Approximate Bayesian Computation for the Analysis of Partial Discharge Data

Kai Hencken[1], Daniele Ceccarelli[1,2], Elsi-Mari Borrelli[1,3] and Andrej Krivda[1]

[1] *ABB Corporate Research Center, Baden-Dättwil, 5405, Switzerland*
kai.hencken@ch.abb.com, andrej.krivda@ch.abb.com

[2] *current address: IRCCS San Raffaele Scientific Institute, Milan, 20127, Italy*
ceccarelli.daniele@hsr.it

[3] *current address: Algorithmiq Ltd, Kanavakatu 3C 00160 Helsinki, Finland*
elsi@algorithmiq.fi

Abstract

Partial Discharges are short breakdowns inside electrical equipment. As they indicate weaknesses of the insulation strength, they are seen as important precursors to a failure of the system. Therefore measurement and analysis of the patterns of instances in time and strength of the discharge are an important tool to analyze the insulation status of electric equipment, that has been addressed already using different methods in the past. In this work we explore how a physics-based stochastic process can be combined with Approximate Bayesian Computation (ABC) as a new way to analyze them. ABC is a method to infer probability distributions of model parameters in cases, where the likelihood is not tractable, but simulations can be done easily. As such it is of interest for complex phenomena or measurement systems, as often found in prognostics applications. Especially the ABC-SMC method was found to be useful here. Real Partial Discharge measurement data was used not only for parameter estimation, but also to do model comparison in order to compare different physical models proposed in the literature.

Nomenclature

$c(t)$	discharge rate above the inception voltage
c_0	constant discharge initialization rate
c_1	time dependent discharge initialization rate
f	line frequency
$M_{i,j}$	i,jth moment of the PRPD pattern rate distribution
Q_k	strength, apparent charge of the discharge event k (pC)
$Q(\theta,\theta')$	kernel function in the ABC-SMC algorithm
t_k	time of the discharge event k
$U_{ext}(t)$	line voltage
U_{inc}	inception voltage of the discharge
$U_{int}(t)$	voltage corresponding to the internal field within the discharge
U_{res}	residual voltage of the discharge
x	measurement data consisting of x_k $k=1,\dots K$
x_k	individual discharge consisting of charge Q_k and phase ϕ_k
γ	proportionality factor between voltage drop and discharge strength
ϵ	accuracy value for the ABC algorithm
ϕ_k	phase of the line voltage at the discharge event k
ν	rate for the reduction of the internal voltage
θ	summary of all discharge model parameters
τ	decay time of the discharge rate

1. Introduction to Simulation Based Inference and Approximate Bayesian Computation

“Approximate Bayesian Computation” (ABC) is the most common method used in the field of “likelihood-free” or “simulation-based” inference. These methods are used, if the statistical model under investigation is easy to simulate from, but the likelihood function, which is at the core of Bayesian or frequentist maximum-likelihood based inference, is not tractable. This means, that it is either not easily accessible to formulate or its numerical evaluation is computationally too demanding.

ABC has gained popularity starting from its inception in the 1990s (Tavaré, Balding, Griffiths, & Donnelly, 1997) in a number of fields ranging from genetics, neural sciences, cosmology and particle physics (Cranmer, Brehmer, & Louppe, 2020). Whereas the underlying principle of ABC stays always the same, a large number of variations of algorithms have been developed in order to address different difficulties or to improve the numerical efficiency (Beaumont, 2019; Sisson, Fan, & Beaumont, 2019). Readily available software packages exist (Csilléry, François, & Blum, 2012; Dutta et al., 2021), simplifying the application, but given the rather easy implementation of Monte Carlo methods in general, a dedicated code was developed in this work.

The main aim of simulation-based or likelihood-free methods is to make use of the possibility to sample a large number of outcomes from the model in order to construct an approximate likelihood or posterior distribution. In the case of ABC one samples both the parameters θ and the measurements x according to the combined probability distribution $(x_n, \theta_n) \sim p(x, \theta)$. This combined distribution is given by $p(x|\theta)p(\theta)$, where $p(\theta)$ is the prior, capturing the knowledge about the parameters and $p(x|\theta)$ the intractable likelihood function underlying the process generating the measured data. By conditioning the samples (x_n, θ_n) on those measurement values x_n that are close to the observed value x, one is able to get samples θ_n from the posterior $\sim p(\theta|x)$. These samples are the basis for further steps in the analysis.

2. Advantage of the Use of Likelihood-Free Inference for Diagnostics and Prognostics Applications

In most applications of prognostics and health management one investigates the possibility to assess the status of technical equipment. In this cases one has the advantage of a good understanding of its functional principle. This is the case, because they are man-made devices, built in order to fulfill a certain function. They are therefore often accompanied by a deep physical or technical understanding. Simulation models are created in many cases as part of the development process, or to analyze the underlying working principle. Such models — including the change of model parameters introducing or leading to faults — are of a high value for good diagnostics or prognostics approaches. Incorporating simulations as part of the algorithms is therefore a good way to capture this knowledge.

In the case of a deterministic simulation model and assuming error-free measurements, the determination of model parameters from the observations or measurements is an inversion problem. In general, this inversion will be ill-posed, meaning that it is numerically unstable to do so. In addition, measurements are in general noisy and even small errors in the measurements can lead to very different and even wrong parameter estimations.

If one is not only interested in a diagnostics, but also a prognostics approach, then for the calculation of the probability of failure ($\mathrm{PoF}(t)$), the determination of the uncertainty of the state of the system, as well as those parameters defining the future dynamics, are required. The Bayesian inference, which is underlying the ABC approach, allows to do so in a consistent way.

The situation becomes even more complex, if the model is of a stochastic nature. Reasons for this, apart from the measurement errors, are often the presence of unknown values or the randomness inherent in some process.

The intractability of the likelihood is in general due to the existence of a large number of hidden, that is unobserved or unknown, parameters, states or values. Within the Bayesian approach one needs to marginalize over them, that is integrate over all possible values, which is numerically impossible in practice. These hidden variables can have different origins: Technical models contain a large number of parameters, which are varying from device to device, but are often not relevant for the degradation state of the system. They will nevertheless influence the way the device is operating and the values of measurements.

