of hybrid-electric vehicles [4]. Innovated power transmission & distribution (T&D) systems, microgrids, and energy storage devices will be developed to ensure efficient and reliable power delivery to maximize the utilization of renewable energy sources, EVs, and DSM tools. However, there exist significant technical challenges in power system operation and control, due to the intermittency of renewable energy resources, buffering effect of energy storage devices, consumer behavior patterns in the context of DSM, and high mobility of EVs. In order to address these challenges, an evolution of the traditional electric power grid to a “smart grid” is underway.

One of the first references to the term “smart grid” is an article published in the September/October 2005 issue of the IEEE Power and Energy Magazine by Amin and Wollenberg, entitled “Toward a smart grid” [5]. Due to the complexity of involved technologies and the variety of visions from stakeholders, the smart grid gives rise to a number of definitions and explanations [6]. The following are a few examples published by authorities such as U.S. Department of Energy (DOE) [7], Independent Electricity System Operator (IESO) [8], and the National Association of Regulatory Utility Commissioners (NARUC) [9].

• DOE definition – “An automated, widely distributed energy delivery network, the smart grid will be characterized by a two-way flow of electricity and information and will be capable of monitoring everything from power plants to customer preferences to individual appliances. It incorporates into the grid the benefits of distributed computing and communications to deliver real-time information and enable the near-instantaneous balance of supply and demand at the device level.”

• IESO definition – “A smart grid is a modern electric system. It uses communications, sensors, automation and
computers to improve the flexibility, security, reliability, efficiency, and safety of the electricity system.”

- **NARUC definition** – “The smart grid takes the existing electricity delivery system and makes it ‘smart’ by linking and applying seamless communications systems that can: 1) gather and store data and convert the data to intelligence; 2) communicate intelligence omnidirectionally among components in the ‘smart’ electricity system; and 3) allow automated control that is responsive to that intelligence.”

Despite the variety in smart grid definitions, we can conclude that smart grid is an electrical grid that uses information and communication technologies to gather information and act accordingly in an automated fashion to improve the efficiency, reliability, economics, and sustainability of electricity production, transmission, distribution, and consumption.

The IEEE 2030 standard on smart grid was introduced in September 2011, which provides guidelines in understanding and defining the interoperability of information and communication technology with the power system, end-user applications, and loads [10]. The smart grid architecture is defined based on the interconnection of an electric power system, a communication system, and an information system, as shown in Fig. 1. In literature, there are several surveys and tutorials on the architectural perspective of the smart grid with respect to the following topics:

- Overview of the smart grid [11] [12];
- Smart grid information system architecture [13];
- Smart grid communication networks [14]–[18];
- Smart grid cyber security and privacy support [19]–[22].

The existing works summarize various approaches to integrate renewable energy sources, energy storage devices, DSM tools, and EVs in the smart grid, and two-way communication techniques for information acquisition and notification, while providing cyber security and privacy support. Further, it is technically challenging to utilize the information acquired through smart grid communications to make optimal decisions on power system planning and operation. Some studies in the area of computational intelligence are presented in [23] for sensing, situational awareness, control, and optimization in the smart grid. Nevertheless, a comprehensive literature review for information management in the smart grid will help to develop new solutions to meet the technical challenges.

In this paper, we focus on the information system of smart grid. Various stochastic information management schemes are surveyed to address the technical challenges on system planning and operation for integrating renewable energy sources, energy storage devices, DSM tools, and EVs. The smart grid system architecture is investigated, based on which the sources of randomness in information management are identified. Component-level stochastic models are further incorporated in the system-level stochastic information management schemes with respect to all domains of the electric power system, including bulk generation, transmission, distribution, and consumption.

The organization of this paper is shown in Fig. 2. Section II describes the smart grid system architecture and briefly introduces the three subsystems in smart grid, with a focus on the technical challenges in information management. In Section III, the component-level stochastic models are presented. Since the system-level stochastic information management is closely related to power system planning and operation functions, Section IV presents an overview of these functions and the associated theories and techniques that can be used for stochastic information management. The state of the art in stochastic information management for bulk generation and transmission systems, distribution systems and microgrids, and DSM is presented in Section V, Section VI, and Section VII, respectively. Due to the unique features of EVs (such as their mobility) in comparison with the traditional electric power system components, we discuss the stochastic information management schemes for EV integration in a separate section, i.e., Section VIII. Section IX concludes this study and discusses open research issues.

II. SMART GRID SYSTEM ARCHITECTURE AND TECHNICAL CHALLENGES IN INFORMATION MANAGEMENT

According to the IEEE 2030 standard [10] and as shown in Fig. 1, the smart grid system architecture is based on an interconnection of three subsystems:
An electric power system which accomplishes the generation, transmission, distribution, and consumption of electricity;

2) A communication system which establishes the connectivity for information exchange among different systems and devices; and

3) An information system which stores and processes data information for decision making on power system operation and management.

Different service providers can participate in the electricity market to provide electricity services to customers and utilities.

In comparison with the traditional electric power systems, more renewable-energy-based distributed generation (DG) units and energy storage devices (including EVs) are integrated in the smart grid. As a result, the traditional electricity consumers are being gradually transformed into electricity “prosumers” who not only consume energy but can also produce energy and feed it to the power grid. Therefore, the basic assumption of unidirectional electricity delivery (from centralized generators to electricity customers) in the traditional electric power system is no longer practical. Bidirectional energy flows need to be established between electricity customers and power distribution systems, as shown in Fig. 1. Moreover, a number of DG units, energy storage devices, and loads in close proximity can be interconnected as a microgrid, which is able to operate in either a grid-connected mode or an islanded mode for reliability enhancement while reducing transmission and distribution losses.

Three kinds of communication networks can be established in the smart grid. A wide area network (WAN) facilitates the communications among bulk generators and transmission facilities for wide-area situational awareness. A neighborhood area network (NAN) or field area network (FAN) supports the communications among distribution substations and field electrical devices for power distribution and microgrid operation. Home area networks (HAN), business area networks (BAN), and industrial area networks (IAN) can be deployed within residential, commercial, and industrial buildings, respectively, for communication among electrical appliances for the DSM purpose. The research and development on smart grid communication networks have been extensively carried out. The smart grid communication network architectures, performance requirements, research challenges, state-of-the-art technologies, development aspects, and experimental studies have been discussed in [14]–[18]. As more and more electric devices in the critical power infrastructure are interconnected via communication networks, cyber security has an immediate impact on the reliability of smart grid. Furthermore, increased connectivity of electrical appliances at the customer side can enable personal information collection, which may invade customer privacy. The cyber security requirements, network vulnerabilities, attack countermeasures, secure communication protocols and architectures, and privacy issues in the future smart grid have been surveyed in recent literature [19]–[22].

Based on information acquired via the communication system, the information system can make optimal decisions on electric power system operation and transmit the control signals via the corresponding communication networks. Although basic information management functionalities are already in place in traditional bulk generation and transmission systems based on the supervisory control and data acquisition (SCADA) systems, developing an advanced information management system in the context of smart grid is technically complex due to the following challenges:

- The output of renewable energy resources is intermittent in nature, which results in large variations in power supply. Although a large body of studies have been carried out to forecast such an uncertain output, the stochastic nature of renewable power generation should be addressed in smart grid planning and operation;

- The buffering effect of energy storage devices not only introduces more state variables in power system operation, but also requires to account for the inter-period buffer state transitions over the entire time frame (which can be up to a week) under consideration. Efficient management schemes should be designed for energy storage devices at a low computational complexity;

- Customer behavior patterns in the presence of DSM are more dynamic than in the traditional electricity grid, which leads to large variations in load demand. The main reason is that the usage of electrical appliances can be shifted over time by electricity customers in response to electricity prices. Moreover, different customers can collaborate with each other to reduce their overall energy bills, based on the information obtained via FAN/NAN communications;

- EV drivers can select different charging locations in response to electricity prices, which can lead to large
variations in charging demand and poor accuracy of charging demand estimation. Further, high EV mobility can result in highly dynamic energy storage capacity of the electric power system, taking account of the random nature in route and/or commute schedules of EV drivers. To address these technical challenges, first we establish proper stochastic models to characterize randomness in renewable power generation, buffering effect of energy storage devices, consumer behavior patterns, and EV mobility. Then, we incorporate the stochastic models in the system-level information management to facilitate smart grid planning and operation.

III. COMPONENT-LEVEL STOCHASTIC MODELS

In this section, we present stochastic models to characterize the randomness in wind and solar power generation, customer energy demand, EV mobility, and component outage. Models for energy storage devices are discussed in comparison with data buffer models in communication networks.

A. Wind Power Generation

Wind speed can be modeled as a random variable following a Weibull distribution [24], with a probability density function (PDF) given by

$$f(v) = \frac{k}{c} \left( \frac{v}{c} \right)^{k-1} e^{-\left( v/c \right)^k} \quad (1)$$

where $c$ and $k$ are the Weibull scale parameter and dimensionless Weibull shape parameter, respectively, indicating the wind strength at the location under consideration and the peak of the wind distribution. The Weibull distribution has a high value of $k$ if the wind speed is very likely to take a certain value. Given a wind turbine, the generation of active power can be represented as a function of the wind speed, which is typically referred to as the power curve. The power curve of a VESTAS 600 kW wind turbine is plotted in Fig. 3 [25].

It is important to incorporate the variation in wind energy during diurnal cycles [24]. The wind energy assessment based on the Weibull distribution and average daily/seasonal wind speeds may not accurately characterize the variation in wind speed probabilities during day and night. This may cause significant over/underestimation of wind power potential when the wind power generation estimation is linked to electricity loads. In order to establish the spatial and temporal correlation in wind power generation, more sophisticated Markov chain models can be used [26]. In a wind farm, the wind speed profile is identical for wind turbines sharing the same row, while the wind speed profile differs across rows [27]. The characteristics should be modeled, such as by reducing the wind speed values from one row to the next, in the direction of the incident wind.

