A General Method for Calibrating Stochastic Radio Channel Models with Kernels

Bharti, Ayush; Briol, François-Xavier; Pedersen, Troels

Published in:
Fourth Workshop on Machine Learning and the Physical Sciences (NeurIPS 2021)

Publication date:
2022

Document Version
Accepted author manuscript, peer reviewed version

A General Method for Calibrating Stochastic Radio Channel Models with Kernels

Ayush Bharti

Department of Computer Science
Aalto University, Finland
ayush.bharti@aalto.fi

François-Xavier Briol

Department of Statistical Science
University College London, UK
f.briol@ucl.ac.uk

Troels Pedersen

Department of Electronic Systems
Aalborg University, Denmark
troels@aau.dk

Abstract

Characterization of the environment in which communication is taking place, termed the *radio channel*, is imperative for the design and analysis of communication systems. Stochastic models of the radio channel are widely used simulation tools that construct a probabilistic model of the radio channel. Calibrating these models to new measurement data is challenging when the likelihood function is intractable. The standard approach to this problem involves sophisticated algorithms for extraction and clustering of multipath components, following which, point estimates of the model parameters can be obtained using specialized estimators. We propose instead an approximate Bayesian computation algorithm based on the maximum mean discrepancy with a kernel careful crafted for this task. The proposed method is able to estimate the parameters of the model accurately in simulations, and has the advantage that it can be used on a wide range of models.

1 Introduction

Stochastic channel models are used to simulate the behavior of the radio channel in order to test the performance of communication and localization systems [22, 25]. Often models are flexible enough to be applied to different scenarios, provided that their parameters can be adjusted accordingly. Adjustment of the model parameters based on data collected from measurement campaigns is called *calibration* (or inference). Calibration is usually challenging since most state-of-the-art stochastic radio channel models have intractable likelihood functions. This renders usual inference techniques such as maximum likelihood estimation or standard Bayesian inference inapplicable. Calibration has been a significant technical challenge to date because existing methods usually depend on the specific model being calibrated. As a result, a new calibration method needs to be developed every time a new model is proposed, which significantly slows down the innovation process.

In this paper, we propose an algorithm based on *approximate Bayesian computation (ABC)* which is able to calibrate models with vastly different mathematical structures. We use the maximum mean discrepancy (MMD) [12] in the ABC framework to compare the distribution of simulated and measured data. The MMD has previously been used for frequentist inference in [8, 9], and in a

Fourth Workshop on Machine Learning and the Physical Sciences (NeurIPS 2021).

Bayesian sense in [10]. Specific ABC methods using kernels include [18, 19, 16, 14], and the MMD has also been used to train generative adversarial networks in [11, 23, 15]. These papers have shown MMD to be a powerful way to represent either data-sets or distributions, and as a result calibrate complex models. They have acted as inspiration for our work, but our algorithm specializes the approach to the problem of calibrating stochastic channel models.

2 Background

Stochastic Channel Model Calibration In this paper, a model is a parametric family of distributions $\{\mathbb{P}_{\boldsymbol{\theta}}\}$ with a p-dimensional parameter vector $\boldsymbol{\theta}$ defined on some Euclidean space. In the case of generative models such as the stochastic channel models, it is straightforward to simulate realizations of $\mathbf{Y}$ from the model, even though the distribution $\mathbb{P}_{\boldsymbol{\theta}}$ is unknown. Calibration then amounts to finding the parameter $\boldsymbol{\theta}$ for which the model output fits the observed data $\mathbf{Y}$ well. For most stochastic radio channel models, $p(\mathbf{Y}|\boldsymbol{\theta})$ is either intractable or cannot be approximated within reasonable computation time. However, likelihood-free inference is possible by comparing simulated data-sets to the observed data. We use the maximum mean discrepancy (MMD) to perform this comparison.

The Maximum Mean Discrepancy (MMD) The MMD is a distance between distributions. To compute it, we first map the distributions to a function space $\mathcal{H}_k$, then use the distance in that space to compare the mapped distributions. The spaces of functions to which we will map distributions are called reproducing kernel Hilbert space (RKHS). We denote the RKHS with $\mathcal{H}_k$, and $\langle \cdot, \cdot \rangle_{\mathcal{H}_k}$ and $\| \cdot \|_{\mathcal{H}_k}$ for its inner product and norm, respectively. Associated to each RKHS, there exists a symmetric and positive definite function $k : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ called a reproducing kernel [4]. This function satisfies two properties: (i) for all $f \in \mathcal{H}_k$, $f(\mathbf{x}) = \langle f, k(\mathbf{x}, \cdot) \rangle_{\mathcal{H}_k}$ (called the reproducing property), and (ii) $k(\mathbf{x}, \cdot) \in \mathcal{H}_k$ for all $\mathbf{x} \in \mathbb{R}^d$. The MMD between $\mathbb{P}$ and $\mathbb{Q}$ embedded in $\mathcal{H}_k$ is defined as the supremum taken over the mean of all functions in the unit ball in an RKHS, i.e. [17]

$$\mathrm{MMD}_k[\mathbb{P}, \mathbb{Q}] = \sup_{\|f\|_{\mathcal{H}_k} \leq 1} \left| \mathbb{E}_{X \sim \mathbb{P}}[f(X)] - \mathbb{E}_{X \sim \mathbb{Q}}[f(X)] \right|. \tag{1}$$