Other hidden variables are unmeasured external influences. These can be environmental or operational factors. As before their value will in general not be important for the state of degradation, but they are influencing the measured quantities.

Finally the measurement principle can have unobserved internal states. An example could be their dependency on earlier measurements, or a probabilistic element in the underlying process. Partial Discharge analysis can be seen as a problem of this type.

All these facts make simulation-based inference methods of interest for diagnostics or prognostics approaches. This is not restricted to the ABC method, discussed here. Methods combining simulations with elements of machine learning are gaining popularity in other fields and are of potential interest in this area as well. This can be interesting, if e.g. only a single remaining useful life (RUL) value should be predicted instead of a full distribution or the simulations need to be accelerated with the help of a surrogate model.

3. Partial Discharge Measurement to Assess the Electric Insulation Status

Electrical equipment — especially in the high voltage area above few kV — is subject to strong electric fields, that are applied over a long time. Insulation material is known to be able to withstand these fields only up to a certain level and changes of material properties with time can lead to a reduction of this critical field strength. Typical examples are material aging, changes due to exposure to water, corrosive gases, dust built-up, crack formation or forming of other defects due to mechanical or thermal stress.

The loss of the electrical insulation capability will ultimately lead to a complete failure of the equipment. This can, in connection with the appearance of electrical arcs, lead to fire or explosions and therefore to the complete destruction of it, resulting also in potential hazards to people. It is therefore of interest to be able to detect the reduction of the insulation strength of a system under voltage and its potential evolution with time.

"Partial Discharge" (PD) describes the phenomena, where the insulation strength is insufficient in only a localized region. The corresponding electrical breakdown will therefore be local only, not bridging the full distance from high voltage to ground. This "partial" electric breakdown is in general extinguished after a short time as the charges flowing during the discharge are reducing the local electrical field to a value below the critical one. The effect of this local field reduction will become weaker with time and due to the change of the externally generated electric field in AC applications, a new discharge will occur after some time.

PD is a rather generic term, which is given to any local electric breakdown. But these are often occurring in different parts of the insulation system, as well as due to different origins: At sharp metallic edges the local field enhancement leads to "corona" discharges; at surfaces the buildup of dust and humidity forms a conductive layer and leads to "surface discharges"; defects inside insulators, often introduced during production, will in general have a lower insulation strength and lead to the formation of "void discharges". Starting from these initial defects, they will further erode the material leading to "treeing" or "tracking" inside or on the surface of the material, which develops and makes the defect worse over time.

According to statistical data, up to 85% of all severe failures of high-voltage equipment can be linked to the presence of partial discharges in those systems. They are one of the main precursors or indicators of the upcoming failure of the electrical insulation capability of a high-voltage system. Using monitoring systems to detect them is one of the most often used methods in high-voltage systems. Both dedicated testing and measurement systems manufacturer, as well as, producers of high-voltage equipment provide a variety of laboratory, off-line and online solutions.

PD measurements are done to assess the quality of individual parts or full system during their production and also as part of the acceptance testing, with limits of the allowed PD activity defined. Whereas this allows to capture defects, that are already present during production or installation, the degradation of the electrical insulation can only be tracked with an online monitoring system.

An assessment of the insulation status is often done using external equipment and at fixed time intervals. Such measurements have the advantage of allowing for expensive, but very accurate equipment to be used, as well as providing some detailed diagnostics results, e.g. by changing the test voltage applied. On the other hand they require an expensive shutdown of the installation and only give a snapshot of the status at that specific time. Online monitoring systems on the other hand provide a continuous measurement, allowing for an early detection, but can also study the evolution with time. But they also need to be rather inexpensive in order to be deployed widely.

The evolution of the partial discharge with time is often summarized by a few key parameters, e.g. a strength and a discharge rate. But an estimate of the time until failure of the electrical insulation cannot be done easily from this. The type of discharge plays a rather important role: whereas a corona discharge can be present in a system for a longer time before a fault occurs, a treeing or surface discharge will often evolve to a full breakdown very fast.

The character of defects and with this the appearance of the discharge will change over time. This has been investigated as a way to infer the remaining lifetime until breakdown for specific defect types in a number of publications, see for example (Montanari, 1995; Wang, Cavallini, Montanari, & Testa, 2010; Lv, Rowland, Chen, Zheng, & Iddrissu, 2017).

In practical applications PD measurement will be disturbed by a number of external phenomena, which need to be distinguished from real PD events. Whereas this can be done in a lab setup with the help of shielding measures, this is not possible in online applications. A major goal of any online PD analysis is therefore to detect them, separate them from other disturbances, but also to classify them according to their origin, the type of discharge or defect, to get further information regarding their severity. Given their importance to avoid severe failures of those expensive systems, a number of detailed analysis approaches have been explored.

4. Basic Principle of Partial Discharge Measurement and Analysis

The local discharge is in general characterized by the time t_k of occurrence and its strength, which is in general measured as "apparent charge" Q_t (International Electrotechnical Commission (IEC), 2000). Due to the stochastic nature of the discharge, the individual events (t_k, Q_k) don't happen at deterministic times or repeat in a systematic way, see Fig. 1.

A full measurement therefore consists of a number of such discharge events $k = 1, \dots N$, that are either recorded continuously or sometimes with gaps within them.

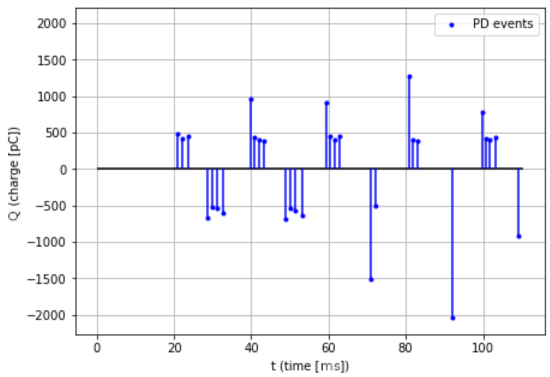

Figure 1. Partial Discharges occur stochastically at times t_k and with varying strengths Q_k. A typical sequence is shown schematically in this plot. A PD measurement consists of capturing a full sequence of such individual discharge events.