B. Solar Power Generation

Solar power generation uses a photovoltaic (PV) system to generate electricity. The output power of a PV system depends on three factors, namely solar cell temperature, solar radiation intensity, and PV system efficiency. Among them, the PV system efficiency depends upon the other two variables. The PV output power, $P_{PV}(t)$, at time $t$ is given by

$$P_{PV}(t) = \epsilon(t) I(t) \quad (2)$$

where $\epsilon(t)$ and $I(t)$ represent the efficiency and radiation intensity, respectively [28]. The efficiency is a function of the radiation intensity and can be calculated as

$$\epsilon(t) = \begin{cases} \frac{\eta_c}{K_c} I(t), & 0 < I(t) < K_c \\ \eta_c, & I(t) \geq K_c \end{cases} \quad (3)$$

where $K_c$ is a threshold of radiation intensity beyond which the efficiency is approximated to be a constant ($\eta_c$). The radiation intensity, $I(t)$, is a sum of deterministic fundamental intensity $I_d(t)$, which is determined by solar altitude angle, and stochastic attenuation amount $\Delta I(t)$ with respect to clouds occlusion and weather effects. The intensity $I_d(t)$ depends on the time of a day and the seasons of a year. The randomness in $\Delta I(t)$ can be modeled by a normal distribution [29]. A typical curve of $I(t)$ is shown in Fig. 4, which follows a quadratic function, neglecting seasonal and sunrise/sunset time effect.

C. Energy Demand

Electricity consumption can be modeled based on a bottom-up technique [30], where the load profile is constructed based on elementary load components such as households or even individual appliances. A simplified bottom-up model is presented in [30], which incorporates the seasonal/hourly and social factors in a probabilistic manner, and can be used to generate realistic domestic electricity consumption profiles on an hourly basis for up to thousands of households. An energy demand model is proposed in [31], taking account the user interactions in real world home energy management. Two prediction algorithms are proposed to estimate the future behavior of a smart home, including a day type model (DTM) and a first order semi Markov model (SMM). The DTM...
assumes a certain regularity of the appliance usage. In model training, the complete record of action sequence is split into day sequences. The days showing a comparable appliance usage are grouped into a specific day type. A decision tree induction technique is used to discover the association rules between contexts and the day types. According to the SMM model, on the other hand, a user action only depends on the previous action and a probability for the transition between two actions. Different from traditional continuous-time Markov models which assume an exponential distribution for the duration of a state transition, the SMM uses an arbitrary distribution for the state duration which is more realistic. In the model training, a count matrix is used to estimate the transition probability between two action types.

D. Vehicle Mobility

The arrival of EVs at a specific charging station follows a Poisson process [32]. This observation conforms with the vehicle mobility models used in communication network research and is verified by experiments [33]. As a result, the inter-arrival time of EVs at a charging station is exponentially distributed. To further capture the spatial and temporal dynamics of an EV traffic flow, fluid traffic theory can be applied. A fluid model is established in [56] based on partial differential equations and the conservation equations of EV traffic flow. When the commute patterns of EV drivers are taken into account, and the conservation equations of EV traffic flow, fluid traffic theory can be applied. A fluid model is typically used, where the internal resistance of a power electronics point of view [38]. A Thevenin-based circuit model is typically used, where the internal resistance of a battery is a non-linear function of the state-of-charge (SOC). As a result, the energy losses in battery charging/discharging operation failure.

Fig. 5: Distribution of the Kitchener metropolitan area commuting distance to work, 2006 [35].

Fig. 6: Three-state PV panel model [28].

E. Component Outage

The random outages and repair process of generators can be modeled by a two-state continuous-time Markov chain [36]. Let $p$ denote the availability probability of a generator and $q = 1 - p$ its unavailability probability, and let $\mu$ and $\lambda$ denote the repair and failure rates of the generator, respectively. Denote the availability of the generator at time $t_0$ and $t$ ($t > t_0$) as $A_{t_0}$ and $A_t$, respectively, and let 1 and 0 represent the up and down status, respectively. Then, the conditional probability $P(t = 0 \mid A_{t_0} = \alpha) (\alpha, \beta \in \{0,1\})$ associated with the availability $\beta$ of the generator at time $t$, given its status $\alpha$ at time $t_0$, is studied in [37]. This model can be applied to hydro and gas generators, and can be potentially extended to model the availability of transmission lines [36].

The failure of a PV panel due to weather effects is modeled in [28], given the fact that a PV panel is more likely to fail in a harsh weather condition (e.g., a lightning storm) in contrast to a normal weather condition. A three-state PV panel model is established in [28] as shown in Fig. 6, where both failure rate ($\lambda$) and repair rate ($\mu$) are taken into account. In Fig. 6, state 1 and state 2 correspond to the states that the radiation is larger and smaller than $K_c$, respectively. State 3 represents an outage in which the PV panel generates no electricity. The PV panel enters state 3 when there is no solar radiation or an operation failure.

F. Energy Storage

Batteries are a widely used means of energy storage. Microscopic battery models are available in literature from a power electronics point of view [38]. A Thevenin-based circuit model is typically used, where the internal resistance of a battery is a non-linear function of the state-of-charge (SOC). As a result, the energy losses in battery charging/discharging and self-discharging (when the battery is stored for a long time) is dependent on the SOC. For each specific battery, the internal resistance needs to be measured to establish a proper microscopic battery model.

To reduce the modeling complexity and facilitate system-level studies, macroscopic battery models can be used [34] [39] [40]. The modeling of a battery is similar to that of a data buffer in communication networks in a sense that the buffering effect can be characterized by certain arrival and departure processes. However, the data buffer models cannot be directly applied because of the electricity
characteristics of batteries. Specifically, the following unique characteristics need to be considered when establishing a macroscopic battery model:

- In each charge and discharge of a battery, a certain amount of energy is lost because of the battery internal resistance and energy conversion loss. The energy loss can be modeled as proportional to the charged/discharged energy based on an average loss rate;
- The lifetime of a battery is shortened after each charging/discharging cycle since the capacity of the battery slowly deteriorates, depending on the depth-of-discharge (DoD). Although the deterioration is almost imperceptible on a daily basis, the loss of the battery value needs to be considered as a cost, which is proportional to the charged (or discharged) energy;
- At each time moment, the battery can be either charged or discharged, but not both. In order to prolong the battery lifetime, the SOC of the battery should not drop below a certain threshold;
- Because of the self-discharge effect, the energy stored in a battery decreases over time.

Quantized values of the above characteristics depend on the types of batteries such as lead-acid, nickel metal hydride, and lithium-ion batteries, and need to be estimated.

In addition to the batteries, there are other types of energy storage devices such as flywheels and heat buffers. Their models are different from that of the batteries. Flywheel stores kinetic energy. The amount of energy stored in a flywheel varies linearly with the moment of inertia and quadratically with the angular velocity \([41]\). An increase in the angular speed increases the energy stored in a flywheel, at the cost of increased energy losses due to higher frictions and thermal losses. On the other hand, a micro combined heat and power (microCHP) unit can be combined with a heat buffer to provide an efficient means for domestic energy storage \([42]\). However, the cost of state transitions (such as the startup cost) needs to be considered in the modeling of a microCHP unit.

IV. SYSTEM-LEVEL STOCHASTIC INFORMATION MANAGEMENT

System-level information management deals with various functions in the planning and operation of an electric power system, such as system planning, system maintenance, unit commitment, economic dispatch, regulation, control, and protection \([43]\). These functions are performed at different time frames, as listed in Table I. The foundation of all planning and operation functions is a power flow analysis. To illustrate the concept of power flow analysis, we used a four-bus power system \([44]\) as an example, with its one-line diagram as shown in Fig. 7. Each bus in the system is deployed at a specific location (i.e., Birch, Elm, Pine, and Maple for buses 1-4, respectively, in Fig. 7) and corresponds to a power generator (for power generation) or a distribution substation at a load center (for power distribution). In Fig. 7, there are two generators G1 and G2 which are connected to bus 1 and bus 4, respectively. Each bus \(i\) in the power system can be described by four scalar parameters, i.e., net active power injections \(P_i\), net reactive power injection \(Q_i\), voltage (magnitude) \(V_i\), and phase angle \(\delta_i\), where the net active and reactive power injections, respectively, equal the active and reactive power generation by generator minus load (denoted by arrow) at the corresponding bus. There are three types of buses in the system:

- A \(PQ\) bus is used to define a load bus, where the net active and reactive power injections \(P_i\) and \(Q_i\) (which equal the negative values of the active and reactive power demand, respectively) are determined by the corresponding load;
- A \(PV\) bus is used to define a generator bus, where the net real power injection \(P_i\) and voltage \(V_i\) are specified by the corresponding generator;
- One of the generator buses in the system should be selected as a \(slack\) bus, where the voltage \(V_i\) and phase angle \(\delta_i\) are used as the system reference. Since the net power injections \(P_i\) and \(Q_i\) are adjustable, the slack bus can balance the active and reactive power in the system and compensate for the losses.

According to the above definition, two parameters are known for each bus in the system while the other two parameters need to be calculated. The buses in the system are connected via a set of transmission lines as shown in Fig. 7. An impedance, \(X_{ij}\), is specified for the transmission line connecting a pair of two buses \(i\) and \(j\).

Power flow analysis is performed to calculate the unknown parameters of each bus in the power system. Based on circuit analysis, power flow equations can be established, which are typically a system of non-linear equations. Since the number of known and unknown parameters are equal in the system, the power flow equations can be solved based on typical methods such as Gauss-Seidel and Newton-Raphson methods \([44]\). Based on the solution, all parameters of all buses can be obtained, which can be further utilized to calculate the active and reactive power flows through each transmission line. For instance, the active power flow \(PF_{ij}\) from bus \(i\) to bus \(j\) (e.g., \(PF_{13}\) from bus 1 to bus 3 in Fig. 7) is given by

\[
PF_{ij} = \frac{V_i V_j}{X_{ij}} \sin(\delta_i - \delta_j). \tag{4}
\]

The power flow on each transmission line should be limited without violating the line flow limit (or thermal limit) of the transmission line. Otherwise, the power generation by generators need to be re-dispatched or re-scheduled to change...
the values of $P_i$’s for PV buses, such that feasible power flow solutions can be obtained. Sometimes, a feasible solution cannot be obtained by merely re-dispatching. In such a case, electric loads need to be curtailed to increase the values of $P_i$’s and $Q_i$’s of PQ buses, which typically causes blackouts for some electricity customers.