Suppose we have access to $\mathbf{X} = \{\mathbf{x}_1, \ldots, \mathbf{x}_{N_X}\} \overset{iid}{\sim} \mathbb{P}$ and $\mathbf{Y} = \{\mathbf{y}_1, \ldots, \mathbf{y}_{N_Y}\} \overset{iid}{\sim} \mathbb{Q}$. Then, an unbiased empirical estimate of $\mathrm{MMD}_k^2[\mathbb{P}, \mathbb{Q}]$ can be obtained as [12]

$$\widehat{\mathrm{MMD}}_k^2[\mathbf{X}, \mathbf{Y}] = \frac{\sum_{i \neq i'} k(\mathbf{x}_i, \mathbf{x}_{i'})}{N_X(N_X - 1)} - \frac{2 \sum_{j=1}^{N_Y} \sum_{i=1}^{N_X} k(\mathbf{x}_i, \mathbf{y}_j)}{N_Y N_X} + \frac{\sum_{j \neq j'} k(\mathbf{y}_j, \mathbf{y}_{j'})}{N_Y(N_Y - 1)}. \tag{2}$$

Note that the choice of kernel k will be key to determine whether the MMD is a suitable probability metric for a given application.

3 Kernels for Radio Channel Measurements

We now construct a kernel for a type of time-series data frequently available from channel measurements. Consider data from a linear, time-invariant radio channel, measured using a vector network analyzer (VNA) in the bandwidth B. The transfer function H_k is measured at K equidistant frequency points resulting in a frequency separation of $\Delta f = B/(K-1)$. The measured signal at each frequency point, Y_n, can be modeled as $Y_n = H_n + W_n$, for $n = 0, 1, \cdots, N_s - 1$, where H_n is the transfer function sampled at the n^{th} frequency and W_n is the complex measurement noise. The additive noise samples are assumed independent and identically distributed (iid) at each frequency point, and are usually modeled as zero-mean circular symmetric complex Gaussian variables with variance σ_W^2.

The time-domain signal, $y(t)$, is obtained by taking the discrete-frequency, continuous-time inverse Fourier transform of Y_n as $y(t) = \frac{1}{N_s} \sum_{n=0}^{N_s-1} Y_n \exp(j 2 \pi n \Delta f t)$, periodic with a period of $t_{\max} = 1/\Delta f$. Multiple realizations of the channel can be obtained by repeating the measurements N_{obs} times, yielding an $N_{\text{obs}} \times N_s$ complex data matrix $\mathbf{Y}$. The data consists of iid realizations from some unknown distribution $\mathbb{Y}$ with state-space $\mathcal{Y}$, where each realisation is a time-series of size N_s.

In order to use the MMD for calibrating stochastic radio channel models, we need a kernel defined on the space of transfer function measurements: $k_{\mathcal{Y}} : \mathcal{Y} \times \mathcal{Y} \to \mathbb{R}$. A significant challenge with this approach is that $\mathcal{Y}$ is usually a high-dimensional space; in fact N_s can be in the order of several

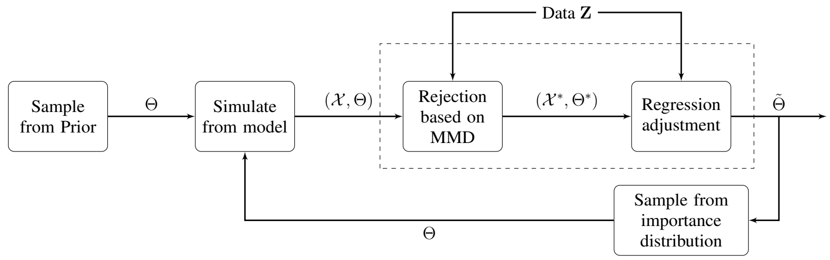

Figure 1: Diagram depicting steps in the proposed kernel-based ABC algorithm

thousands. To tackle this issue, we construct a kernel specifically tailored to transfer function measurements. We base the kernel on the temporal moments of $y(t)$, defined as

$$m^{(i)} = \int_0^{t_{\max}} t^i |y(t)|^2 \mathrm{d}t, \quad i = 0, 1, 2, \dots, I. \tag{3}$$