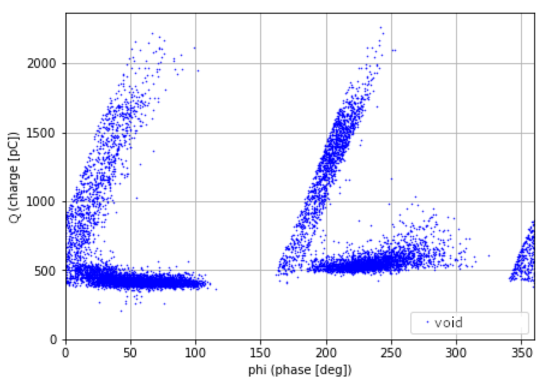

Figure 2. A typical PRPD pattern of a void discharge. The points are individual discharge events (ϕ_k, Q_k) of the phase and discharge strength.

As the AC voltage $U_{ext}(t)$ plays a dominant role in reaching the critical electric field level of the defect, its value at the moment of the discharge $U_k = U_{ext}(t_k)$ or alternatively the phase $\phi_k = \phi(t_k)$ are recorded as well. A number of methods have been proposed to analyze this sequence of individual PD events. The most commonly used one is the "Phase Resolved Partial Discharge" (PRPD) analysis. In this, one plots the points (ϕ_k, Q_k) in a two-dimensional scatter plot, or converting them into a density plot. As the partial discharge is driven by the applied AC line voltage, one expects that the pattern formed in these graphs reflects the nature of the discharge. In addition, depending on the measurement approach used, one is not able to measure the polarity of the discharge Q_k. Therefore we will in the following assume that only the absolute value of the discharge is available as Q_k. A typical PRPD plot is shown in Figure 2 of a void, and in Figure 3 of a corona discharge.

Different PD types, but also different external disturbances,

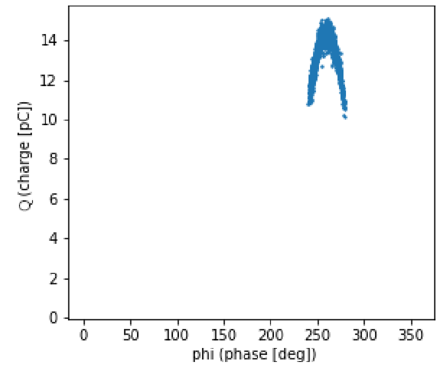

Figure 3. A typical PRPD pattern of a corona discharge. Compared to the void discharge, at lower voltage levels the corona discharges occur only during the second half of the cycle, in this case for negative voltages.

lead to different patterns in the PRPD diagram, which are used as the basis for further classification of the defect. The patterns are converted into features, summarizing the pattern itself. A common approach is to bin the data in the phase direction and use the distribution of the number of points, the average, as well as, the maximal discharge strength as features, which are further characterized by typical statistical measures like the skewness and kurtosis (Krivda, 1995b, 1995a). Other approaches interpret the density as an image and make use of methods from image classification.

Partial Discharge Classification has been investigated using almost all approaches developed in machine learning in the past; reviews are given, e.g. in (Danikas, Gao, & Aro, 2003; Sahoo, Salama, & Bartnikas, 2005; Ma, Chan, Saha, & Ekanayake, 2013; Raymond, Illias, Bakar, & Mokhlis, 2015; Barrios, Buldain, Comech, Gilbert, & Orue, 2019; Lu, Chai, Sahoo, & Phung, 2020).

Due to the complex nature of the underlying stochastic process, as well as the large variety of defects and therefore potential patterns, most analysis or classification approaches tend to use data-driven algorithms, based on collecting data from a large number of different discharges either in the lab or in the field and training the algorithm with them.

On the other hand, models to describe partial discharges from a microscopic or physical approach have been explored in the literature as well, see e.g. (Niemeyer, 1995; Cavallini & Montanari, 2006; Callender, Golosnoy, Rapisarda, & Lewin, 2018; Callender & Lewin, 2020).

In many cases the motivation was to develop a deeper understanding of the processes at work. But there were also attempts to use them as a basis for more model-based PD analysis approaches (Heitz, 1999; Altenburger, Heitz, & Timmer, 2002; Cavallini & Montanari, 2006; Patsch & Berton, 2002). In this work we are picking up this way of PD analysis and combining it with modern statistical inference methods.

5. Physics-Based Partial Discharge Modeling

A simple model to describe the creation of partial discharges is given in this section, following in major parts the one proposed in (Heitz, 1999). This model requires only a rather small set of parameters, which helps with the inference.

In order to reduce the complexity of the approach, we restrict ourselves to a symmetric description, that is, the processes during the positive and negative voltage half-cycles are assumed to be described by the same set of parameters. This leads to a symmetric PRPD pattern, which is not in agreement with many measurements. A more general model allows them to be different for the two half-cycles, doubling their number, but this is needed in order to describe (strongly) asymmetric patterns. As we are mostly exploring the applicability of the ABC approach here, the restriction to symmetric ones seems to be justified.

The discharge within any defect will depend on the local electric field strength. As this field is not directly accessible to us, but can be assumed to be proportional to some related voltage across them, the model converts all electric fields into corresponding voltages. For example, the field inside the defect generated by the applied external line voltage is proportional to $U_{ext}(t) = U_0 \cos(2\pi f t)$. It is convenient to normalize all voltages in the system to be proportional to the amplitude of this line voltage, setting $U_0 = 1$, that is

$$U_{ext}(t) = \cos(2\pi f t). \tag{1}$$

A discharge occurs due to the electrical field exceeding a critical level, corresponding to an "inception voltage" U_{inc}. But the discharge will not occur instantaneously when the total voltage reaches this level; only with a certain probability per time unit, that is a rate $c(t)$. This rate is assumed to originate from two different sources, related to the availability of "seed electrons" assumed to trigger the start of the discharge: A constant source of seed electrons, giving a constant rate c_0 and a time dependent one, with seed electrons being produced during the last discharge, but recombining and therefore disappearing with time. The rate connected to these is given as $c_1 \exp(-t/\tau)$. The total rate is therefore

$$c(t) = \Theta(U_{tot}(t) - U_{inc})\,(c_0 + c_1 \exp(-t/\tau)) \tag{2}$$

with the Heaviside or indicator function Θ being one if U_{tot} is larger than U_{inc} and zero otherwise.