All the decisions on power generation scheduling and load curtailment should be made by an information management system for power system operation. The basic requirement of power system operation is to balance the amount of electric power production and consumption at each time instant, while satisfying the power system constraints such as the capacity limit (i.e., maximum active and reactive power generation) of each generator and the line flow limit of each transmission line. In this paper, we focus on the following functions of information management for power system operation:

- **Unit commitment** – The unit commitment problem can be stated as finding the optimal decision on which power generator should be on-line (or active) over time, which minimizes the operation cost of the system. Three kinds of costs need to be considered, i.e., fixed cost of on-line generator, power generation cost, and power generator startup cost [45]. Consider the example in Fig. 7. If the load demand is low and can be fully supported by generator G1 at a low power generation cost, generator G2 can be shut down to avoid an extra operation cost incurred by the fixed cost of on-line generator. The unit commitment decision should base on system load dynamics, since it is not economical to frequently start up and shut down a power generator because of the startup cost;

- **Economic dispatch** – Economic dispatch makes short-term decisions on the optimal power generation of each on-line power generator in the system to meet the load demand at a minimum cost, while satisfying power system constraints to ensure reliable power system operation. A second-order cost function $C_g(\cdot)$ is typically used to represent the power generation cost of a generator ($g$) [46], given by

$$C_g(P_g) = a_g P_g^2 + b_g P_g + c_g, \ g \in \mathcal{G}$$

(5)

where $\mathcal{G}$ is the set of generators in the system (e.g., $\mathcal{G} = \{G1, G2\}$ in Fig. 7), $P_g$ is the active power output of generator $g$, and $a_g$, $b_g$, and $c_g$ are the generation cost coefficients. Consider the example in Fig. 7. If the load demand is high and needs to be shared among the two generators, there may exist an optimal tradeoff point between the power generations by the two generators (without violating the power system constraints) due to the nonlinearity of the cost function (5), which corresponds to an optimal economic dispatch decision;

- **Power generation control** – Based on the economic dispatch decisions, power generation control (also referred to as the automatic generation control in traditional electric power systems) can be performed to adjust the output of generators in a power system, in response to instant changes in the load [47]. Since power generation and load demand should be balanced closely in a power system, frequent adjustments to the outputs of generators are necessary. The adjustments can be performed based on system frequency, which increases if there is more power generation than load demand, and vice versa.

Another critical function of the information management is system planning, which aims at finding the optimal combination, design, and sizing of energy sources and energy storage devices to meet the future electricity demand at a minimum lifecycle cost, while taking into account the environmental issues [48].

Despite a rich literature on the planning and operation of traditional electric power systems, the proposed schemes cannot be directly applied to the future smart grid with a deep penetration of renewable energy sources, energy storage devices, DSM tools, and EVs. Specifically, the randomness in renewable power generation, buffering effect of energy storage devices, consumer behavior patterns in the context of DSM, and high mobility of EVs should be considered. To address this problem, system-level stochastic information management schemes should be developed by incorporating the component-level stochastic models discussed in Section III into the planning and operation of different domains of the electric power system, including bulk generation and transmission, distribution, and customers. Stochastic modeling, optimization, and control techniques are studied recently in literature for the system-level stochastic information management, which have a potential for application in the future smart grid. A brief summary of the basic theories and techniques is given below:

- **Convolution technique** – Given two random variables $X$ and $Y$ with PDFs $f_X(x)$ and $f_Y(y)$, respectively, in a linearized system, the PDF of an output random variable $Z = X + Y$ can be calculated as $f_Z(z) =
be modeled as a queuing system, the queueing theory can be applied for EV charging station planning and operation [56]. Moreover, an energy storage device can be modeled as a queue based on an analogy between the energy stored in the device and the number of customers in a queue [57];

- **Stochastic inventory theory** – The theory is concerned with the optimal design of an inventory (or storage) system to minimize its operation cost. Different from the queueing models, the ordering (or arrival) process of an inventory can be regulated. The inventory theory studies the optimal decision making process in terms of when and how much to replenish the inventory based on the stochastic information of future demands. It can be applied for energy storage device operation based on an analogy between the energy storage and inventory level [34];

- **Monte Carlo simulation (MCS)** – MCS generates scenarios according to certain distributions of the random variables in the system. A deterministic problem (e.g., power flow analysis) is solved for each scenario [58] [59]. The system performance metrics are evaluated based on the solutions of the deterministic problems and the probability that each of the scenarios is generated. MCS can be applied to performance evaluation of stochastic information management in electric power systems when high accuracy can be achieved in system dynamics modeling. Although MCS is computationally expensive, the results obtained via MCS can be used as the benchmarks to evaluate the performance of other stochastic information management techniques such as the convolution technique and moment estimation technique.

Various stochastic information management schemes are proposed in literature based on these basic theories and techniques and their variations and/or modifications, to be discussed in details in the following sections. The literature associated with each of domains in the electric power system (in terms of bulk generation and transmission, distribution, and consumption) is summarized according to the functions of smart grid planning and operation as listed in Table I.

### V. BULK GENERATION AND TRANSMISSION SYSTEMS

The bulk generation and transmission systems for smart grid are mostly evolved from those of the traditional electric power grid, while more renewable energy sources and advanced information and communication systems are incorporated. An overview of the bulk generation and transmission systems is given in Fig. 8, which is based on the Ontario case [60]. Specifically, the bulk generation refers to the generators of electricity in bulk quantities, including both conventional and renewable energy sources such as nuclear, coal, gas, solar, and wind. Since electric power is typically generated at a relatively low voltage like 30 kilovolt (kV), step-up transformers are used to increase the voltage and transfer the electric power to the high-voltage (e.g., 230/500 kV) transmission lines, such that electricity can be transmitted at low losses. Through long-distance transmissions (typically tens or hundreds of kilometers), the electric power reaches the distribution substations
A. Probabilistic Power Flow

One prerequisite information of the traditional power flow analysis is the net active power injection $P_i$ of each generator bus (i.e., PV bus) $i$, as shown in Fig. 7, which is based on the economic dispatch decisions. However, because of the potential high penetration of renewable energy sources in the future smart grid, the value of $P_i$ becomes a random variable which depends on weather conditions. As a result, the traditional power flow analysis needs to be extended to a probabilistic power flow (PPF) analysis. The PPF is a technique to derive the probability distribution of the output variables of power flow analysis such as bus voltages and line flows, given that the input variables such as power generation and load are represented by random variables following certain distributions. The PPF analysis was originally proposed to address the randomness in load demand, and is recently extended to investigate the randomness in renewable power generation of an electric power system.

MCS combined with simple random sampling (SRS) is a popular method in literature for solving PPF problems with load uncertainties [58]. The original technique can be extended to take into account the stochastic nature of DG output [59], where the uncertainties in both locations and on/off state of the DG units are incorporated in the problem formulation, and a Newton-Raphson method can be used to solve the power flow equations. In order to calculate the correlation between the stochastic inputs, a multidimensional stochastic dependence structure can be used in MCS [61], where the mutual dependence is addressed by either stochastic bounds or a joint normal transform method.

Given a large sample size, the MSC with SRS can provide accurate solutions for PPF problems, but at the cost of a heavy computational burden. In order to address this problem, a stratified sampling technique - Latin hypercube sampling (LHS), with random permutation, can be used [62] [63]. However, when the LHS is used to solve multivariate input random problems, the accuracy is affected by the correlations between samples of different input random variables. In order to minimize the undesired correlations between samples to improve the accuracy of a PPF solution, an efficient sampling method, namely the LHS combined with Cholesky decomposition (LHS-CD) method, can be applied [64]. The probabilistic distributions of input random variables can be well captured by the LHS, while the undesired correlations between samples of different input random variables are reduced by Cholesky decomposition. To better characterize the correlated wind speeds of different wind farms, an extended Latin hypercube sampling algorithm can be used to solve PPF problems [65]. By employing rank numbers of the sampling points to generate correlated wind speed samples for different wind farms, negative wind speed values can be avoided during the transformation from uncorrelated samples to correlated samples, which improves the sampling accuracy. To further reduce the computational complexity of MCS, the scenario reduction technique can be used. A probabilistic distribution load flow (PDLF) algorithm is presented in [25] to study the effect of connecting a wind turbine to a distribution system.
TABLE II: Methodology for stochastic information management in bulk generation and transmission systems.

<table>
<thead>
<tr>
<th>Applications</th>
<th>Theory or Technique</th>
<th>Variations and/or Modifications</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MCS</td>
<td>MCS with SRS [58] [59] [61]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MCS with LHS [62] [63]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MCS with LHS-CD [64]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MCS with extended LHS for wind farms [65]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MCS with classification based scenario reduction [25]</td>
</tr>
<tr>
<td>PPF analysis</td>
<td>Convolution technique</td>
<td>Convolution technique with Fast Fourier Transform [66]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Convolution technique with Von Mises method [67]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Convolution technique with combined Cumulants and Gram-Charlier expansion theory [68]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Combined convolution technique and MCS [69]</td>
</tr>
<tr>
<td></td>
<td>Interval based technique</td>
<td>Interval arithmetic [50]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Affine Arithmetic method [70] [71]</td>
</tr>
<tr>
<td></td>
<td>MCS</td>
<td>MCS with scenario tree model [75] [76]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MCS with forward selection based scenario reduction [77]</td>
</tr>
<tr>
<td>Unit commitment</td>
<td>Interval based technique</td>
<td>Comparison with MCS based optimization techniques [72]</td>
</tr>
<tr>
<td></td>
<td>Dynamic programming</td>
<td>Partially observable Markov decision process [78]</td>
</tr>
<tr>
<td>Economic dispatch</td>
<td>MCS</td>
<td>MCS with classification based scenario reduction [80]</td>
</tr>
<tr>
<td></td>
<td>Moment estimation</td>
<td>First-order second-moment method [51]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Point estimation method [81]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Two-point estimation method [82]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Extended point estimation method with dependent input random variables [83]</td>
</tr>
<tr>
<td></td>
<td>Dynamic programming</td>
<td>Multi-timescale scheduling [52]</td>
</tr>
<tr>
<td>Power generation control</td>
<td>Stochastic control</td>
<td>Adaptive critic design [53] [84] [85]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Two-level stochastic control [87]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Kalman-Bucy filter [88]</td>
</tr>
<tr>
<td></td>
<td>Stochastic game</td>
<td>Zero-sum stochastic game [54]</td>
</tr>
<tr>
<td>Wide area measurement</td>
<td>State estimation</td>
<td>Discrete algebraic Riccati equation [55]</td>
</tr>
<tr>
<td></td>
<td>Stochastic control</td>
<td>Uncertainty propagation theory [94]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Adaptive critic design [99]</td>
</tr>
</tbody>
</table>

For scenario reduction, the original wind speed levels are reclassified into a reduced number of levels by re-defining the ranges of wind speed. The reduced scenarios are incorporated in MCS to solve the PDLF problem. Although the PDLF problem is formulated for power distribution systems, the scenario reduction technique is general and can be applied to bulk generation and transmission systems.