Referring to [7, 5], the first few moments are well modeled by a log-normal distribution. Thus, taking the entry-wise logarithm $z^{(i)} = \ln m^{(i)}$ brings the moments to the same scale and gives an approximately Gaussian vector $\mathbf{z} = [z^{(0)}, \dots, z^{(I-1)}]$. Alternatively, an anisotropic kernel could be used to account for the difference of scales of the temporal moments. Multiple channel realizations yield $\mathbf{Z} = (\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_{N_{\text{obs}}})$. Define the mapping $A_I : \mathcal{Y} \to \mathbb{R}^I$ from $\mathcal{Y}$ to the I-dimensional space of log temporal moments. We propose to construct a kernel $k_{\mathcal{Y}}$ for transfer function data as

$$k_{\mathcal{Y}}(\mathbf{y}, \mathbf{y}') := k_{\text{SE}}(A_I(\mathbf{y}), A_I(\mathbf{y}')), \quad \text{for all } \mathbf{y}, \mathbf{y}' \in \mathcal{Y}, \tag{4}$$

where $k_{\text{SE}}(\mathbf{x}, \mathbf{x}') = \exp(-\|\mathbf{x} - \mathbf{x}'\|_2^2 / l^2)$ is the squared-exponential kernel in dimension I. Given this choice, we have $\widehat{\text{MMD}}^2_{k_{\mathcal{Y}}}[\mathbf{Y}, \mathbf{Y}_{\text{sim}}] = \widehat{\text{MMD}}^2_{k_{\text{SE}}}[\mathbf{Z}, \mathbf{X}]$, where $\mathbf{X}$ is the simulated log temporal moments data-set. Following [12], we propose to set the lengthscale using the median heuristic $l = \sqrt{\text{med}/2}$, where med denotes the median of the set of squared two-norm distances $\|\mathbf{x}_i - \mathbf{x}_j\|_2^2$ for all pairs of distinct data points in $\mathbf{X}$. This choice of l scales the kernel with the spread of the data, and is robust to outliers.

4 Kernel-based Approximate Bayesian Computation Method

With this kernel at hand, we propose to calibrate a stochastic radio channel model using the ABC approach based on MMD depicted in Fig. 1[1]. We employ the Population Monte Carlo (PMC) ABC method [2] to iteratively refine our approximation of the ABC posterior. At the end of each iteration, we perform local-linear regression adjustment [3] to further improve the posterior approximation.

First, M independent parameter samples $\Theta = (\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_M)$ are drawn from the prior $p(\boldsymbol{\theta})$. For each $i \in \{1, \dots, M\}$, we then simulate a data-set $\mathbf{X}_i \sim \mathbb{P}_{\boldsymbol{\theta}_i}$ and compute the MMD (based on the kernel in (4)) between $\mathbf{X}_i$ and $\mathbf{Z}$. The parameter samples resulting in the M_ϵ smallest MMD values are then accepted and the rest is discarded.

The accepted samples are then adjusted using local-linear regression [3] to improve the posterior approximation. Given the accepted set $\{(\mathbf{s}_i, \boldsymbol{\theta}_i)\}_{i=1}^{M_\epsilon}$ where $\mathbf{s}$ is a vector of summary statistics of $\mathbf{X}$ (in our case, we take the temporal moments), the i^{th} accepted parameter sample is adjusted as $\tilde{\boldsymbol{\theta}}_i = \boldsymbol{\theta}_i^* - (\mathbf{s}_i - \mathbf{s}_{\text{obs}})^\top \hat{\boldsymbol{\beta}}$ where

$$\hat{\boldsymbol{\beta}} := \arg\min_{\boldsymbol{\alpha}, \boldsymbol{\beta}} \sum_{i=1}^{M_\epsilon} \left[\boldsymbol{\theta}_i^* - \boldsymbol{\alpha} - (\mathbf{s}_i - \mathbf{s}_{\text{obs}})^\top \boldsymbol{\beta} \right]^2 \mathcal{W}_{\left(\widehat{\text{MMD}}^2_{k_{\text{SE}}}[\mathbf{X}_i, \mathbf{Z}]\right)}.$$

Here, $\mathcal{W}$ is the Epanechnikov function, $\mathcal{W}_{(\delta)} = 1 - (\delta/\delta_{\max})^2$ for $|\delta| \le \delta_{\max}$ and zero otherwise. The regression adjustment therefore gives the adjusted parameter values $\tilde{\Theta} = (\tilde{\boldsymbol{\theta}}_1, \dots, \tilde{\boldsymbol{\theta}}_{M_\epsilon})$.