The corresponding total voltage inside the defect is given by the external one U_{ext} together with the one produced by the charges created by the individual discharges $U_{int}(t)$, that is

$$U_{tot}(t) = U_{ext}(t) + U_{int}(t). \tag{3}$$

Similarly to the reduction of seed electrons, it is assumed, that if no discharge happens, the internal charges and therefore

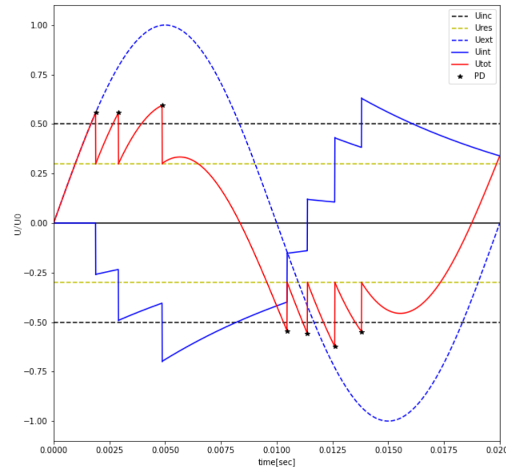

Figure 4. The evolution of the external ($U_{ext}(t)$), total (U_{tot}), and internal (U_{int}) voltage is shown during one cycle. If the total electric field exceeds the inception voltage U_{inc} a discharge can occur with some rate $c(t)$. Such a discharge reduces the voltage to the residual one (U_{res}). Without a discharge the internal voltage is reduced in time.

$U_{int}(t)$ will be reduced with time, described by a decay rate ν, that is

$$\frac{dU_{int}(t)}{dt} = -\nu U_{int}(t) \tag{4}$$

If, on the other hand, a discharge is initiated, charges are flowing and reduce the electrical field and therefore the corresponding voltage U_{tot}. The discharge happens during a short time, until the total electric field, that is the corresponding voltage, reaches the "residual voltage" U_{res}, where it is extinguished

$$U^+_{tot}(t) = U_{ext}(t) + U^+_{int}(t) = U_{res}. \tag{5}$$

Here U^+ denotes the value immediate after the discharge. The total strength of the discharge is assumed to be proportional to the change in voltage

$$\Delta U = U^-_{tot}(t) - U^+_{tot}(t) = U^-_{int}(t) - U^+_{int}(t) \tag{6}$$

where U^- denotes the value immediately before the discharge. The measured discharge strength is given by

$$Q = \gamma \Delta U \tag{7}$$

where the proportionality factor γ is a property of the defect and geometry.

The typical change of U_{tot} and U_{int} with time and the relation with the occurrence of a discharge is shown in Fig. 4.

This description is sufficient to create a simulation code for the generation of discharge event sequences (Q_k, t_k) by following the evolution of the different voltages over time and selecting whether a discharge happens based on the value of $c(t)$. The likelihood function on the other hand is more difficult to derive, especially for more complex models.

The model has overall seven parameters that need to be determined $\theta = (U_{inc}, U_{res}, c_0, c_1, \tau, \nu, \gamma)$. Despite its rather simple form, the model is able to create quite a number of different PRPD patterns, therefore it can be seen as a model to cover different types of PD defects. The main reason for the variability of patterns is the fact, that there are three time scales in the model: one is related to the AC voltage, corresponding e.g. to 50Hz, one to the reduction of $c(t)$, given by τ, and finally one to the change of $U_{int}(t)$, given by ν. Depending on their respective values or relative sizes the sequence of discharge can vary, e.g. leading to cases with many discharges happening during a half-cycle to some, where a long time between consecutive discharges is present.

6. BASIC PRINCIPLE OF ABC

The basic principle of ABC can be described in a rather simple way: One assumes to have a prior distribution $p(\theta)$ with the capability to generate easily samples from it. Using these model parameters θ_n, it is again assumed that one can generate measurement samples x_n, even though $p(x|\theta)$ is in general not available. The combination (x_n, θ_n) are samples generated from the combined probability density function $p(x, \theta)$.

If one selects from these pairs only those n with $x_n = x$, that is which agree with the observed value x, it is easy to see that the corresponding θ_n are distributed according to the posterior distribution $p(\theta|x)$. In practice such an algorithm is not usable, especially when the measurements x is continuous or high-dimensional, as the agreement $x_n = x$ is hardly ever fulfilled.

Instead one introduces a distance measure between two measurements x and x': $\rho(x, x')$ and the requirement of $x = x'$ is relaxed to

$$\rho(x, x') \leq \epsilon \tag{8}$$

with ϵ chosen sufficiently small. As the agreement between x and x' is no longer exact, the posterior distribution will be approximate as well.

In many cases the construction of the distance ρ is done with the help of some summary statistics $t(x)$ and the distance is then defined with respect to them, giving

$$\rho(x, x') = \rho(t(x), t(x')) \leq \epsilon. \tag{9}$$

This leads to the simplest approach, the “ABC rejection algorithm”, as given in Algorithm 1.

The value of ϵ can often be chosen by running the algorithm with a decreasing series of values until the approximate posterior distribution does not change significantly.