Mathematical analysis is another important approach to solving PPF problems. The convolution technique is typically used based on linearized power flow equations, such that the output random variables (e.g., line flows and bus voltages) can be represented by a linear combination of input random variables in terms of the power injection at each bus [49]. However, the computational complexity of the convolution technique is high when the system is large. Improvement over the convolution technique can be made based on Fast Fourier Transform [66], Von Mises method [67], and combined Cumulants and Gram-Charlier expansion theory [68]. The convolution technique can also be combined with MCS to solve PPF problems [69]. The PDF of a requested dependent generation (RDG) random variable can be obtained by the convolution technique since the variables involved are independent or linearly dependent. Then, the realizations of the RDG random variable are generated via MCS, based on which deterministic power flow equations are solved to obtain bus voltages, phase angles, and line flows.

In literature, interval based techniques are applied to solve PPF problems. Interval arithmetic can be used to provide strict bounds to the solution of PPF problems, where the interval linear power flow equations are solved by either explicit inverse of matrices or by iterative methods [50]. However, the solution accuracy is limited because of the linearization process. In order to address this problem, an Affine Arithmetic based method can be used to represent the uncertain variables in an affine form [70]. The method is further extended in [71] in a way that a mixed complementarity problem is developed to solve the deterministic power flow problem, considering reactive power limits and voltage recovery. Then, the intervals of power flow variables are obtained based on the Affine Arithmetic method. In comparison with MCS, the Affine Arithmetic method is faster and does not need any information regarding the probability distribution of random variables.
However, the estimate bounds of the power flow variables are relatively conservative.

B. Unit Commitment

The traditional unit commitment problem becomes significantly complicated when renewable energy sources and DSM tools are incorporated in the system, since new dimensions of randomness should be considered in the unit commitment decision making. Specifically, the power system needs to have a plan for alternative backup generation in a case that the day-ahead forecast of renewable power generation is not consistent with the actual realization, or the real-time power consumption deviates greatly from the load forecast in the presence of DSM tools. With a high penetration of renewable energy sources, the dependency of power systems on renewable energy sources can result in additional supply risks associated with the variability of renewable power generation. On the other hand, when DSM is widely adopted by electricity customers, the inaccuracy of price-sensitive load forecast may pose risks on real-time generation/load balance [72]. To address these challenges, the traditional unit commitment schemes need to be extended to incorporate the randomness in renewable power generation and customer behavior patterns in the presence of DSM.

A stochastic unit commitment scheme can be used to schedule various power resources such as DG units, conventional thermal generation units, and DSM tools [73]. The DSM tools, interruptible loads, DG units, and conventional thermal generators can be used to provide reserves to compensate for the randomness in DG output and load demand. The resources connected to the distribution system can participate in wholesale electricity market through aggregators based on communication technologies. In [74], an optimal strategy for the declaration of day-ahead generation availability is investigated for the Availability Based Tariff regime in India. The expected revenue of the generator is maximized by considering various stochastic parameters, such as the availability of the generation unit, unscheduled interchange, and load. The state transition of the generation unit is modeled as a Markov process, while the unscheduled interchange and load are modeled using discrete probability distributions and are related to grid frequency. An iterative approach based on MCS and SRS is performed to solve the problem.

To reduce the computational complexity of MCS, scenario reduction techniques are proposed in literature to solve stochastic unit commitment problems. A stochastic decomposition method can be applied to solve a large-scale unit commitment problem with future random disturbances to minimize the average generation cost [75]. The random disturbances (or outages) in the system are modeled as a scenario tree, which is constructed based on an either fully or partially random variant method. For the deterministic problem with respect to each scenario, an augmented Lagrangian technique can be applied, which provides satisfactory convergence properties. On the other hand, the traditional electricity market clearing schemes cannot fully integrate the stochastic nature of renewable power generation [76]. To address this problem, a short-term forward electricity market clearing problem is formulated in [76] based on a stochastic security framework, where a scenario tree is used to model the net load forecast error. To reduce the computational complexity, unlikely inter-period transitions are not included in the scenario tree. In comparison with the traditional deterministic approaches based on worst-case scenario wind and demand conditions, the stochastic approach puts higher weights to the conditions which are more likely to happen. As a result, the economic performance of the market is improved via taking advantage of the freely-available wind power by reducing reserve scheduling and classic hydrothermal generation unit commitment costs. The impact of intermittent wind power generation on short-term power system operation in terms of electricity market prices, social welfare, and system capacity is investigated in [77]. The MCS is used to generate scenarios of wind power generation, while a forward selection algorithm is applied to obtain a reduced set of scenarios. The reduced scenarios are incorporated into the unit commitment problem formulation under a locational marginal price (LMP) based electricity market settlement and an economic dispatch model.

In [72], a comparison between MCS based and interval based optimization techniques is presented in the context of stochastic security-constrained unit commitment (stochastic SCUC). The stochastic SCUC problem is formulated as a mixed-integer programming (MIP) problem and solved based on the two techniques. The uncertainty of wind power generation is considered. In the MCS, a large number of scenarios are generated to simulate wind speed uncertainty, which follows the Weibull distribution with an autocorrelation factor and diurnal pattern. For the interval based approach, the optimization problem for the base case without wind power generation uncertainty is updated with respect to the lower and upper bounds of wind power generation. It is shown that the MCS based optimization is not sensitive to the number of scenarios, at the cost of a high computational complexity. On the other hand, the interval based optimization has a lower computational complexity. However, how to determine the uncertainty interval is critical for obtaining the optimal solution.

Dynamic programming can be used to address the stochastic unit commitment problem [78]. The basic assumption in the problem formulation is that the renewable power generation can be characterized based on a hidden Markov model, while the stochastic power demand can be modeled by a Markov-modulated Poisson process. Structural results are derived by transforming the unit commitment problem as a partially observable Markov decision process.

C. Economic Dispatch and Optimal Power Flow

Due to the large-scale integration of renewable energy sources, traditional economic dispatch schemes, which rely on an accurate forecast of power generation and load demand, cannot be directly applied in the future smart grid. As discussed in Section III, the randomness in renewable power generation is characterized based on stochastic models. Without taking into account the randomness, traditional eco-
D. Power Generation Control

In real-time power system operation, the power generation and load demand should be balanced closely. However, when the penetration rate of renewable energy sources is high, significant power flow redistributions in power transmission may occur in a relatively short period of time. Specifically, a large increment or decrement of renewable power generation at one bus may cause a temporary generation-demand imbalance, followed by the generation adjustments at other buses and a redistribution of power flows across the electric power system. Because of the limited capability of automatic generation control in a traditional electric power system, transmission line overloading and bus over-/under- voltage may occur [53]. Stochastic control techniques should be developed to address this problem by taking into account the randomness in renewable power generation.

To provide a coordinating control solution to multiple grid-connected energy systems, dynamic stochastic optimal power flow (DSOPF) control strategies can be used [53] [84] [85]. The DSOPF controller is to replace the traditional automatic generation control and secondary voltage control, while providing nonlinear optimal control to the system-wide AC power flow. A DSOPF control algorithm is based on the conceptual framework of adaptive critic design [86] to incorporate prediction and optimization over power system stochastic disturbances. In this way, system analytical models are not required in the optimal controller design. To further investigate the potential of the DSOPF control algorithm for large power systems, a 70-bus test system with large wind plants is developed in [87] based on a two-level DSOPF control scheme. The lower-level area DSOPF controllers control their own area power networks, while the top-level global DSOPF controller coordinates the area controllers by adjusting the inter-area tie-line power flows. In this way, the control and computational load is distributed to multiple area DSOPF controllers, which can facilitate a practical application of the DSOPF controller in a large power network.

A non-stationary Markov chain can be used to model the time transient household load in the smart grid, where the time variant parameters of the Markov chain are estimated based on a maximum likelihood estimator [88]. Based on the load mode, a Kalman-Bucy filter based load tracking scheme can be applied for utility-maintained central power plant to ensure grid reliability, under time-varying load demand and renewable power generation.

The impact of communication systems on power generation control is discussed in [54]. Specifically, if wireless communication systems are used for wide area system monitoring and control, a jammer can send strong interference to jam the data transmission to cause denial-of-service attacks. Multiple channels can be used to avoid jamming interference [54]. The jamming and anti-jamming are modeled as a zero-sum stochastic game, while a quadratic function can be used as the payoff function to facilitate the linear quadratic Gaussian (LQG) control in the power system.
E. Wide Area Measurement

In traditional bulk generation and transmission systems, wide area measurement is mainly performed by remote terminal units (RTUs) of the SCADA system [89]. The most commonly used measurements include active/reactive power flow along transmission lines, active/reactive power injection of buses, and the voltage magnitude of buses. The measurement and control are performed once every a few seconds or even longer. In the future smart grid, with the advancement of clock synchronization via the global positioning system (GPS), phasor measurement units (PMUs) can achieve more accurate and timely (typically 30 samples per second) measurements in comparison with the traditional RTUs. Accordingly, two additional measurements in terms of voltage and current phasors (i.e., the phase angles and magnitudes) of buses and along transmission lines, respectively, can be obtained. The primary benefits of a PMU-enabled wide area monitoring system include [90]:

• Providing early warning of deteriorating system conditions based on which the operators can take corrective actions;
• Providing wide-area system visibility such that the cascading effect of disturbances can be limited;
• Improving transmission reliability and allowing for immediate post-disturbance analysis based on monitoring data.

Power system control schemes can be designed by leveraging wide area monitoring [91] [92]. However, due to the randomness of renewable power generation in the future smart grid, stochastic modeling and optimization techniques should be used for the placement and operation of PMUs.

