A Bayesian Inference Approach to Unveil Supply Curves in Electricity Markets

Lesia Mitridati, Pierre Pinson, *Senior Member, IEEE*

***Abstract*—With increased competition in wholesale electricity markets, the need for new decision making tools for strategic producers has arisen. Optimal bidding strategies have traditionally been modeled as stochastic profit maximization problems. However, for producers with non-negligible market power, modeling the interactions with rival participants in a market is fundamental. This can be achieved through equilibrium and hierarchical optimization models. The efficiency of these methods rely on the strategic producer's ability to model rival participants' behavior and supply curve. But a substantial gap remains in the literature on modeling this uncertainty. In this study we propose a Bayesian inference approach to reveal the aggregate supply curve of rival participants in electricity markets. The algorithm proposed relies on Markov Chain Monte Carlo and Sequential Monte Carlo methods. The major appeal of this approach is that it provides a complete model of the uncertainty of the supply curve. We show on a realistic case study that we are able to approximate accurately the aggregate supply curve. Finally we show how this piece of information can be used by a price-maker producer in order to devise an optimal bidding strategy.**

***Index Terms*—Bayesian inference, Sequential Monte Carlo, Markov Chain Monte Carlo, strategic bidding**

NOMENCLATURE

Indexes

d	Index for days $(d = 1, ..., D)$
h	Index for hours $(h = 1, ..., 24)$
g	Index for generation blocks $(g = 1, ..., G)$
w	Index for wind producers $(w = 1, ..., W)$
m	Index for Markov Chain Monte Carlo iterations $(m = 1, ..., M)$
i	Index for Sequential Monte Carlo particles $(i = 1, ..., N)$

Market Clearing Variables and Parameters

$\lambda_{d,h}^{spot}$	Spot price at hour h of day d
$P_{g,d,h}$	Power output of generator g at hour h of day d
$P_{w,d,h}$	Power output of wind producer w at hour h of day d
$\overline{P}_{g,d,h}$	Maximum prower output for generator g at hour h of day d
$\overline{P}_{w,d,h}$	Wind production available for producer w at hour h of day d
$P_{g,d}^{0}$	Initial production of generator g at beginning of day d
$L_{d,h}$	Electricity load at hour h of day d
$\alpha_{g,d,h}$	Price offer parameter of generator g at hour h of day d
$\beta_{g,d,h}$	Price offer parameter of generator g at hour h of day d
$\overline{R}_g$	Ramp-up limit of generator g
$\underline{R}_g$	Ramp-down limit of generator g

Hidden Markov Model States and Parameters

$\Lambda_{g,d}$	Random vector of latent states of generator g in day d
Λ_d	Random vector of latent states of all generators in day d
Y_d	Random vector of observable states in day d
Θ	Random vector of static model parameters
$\mu_{g,\theta}(.)$	Initial density of the latent states $\Lambda_{g,1:D}$
$f_{g,\theta}(.)$	Transition function of the latent states $\Lambda_{g,1:D}$
$\mu_{\theta}(.)$	Initial density of the latent states $\Lambda_{1:D}$
$f_{\theta}(.)$	Transition function of the latent states $\Lambda_{1:D}$
$\mathcal{L}_{\theta}(.)$	Observations likelihood
$\pi(.)$	Prior density of the static parameters Θ

Markov Chain Monte Carlo Algorithm Notations

$q\,(.)$	Proposal density
$\alpha\,(.)$	Acceptance probability
$K\,(.)$	Markov chain transition kernel
M_0	Number of burn-in iterations

Sequential Monte Carlo Algorithm Notations

$q\,(.)$	Importance function
$w_d^{(i)}$	Normalized importance weight of the particle $\lambda_d^{(i)}$
A_d^i	Ancestral index of the particle $\lambda_d^{(i)}$
$\hat{p}(.)$	Monte Carlo approximation of the target density $p(.)$

I. INTRODUCTION

THE liberalization of the power sector has led to the development of wholesale electricity markets. Due to the increased competition in the generation and retail sectors, new decision-making tools for investment and optimal offering have been developed. These new methods, allowing strategic producers to model other participants' behavior, mainly rely on game theory, equilibrium models and hierarchical optimization [1]. In addition, in order to cope with the growing uncertainty resulting from the large penetration of renewable

The authors are partly supported by the Danish Innovation Fund through the projects 5s - Future Electricity Markets (12-132636/DSF) and CITIES (DSF-1305-00027B), as well as EUDP through the project EnergyLab Nordhavn (EUDP 64015-0055).

energy sources and the fragmentation of the generation sector, optimization under uncertainty tools, such as stochastic and robust optimization [2], are necessary. However, the efficiency of these methods relies on an accurate modeling of uncertainty sources. Although the issues of wind, prices and demand forecasting have been extensively addressed in the literature [3], little work has been done on modeling the uncertainty from rival participants' behavior.

The issue of strategic bidding has traditionally been addressed in the literature as a profit maximization problem under uncertainty. Reference [4] formulated the optimal bidding strategy of a producer on a wholesale market as a robust optimization problem, introducing uncertainty in market prices through confidence intervals. Reference [5] modeled the strategic offering of a wind power producer, for different risk preferences, as a two-stage stochastic optimization problem. Uncertainty in wind production, day-ahead and real-time prices is introduced through a finite number of scenarios. However, for participants with non-negligible market power, modeling the interactions with rival participants is essential. The optimal bidding strategy of a price-maker wind producer in the day-ahead and balancing markets has been formulated as a stochastic hierarchical problem [6], [7]. In this approach the market clearing is explicitly formulated as a constraint of the producer's profit maximization problem. These papers showed a significant improvement in bidding performances, provided rival participants' aggregate supply curve is known. Furthermore, the uncertainty of the supply curve has been accounted for in [8] through a finite number of scenarios. Similarly, the authors in [9] modeled the heat dispatch problem in the Copenhagen area as a stochastic hierarchical optimization problem in which combined heat and power plants are price-makers in the day-ahead electricity market. Yet, none of the aforementioned papers provided a model of the uncertainty of the supply curve.

These studies suggest that an accurate model of the uncertainty of the aggregate supply curve is a key piece of information for a strategic producer in order to achieve informed decisions. However, information on rival participants' bids is seldom available due to market operators' confidentiality policies. To the best of our knowledge only one paper proposed a method for inferring these bids [10]. Rival participants' offering bids on an electricity market are revealed using an inverse optimization problem. This approach provides an accurate estimation of the marginal cost of each participant, provided the dispatched production of each generation block and all the generators' technical characteristics are perfectly known. In practice these assumptions are quite restrictive. Additionally, this approach does not provide a model of the uncertainty of the offering bids.