The disadvantage of this algorithm is, that it requires a large number of runs to find a good value of ϵ. In addition, the sampling of θ_n is done with respect to the prior distribution only. This leads to an overall rather inefficient approach. Therefore more refined algorithm, based e.g. on MCMC and SMC have been proposed in the literature (Sisson et al., 2019). In this work we have used the ABC-SMC approach as described in (Toni, Welch, Strelkowa, Ipsen, & Stumpf, 2009). This algorithm uses a decreasing sequence of ϵ_t, either predefined or dynamically adjusted at each step $t = 1, \ldots T$. Details of the algorithm are given in Algorithm 2.

```
for n = 1, ..., N do
    do
        Sample θ* ~ p(θ)
        Sample x* ~ p(x|θ*))
        Calculate D = ρ(x, x*)
    while D > ε;
    θ_n = θ*
end
```

Algorithm 1: The ABC rejection algorithm.

```
Sample θ_n^0 ~ p(θ), n = 1, ..., N
Initialize w_n^0 = 1/N, n = 1, ..., N
for t = 1, ..., T do
    for n = 1, ..., N do
        do
            Sample θ* from θ^{t-1} using a multinomial
              distribution with weights w^{t-1}
            Perturb θ* to a new θ** ~ Q(θ, θ*)
            Sample x* from p(x|θ**)
            Calculate D = ρ(x, x*)
        while D > ε_t;
        θ_n^t = θ**
    end
    Calculate new weights w^t using
```

$$w_n^t = \frac{p(\theta_n^t)}{\sum_{n'=1}^{N} w_{n'}^{t-1} Q(\theta_n^t | \theta_{n'}^{t-1})}$$

```
    and normalize them
end
```

Algorithm 2: The ABC-SMC algorithm.

In the initial step a starting population of θ_n from the prior distribution is chosen and some equal initial weights are initialized. In each subsequent step a value from the current population of parameters is chosen according to its corresponding weight and perturbed using some kernel function $Q(\theta, \theta')$. A value of x is simulated and the new value of θ is kept, if the acceptance criterion $\rho(x, x') \leq \epsilon_t$ is fulfilled. This process is repeated until N new values of parameters θ_n are obtained. The weights are adjusted to account for the prior distribution and the probability of parameters to be chosen based on the former weights and the kernel function. The value of ϵ_t is reduced and the process repeats.

The main advantage of this approach is that “good” values θ_n

are kept and that an automatic reduction of the ϵ_t is done. Refinements with respect to choosing this sequence in an adaptive way and also in order to adjust the kernel with time to improve its width have been proposed in the literature and were implemented and tested as well. Other improvements are possible by keeping only the best results out of a much larger population of parameter values in each step.

Overall these improvement were found to reduce the computational effort and avoid a degeneration of the samples, that is often seen in SMC algorithm (Doucet, De Freitas, Gordo, & James, 2001).

7. Application of ABC to Partial Discharge Data

In order to apply the ABC algorithm to identify the parameters $\theta = (U_{inc}, U_{res}, c_0, c_1, \tau, \nu, \gamma)$ of the underlying partial discharge model, some additional steps are needed. The basic ABC algorithm requires a measure of the closeness of two discharge patterns x and x', where each

$$x = \{(t_k, Q_k)\}_{k=1,\dots K} \tag{10}$$

consists of a sequence of K PD discharge events. Depending on how the measurement is performed, the number of events can be different between x and x', e.g. if measurements are done during a pre-specified length of time T, or alternatively can be the same, if a fixed number of events is calculated, but this means that the total measurement time T is not the same between them. The total rate of discharges K/T gives important information about the PD type or the parameters of the model. In most practical application of PRPD analysis this is however ignored, and the discharge probability distribution is analyzed instead.

A number of distance measures $\rho(x, x')$ can be chosen in order to compare the two-dimensional point clouds. If one normalizes the PRPD patterns to correspond to a probability density, as commonly done, statistical distances can be used, but they are not suitable for a comparison of the rate densities. Measures could also be based on the features extracted from the PRPD plot, as described in Sec. 4. These have the same disadvantages of being insensitive to the total rate, unless this rate is added to the list of features and also not being comparable in their values. They can also not be extended in a systematic way in most cases.

In this work an approach based on moments of the two-dimensional distribution is used. They are defined both in the phase and discharge-strength direction. As the phase is circular, an expansion in terms of a trigonometric series is done, whereas in the discharge direction, for which we assume that only the absolute value of the discharge strength is measured, power moments are used. The most basic definition of the moments is

$$M_{i,j} = \frac{1}{T} \sum_k (Q_k)^i \,[\cos, \sin]\,(j\phi_k) \tag{11}$$

but for computational efficiency other definitions, e.g normalizing the power series, or using powers of the base trigonometric function instead of the normal Fourier series have been used as well for comparison. Overall no major differences were found with respect to the posterior distribution calculated. Moments up to $i, j = 5$ were used to reduce the computational effort.

The distance measure used is a weighted Euclidean norm

$$\rho(M, M') = \sqrt{\sum_{i,j} w_{i,j} \left(M_{i,j} - M'_{i,j}\right)^2}. \tag{12}$$

A convenient choice of the weights $w_{i,j}$ was used by bootstrap resampling the measured events and using the (inverse of the) estimated variance of each moment for it. In this way the natural variation of each moment is used and the norm definition is independent of the simulated results.

For the application of ABC to the model uniform priors were chosen for U_{inc} and U_{res}, as these values are required to be positive and are bounded, as too large values make it impossible to have PD events. For all other parameters, which are required to be positive, suitable gamma priors were chosen.

8. Results

The method was applied to both simulated (synthetic) and real data. For simplicity only symmetric PRPD patterns were studied here; the extension to asymmetric ones could be addressed in the same way, but given the fact that the number of parameters almost doubles, this was not done. As a real example, data from a surface discharge measurement was used and was symmetrized before its use.

To confirm in a first step that the approach is suitable in principle and is able to reconstruct the parameters used in the generation of the data, it was applied to synthetic data, generated by the same PD model as used for the analysis. Results of the PRPD pattern for the initial and the final parameters are shown in Fig. 5. The visual convergence of the PRPD patterns is clearly visible. A comparison of the estimated parameters being close to the one used for the simulation is shown in Fig. 6 for one parameter as example. The parameter γ becomes more narrow as the SMC algorithm proceeds.