In literature, there is a large body of research on optimal placement of PMUs, aiming at ensuring power system observability with the minimum number of PMUs and at determining the locations of the PMUs. The discrete algebraic Riccati equation can be used for a quantitative measure of the steady-state covariance of dynamic state estimation uncertainties [55]. Then, the PMU configuration with the least expected uncertainty is selected among many alternatives, where each alternative ensures the network observability with the minimum number of PMUs [93]. The uncertainty propagation theory can be used to assign appropriate weight factors for both conventional and PMU measurements in a hybrid state estimator [94]. This approach helps to obtain accurate state estimation with a small variance in the presence of random measurement errors, and can facilitate various energy management system applications. The greedy randomized adaptive search procedure can be combined with Monte Carlo simulation for PMU placement to record voltage sag magnitudes for fault location in distribution system [95]. The procedure minimizes the error in the distance between the true fault location and predicted fault location. The PMU placement problem can also be addressed based on an information-theoretic approach, which adopts Shannon entropy as a measure of uncertainties in the system states to qualitatively assess the information gain from PMU measurements [96]. In [97], an ant colony optimization technique is used to solve the PMU placement problem, and the convergence speed is improved by introducing stochastic perturbing progress. On the other hand, wide-area measurement can facilitate the operators in enhancing power system quality and control. An analysis of frequency quality, such as the total duration of under frequency and its correlation with time, is discussed in [98]. The analysis can be helpful for frequency control in the presence of electricity markets and increased use of renewable energy sources. Further, DSOOPF controllers based on wide-area measurements can incorporate adaptive critic design to provide nonlinear optimal control of power generator [99].

VI. DISTRIBUTION SYSTEMS AND MICROGRID

The distribution system is a part of electric power system that delivers the electric energy to consumers. The units below the step-down transformer station in Fig. 8, including distribution lines (27.6 and 13.8 kV) and pole-mounted transformers, illustrate the distribution system. Distribution networks usually have radial or looped feeder line configuration for power distribution as opposed to meshed configuration (i.e. redundant connections) of transmission network [100] [101]. Traditionally, a distribution system was not designed for the connection of power generating stations. On the other hand, the smart grid is anticipated to organically move from traditional centralized generation to a DG approach [11]. DG unit is an electric energy source connected directly to the distribution network [101]. Synchronous generator, asynchronous generator and power electronic converter interface are three basic generation technologies ranging from kilowatt (kW) to few Megawatt (MW) generation capacity in DG. The addition of DG units in the distribution system has impacts on the following aspects [101]:

• A DG unit increases the voltage variation in the distribution system when its operation is not coordinated with local loads;
• DG units can supply energy to local loads to help reduce T&D losses by decreasing the amount of energy drawn from the main grid (or utility grid);
• A sudden and large variation of DG unit outputs can cause voltage flickering, while the use of power electronic devices in the DG units can introduce harmonics in the distribution system, thereby degrading power quality;
• The protection system needs modification in overcurrent protection because of the changes in power flow caused by DG units;
• System reliability can be enhanced when DG units are used as back up energy sources.

In order to accommodate an integration of DG into the distribution system, the system approach is commonly known as ‘microgrid’.

Microgrid is an emerging system approach to integrate the DG units, storage, loads, and their control into a single sub-system as a controllable unit operating in either grid connected or islanded mode [102], thereby realizing a low-emission and energy efficient system. A typical architecture of a microgrid, as illustrated in Fig. 9, is assumed to have three feeders (A, B, and C) with radial feeder line configuration to transfer
High penetration of renewable energy resources with intermittent generation, random outages of components such as distribution lines, and random demands from consumers are the major factors for the introduction of stochastic phenomena in modeling a microgrid. The planning and operation of microgrids with consideration of such randomness are important and challenging. A summary of the stochastic information management schemes for distribution systems and microgrids is given in Table III.

### A. Microgrid Planning

Microgrid planning refers to making a decision on mixture of DER and its sizing under economical, environmental, and reliability considerations over a span of years [48]. Microgrids can exist in different forms with a unique objective. For example, a remotely located microgrid needs to operate in an isolated manner, an industrial microgrid needs to serve critical loads, and a utility microgrid needs to facilitate the main grid [106]. In order to fulfill its objective, each form of microgrids will have a unique combination of DER. A utility microgrid may survive with only renewable energy sources and batteries with support of utility supply. On the other hand, a remotely located microgrid cannot operate with only renewable energy sources and batteries. Without a continuous energy supplier, it would fail to maintain the required level of SOC in batteries. Hence, it needs to be served by dispatchable sources, such as microhydro and diesel generators.

An economical consideration refers to reduction in various costs such as fuel cost, electricity cost, cost of load curtailment, and incentives (negative cost) of supplying energy back to utility. Similarly, an environmental consideration refers to low emission of GHG (due to the use of fossil fuel in generation). The reliability is usually measured through various reliability indices such as system average interruption frequency index (SAIFI), system average interruption duration index (SAIDI), customer average interruption frequency index (CAIFI), expected energy not supplied (EENS), and loss of load expectation (LOLE), which also act as performance indices in microgrid planning.

A random output of renewable energy sources can be modeled as discussed in Section III and the randomness of system component (such as a DG unit and a section of microgrid) failure and repair can be modeled with independent and exponentially distributed time-to-failure and time-to-repair [107] or with double-Weibull distribution as discussed...
in [108]. After such modeling, the MCS with SRS can be used to generate scenarios with different combinations of DER over a span of one year [107] [108] [29], for instance. The scenario reduction technique can be used to reduce the number of scenarios for computational efficiency [109]. Based on the generated scenarios, different reliability indices, total cost and total emission can be computed thereby deciding the right combination and sizing (such as power rating and energy rating) of the DER. In addition, a decision on adding protection devices (such as circuit breakers) can be made according to the analysis of its impact on reliability [108].

In [48], software HOMER is used to solve the microgrid planning problem, to find out different combinations and sizes of DER units (such as diesel, solar, microhydro, and batteries) for the least cost microgrid, taking into account environmental impact. An evaluation methodology of reliability with consideration of pure stochastic generation and influence of supply-to-load correlation is demonstrated in [107]. An MIP problem is formulated in [109] for economically optimal energy storage system sizing. These studies demonstrate that the microgrid planning can aim at the sizing of particular DER, fulfilling economical and reliability objectives separately.

B. Microgrid Operation

Microgrid operation usually aims at reducing an overall cost by providing optimal schedule and coordination between DER and load. Similar to the microgrid planning, microgrid operation focuses on obtaining economical and environmental benefits, and achieving power quality and reliability. The power quality and reliability is commonly measured with parameters, such as SAIDI, SAIFI, and CAIDI, which capture the outage of components due to voltage fluctuations. Microgrid operation should maintain the supply and demand balance instantaneously and economically over a time horizon for power quality and reliability, under system component physical constraints (such as voltage limit, line flow limit). As the microgrid operation is a time process with uncertainties, the model predictive control with dynamic programming can be used to optimize the future behavior with uncertainties (handled by stochastic dynamic programming) [112]. Similarly, a stochastic optimization problem can be formulated to minimize the average cost of energy over all random scenarios. The random scenarios can be represented by distribution functions of random sources, randomly generated demands, and renewable generation based on the distribution of uncertainties (such as i.i.d. and Gaussian [111]), and random outages of components (modeled by two-state Markov-chain with failure and repair rate [110]). The scenarios are generated using MCS and scenario reduction techniques to bundle a large number of close scenarios (in terms of statistical metrics) into a small number of scenarios with corresponding probabilities.

A stochastic security-constrained unit commitment problem can be formulated based on MIP to reduce DG cost, including startup and shutdown costs, cost of energy supplied from the main grid, and opportunity cost due to microgrid load curtailment [110]. Integrated scheduling, and control of supply and demand by capturing its randomness [111], are examples of microgrid operation optimization, incorporating randomness.

C. Energy Storage Management

An addition of an energy storage device in the power system can 1) enhance system reliability by supporting the local load during outage of power generation, and 2) reduce energy cost by drawing energy from the grid when electricity price is low, and by feeding energy back to the grid and/or supplying the local demand when electricity price is high. Operation analysis of energy storage devices needs to capture the temporal dependency, in which the current state of energy storage devices depends upon previous states.

Analytical frameworks are presented in literature to evaluate the impact of energy storage on the distribution system. A probabilistic modeling framework [113] is developed for active storage devices, which not only can consume but also can supply electricity to a power system. The bounds on the probability of a load curtailment event are derived based on asymptotic probability theory via limited observable characteristics of the devices. A Karhunen-Loeve framework can be used to model the solar radiation intensity to characterize the PV unit output under a variety of conditions and at different geographical locations [114]. The capacity of energy storage devices is represented by a deterministic model, using an artificial neural network to estimate the capacity reduction over time. Given an appropriate stochastic load model, the MCS can be used to evaluate the probabilistic behavior of the system. Queuing models are developed for PV panels with energy storage [57]. The arrival process of queue corresponds to the non-stationary solar irradiation, while the departure process of the queue represents the energy selling to the grid or used by local loads. The $G(t)/G(t)/\infty$ queuing analysis is conducted for performance evaluation.

To achieve optimal operation of energy storage devices, stochastic control schemes should be developed. Approximate dynamic programming (ADP) driven adaptive stochastic control (ASC) for the smart grid is studied in [115]. A specific application of economic dispatch is investigated, where the DG unit is linked to an energy storage device. Since a multidimensional control variable is involved in the ASC problem formulation, an ADP algorithm is developed to solve the problem, which achieves performance close to the optimal at a low computational complexity. The energy storage device operation problem is further studied by considering the variations in wind, load demand, and electricity prices. It is shown that the ADP scheme is efficient in solving high dimensional energy allocation problems, provided that the basis functions of approximate policy iterations are properly selected.

VII. DEMAND SIDE MANAGEMENT

It was 1980s when the ERPI introduced the DSM publicly, as an energy crisis started to emerge [116] [117]. The DSM provides a basis of adjusting the consumption level to provide instantaneous balance of generation and demand. The DSM, also known as energy demand management, represents a large group of schemes (such as load management, energy efficiency, energy saving, and smart pricing) adopted by utilities that motivate the consumers to change their energy usage patterns to achieve better economy and load factor
TABLE IV: Methodology for stochastic information management in DSM.