That is why we propose a more flexible method, allowing us to approximate the aggregate supply curve and providing a complete model of the uncertainty, based only on publicly available information, namely, spot prices, electricity loads, and technical characteristic of generation units. Bayesian inference enables us to update our prior belief on unobservable variables, modeled as a probability distribution, as new information is acquired through the observed variables. Since exact inference is often computationally intractable we focus on approximate stochastic inference methods, such as Markov Chain Monte Carlo (MCMC) [11] and Sequential Monte Carlo (SMC) [12], [13]. These methods have found various applications in finance [14], multi-target tracking [15] and meteorology [16], [17]. These methods approximate the probability distribution of the unknown variables by building iteratively a large sample, approximately distributed according to this distribution. A major appeal of this approach is that it provides an estimator of the target distribution from which it is easy to sample.

In view of the state-of-the-art described above, the contributions of this paper are twofold. Our first contribution is methodological. We propose a Particle MCMC algorithm for revealing the aggregate supply curve in a day-ahead electricity market. This method solely requires observations on the spot prices and total electricity traded, and publicly available information on the technical characteristics of the generation units. It provides a complete model of the uncertainty of the supply curve. Our second contribution is to show the accuracy of the proposed method through a realistic case study. In order to generate realistic market data we simulate a market clearing for a certain number of days, using known offering bids and demand profiles. The market outcomes are used as observations for the inference problem. This way we can compare the results of the inference algorithm to the known bids of the generators.

The structure of this paper is the following. Section II introduces the problem and its formulation as a Bayesian inference problem. The Bayesian inference algorithm proposed is detailed in Section III. Section IV presents an application of the proposed algorithm on a realistic case study. Finally Section V concludes the paper and gathers perspectives regarding future works.

II. Inference Problem Formulation

A. Market Framework

We consider a strategic producer participating in a day-ahead electricity market. We propose a market clearing model similar to European electricity markets. This model is used for two purposes: to build a coherent case study in Section IV, and in the inference algorithm to simulate theoretical market outcomes. For inference purposes, we only need to approximate the market mechanism. Hence, we can make a number of simplifying assumptions.

Participants submit their offers to the market operator for each hour of the following day. They can place two types of offers: single hourly or block orders, linking different hours of the day. For single hourly orders, each participant specifies a combination of prices $C_{g,d,h}$ and quantities $\overline{P}_{g,d,h}$ that it is willing to trade at each hour. And wind producers are assumed to offer their production at zero marginal cost. As a large majority of the energy is traded through single hourly orders we neglect block orders. Instead, we introduce ramping constraints to capture inter-temporal dependencies in the offers of the participants. This assumption allows us to formulate the market clearing problem without introducing binary variables and to directly compute the prices $\lambda_{d,h}^{spot}$ as the dual variables of the power balance equations.

In order to model price-dependent hourly offers, we approximate stepwise offering curves by linear functions

$$C_{g,d,h}\left(P_{g,d,h}\right) = \alpha_{g,d,h} P_{g,d,h} + \beta_{g,d,h}. \tag{1}$$

This assumption allows us to model each participant's offering curve using only two price parameters $\alpha_{g,d,h}$ and $\beta_{g,d,h}$, instead of dividing it into numerous stepwise blocks. And since in practice the competition on the demand side is fairly limited, we assume a completely inelastic demand. As a result, the market clearing for a specific day d can be formulated as

$$\min_{\Omega_d} \quad \sum_{g,h} \left(\alpha_{g,d,h} P_{g,d,h} + \beta_{g,d,h}\right) P_{g,d,h} \tag{2a}$$

$$\text{s.t.} \quad \sum_g P_{g,d,h} + \sum_w P_{w,d,h} = L_{d,h} \quad \forall h \quad : \lambda_{d,h}^{spot} \tag{2b}$$

$$0 \leq P_{g,d,h} \leq \overline{P}_{g,d,h} \qquad \forall g, h \tag{2c}$$

$$0 \leq P_{w,d,h} \leq \overline{P}_{w,d,h} \qquad \forall w, h \tag{2d}$$

$$\underline{R}_g \leq P_{g,d,h+1} - P_{g,d,h} \leq \overline{R}_g \qquad \forall g, h \tag{2e}$$

$$\underline{R}_g \leq P_{g,d,1} - P_{g,d}^0 \leq \overline{R}_g \qquad \forall g \tag{2f}$$

where $\Omega_d = \{P_{g,d,h}, P_{w,d,h} : \forall g, w, h\}$. The objective of this quadratic optimization problem is to minimize the production cost in (2a), subject to power balance equations (2b), power output limits (2c), (2d), and ramping constraints (2e), (2f).

B. Inference Aim

The price offers of the participants in the day-ahead electricity market and the aggregate supply curve are not usually disclosed by the market operator. Our aim is to model the uncertainty of the aggregate supply curve, based on publicly available information. In order to model the market clearing and the price formation we first split the aggregate supply curve into generation blocks. These blocks can be of arbitrary size, e.g. 1MW, or represent a single generation unit or a group of generation units with similar characteristics.

The market clearing mechanism described in Section II-A can be modeled as a Hidden Markov Model (HMM) [18], in which the unobservable latent states are the offering bids of the generation blocks and the observed states are the spots prices and electricity loads. In practice available observations may vary depending on the transparency policies of the specific markets. For instance, in [10] it is assumed that spot prices and accepted production for each generation block can be obtained from the market data. However, this assumption limits the applicability of the proposed method. That is why we consider that only spot prices and electricity loads are disclosed. Fig. 1 shows the structure of this HMM.

Each generation block g fixes its price offer $\lambda_{g,d} = [\alpha_{g,d,h}, \beta_{g,d,h} : h = 1, ..., 24]$ for each hour of the following day. For simplicity we assume that generators solely exercise their strategic behavior by altering their offering prices and offer their full capacity on the market, i.e. $\overline{P}_{g,d,h} = \overline{P}_g$ for all days and hours of the simulation period . Consequently, we model the maximum power output of each generator as a static parameter. Although this assumption is quite restrictive, it can be relaxed by treating the quantities offered to the market as latent variables and infer their distribution as well.

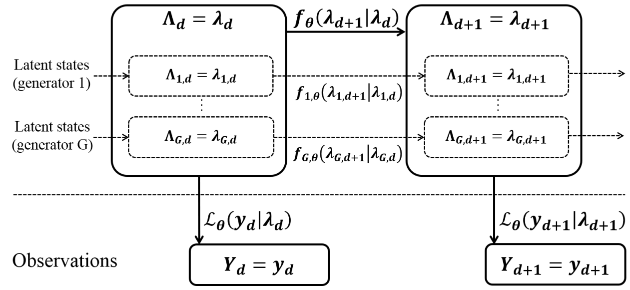


Fig. 1. HMM structure of the market clearing mechanism

The unknown latent states $\lambda_{g,d}$ are modeled as random vectors, denoted $\Lambda_{g,d}$. Each generation block g updates its offer from one day to the following, using a so-called introspection process. It is not required for us to know the nature of this introspection process. We model the stochastic process $\Lambda_{g,1:D} = \{\Lambda_{g,d} : d = 1, ..., D\}$ as a first-order Markov process. Our prior knowledge of the latent states and the introspection process is summarized by the initial density

$$\mu_{g,\theta}\left(\lambda_{g,1}\right) = \mathbb{P}\left(\Lambda_{g,1} = \lambda_{g,1} \mid \Theta = \theta\right) \tag{3}$$

and transition function

$$\begin{aligned} &f_{g,\theta}\left(\lambda_{g,d+1} \mid \lambda_{g,d}\right) \\ &= \mathbb{P}\left(\Lambda_{g,d+1} = \lambda_{g,d+1} \mid \Lambda_{g,d} = \lambda_{g,d}, \Theta = \theta\right). \end{aligned} \tag{4}$$

The transition function and initial density also depend on a set of static model parameters. These parameters are also assumed unknown and modeled as a random vector Θ. The prior density $\pi(\theta) = \mathbb{P}(\Theta = \theta)$ represents our prior knowledge of these parameters.