In the next step, the algorithm was applied to real data. The comparison of the PRPD data was found to be less convincing than in the synthetic case. A more thorough look showed, that there were some difficulties in describing the sharp drop of PD events at the lower end of the pattern.

A deeper investigation revealed that the experimental data

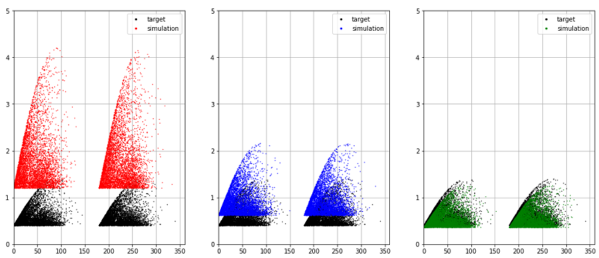

Figure 5. Convergence of the PRPD pattern to the synthetic one during the progression of the ABC-SMC algorithm. Shown are the patterns due to the mean value of the parameter.

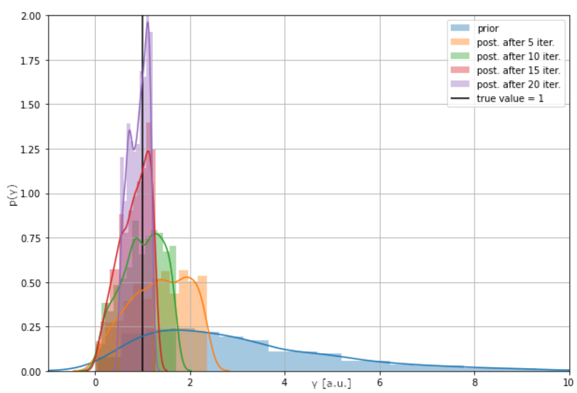

Figure 6. Convergence of the parameter γ to the one used in the generation of the data during the progression of the ABC-SMC algorithm.

had a cutoff introduced, that was not coming from the discharges per se, but from the measurement system, in order to suppress noise. As this threshold is not related to any physical property of the discharge, the PD model will in general not obey it.

The introduction of this threshold on the other hand can be very easily done into the simulation, that is, discharge events are generated and those with a strength below the threshold are removed before the moments are calculated. This shows the very flexible nature of any simulation-based inference, as such measurements effects can be considered in an easy way. It was found, that the ABC algorithm improves the agreement with the measurement in the PRPD plot, if the value is assumed to be known.

We further tested, whether the approach is able to cope with data, where an unknown threshold level exists. This may occur when some automatic pre-processing unit will introduce it without further knowledge of its value. The threshold is introduced as an additional parameter in the model and determined in the same way as the other ones. The ABC algorithm was found to be able to correctly estimate the threshold level. A comparison of the PRPD patterns finally found is shown in

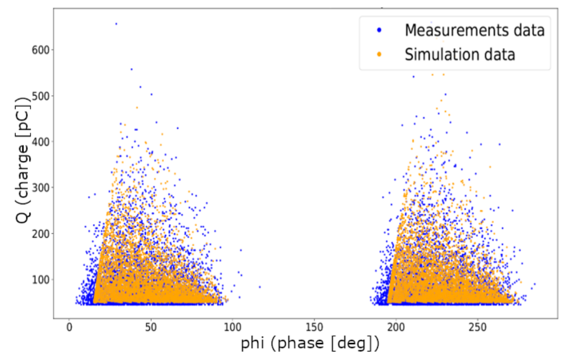

Figure 7. A comparison of a real surface discharge pattern with a cut-off and the result of the ABC algorithm. The threshold was not fixed to the known value in this case, but was determined by the ABC algorithm as well.

Fig. 7. The good agreement at the lower cutoff shows that the introduction of any measurement effects is important to get a good result.

ABC allows even to do further statistical analysis. In Bayesian analysis model comparison can be done in a natural way, using e.g. Bayes factors, or determining the posterior probabilities of e.g. two different models. This can also be implemented in ABC, as described e.g. in (Toni, 2011). This model comparison can be applied to do classification by selecting e.g. priors for the parameters, that are compatible with only a specific type of discharge. Alternatively, it was used here in order to compare different PD models. There have been discussions about the details of the mechanism to generate a specific feature of void discharges, called the "rabbit-ear", in the literature (Cavallini & Montanari, 2006; Niemeyer, 1995). ABC can be used to compare the two models and decide, which one is more likely to explain the data, even though they have different parameters, as well as even different numbers of them.

The convergence of the probability of each model given the measurement is shown in Fig. 8, whereas a comparison of the predicted PRPD for each model is shown in Fig. 9. Whereas the results are too early to draw already strong conclusions, this shows the potential of ABC even beyond the parameter estimation capabilities.

9. Conclusion

This work has shown, that partial discharge analysis can be done with the help of a physical model of the discharge combined with the use of ABC as a simulation-based inference method. A simple physical model was presented and the summary statistics and distance measure needed to define the closeness of the simulated pattern with the measured one was discussed. Starting from the simplest ABC algorithm the use

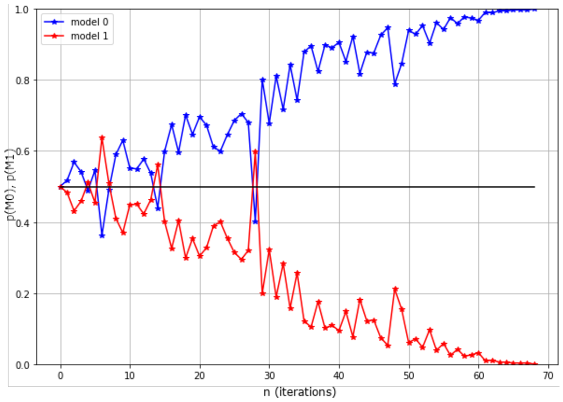

Figure 8. Posterior model probability of two simulations model explaining the origin of the "rabbit ear" phenomena in void discharges. It can be seen that after a number of iterations "model 0" is more likely than "model 1" to explain the PRPD pattern.

of more efficient ones, based on SMC and sequences of accuracy measures is discussed.