<table>
<thead>
<tr>
<th>Applications</th>
<th>Theory or Technique</th>
<th>Variations and/or Modifications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Applications</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Demand response</td>
<td>MCS</td>
<td>MCS with SRS [121]</td>
</tr>
<tr>
<td></td>
<td>Interval based technique</td>
<td>Robust optimization [121]</td>
</tr>
<tr>
<td></td>
<td>Dynamic programming</td>
<td>Markov decision process [124] [125]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Stochastic dynamic programming [126]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Approximate dynamic programming [127]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Lyapunov optimization [120]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Risk of loss minimization [129]</td>
</tr>
<tr>
<td></td>
<td>Stochastic game</td>
<td>Cooperative game [130]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(N)-person nonzero-sum stochastic differential game [131]</td>
</tr>
<tr>
<td>Load prediction</td>
<td>Stochastic control</td>
<td>Kalman filter [122] [123]</td>
</tr>
<tr>
<td>Cellular network and IDC operation</td>
<td>Queueing theory</td>
<td>(M/M/c/c) queue [134]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(M/M/c) queue [135]</td>
</tr>
<tr>
<td></td>
<td>Dynamic programming</td>
<td>Lyapunov optimization [136]</td>
</tr>
</tbody>
</table>

(defined as the average load divided by the peak load). Energy efficiency and energy saving schemes address long term issues of environmental impact, whereas load management schemes, commonly referred to as demand response, addresses short term issues of supply and demand balance [117]. Utilities view DSM as load shaping objectives which comprises of six fundamental categories, namely for peak clipping, valley filling, load shifting, strategic conservation, strategic load growth, and flexible load shape, as shown in Fig. 10 [116]. Among them, peak clipping, which reduces system peak loads, is achieved by direct load control. Valley filling, which builds up off-peak load, can be achieved by electrification such as EV charging during night time. Load shifting is to shift loads from on-peak to off-peak hours. Strategic conservation and strategic load growth are general decrement and increment in sales, respectively. Flexible load, related to reliability, is the willingness of customers over variations in quality of services by possessing interruptible or curtailable loads, but with certain incentives. The Federal Energy Regulatory Commission (FERC) defines demand response as: “Changes in electric usage by end-use customers from their normal consumption patterns in response to changes in the price of electricity over time, or to incentive payments designed to induce lower electricity use at times of high wholesale market prices or when system reliability is jeopardized.” Beside the load management program or demand response, the authors of [118] discuss the smart pricing that motivates customers to consume wisely and efficiently, which is beneficial for both customers and utilities economically and environmentally. In this section, we focus our discussion on demand response which is the major component of DSM in the future smart grid. The demand response components such as dynamic demand-sensitive real-time pricing in the electricity market, uncertain human behavior patterns [119], intermittent generation sources [120], and stochastic noise in energy metering devices or sensors are the factors introducing stochastic nature into the system. Accordingly, stochastic information management schemes need to be developed, and some related works in literature are summarized in Table IV.

Based on the preferences in residential appliances, the operation tasks in demand response can be categorized into deferrable/non-deferrable and interruptible/non-interruptible ones [121], while Gaussian distributions can be used to model the dynamics in real-time electricity prices. To reduce the computational complexity of MCS for real-time demand response, an interval based technique - robust optimization can be used [121] to determine price uncertainty intervals for simulating the real-time price uncertainty. The worst case scenario with respect to the uncertainty intervals is considered, given that the electricity prices can be uncertain in a given number of time slots. Although the robust optimization can achieve lower computational complexity, the electricity bill of the residential customer is higher in comparison with that of the MCS scheme with a higher computational complexity.

Kalman filtering is widely used for load prediction in the process of demand response. An efficient interaction infrastructure between utility and distributed customers is proposed in [122], where each interaction cycle includes demand response and stochastic tracking control of conventional generation facilities. The sum of load demand signals serves as the reference signal which needs to be tracked by conventional generation. A Kalman filter based prediction scheme is developed to compensate for the delays of different customer load demand signals. A uncertainty-aware minority-game based energy management system can be used for energy resource allocation in smart buildings with solar power generation and main grid connection [123]. With multiple agents deployed in the building and each agent corresponding to a smart con-
controller, the stochastic noise from energy meters/sensors can be reduced via Kalman filtering. Based on supervised learning of uncertain energy profiles, the agents play a modified minority game to allocate the limited solar energy resource.

Dynamic programming can be used to optimize the demand response processes. Specifically, a Markov decision process can be used to model individual devices participating in demand response markets [124]. Four types of devices can be defined, i.e., optional loads that can be curtailed (e.g. light dimming), deferrable loads that can be delayed (e.g. dishwashers), controllable loads with inertia (e.g. thermostatically-controlled loads), and storage devices that can alternate between charging and generating. The optimal price-taking control strategy can be derived based on the Markov decision process model. In the smart grid, the capital expenditure of using communication networks for demand response may be nonnegligible [125]. In such a case, the overall operation cost of demand response should be jointly optimized with the communication cost.

Based on the LMP electricity market model, the problem of determining when to inquire the power price can be formulated as a Markov decision process [125]. Dynamic programming is performed to obtain the optimal strategy while a myopic approach can be used to achieve low computational complexity. In [126], a stochastic dynamic programming problem is formulated for energy usage in a micro-scale smart grid system with a goal of optimizing a finite horizon cost function which reflects both the cost of electricity and comfort/lifestyle. The model is extended in [127], assuming that the key models and forecasts are unknown and implicitly learned via a softmax algorithm with neighborhood updating. The algorithm implements ADP to reduce the dependencies on models and forecasting. In order to reduce the computational complexity in dynamic programming with respect to a large state space in demand response (typically referred to as the curse of dimensionality [128]), a Lyapunov optimization technique can be used to develop simple energy allocation algorithms [120], and an upper bound of the objective function in optimal power scheduling can be used to minimize the risk of loss for each electricity customer [129].

Stochastic games can be formulated to model the competition among electricity customers. A dynamic pricing scheme is typically used as an incentive for customers to achieve an aggregated load profile suitable for utilities. A cooperative game approach can be applied to reduce total cost and peak-to-average ratio of the system when customers can share all their load profiles [130]. On the other hand, when customers have access only to the total load of the system, distributed stochastic strategies need to be developed to exploit this information for overall load profile improvement. In [131], a dynamic game is used to model the distributed demand side management. A two-layer optimization framework is established, where the appliances of different players (e.g., households) are scheduled for energy consumption at the lower level, while the interaction among different players in their demand responses is captured through the market price. An $N$-person nonzero-sum stochastic differential game with a feedback Nash equilibrium can be used for the two-layer optimization framework.

Due to the ever-increasing demand of mobile Internet services and cloud computing, the energy bills associated with the massive power consumption of cellular networks and Internet data centers (IDCs) have laid a heavy burden upon the operators. As a result, how to reduce energy consumption for cellular networks and IDCs has attracted considerable attention recently [132] [133]. To achieve energy saving, the base stations of a cellular network and the servers in an IDC can be strategically switched off. However, different from traditional electric loads, when a cellular network or IDC is powered by the smart grid, quality of service (QoS) requirements of users should be considered in addition to the real-time electricity price. Queueing analysis is widely used in literature to model the QoS provisioning to users. An $M/M/c/c$ queueing model is developed in [134], and the Erlang-B formula is used to calculate the service blocking probability. To ensure acceptable QoS in the cells whose base stations have been switched off, coordinated multipoint (CoMP) technology is used. A Stackelberg game with two levels (i.e., a cellular network level and a smart grid level) is formulated for the active base stations to decide on which retailers to procure electricity from and how much electricity to procure. On the other hand, to quantify the service level agreement (SLA) of each server in an IDC, the steady-state results of an $M/M/c/c$ queue can be employed [135]. A bi-level programming problem is investigated to minimize the operation risk of IDCs against the uncertainties in dynamic workload and time-varying electricity prices. The operation costs of an IDC can be further reduced by utilizing the uninterrupted power supply (UPS) units as energy storage devices [136]. The Lyapunov optimization technique is used to develop an online control algorithm to minimize the time average cost of an IDC. The algorithm has a lower computational complexity in comparison with the dynamic programming approach and does not require any prior knowledge of the statistics of workload or electricity price, which is suitable for real-time IDC operation in the presence of workload and pricing uncertainties.

VIII. ELECTRIC VEHICLE INTEGRATION

With a fast-growing EV penetration rate, the charging demand is expected to constitute a significant portion of the total power demand in the future smart grid. On the other hand, the battery storage of EVs can be better utilized to potentially improve the efficiency and reliability of electricity delivery via V2G systems. A key feature of the V2G systems is a bidirectional energy delivery mechanism which enables the EV to either draw energy from or feed energy back to the grid. Different from traditional stationary energy storage systems, the main issue in efficiently managing EV charging demand and utilizing EV batteries for energy storage is the highly dynamic vehicle mobility. Although domestic EV charging demands can be well estimated based on the commute patterns of EV owners which are relatively stable [137], the charging station planning and operation are a relatively challenging issue due to the uncertainty in EV arrivals to a charging station and the randomness in EV energy demands. In order to optimize EV charging infrastructure, the charging demand
should be forecasted based on EV mobility statistics. On the other hand, the energy storage of EVs can be used by the electric power system via V2G systems. In the presence of EV mobility, stochastic models need to be established to characterize the V2G system capacity. Table V provides a summary of the stochastic information management schemes in literature for EV integration.

### A. Charging Station Planning

Various issues need to be considered when planning an EV charging station, such as location and charging infrastructure selection, aggregated charging demand estimation, metering, and safety related issues [138]. An accurate estimation of the aggregated charging demand is critical for the utility to evaluate the transmission capability of the existing system for the electricity delivery to the charging station. If available transmission capacity is inadequate, an upgrade of the existing system is needed.

Queueing theory can be used to analyze the aggregated charging demand. Each charging station is modeled as a queue, while the vehicles are modeled as customers in the queue. According to the component-level models in Section III, the arrivals of EVs at a specific charging station follow a Poisson process. Given a certain capacity of the charging infrastructure (e.g., a maximum charge power of 1.44 kW, 3.3 kW, and 150 kW for levels 1, level 2, and level 3 charging infrastructures according to [139]), the charging time of each EV is directly determined by the energy demand, which can be modeled by an exponential distribution. According to the stationary distribution derived based on the queueing analysis, key performance metrics of the charging station can be obtained, such as the probability distribution of the aggregated vehicle charging demand, vehicle waiting time, and charging blocking probability.