As we are interested in inferring the aggregate supply curve and not the individual bids of each generation block, we consider directly the random vectors $\Lambda_d = [\Lambda_{g,d} : g = 1, ..., G]$ and the stochastic process $\Lambda_{1:D} = \{\Lambda_d : d = 1, ..., D\}$, modeling the aggregate latent states of all generation blocks. As participants cannot observe the bids of their rivals, the stochastic processes $\Lambda_{g,1:D}$ are assumed mutually independent. As a result, the stochastic process $\Lambda_{1:D}$ is a Markov process and our prior knowledge of the aggregate latent states is summarized by the initial density and transition function

$$\begin{cases} \mu_\theta\left(\lambda_1\right) = \prod_g \mu_{g,\theta}\left(\lambda_{g,1}\right), \\ f_\theta\left(\lambda_{d+1} \mid \lambda_d\right) = \prod_g f_{g,\theta}\left(\lambda_{g,d+1} \mid \lambda_{g,d}\right). \end{cases} \tag{5}$$

The realizations of the stochastic process $\Lambda_{1:D}$ can not be observed directly, however we can refine our prior knowledge through observations of the spot prices and electricity loads, denoted for simplicity $y_d = \left[\lambda_{d,h}^{spot}, L_{d,h} : h = 1, ..., 24\right]$. In HMMs these observable variables are also modeled as random vectors Y_d and assumed conditionally independent given the latent states. Hence the observations likelihood can be expressed as

$$\begin{aligned}&\mathcal{L}_\theta\left(y_d \mid \lambda_d\right)\\&=\mathbb{P}\left(Y_d=y_d \mid \Lambda_{1:D}=\lambda_{1:D}, Y_{1:d-1}=y_{1:d-1}, \Theta=\theta\right)\\&=\mathbb{P}\left(Y_d=y_d \mid \Lambda_d=\lambda_d, \Theta=\theta\right).\end{aligned} \tag{6}$$

Our aim is to infer the joint posterior density of the latent states and parameters conditionally on the observations

$$p\left(\lambda_{1:D}, \theta \mid y_{1:D}\right)=\mathbb{P}\left(\Lambda_{1:D}=\lambda_{1:D}, \Theta=\theta \mid Y_{1:D}=y_{1:D}\right). \tag{7}$$

As exact inference in HMMs is often intractable, we propose an approximate stochastic inference algorithm that targets the joint posterior density by iterative sampling.

III. Bayesian Inference Algorithm

A. Gibbs Sampler

It is not possible to sample directly from the target density $p\left(\lambda_{1:D}, \theta \mid y_{1:D}\right)$. But Bayes' formula provides an expression of the posterior density in function of the observations likelihood and prior density

$$\begin{aligned}&p\left(\lambda_{1:D}, \theta \mid y_{1:D}\right)\\&\propto \pi(\theta)\, \mu_\theta\left(\lambda_1\right) \prod_{d=1}^{D} \mathcal{L}_\theta\left(y_d \mid \lambda_d\right) f_\theta\left(\lambda_{d+1} \mid \lambda_d\right),\end{aligned} \tag{8}$$

where $\propto$ represents the proportionality operator. As a result, Markov Chain Monte Carlo (MCMC) methods allow us to approximate this target density by generating a correlated sequence of samples $\left\{\lambda_{1:D}^{(m)}, \theta^{(m)}: m=1,...,M\right\}$ using a Markov process. That is, at each iteration $(m+1)$ the updated values are drawn from a Markov transition kernel $K\left(\lambda_{1:D}^{(m+1)}, \theta^{(m+1)} \mid \lambda_{1:D}^{(m)}, \theta^{(m)}\right)$. Under certain assumptions on the transition kernel (irreducibility, aperiodicity and invariance) the Markov chain will converge to the target density [11]. That is, after a transient phase, the realized states $\left\{\lambda_{1:D}^{(m)}, \theta^{(m)}: m=M_0,...,M\right\}$ will mimic samples drawn from the target density.

MCMC methods mainly rely on the Metropolis-Hastings (MH) algorithm [19]. At iteration $(m+1)$ the MH update involves drawing a candidate value from a proposal density $q\left(\lambda_{1:D}^{*}, \theta^{*} \mid \lambda_{1:D}^{(m)}, \theta^{(m)}\right)$. The candidate is accepted with probability

$$\begin{aligned}&\alpha\left(\lambda_{1:D}^{*}, \theta^{*} \mid \lambda_{1:D}^{(m)}, \theta^{(m)}\right)\\&=\min\left[1, \frac{p\left(\lambda_{1:D}^{*}, \theta^{*} \mid y_{1:D}\right) q\left(\lambda_{1:D}^{(m)}, \theta^{(m)} \mid \lambda_{1:D}^{*}, \theta^{*}\right)}{p\left(\lambda_{1:D}^{(m)}, \theta^{(m)} \mid y_{1:D}\right) q\left(\lambda_{1:D}^{*}, \theta^{*} \mid \lambda_{1:D}^{(m)}, \theta^{(m)}\right)}\right],\end{aligned} \tag{9}$$

otherwise the Markov chain remains at $\lambda_{1:D}^{(m+1)}=\lambda_{1:D}^{(m)}$ and $\theta^{(m+1)}=\theta^{(m)}$. By construction the transition kernel of this Markov chain satisfies the detailed balance equation and the MH algorithm admits $p\left(\lambda_{1:D}, \theta \mid y_{1:D}\right)$ as invariant distribution. As a result, under weak assumptions on the proposal density, asymptotic convergence is guaranteed [11].