An interesting feature of simulation based inference is the possibility to incorporate quite easily phenomena, which are difficult to define in a probabilistic framework. As an example the measurement threshold, which is often applied in practice to reduce the amount of noise, was incorporated easily and it was shown that inferring not only the PD model parameters, but also the threshold itself was possible. Finally the use of ABC for model comparison was shown, which has a direct link to PD classification. Alternatively, it can be used as a method to distinguish between different PD models discussed in the literature.

Simulation based inference of partial discharge can be seen as a follow-up of research activities done in the 90s, in contrast to data-driven methods explored more recently. One limitation at that time was clearly the available computing power, which required quite specialized techniques. With the increase of computing power, but also the development of modern inference techniques this approach to PD analysis seems to be within reach today.

Simulation-based inference should be seen as an interesting approach to diagnostics and prognostics beyond the application to PD analysis. The detailed knowledge of the technical devices monitored are an import input, that should be taken into account. In addition, the uncertainty quantification of the parameters extracted, but also the large number of "hidden variables" that are not relevant for the degradation of the system, can both be dealt with. Therefore the application of this method to other systems is clearly of interest, and should be explored in the future.

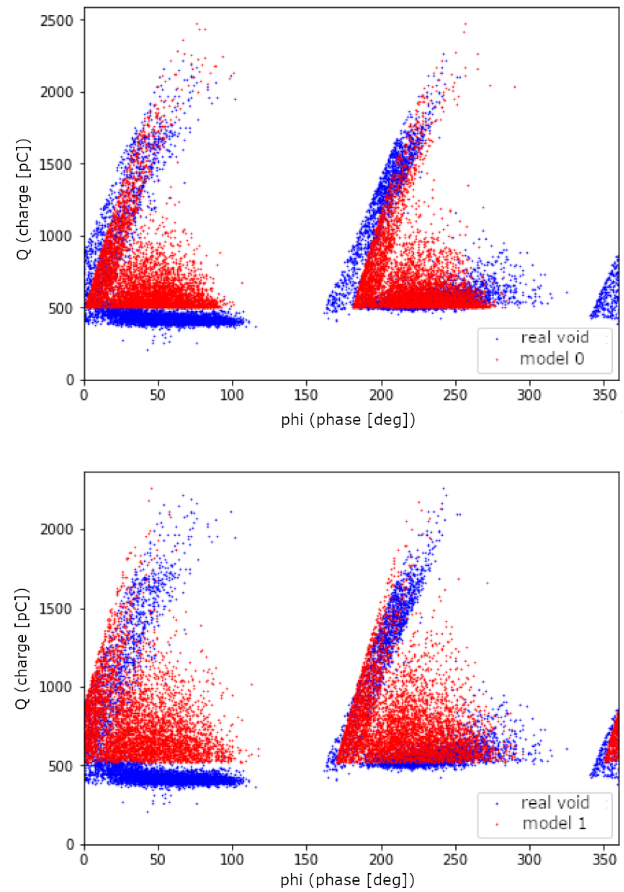

Figure 9. A comparison of the PRPD pattern as generated by the two models in the final iteration and compared to the real data. The upper plots shows the result for model0, the lower one for model 1.

Acknowledgment

This work has profited from a large number of discussions with different people within the research center and within ABB. Especially the insight of Yannick Maret at ABB on signal processing and data analysis, as well as Thomas Christen and Jörg Lehmann, both now at Hitachi Energy, into the foundations of electrical insulation, partial discharges, and the theory of stochastic processes have helped at various stages of the work.

References

Altenburger, R., Heitz, C., & Timmer, J. (2002). Analysis of phase-resolved partial discharge patterns of voids based on a stochastic process approach. *Journal of Physics D: Applied Physics*, *35*(11), 1149.

Barrios, S., Buldain, D., Comech, M. P., Gilbert, I., & Orue, I. (2019). Partial discharge classification using deep learning methods—survey of recent progress. *Energies*, *12*(13), 2485. doi: 10.3390/en12132485

Beaumont, M. A. (2019). Approximate bayesian computation. *Annual Review of Statistics and Its Application*, *6*(1), 379-403.

Callender, G., Golosnoy, I. O., Rapisarda, P., & Lewin, P. L. (2018). Critical analysis of partial discharge dynamics in air filled spherical voids. *Journal of Physics D: Applied Physics*, *51*(12), 125601.

Callender, G., & Lewin, P. L. (2020). Modeling partial discharge phenomena. *IEEE Electrical Insulation Magazine*, *36*(2), 29–36.

Cavallini, A., & Montanari, G. C. (2006). Effect of supply voltage frequency on testing of insulation system. *IEEE transactions on Dielectrics and Electrical Insulation*, *13*(1), 111–121.

Cranmer, K., Brehmer, J., & Louppe, G. (2020). The frontier of simulation-based inference. *Proceedings of the National Academy of Sciences*, *117*(48), 30055-30062.

Csilléry, K., François, O., & Blum, M. G. (2012). abc: an R package for approximate bayesian computation (abc). *Methods in ecology and evolution*, *3*(3), 475–479.

Danikas, M. G., Gao, N., & Aro, M. (2003). Partial discharge recognition using neural networks: A review. *Electrical Engineering*, *85*(2), 87–93.

Doucet, A., De Freitas, N., Gordo, & James, N. (2001). *Sequential monte carlo methods in practice* (Vol. 1) (No. 2). Springer.

Dutta, R., Schoengens, M., Pacchiardi, L., Ummadisingu, A., Widmer, N., Künzli, P., ... Mira, A. (2021). ABCpy: A high-performance computing perspective to approximate bayesian computation. *Journal of Statistical Software*, *100*(7), 1–38.

Heitz, C. (1999). A generalized model for partial discharge processes based on a stochastic process approach. *Journal of Physics D: Applied Physics*, *32*(9), 1012.