Multi-server queues are widely used in existing research. An EV charging demand model is presented in [56] for a rapid charging station at highway exit. Different from most previous studies which assume a fixed charging location and fixed charging time for each EV, the model captures the spatial and temporal variations of EV charging demands. The arrival rate of discharged vehicles is estimated based on fluid dynamic model. Then, an $M/M/c$ queueing analysis is performed for a Poisson arrival process, exponential charging times, and $c$ identical chargers at a charging station, such that at most $c$ vehicles can be charged simultaneously. Based on the stationary distribution of the queue, the average charging demand and expected number of busy chargers can be determined. Similar queueing analysis is used in [140], where the EV charging station is considered as a specific case of a microgrid with controllable loads (electric vehicles), storage devices, and grid interconnection. Investment, operational costs, physical constraints, and different electricity pricing strategies can be investigated in the planning problem. The $M/M/c$ queueing analysis can also be applied to charging station planning on urban trunk roads [141]. The number of chargers within a charging station is optimized with respect to a weighted sum of charging service cost and customer waiting time. For a limited waiting space, an $M/M/c/N$ queueing analysis can be performed, where $c$ is number of chargers and $N$ represents the sum of the number of chargers and waiting locations [142]. Then, the performance metrics of the charging station can be evaluated, such as the utilization of chargers (i.e., the ratio of the average number of charging and waiting EVs to the total number of chargers and waiting locations), the time of waiting, and customer charging blocking probability can be evaluated, based on which the number of chargers can be optimized. Similar queueing models with truncation is used in [143] to estimate the EV charging demand for different charging technologies, i.e., fast charging and battery switching.

Local energy storage devices can be used to improve the charging station service, where EV charging demand can be satisfied by either an electric power grid or a local energy storage device [144]. When the EV charging demand is high and cannot be satisfied by the electric power grid due to its transmission capacity limit, the energy in the local energy storage device can be utilized to support additional charging demand. A two-dimensional continuous-time Markov chain model can be used for the performance analysis of such a

### TABLE V: Methodology for stochastic information management in EV integration.

<table>
<thead>
<tr>
<th>Applications</th>
<th>Theory or Technique</th>
<th>Variations and/or Modifications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Charging station planning</td>
<td>Queueing theory</td>
<td>$M/M/c$ queue [56] [140] [141]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$M/M/c/N$ queue [142] [143]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Two-dimensional Markov chain [144]</td>
</tr>
<tr>
<td>Load flow analysis</td>
<td>Queueing theory</td>
<td>$M/M/\infty$ queue [145]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$M/M/c$ queue [32] [146]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$M/M/c/k/N_{max}$ queue [32]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$M/c/GI/\infty$ queue [147]</td>
</tr>
<tr>
<td>Charging demand coordination</td>
<td>Queueing theory</td>
<td>$M/M/\infty$ queue [148]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$M/M/c$ queue [149]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GI/D/1 queue [150]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Multi-queue system [151]</td>
</tr>
<tr>
<td>V2G system Operation</td>
<td>Stochastic inventory theory</td>
<td>$M/M/c$ queue [154]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Modified backward iteration [34]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Policy adjustment [158]</td>
</tr>
</tbody>
</table>
charging station, where the two dimensions correspond to the number of vehicles that can be charged simultaneously by the station and the energy stored in the energy storage unit, respectively. The quality of service (e.g., charging blocking probability of EVs) can be evaluated, which provides useful measures for the charging station and local energy storage device sizing.

B. Power Flow Analysis with EV Charging Demand

Given sufficient transmission capacity of the electric power system, the charging demand of EVs can be satisfied. However, it is important to evaluate the impact of EV charging on the electric power system. According to the power flow analysis in Section V, EV charging demand can affect the net active power injection of each bus and cause voltage deviations in the system, which may result in unstable power system operation. To characterize the impact of EV charging on the electric power system, queueing theory is widely used.

If the number of chargers at a charging station is sufficiently large, an $M/M/\infty$ queueing analysis can be applied to estimate EV charging demand [145]. The queueing model and wind power generation model can be incorporated into a probabilistic constrained load flow study. In [146], the electric vehicle charging demand is modeled as a $PQ$ bus while taking into account the capacity limit of the load bus. The randomness in the charging demand is characterized in an $M/M/c$ queueing analysis, where $c$ is related to the charging capacity of the load bus. The charging demand of the $PQ$ bus is given in a closed-form representation of charging time. The MCS can be used for load flow analysis, where the number of charging EVs is randomly generated based on the stationary distribution of the queue. It is observed in [32] that the charging station and residential community should be modeled in different ways, i.e., by an $M/M/c$ queue and an $M/M/c/k/N_{\text{max}}$ queue, respectively. In the latter scenario, since the charging slots are generally privately owned or shared only by the residents, the maximum number of customers being served or waiting in the queue is limited to $k$, while the maximum number of possible customers to be served is $N_{\text{max}}$. The stationary distributions of the two queues are used to facilitate a PPF analysis.

Despite the convenience of queueing analysis, the assumption of a constant vehicle arrival rate may lead to inaccurate studies since vehicle arrival is semi-periodic in nature [147]. For instance, the vehicles arrive more frequently during the evening and early night hours on each day. An $M_{t}/GI/\infty$ queue can be used to address this problem, where vehicles arrive with a time-dependent rate and are served according to a general service time [147].

C. EV Charging Demand Coordination

Queueing theory can be applied to facilitate the EV charging demand coordination. A straightforward way is to set a maximum number of available chargers ($c$) such that the original $M/M/\infty$ queue becomes an $M/M/c$ queue. Taking advantage of wireless communications, the charging demand of an EV which is physically connected to a charger can be deferred to reduce the peak demand of the grid, if the available charging sockets are fully occupied [149]. A $GI/D/1$ queueing model is used in [150], where a general arrival process (specifically, a Gaussian process) is used to capture different power consumption profiles of different EVs. A controllable deterministic service process is used to model the threshold of aggregated charging power specified by the utility. Based on the queueing analysis, the smart grid can predict the occurrence of an outage at the start of an epoch, based on which the load shedding decisions are made to defer the charging process of some EVs by limiting the total charging capacity.

Another way of controlling the charging demand is to control the arrival rate of EVs at a charging station. Without EV charging demand coordination, an $M/M/\infty$ queue can be used to estimate the probability of a distribution system overloading [148]. The model is further extended to a variable-rate version such that the arrival rate of EVs at the charging station is a (controllable) function of the number of charging EVs. Control algorithms are developed to adjust the arrival rate of EVs such that the utilization of distribution system capacity can be maximized while maintaining a negligible probability of overloading. The reliability of the control algorithms on the communication network is minimal since only rate-limited, one-way broadcast communication is needed for the notification of the adjustment of the arrival rate of EVs. A multi-queue system is used in [151] to model a group of charging stations, where each charging station is modeled as an $M/M/1$ queue. Based on the queueing analysis, the arrival rate of each queue is optimized. Price control methods can be developed to find the optimal arrival rate.

D. Vehicle-to-Grid (V2G) System Operation

Two kinds of services can be provided by V2G systems [152] [153]. The ancillary services are used to ensure short-term supply-demand balances in the electric power system. Since the imbalances are temporary and small-scale in nature, the ancillary services may not necessarily involve energy delivery. On the other hand, the load shaving services use the energy stored in vehicle batteries to compensate for the peak load of the power system. From the vehicle owners point of view, the energy cost can be relatively reduced by drawing cheap energy from the grid, and vice versa. Since a significant amount of energy transactions may deplete EV batteries, providing efficient load shaving services is a more challenging issue for the stochastic information management in a V2G system.

Different from the EV charging, both power demand and supply should be estimated for V2G load shaving services. The $M/M/c$ queueing model developed in [146] can be extended to model both EV demand and supply in a V2G system [154]. The discharging time of an EV for V2G service provisioning is modeled as an exponentially distributed random variable, provided that a certain amount of energy is reserved in the EV battery for a commute purpose. The amount is considered to be pre-determined and insensitive to electricity price. The amount of energy reservation can be facilitated by analyzing the average commute energy demand of an EV [155] [156].
However, this kind of estimates can lead to suboptimal solutions for load shaving services. A practical example is given in [157] where a vehicle driver may undertake an unexpected journey and the commute energy demand depends on the actual traffic condition. As a result, a significantly large amount of energy should be reserved to address the uncertainty [157]. In [34], the traditional energy store-and-deliver mechanism for stationary battery management is extended to an energy store-carry-and-deliver mechanism for EV battery management. The energy cost minimization problem under time-of-use electricity pricing is mathematically formulated. Based on the stochastic inventory theory, a state-dependent double-threshold policy is proved to be optimal. A modified backward iteration algorithm based on estimated statistics of plug-in hybrid electric vehicle (PHEV) mobility and energy demand can be used to facilitate practical applications [34]. Stochastic inventory theory can also be used to solve a multi-vehicle aggregator design problem by considering the power system constraints. A policy adjustment scheme is developed in [158] to adjust the two thresholds of the optimal policy adopted by each PHEV, such that the aggregated recharging and discharging power constraints of the power system can be satisfied, while minimizing the incremental cost (or revenue loss) of PHEV owners.

For V2G ancillary services, the energy constraint for V2G system based frequency regulation is related to the SOC of the EV battery. The energy deviation caused by a single regulation signal is obtained in [159], based on which a probabilistic distribution of successful regulation is estimated. Random walk theory is employed for stochastic analysis of the distribution. The distribution is averaged to form a weight function, so that it can be associated with a cost function to rate the current value of regulation.

IX. Conclusions and Discussions

In this paper, we have presented an overview of the state of the art on stochastic information management for the smart grid. Component-level stochastic models are investigated to characterize the sources of randomness in the smart grid. The models are further incorporated in the system-level stochastic information management schemes to facilitate the planning and operation of bulk generation and transmission systems, distribution systems, and customer appliances, and to facilitate the integration of renewable energy sources, energy storage devices, DSM tools, and EVs.

Most of the existing stochastic information management schemes evolve from those for traditional power system planning and operation. As a result, they do not provide effective or efficient solutions to handle larger system dynamics in the future smart grid. There are many open research issues:

- **Optimal energy storage device operation** – To reduce the computational complexity in energy storage device management, a suboptimal ADP technique can be used. However, the policy and value function approximation mechanisms incorporated in the ADP technique require basic knowledge about the structure of the optimal control policy. Based on our preliminary studies [34] [158], the stochastic inventory theory is a powerful tool to characterize the optimal operation policy of a single battery, based on an analogy between the SOC of a battery and the stock level of an inventory. The optimal operation policy of a battery follows a double-threshold policy, with the thresholds corresponding to battery charging and discharging, respectively. However, how to apply the stochastic inventory theory to establish optimal operation over other energy storage devices with start-up cost needs further investigation. One possible approach is to model the cost as a fixed ordering cost in the stochastic inventory model. The corresponding optimal threshold policy is of an \((s, S)\) type [160], where the energy-level thresholds correspond to decisions on whether or not to start the energy storage device respectively. Further, it is critical to extend the ADP policies to the control of multiple energy storage devices which may coexist in an electric power system. New approximation techniques need to be developed based on the fact that the value function of an inventory model based on the \((s, S)\) policy is a \(K\)-convexity function. Accordingly, new ADP policies should be developed based on the approximation for computational complexity reduction of energy storage device operation.