The Gibbs sampler algorithm is a special case of the MH algorithm, for which the parameters and latent states are updated alternatively using their posterior densities as proposal densities [20], such that

$$\theta^{(m)} \sim p\left(. \mid \lambda_{1:D}^{(m-1)}, y_{1:D}\right), \tag{10a}$$

$$\lambda_{1:D}^{(m)} \sim p\left(. \mid y_{1:D}, \theta^{(m)}\right). \tag{10b}$$

It results from (9) that the acceptance probability in the Gibbs sampler is always equal to one. In most cases, it is easy to sample directly from $p\left(\theta \mid \lambda_{1:D}^{(m-1)}, y_{1:D}\right)$. Assuming that we can sample from $p\left(\lambda_{1:D} \mid y_{1:D}, \theta^{(m)}\right)$, the Gibbs sampler algorithm builds iteratively a sequence of samples $\left\{\lambda_{1:D}^{(m)}, \theta^{(m)}: m=M_0,...,M\right\}$ approximately distributed according to the target density.

B. Sequential Monte Carlo Approximation

For a given value $\theta^{(m)}$ of the parameters it is generally not possible to sample directly from the posterior density $p\left(\lambda_{1:D} \mid y_{1:D}, \theta^{(m)}\right)$. The authors in [21] suggest approximating it using a Sequential Monte Carlo (SMC) algorithm. The general idea behind SMC algorithms is to build iteratively a large cloud of equally-weighted particles $\left\{\lambda_{1:D}^{(i)}: i=1,...,N\right\}$ approximately distributed according to the posterior density $p\left(\lambda_{1:D} \mid y_{1:D}, \theta^{(m)}\right)$.

For that, we divide the task by sequentially approximating the densities $p\left(\lambda_{1:d} \mid y_{1:d}, \theta^{(m)}\right)$ for $d=1,...,D$. At each step d, we consider the cloud of equally-weighted particles $\left\{\lambda_{1:d}^{(i)}: i=1,...,N\right\}$ approximately distributed according to the density of interest $p\left(\lambda_{1:d} \mid y_{1:d}, \theta^{(m)}\right)$. The factorization

$$\begin{aligned}&p\left(\lambda_{1:d+1} \mid y_{1:d+1}, \theta^{(m)}\right)\\&\propto p\left(\lambda_{1:d} \mid y_{1:d}, \theta^{(m)}\right) f_{\theta^{(m)}}\left(\lambda_{d+1} \mid \lambda_d\right) \mathcal{L}_{\theta^{(m)}}\left(y_{d+1} \mid \lambda_{d+1}\right)\end{aligned} \tag{11}$$

suggests propagating each particle $\lambda_{1:d}^{(i)}$ to the following day using an importance sampling-resampling mechanism. N offspring particles $\tilde{\lambda}_{d+1}^{(i)}$ are sampled from a so-called importance function $q\left(\tilde{\lambda}_{d+1}^{(i)} \mid \lambda_d^{(i)}, y_{d+1}, \theta^{(m)}\right)$, and then resampled based on their relative importance weights,

$$\begin{aligned}w_{d+1}^{(i)} &\propto \frac{p\left(\lambda_{1:d}^{(i)}, \tilde{\lambda}_{d+1}^{(i)}, y_{1:d+1} \mid \theta^{(m)}\right)}{p\left(\lambda_{1:d}^{(i)}, y_{1:d} \mid \theta^{(m)}\right) q\left(\tilde{\lambda}_{d+1}^{(i)} \mid \lambda_d^{(i)}, y_{d+1}, \theta^{(m)}\right)}\\&\propto \frac{f_{\theta^{(m)}}\left(\tilde{\lambda}_{d+1}^{(i)} \mid \lambda_d^{(i)}\right) \mathcal{L}_{\theta^{(m)}}\left(y_{d+1} \mid \tilde{\lambda}_{d+1}^{(i)}\right)}{q\left(\tilde{\lambda}_{d+1}^{(i)} \mid \lambda_d^{(i)}, y_{d+1}, \theta^{(m)}\right)}.\end{aligned} \tag{12}$$

In practice we implement a stratified resampling mechanism, for which we sample N independent random variables from the uniform distributions

$$u^{(i)} \sim \mathcal{U}\left(\frac{i-1}{N}, \frac{i}{N}\right] \quad \text{for} \quad i=1,...,N. \tag{13}$$

We then define the ancestral indexes $A^i_{d+1} = j$ and the resampled particles $\lambda^{(i)}_{d+1} = \tilde{\lambda}^{(j)}_{d+1}$, such that

$$\sum_{k=1}^{j-1} w^{(k)}_{d+1} < u^{(i)} \leq \sum_{k=1}^{j} w^{(k)}_{d+1}. \tag{14}$$

This resampling mechanism mitigates the variance introduced by standard multinomial resampling mechanisms, by allowing the particles with the lowest importance weights to be resampled at most once [22]. The cloud of N equally-weighted particles $\left\{\lambda^{(i)}_{1:d+1} : i = 1, ..., N\right\}$ is now considered approximately distributed according to the probability density $p\left(\lambda_{1:d+1} \mid y_{1:d+1}, \theta^{(m)}\right)$ and propagated to the following step. The appeal of this approach is that it provides an estimator of the posterior density from which it is easy to sample,

$$\hat{p}\left(\lambda_{1:D} \mid y_{1:D}, \theta^{(m)}\right) = \frac{1}{N} \sum_{i=1}^{N} \delta\left(\lambda^{(i)}_{1:D} - \lambda_{1:D}\right), \tag{15}$$

where $\delta(.)$ represents the Dirac delta function.

However, Reference [21] showed that an approximate Gibbs sampler, updating the states $\lambda^{(m+1)}_{1:D}$ by sampling from this Monte Carlo estimator does not admit the target density $p\left(\lambda_{1:D}, \theta \mid y_{1:D}\right)$ as invariant density. Instead, it is necessary to use a conditional SMC algorithm to approximate the posterior density $p\left(\lambda_{1:D} \mid y_{1:D}, \theta^{(m)}\right)$ at each iteration $m \geq 1$. The particles $\left\{\lambda^{(i)}_{1:D} : i = 1, ..., N\right\}$ are generated conditionally on the reference trajectory $\lambda^{(m-1)}_{1:D}$, associated with the ancestral lineage $A^{m-1}_{1:D}$. The conditioning on the reference trajectory is implemented by ensuring that this path survives all the sampling and resampling steps and generating the remaining $N-1$ particles the usual way. Informally, the introduction of a reference trajectory can be thought of as guiding the sampled particles to a relevant region of the space. In practice, it ensures that the transition kernel leaves the target distribution invariant for any number of particles $N \geq 2$ [23].