International Electrotechnical Commission (IEC). (2000). *High-voltage test techniques - partial discharge measurements* (Tech. Rep. Nos. 60270:2000+AMD1:2015 CSV). Geneva, Switzerland: IEC.

Krivda, A. (1995a). Automated recognition of partial discharges. *IEEE Transactions on Dielectrics and Electrical Insulation*, *2*(5), 796–821.

Krivda, A. (1995b). *Recognition of discharges: Discrimination and classification* (PhD thesis). Technical University Delft.

Lu, S., Chai, H., Sahoo, A., & Phung, B. T. (2020). Condition monitoring based on partial discharge diagnostics using machine learning methods: A comprehensive state-of-the-art review. *IEEE Transactions on Dielectrics and Electrical Insulation*, *27*(6), 1861–1888. doi: 10.1109/TDEI.2020.009070

Lv, Z., Rowland, S. M., Chen, S., Zheng, H., & Iddrissu, I. (2017). Evolution of partial discharges during early tree propagation in epoxy resin. *IEEE Transactions on Dielectrics and Electrical Insulation*, *24*(5), 2995-3003.

Ma, H., Chan, J. C., Saha, T. K., & Ekanayake, C. (2013). Pattern recognition techniques and their applications for automatic classification of artificial partial discharge sources. *IEEE Transactions on Dielectrics and Electrical Insulation*, *20*(2), 468–478.

Montanari, G. (1995). Aging and life models for insulation systems based on pd detection. *IEEE Transactions on Dielectrics and Electrical Insulation*, *2*(4), 667-675.

Niemeyer, L. (1995). A generalized approach to partial discharge modeling. *IEEE transactions on Dielectrics and Electrical insulation*, *2*(4), 510–528.

Patsch, R., & Berton, F. (2002). Pulse Sequence Analysis - a diagnostic tool based on the physics behind partial discharges. *Journal of Physics D: Applied Physics*, *35*(1), 25–32.

Raymond, W. J. K., Illias, H. A., Bakar, A. H. A., & Mokhlis, H. (2015). Partial discharge classifications: Review of recent progress. *Measurement*, *68*, 164–181. doi: 10.1016/j.measurement.2015.02.032

Sahoo, N., Salama, M., & Bartnikas, R. (2005). Trends in partial discharge pattern classification: A survey. *IEEE Transactions on Dielectrics and Electrical Insulation*, *12*(2), 248–264. doi: 10.1109/TDEI.2005.1430395

Sisson, S. A., Fan, Y., & Beaumont, M. A. (Eds.). (2019). *Handbook of approximate bayesian computation*. CRC Press.

Tavaré, S., Balding, D. J., Griffiths, R. C., & Donnelly, P. (1997, 02). Inferring Coalescence Times From DNA Sequence Data. *Genetics*, *145*(2), 505-518.

Toni, T. (2011). Abc smc for parameter estimation and model selection with applications in systems biology. *Nature Precedings*, 1–1.

Toni, T., Welch, D., Strelkowa, N., Ipsen, A., & Stumpf, M. P. H. (2009). Approximate bayesian computation scheme for parameter inference and model selection in dynamical systems. *Journal of Royal Society Interface*, *6*, 187–202.

Wang, L., Cavallini, A., Montanari, G. C., & Testa, L. (2010). Patterns of partial discharge activity in xlpe: From inception to breakdown. In *2010 10th ieee international conference on solid dielectrics* (p. 1-4).

BIOGRAPHIES

Kai Hencken is a senior principal scientist for "Physical and Statistical Modeling" at the ABB Corporate Research Center, Baden-Dättwil, Switzerland. He received the Diploma in Physics (1990) and the Ph.D in Theoretical Physics (1994) from the U Basel. He was a post-doc at the Institute for Nuclear Theory, U Washington, Seattle, USA from 1995 to 1997 and at the University of Basel from

1997 until 2005, where he received his Habilitation in 2000 and is a lecturer since. In 2005 he joined the theoretical physics group at the ABB Corporate Research Center. His research interests are the combination of physical modeling with advanced statistical methods to solve problems related to industrial devices. A strong focus is on developing diagnostics and prognostics approaches.

Daniele Ceccarelli is a Data Scientist at Ospedale San Raffaele, Milano, Italy. He receiced the Bachelor in Mathematical Engineering in 2018 and the Master in Mathematical Engineering in 2021 from the Politecnico di Milano, Milano, Italy. In 2021 he worked as Intern at the ABB Corporate Research Center, Baden-Dättwil, Switzerland. His research interests are Bayesian statistics, Data Science and Physics-Informed Machine Learning.

Elsi-Mari Borrelli is a Lead Quantum Biology Researcher at Algorithmiq since 2022. She received the Master in Physics in 2009 and a Ph.D. in Theoretical Physics in 2013 from the University of Turku, Turku, Finland. She was a post-doc at Aalto University, Aalto, Finland from 2013 to 2015 and at the University of Turku, Turku, Finland from 2015-2017. From 2017 to 2021 she worked at the ABB Corporate Research Center, Baden-Dättwil, Switzerland. Her research interests range from advanced analytics methods for industrial devices to application of quantum algorithms for complex optimization problems.

Andrej Krivda is a principal scientist at the ABB Corporate Resarch Center, Baden-Dättwil, Switzerland. He received the M.Sc. degree from the Kosice University of Technology, Kosice, Slovakia, and the Ph.D. degree from the Delft University of Technology, Delft, The Netherlands. He was with the Royal Institute of Technology, Stockholm, Sweden, for one year, where he worked on automated recognition of partial discharge patterns in electrical power generators. He was four years with the Queensland University of Technology, Brisbane, Australia, where he studied outdoor insulation for overhead power lines. Since 2001, he has been with ABB Switzerland Ltd., Corporate Research, Baden-Dättwil, Switzerland. His current research interests include partial discharges, dielectric spectroscopy, outdoor insulation and intelligent systems for recognition of digital partial discharge patterns. Dr.Krivda was awarded the Power Engineering Journal and the Journal of Electrical and Electronics Engineering, Australia, premiums for his work on overhead insulated mains in 2000.