- **Planning of interactive charging stations** – Queueing theory can be used for the planning of a single charging station. However, for a distribution system with more than one charging station, the single queue models are no longer applicable. Interactions among different charging stations depend on vehicle owners’ response to charging prices and can affect the power flow in the distribution system. Queueing network models should be developed to address this problem, where each charging station is represented by one queue in the queueing network. The interaction among different charging stations can be modeled by the routing probabilities among different queues, which depend on EV user responses to charging prices. According to studies on vehicular communication networks, the BCMP queueing network can provide a good estimation of vehicle traffic [161]. The product form solution of the stationary distribution of the BCMP queueing network can help reduce the computational complexity. Based on the queueing network model, the probability distribution of the number of EVs in each charging station can be obtained. Consequently, the probability distribution of EV charging demand at each charging station can be obtained, which can be incorporated in PPF to facilitate the planning of charging stations.

- **EV charging station selection** – Vehicular communication networks can assist EV charging station selection. EVs can report their locations and path selection decisions via vehicle-to-infrastructure (V2I) communications with the roadside units (RSUs) which are connected to a vehicle traffic sever. The optimal path of each EV is calculated by the server and transmitted to the EV via RSUs. When multiple charging stations are deployed within the same distribution system, the maximum loading of the charging stations is correlated with each other based on power
flow analysis. The EV charging station selection should depend on not only the traffic statistics but also the power flow of the electric power system. Our recent research has shown that the optimal EV charging station selection problem can be simplified by establishing a linear relation between the loading and voltage of each distribution system bus [162]. Then, the Lagrange duality optimization techniques can be used to solve the associated optimization problem:

- **Information security and privacy support** – Most existing research assumes that the stochastic information provided via smart grid communications is authentic. However, if some malicious nodes in the network inject bad data [163] [164], the power system operation via stochastic information management can be interrupted, as the power generation/demand cannot longer be balanced and the frequency of the power system deviates from the nominal frequency. Several existing works have analyzed and modeled bad data injection attacks and presented corresponding defensive strategies. For example, to address the malicious meter inspection (MMI) problem, a tree-based inspection algorithm is exploited and analyzed [165], while the attack strategies and countermeasures are introduced for the bad data attacks on smart grid state estimation [166] [167]. However, there are many new types of bad data injection attacks that have not been tackled. To fill this gap, cognitive bad data detection techniques based on machine learning should be developed to not only identify but also address the new types of attacks. On the other hand, customers’ private information (in terms of energy consumption and EV mobility statistics) is needed for stochastic information management. Once unauthorized entities access the private data, customers’ privacy is violated [168]–[171]. Therefore, customers should be able to grant access to their data so that only authorized entities can decrypt and read the specific data. To this end, the development of multi-authority and ciphertext-policy attribute based encryption (CP-ABE) techniques to enable access control in the future smart grid [172] can be an interesting future research direction.

- **Joint system planning** – Consider the microgrid as an example, which is evolved from traditional power distribution systems and thus is cost-sensitive in nature. For this reason, microgrid planning should take into account not only the expenses on power system assets, but also the cost of establishing a communication network. In order to capture the impact of communication network in microgrid, the interaction between communication system and electric power system needs to be studied based on a stochastic approach. For instance, a WiFi or ZigBee network can be used for low-cost installation and operation on a license-free radio frequency band. However, the performance of power system operation may degrade because of a system status report delay, which is a random variable depending on the medium access control and data traffic load [173] [174]. On the other hand, a cellular network with dedicated radio resources for a low communication delay can be used to improve the efficiency and reliability of power system operation. However, the operation cost by using a cellular network on a licensed radio frequency band is non-negligible [175]. Our recent research has shown a tradeoff between the operation cost of a cellular network and power generation cost in a microgrid based on economic dispatch [176]. A joint planning of the communication system and electric power system over a long time frame is needed to facilitate the deployment of future smart grid.

In summary, extensive R&D efforts are required to develop stochastic information management schemes to facilitate smart grid planning and operation to achieve efficiency, reliability, economics, and sustainability in electricity production and distribution. The research is interdisciplinary in nature and calls for a close collaboration among the researchers from both information/communication system discipline and power/energy system discipline.

**ACKNOWLEDGMENT**

The authors would like to thank Professor Kankan Bhattacharya from the Electricity Market Simulation and Optimization Lab at the University of Waterloo and Wei Wang for valuable discussions.

**REFERENCES**

[10] IEEE Std 2030-2011, “Guide for smart grid interoperability of energy technology and information technology operation with the electric power system (EPS), and end-use applications and loads,” 2011.
Hao Liang (S’09-M’14) received his Ph.D. degree in Electrical and Computer Engineering from the University of Waterloo, Canada, in 2013. He is currently a postdoctoral research fellow at the Broadband Communications Research (BBCR) Lab and Electricity Market Simulation and Optimization Lab (EMSOL) at the University of Waterloo. His research interests are in the areas of smart grid, wireless communications, and wireless networking. He is a recipient of the Best Student Paper Award from IEEE 72nd Vehicular Technology Conference (VTC Fall-2010), Ottawa, ON, Canada.

Dr. Liang was the System Administrator of IEEE Transactions on Vehicular Technology (2009-2013). He served as a Technical Program Committee (TPC) Member for major international conferences in both information/communication system discipline and power/energy system discipline, including IEEE International Conference on Communications (ICC), IEEE Global Communications Conference (Globecom), IEEE VTC, IEEE Innovative Smart Grid Technologies Conference (ISGT), and IEEE International Conference on Smart Grid Communications (SmartGridComm).

Amit Kumar Tamang received the M.E. (2012) degree in Information and Communication Technologies from Asian Institute of Technology, Thailand and the B.E. (2007) degree in Electronics and Communications from Tribhuvan University, Nepal. He was a lecturer in Tribhuvan University and Kathmandu University in Nepal during 2008-2009 and 2012-2013 respectively. Since 2013, he has been working toward a M.A.Sc. degree at the Department of Electrical and Computer Engineering, University of Waterloo, Canada. Mr. Tamang’s current research interest includes electric vehicle integration into the microgrid in Smart Grid.

Weihua Zhuang (M’93-SM’01-F’08) has been with the Department of Electrical and Computer Engineering, University of Waterloo, Canada, since 1993, where she is a Professor and a Tier I Canada Research Chair in Wireless Communication Networks. Her current research focuses on resource allocation and QoS provisioning in wireless networks, and smart grid. She is a co-recipient of the best paper awards from IEEE Globecom 2013 and 2012, the IEEE International Conference on Communications (ICC) 2007 and 2012, IEEE Multimedia Communications Technical Committee in 2011, IEEE Vehicular Technology Conference (VTC) Fall 2010, IEEE Wireless Communications and Networking Conference (WCNC) 2007 and 2010, and the International Conference on Heterogeneous Networking for Quality, Reliability, Security and Robustness (QShine) 2007 and 2008. She received the Outstanding Performance Award 4 times since 2005 from the University of Waterloo, and the Premiers Research Excellence Award in 2001 from the Ontario Government. Dr. Zhuang was the Editor-in-Chief of IEEE Transactions on Vehicular Technology (2007-2013), the Technical Program Symposia Chair of the IEEE Globecom 2011, and an IEEE Communications Society Distinguished Lecturer (2008-2011). She is a Fellow of the IEEE, a Fellow of the Canadian Academy of Engineering (CAE), a Fellow of the Engineering Institute of Canada (EIC), and an elected member in the Board of Governors of the IEEE Vehicular Technology Society (2012-2014).

Xuemín (Sherman) Shen (IEEE M’97-SM’02-F’09) received the B.Sc.(1982) degree from Dalian Maritime University (China) and the M.Sc. (1987) and Ph.D. degrees (1990) from Rutgers University, New Jersey (USA), all in electrical engineering. He is a Professor and University Research Chair, Department of Electrical and Computer Engineering, University of Waterloo, Canada. He was the Associate Chair for Graduate Studies from 2004 to 2008. Dr. Shen’s research focuses on resource management in interconnected wireless/wired networks, wireless network security, social networks, smart grid, and vehicular ad hoc and sensor networks. He is a co-author/editor of ten books, and has published more than 600 papers and book chapters in wireless communications and networks, control and filtering. Dr. Shen served as the Technical Program Committee Chair/Co-Chair for IEEE InfoCom’14, IEEE VTC’10 Fall, the Symposium Chair for IEEE ICC’10, the Tutorial Chair for IEEE VTC’11 Spring and IEEE ICC’08, the Technical Program Committee Chair for IEEE Globecom’07, the General Co-Chair for Chincocom’07 and QShine’06, the Chair for IEEE Communications Society Technical Committee on Wireless Communications, and P2P Communications and Networking. He also serves/served as the Editor-in-Chief for IEEE Network, Peer-to-Peer Networking and Application, and IET Communications; a Founding Area Editor for IEEE Transactions on Wireless Communications; an Associate Editor for IEEE Transactions on Vehicular Technology, Computer Networks, and ACM/Wireless Networks, etc.; and the Guest Editor for IEEE JSAC, IEEE Wireless Communications, IEEE Communications Magazine, and ACM Mobile Networks and Applications, etc. Dr. Shen received the Excellent Graduate Supervision Award in 2006, and the Outstanding Performance Award in 2004, 2007 and 2010 from the University of Waterloo, the Premier’s Research Excellence Award (PREA) in 2003 from the Province of Ontario, Canada, and the Distinguished Performance Award in 2002 and 2007 from the Faculty of Engineering, University of Waterloo. Dr. Shen is a registered Professional Engineer of Ontario, Canada, an IEEE Fellow, an Engineering Institute of Canada Fellow, a Canadian Academy of Engineering Fellow, and a Distinguished Lecturer of IEEE Vehicular Technology Society and Communications Society.