Additionally, the importance function has a great influence on the convergence speed of SMC algorithms. It is usually recommended to set the posterior density $p\left(\lambda_{d+1} \mid \lambda_d, y_{d+1}, \theta^{(m)}\right)$ as importance function [21]. As it is not possible to sample directly from the posterior density, we use the transition function of the latent Markov chain $f_{\theta^{(m)}}\left(\tilde{\lambda}_{d+1} \mid \lambda_d\right)$ as importance function. As a result, the importance weights defined in (12) only depend on the observations likelihood

$$w^{(i)}_{d+1} \propto \mathcal{L}_{\theta^{(m)}}\left(y_{d+1} \mid \tilde{\lambda}^{(i)}_{d+1}\right). \tag{16}$$

This method relies on the assumption that the observations likelihood $\mathcal{L}_{\theta^{(m)}}$ is known and can be computed pointwise. In our model $\mathcal{L}_{\theta^{(m)}}$ is difficult to express analytically due to the complex HMM structure. And designing a very uninformative likelihood function would reduce the efficiency of the resampling step in the SMC algorithm. Another solution, inspired by [16] and [17], is to simulate for each particle $\tilde{\lambda}^{(i)}_{d+1}$ a theoretical market outcome $y^{(i)}_{d+1} = \left[\lambda^{spot(i)}_{d+1,h}, L^{(i)}_{d+1,h} : \forall w, g, h\right]$ using the market clearing model in Section II-A. This way we can define the observations likelihood and importance weights in function of the simulated markets outcomes. Additionally, since power balance $L^{(i)}_{d+1,h} = L_{d+1,h}$ is enforced for each particle, we can express the importance weights only in function of the observed and simulated spot prices without loss of generality

$$w^{(i)}_{d+1} \propto \mathcal{L}_{\theta^{(m)}}\left(\lambda^{spot}_{d+1} \mid \lambda^{spot(i)}_{d+1}\right). \tag{17}$$

The conditional SMC algorithm described above is detailed in Algorithm 1.

Algorithm 1 Conditional SMC algorithm at iteration $m \geq 1$

Let $\lambda^{(m-1)}_{1:D}$ be the reference path, associated with the ancestral lineage $A^{m-1}_{1:D}$
for $d = 1$ **do**
  For $i \neq A^{m-1}_1$, sample the $N-1$ particles $\tilde{\lambda}^{(i)}_1 \sim \mu_{\theta^{(m)}}(.)$
  For $i = A^{m-1}_1$, set the particle $\tilde{\lambda}^{(i)}_1 = \lambda^{(m-1)}_1$
  Simulate a market clearing for each particle $\tilde{\lambda}^{(i)}_1$ and compute the theoretical market outcomes $y^{(i)}_1$
  Compute the importance weights $w^{(i)}_1$ for all particles
  For $i \neq A^{m-1}_1$, resample the $N-1$ particles $\lambda^{(i)}_1$ and their ancestral indexes A^i_1 from the weighted sample $\left\{\tilde{\lambda}^{(i)}_1 : i = 1, ..., N\right\}$
  For $i = A^{m-1}_1$, set the resampled particle $\lambda^{(i)}_1 = \lambda^{(m-1)}_1$ and $A^i_1 = A^{m-1}_1$
end for
for $d = 2, ..., D$ **do**
  For $i \neq A^{m-1}_d$, sample the $N-1$ particles $\tilde{\lambda}^{(i)}_d \sim f_{\theta^{(m)}}\left(. \mid \lambda^{(i)}_{d-1}\right)$
  For $i = A^{m-1}_d$, set the particle $\tilde{\lambda}^{(i)}_d = \lambda^{(m-1)}_d$
  Simulate a market clearing for each particle $\tilde{\lambda}^{(i)}_d$ and compute the theoretical market outcomes $y^{(i)}_d$
  Compute the importance weights $w^{(i)}_d$ for all particles
  For $i \neq A^{m-1}_d$, resample the $N-1$ particles $\lambda^{(i)}_d$ and their ancestral indexes A^i_d from the weighted sample $\left\{\tilde{\lambda}^{(i)}_d : i = d, ..., N\right\}$
  For $i = A^{m-1}_d$, set the resampled particle $\lambda^{(B_d)}_d = \lambda^{(m-1)}_d$ and $A^i_d = A^{m-1}_d$
  Set $\lambda^{(i)}_{1:d} = \left\{\lambda^{(A^i_d)}_{1:d-1}, \lambda^{(i)}_d\right\}$ and $A^i_{1:d} = \left\{A^{A^i_d}_{1:d-1}, A^i_d\right\}$
end for

C. Particle Gibbs Sampler

Based on the Gibbs sampler and the conditional SMC algorithm described above, we propose a so-called Particle Gibbs Sampler (PGS) targeting the joint posterior density $p\left(\lambda_{1:D}, \theta \mid y_{1:D}\right)$. At each iteration $m \geq 1$, the parameters are updated using their posterior density $p\left(\theta^{(m)} \mid \lambda^{(m-1)}_{1:D}, y_{1:D}\right)$ as importance function. A conditional SMC algorithm (Algorithm 1) is run to generate a cloud of particles $\left\{\lambda^{(i)}_{1:D} : i = 1, ...N\right\}$ approximately distributed according to the posterior density $p\left(\lambda_{1:D} \mid y_{1:D}, \theta^{(m)}\right)$. The latent states are updated by sampling from the Monte Carlo estimator defined in (15). In practice this is realized by sampling an index i_0 from the discrete uniform distribution $\mathcal{U}\{1, N\}$ and setting the states $\lambda_{1:D}^{(m)} = \lambda_{1:D}^{(i0)}$, associated with the ancestral lineage $A_{1:D}^{m} = A_{1:D}^{i_0}$. The PGS described above is detailed in Algorithm 2.

Algorithm 2 Particle Gibbs Sampler

Set arbitrary the initial parameters $\theta^{(0)}$, states $\lambda_{1:D}^{(0)}$ and ancestral lineage $A_{1:D}^{0}$
for $m = 1, ..., M$ **do**
Sample the updated value of the parameters θ from their posterior density: $\theta^{(m)} \sim p\left(. \mid \lambda_{1:D}^{(m-1)}, y_{1:D}\right)$
Run a conditional SMC algorithm targeting the posterior density $p\left(\lambda_{1:D} \mid y_{1:D}, \theta^{(m)}\right)$, conditionally on the reference path $\lambda_{1:D}^{(m-1)}$ associated with the ancestral lineage $A_{1:D}^{m-1}$ (Algorithm 1)
Sample the updated value of the latent states from the Monte Carlo approximation: $\lambda_{1:D}^{(m)} \sim \hat{p}\left(. \mid y_{1:D}, \theta^{(m)}\right)$ (the ancestral lineage $A_{1:D}^{m}$ is also implicitly sampled)
end for

IV. Case Study

We build a realistic case study by simulating a day-ahead market clearing over thirty days. Based on the observations on spot prices and total production from this simulation, we then apply the algorithm presented above to infer the aggregate supply curve of this system. Finally the estimated aggregate supply curve is compared to the data initially assumed.

A. Case Study Setup

We consider a system with seven thermal generators and one wind producer. The maximum power output, initial production and ramping limits of each unit are derived from [24] and collated in Table I. We generate randomly the bids of each generator by sampling independently for each hour of the 30-day simulation period the price parameters according to the normal distributions

$$\begin{cases} \alpha_{g,d,h} \sim \mathcal{N}\left(\alpha_g^0, \omega_{\alpha_g}^2\right) \ \forall d, h, \\ \beta_{g,d,h} \sim \mathcal{N}\left(\beta_g^0, \omega_{\beta_g}^2\right) \ \forall d, h. \end{cases} \tag{18}$$

The mean values (α_g^0, β_g^0) and the standard deviations (ω_{α_g}, ω_{β_g}) of these normal distributions are collated in Table I.

TABLE I
Generation units parameters

	G_1	G_2	G_3	G_4	G_5	G_6	G_7
$\overline{P}_g$	400	400	300	350	150	400	300
$P_{g,1}^0$	400	400	0	0	0	0	0
$\overline{R}_g$	20	20	60	80	80	180	300
$\underline{R}_g$	-20	-20	-60	-80	-80	-180	-300
β_g^0	5.4	6.1	10.5	10.9	13.3	20.7	26.1
ω_{β_g}	0.5	0.5	1.0	1.0	1.0	2.0	2.0
α_g^0	0.005	0.005	0.011	0.011	0.015	0.03	0.05
ω_{α_g}	0.001	0.001	0.0015	0.0015	0.002	0.002	0.002

We use the load profile data provided in [24] and resize it to match our system. We then introduce load factors for three representative days: base day, shoulder day and peak day and generate randomly the load profiles for thirty days. We consider a single wind farm with an installed capacity of 200MW and we use historic wind production factors from Nord Pool. Finally, using this data we simulate a market clearing for the time period considered. Fig. 2 shows the data used for this simulation and the results of the market clearing.

B. Implementation of the PGS Algorithm

We now assume that electricity spot prices and loads for each hour of the simulation period are directly available from market data. In addition, wind production for each hour of the simulation period can easily be estimated using historical meteorological data. Finally, we assume that the technical characteristics, i.e power output and ramp limits, of the generation blocks are perfectly known. In most cases this information can be accessed through publicly available data. This assumption can be relaxed by including these parameters into the set of unknown parameters θ and estimating their posterior density.

Our prior knowledge on the latent states $\alpha_{g,1:D}$ and $\beta_{g,1:D}$ is characterized by their prior densities. When no information is available, it is recommended to set an uninformative prior density that covers the whole range of possible values. For each generation block g, we model the initial density and transition function of the Markov chain $\alpha_{g,1:D}$ as Normal distributions, such that

$$\begin{cases} \alpha_{g,1} \mid \mu_{\alpha_g}, \sigma_{\alpha_g} \sim \mathcal{N}\left(M_{\alpha_g}, \Sigma_{\alpha_g}\right), \\ \alpha_{g,d+1} \mid \alpha_{g,d}, \sigma_{\alpha_g} \sim \mathcal{N}\left(\alpha_{g,d}, \Sigma_{\alpha_g}\right) \ \forall d, \end{cases} \tag{19}$$

where $\Sigma_{\alpha_g} = \sigma_{\alpha_g}^2 I_{24}$ is the covariance matrix and $M_{\alpha_g} = \left[\mu_{\alpha_g} : h = 1, ..., 24\right]$ the mean vector of these Normal distributions. In addition, the static parameters μ_{α_g} and $h_{\alpha_g} = \frac{1}{\sigma_{\alpha_g}^2}$ are also considered unknown. A standard scheme is to use an independent Normal-Gamma prior to describe them, such that

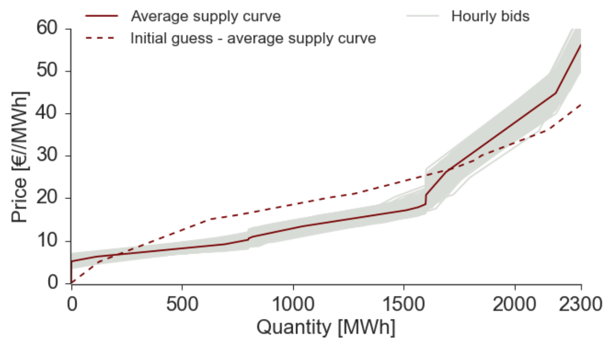
$$\begin{cases} \mu_{\alpha_g} \sim \mathcal{N}\left(m_{\alpha_g}, V_{\alpha_g}\right), \\ h_{\alpha_g} \sim \Gamma\left(a_{\alpha_g}, b_{\alpha_g}\right), \end{cases} \tag{20}$$

where a_{α_g} is the shape parameter and b_{α_g} the inverse scale parameter of this Gamma distribution. By analogy, we define Normal initial densities and transition functions for the latent states $\beta_{g,1:D}$, and Normal-Gamma prior densities for the static model parameters μ_{β_g} and h_{β_g}. The selected values for the hyper-parameters are collated in Table II.

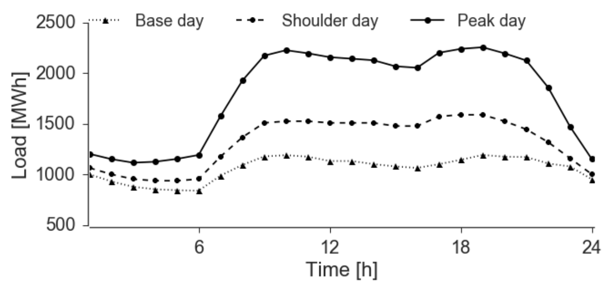
The observations likelihood is defined conditionally on the simulated spot prices $\hat{\lambda}_d^{spot(i)}$. For simplicity we use a normal distribution

$$\lambda_d^{spot} \mid \lambda_d^{spot(i)}, \sigma^{spot} \sim \mathcal{N}\left(\lambda_d^{spot(i)}, \Sigma^{spot}\right) \tag{21}$$

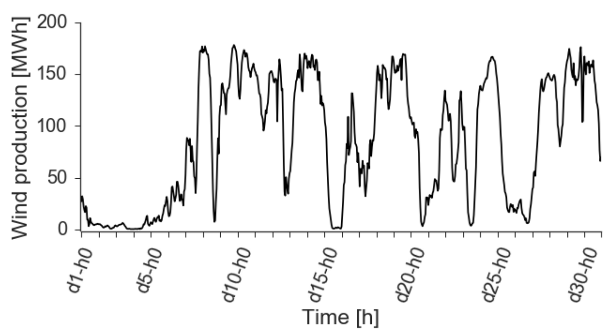
where $\Sigma^{spot} = \sigma^{spot^2} I_{24}$ is the covariance matrix of this Normal distribution. We set arbitrary $\sigma^{spot} = 2$.



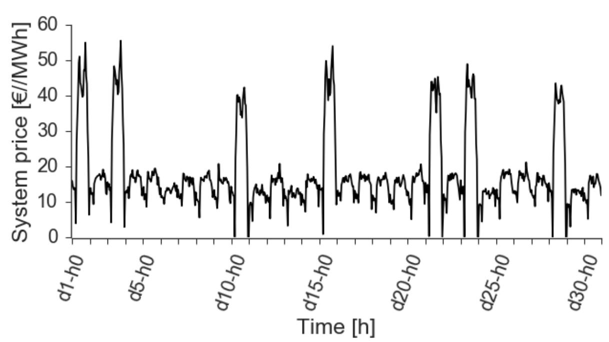
(a) Aggregate supply curves for each hour of the simulation period



(b) Load for 3 representative days



(c) Wind production for each hour of the simulation period



(d) Electricity spot prices for each hour of the simulation period

Fig. 2. Case Study Setup and market clearing outcomes

The static parameters' posterior densities can be updated at each MCMC iteration m based on the expression of the observations likelihood and prior densities, such that

$$\begin{cases} \mu_{\alpha_g}^{(m)} \mid y_{1:D}, h_{\alpha_g}^{(m-1)} \sim \mathcal{N}\left(\bar{m}_{\alpha_g}, \bar{V}_{\alpha_g}\right) \\ h_{\alpha_g}^{(m)} \mid y_{1:D}, \mu_{\alpha_g}^{(m)} \sim \Gamma\left(\bar{a}_{\alpha_g}, \bar{b}_{\alpha_g}\right) \end{cases} \tag{22}$$

TABLE II
PRIOR DENSITIES: STATIC HYPER-PARAMETERS

	G_1	G_2	G_3	G_4	G_5	G_6	G_7
m_{β_g}	0	5	15	15	15	20	30
V_{β_g}	100	100	100	100	100	100	100
m_{α_g}	0.01	0.01	0.01	0.01	0.01	0.02	0.02
V_{α_g}	0.01	0.01	0.01	0.01	0.01	0.01	0.01
a_{β_g} $[10^{-4}]$	1	1	1	1	1	1	1
b_{β_g} $[10^{-4}]$	1	1	1	1	1	1	1
a_{α_g}	0.1	0.1	0.1	0.1	0.1	0.1	0.1
b_{α_g} $[10^{-7}]$	5	5	5	5	5	5	5

where

$$\begin{cases} \frac{1}{\bar{V}_{\alpha_g}} = \frac{1}{V_{\alpha_g}} + h_{\alpha_g}^{(m-1)}, \\ \bar{m}_{\alpha_g} = \bar{V}_{\alpha_g}\left(\frac{m_{\alpha_g}}{V_{\alpha_g}} + h_{\alpha_g}^{(m-1)} \alpha_{g,1}^{(m-1)}\right), \\ \bar{a}_{\alpha_g} = a_{\alpha_g} + \frac{D}{2}, \\ \bar{b}_{\alpha_g} = b_{\alpha_g} + \frac{\left(\alpha_{g,1}^{(m-1)} - \mu_{\alpha_g}^{(m)}\right)^2 + \sum_{t=2}^{T}\left(\alpha_{g,t}^{(m-1)} - \alpha_{g,t-1}^{(m-1)}\right)^2}{2}. \end{cases} \tag{23}$$

As suggested in the literature, an adequate number of iterations to achieve convergence of the MCMC algorithm is $M = 2000$. And we set the number of particles $N = 2000$.

C. Results

After a certain number of iterations ($M_0 = 500$) the correlated sequence $\left\{\lambda_{1:D}^{(m)}, \theta^{(m)} : m = M_0, ..., M\right\}$ is considered approximately distributed according to the target distribution $p\left(\lambda_{1:D}, \theta \mid y_{1:D}\right)$.

Fig. 3 shows the estimated aggregate supply curve and the 2σ confidence interval, based on the mean values of the static parameters. The algorithm is able to accurately approximate the average supply curve, and to model the uncertainty. However, for electricity loads lower than 500MWh the error of the algorithm is significant due to the lack of observations for this part of the supply curve.

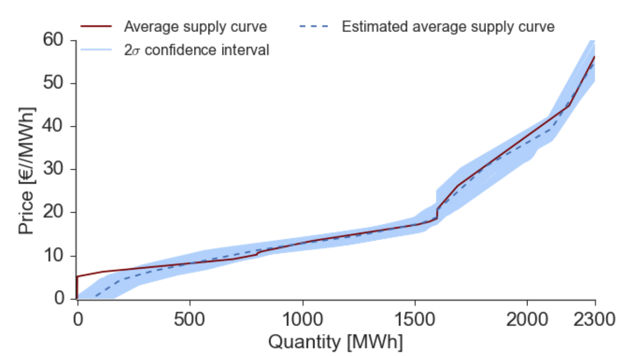


Fig. 3. Estimated average supply curve and 2σ confidence interval

Table III shows the estimated means of the static model parameters for each generation block. The PGS algorithm fails to identify separate generation blocks when their true marginal costs are too similar. This was to be expected because we assumed no information on the electricity dispatch of each block and the marginal producer at each time step. However this information is not necessary in order to estimate the aggregate supply curve.

TABLE III
POSTERIOR DENSITY OF THE STATIC MODEL PARAMETERS

	G_1	G_2	G_3	G_4
$\hat{\mu}_{\beta_g}$	-2.5	4.3	11.2	6.4
$\hat{\sigma}_{\beta_g}$	1.3	0.9	0.7	1.0
$\hat{\mu}_{\alpha_g}$	0.016	0.012	0.010	0.018
$\hat{\sigma}_{\alpha_g}$	0.0024	0.0013	0.0019	0.0015
	G_5	G_6	G_7	
$\hat{\mu}_{\beta_g}$	12.6	20.6	31.1	
$\hat{\sigma}_{\beta_g}$	0.8	1.3	1.1	
$\hat{\mu}_{\alpha_g}$	0.021	0.023	0.040	
$\hat{\sigma}_{\alpha_g}$	0.0020	0.0017	0.0029	

For a price-maker generator participating in the day-ahead market, the estimated value of the average static parameters in Table III can be used to generate scenarios for the aggregate supply curves for each hour of the following day,

$$\begin{cases} \alpha_g, d+1, h \sim \mathcal{N}\left(\hat{\mu}_{\alpha_g}, \hat{\sigma}_{\alpha_g}\right) & \forall g, h \\ \beta_g, d+1, h \sim \mathcal{N}\left(\hat{\mu}_{\beta_g}, \hat{\sigma}_{\beta_g}\right) & \forall g, h \end{cases} \tag{24}$$

V. Conclusion

The Particle Gibbs Sampler algorithm proposed in this paper provides a complete model of the uncertainty of the aggregate supply curve, through an estimate of its posterior probability density. We showed on a realistic case study that this method allows us to accurately approximate the aggregate supply curve. Moreover, it provides an estimator of the posterior distribution of the static model parameters from which it is easy to sample. These results are particularly valuable for price-maker participants in an electricity market, in order to generate scenarios of the supply curve and devise optimal bidding strategies.

This work opens up various opportunities for future research. First of all, the model proposed in this paper can be generalized to accommodate different market frameworks by adapting the market clearing mechanism introduced in Section II. In particular, it is possible to study unit commitment problems by including network and operational constraints. The technical characteristics of the generators, can also be treated as unknown static parameters. And more generally, competition in quantity can be introduced by modeling quantities offered to the market as latent states. Furthermore, the computational cost is a major limitation for implementing SMC and MCMC algorithms to high-dimensional inference problems. Certain steps of the algorithm, such as sampling and resampling, can be parallelized. Additionally, alternative inference methods could be investigated to address this issue. It is sometime suggested that combining deterministic and stochastic inference techniques, also called Rao-Blackwellisation, can improve greatly the computational time of MCMC algorithms [25]. Similarly, methods combining variational approximation and MCMC could allow improving convergence speed of standard MCMC algorithms [26]. Finally, the method proposed in this paper can be generalized by investigating nonparametric Bayesian inference methods, allowing us to reduce the number of assumptions on the underlying states and parameters [27].

References

[1] S. A. Gabriel, A. J. Conejo, J. D. Fuller, B. F. Hobbs, and C. Ruiz, *Complementarity Modeling in Energy Markets*. New York, NY, USA: Springer, 2012.

[2] A. J. Conejo, M. Carrión, and J. M. Morales, *Decision Making Under Uncertainty in Electricity Markets*. New York, NY, USA: Springer, 2010.

[3] J. M. Morales, A. J. Conejo, H. Madsen, P. Pinson, and M. Zugno, *Integrating Renewables in Electricity Markets: Operational Problems*. New York, NY, USA: Springer, 2014.

[4] L. Baringo and A. J. Conejo, "Offering strategy via robust optimization," *IEEE Trans. Power Syst.*, vol. 26, pp. 1418–1425, Aug. 2011.

[5] J. M. Morales, A. J. Conejo, and J. Pérez-Ruiz, "Short-term trading for a wind power producer," *IEEE Trans. Power Syst.*, vol. 25, pp. 554–564, Feb. 2010.

[6] L. Baringo and A. J. Conejo, "Strategic offering for a wind power producer," *IEEE Trans. Power Syst.*, vol. 28, pp. 4645–4654, Nov. 2013.

[7] M. Zugno, J. M. Morales, P. Pinson, and H. Madsen, "Pool strategy of a price-maker wind power producer," *IEEE Trans. Power Syst.*, vol. 28, pp. 3440–3450, Aug. 2013.

[8] L. Baringo and A. J. Conejo, "Offering strategy of wind-power producer: A multi-stage risk-constrained approach," *IEEE Trans. Power Syst.*, vol. 31, pp. 1420–1429, Mar. 2016.

[9] L. Mitridati and P. Pinson, "Optimal coupling of heat and electricity systems: A stochastic hierarchical approach," in *2016 Int. Conf. Probabilistic Methods Applied to Power Syst. (PMAPS)*. IEEE, 2016.

[10] C. Ruiz, A. J. Conejo, and D. J. Bertsimas, "Revealing rival marginal offer prices via inverse optimization," *IEEE Trans. Power Syst.*, vol. 28, pp. 3056–3064, Aug. 2013.

[11] C. Andrieu, N. De Freitas, A. Doucet, and M. I. Jordan, "An introduction to MCMC for machine learning," *Mach. Learning*, vol. 50, no. 1, pp. 5–43, Jan. 2003.

[12] P. Del Moral, A. Doucet, and A. Jasra, "Sequential Monte Carlo samplers," *J. Roy. Stat. Soc. B*, vol. 68, no. 3, pp. 411–436, June 2006.

[13] A. Doucet, S. Godsill, and C. Andrieu, "On sequential Monte Carlo sampling methods for Bayesian filtering," *Stat. and Comp.*, vol. 10, no. 3, pp. 197–208, July 2000.

[14] B. Eraker, "MCMC analysis of diffusion models with application to finance," *J. Bus. & Econ. Stat.*, vol. 19, no. 2, pp. 177–191, 2001.

[15] F. Gustafsson, F. Gunnarsson, N. Bergman, U. Forssell, J. Jansson, R. Karlsson, and P.-J. Nordlund, "Particle filters for positioning, navigation, and tracking," *IEEE Trans. Signal Process.*, vol. 50, pp. 425–437, Feb. 2002.

[16] H. Järvinen, M. Laine, A. Solonen, and H. Haario, "Ensemble prediction and parameter estimation system: the concept," *Quart. J. Roy. Meteorological Soc.*, vol. 138, no. 663, pp. 281–288, Jan. 2012.

[17] M. Laine, A. Solonen, H. Haario, and H. Järvinen, "Ensemble prediction and parameter estimation system: the method," *Quart. J. Roy. Meteorological Soc.*, vol. 138, no. 663, pp. 289–297, Jan. 2012.

[18] Z. Ghahramani, "An introduction to hidden Markov models and Bayesian networks," *Int. J. Pattern Recognition and Artificial Intell.*, vol. 15, no. 01, pp. 9–42, Feb. 2001.

[19] W. K. Hastings, "Monte Carlo sampling methods using Markov chains and their applications," *Biometrika*, vol. 57, no. 1, pp. 97–109, 1970.

[20] A. F. Smith and G. O. Roberts, "Bayesian computation via the Gibbs sampler and related Markov Chain Monte Carlo methods," *J. Roy. Stat. Soc. B*, vol. 55, no. 1, pp. 3–23, 1993.

[21] C. Andrieu, A. Doucet, and R. Holenstein, "Particle Markov Chain Monte Carlo methods," *J. Roy. Stat. Soc. B*, vol. 72, no. 3, pp. 269–342, June 2010.

[22] R. Douc and O. Cappé, "Comparison of resampling schemes for particle filtering," in *Proc. 4th Int. Symp. Image and Signal Process. and Anal. (ISPA)*. IEEE, 2005, pp. 64–69.

[23] F. Lindsten, R. Douc, and E. Moulines, "Uniform ergodicity of the particle Gibbs sampler," *Scandinavian J. Stat.*, vol. 42, no. 3, pp. 775–797, Sept. 2015.

[24] C. Ordoudis, P. Pinson, J. M. Morales, and M. Zugno, *An Updated Version of the IEEE RTS 24-Bus System for Electricity Market and Power System Operation Studies*. Technical University of Denmark, 2016.

[25] G. Casella and C. P. Robert, "Rao-Blackwellisation of sampling schemes," *Biometrika*, vol. 83, no. 1, pp. 81–94, 1996.

[26] N. De Freitas, P. Højen-Sørensen, M. I. Jordan, and S. Russell, "Variational MCMC," in *Proc. 17th Conf. Uncertainty in Artificial Intell.* Morgan Kaufmann Publishers Inc., 2001, pp. 120–127.

[27] S. G. Walker, P. Damien, P. W. Laud, and A. F. Smith, "Bayesian nonparametric inference for random distributions and related functions," *J. Roy. Stat. Soc. B*, vol. 61, no. 3, pp. 485–527, 1999.