The ABC method then draws a new set of M parameter samples from $(\tilde{\boldsymbol{\theta}}_1, \dots, \tilde{\boldsymbol{\theta}}_{M_\epsilon})$ in a sequential Monte Carlo fashion [2]. These new samples form the prior distribution for the next iteration of

[1] The source code is available at `https://github.com/bharti-ayush/Kernel-based-ABC`.

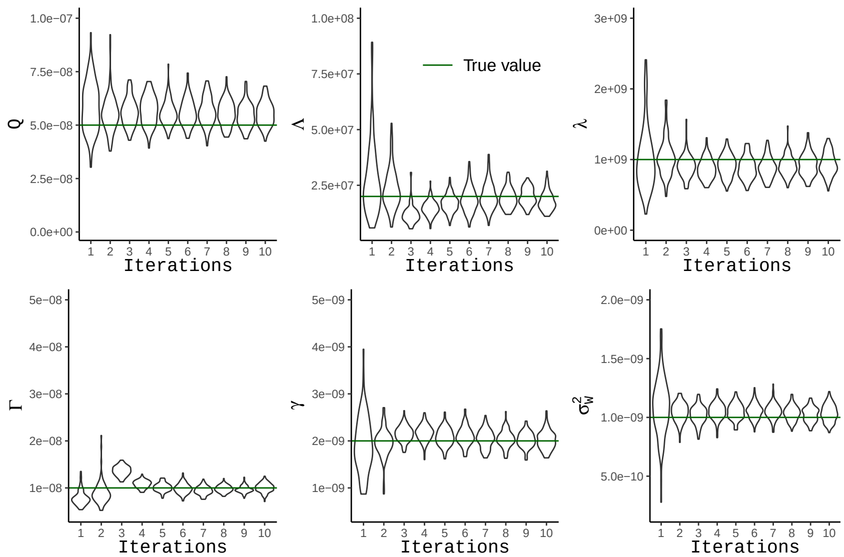

Figure 2: Violin plots of ABC posterior samples of Saleh-Valenzuela model parameters as a function of PMC iterations. Note that a violin plot is similar to a box plot with the addition of a rotated kernel density plot on each side. The dark green lines denote the true parameter values $\boldsymbol{\theta}_{\text{true}} = [5 \times 10^{-8}, 2 \times 10^{7}, 10^{9}, 10^{-8}, 2 \times 10^{-9}, 10^{-9}]^\top$. Settings: $B = 4$ GHz, $N_s = 801$, and $t_{\max} = 200$ ns, $I = 4$, $M = 2000$, $M_\epsilon = 100$.

the algorithm, where they are used to generate simulated data from the model again and perform regression adjustment. The set of parameters in the initial iteration, $\tilde{\Theta}^{(1)} = (\tilde{\boldsymbol{\theta}}_1^{(1)}, \dots, \tilde{\boldsymbol{\theta}}_{M_\epsilon}^{(1)})$, are assigned equal weights. The next set of parameters is obtained by drawing M values from $\tilde{\Theta}^{(1)}$ and perturbing these according to the proposal $\varphi(\boldsymbol{\theta}; \tilde{\boldsymbol{\theta}}, \boldsymbol{\Sigma}) = \mathbb{1}(\boldsymbol{\theta} \in \mathcal{R}) \exp(-\frac{1}{2}(\boldsymbol{\theta} - \tilde{\boldsymbol{\theta}})^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta} - \tilde{\boldsymbol{\theta}}))$ where $\mathbb{1}$ is an indicator function, $\mathcal{R} \subset \mathbb{R}^p$ is the prior range, and $\boldsymbol{\Sigma}$ is a diagonal matrix with variances $\sigma_j^2 > 0$ corresponding to parameter θ_j along the diagonal. We set the diagonal elements of $\boldsymbol{\Sigma}$ to twice the empirical variance of the adjusted parameter samples.

At iteration t, the set of samples $\Theta^{(t)}$ is used to simulate $(\mathbf{X}_1^{(t)}, \dots, \mathbf{X}_M^{(t)})$ from the model for MMD computation and regression adjustment. Here, the weight of each parameter sample is selected as

$$w_j^{(t)} \propto p\left(\boldsymbol{\theta}_j^{(t)}\right) \Big/ \sum_{i=1}^{M_\epsilon} w_i^{(t-1)} \varphi\left(\boldsymbol{\theta}_j^{(t)}; \tilde{\boldsymbol{\theta}}_i^{(t-1)}, \boldsymbol{\Sigma}^{(t-1)}\right), \qquad \text{for } j = 1, \dots, M_\epsilon. \tag{5}$$

The adjusted parameter values after iteration T are now samples from the ABC posterior.

5 Simulation Experiment

To study the performance of this algorithm, we focus on the model of Saleh and Valenzuela [21] who defined the transfer function as $H_k = \sum_l \sum_p \beta_{pl} \exp\left(-j2\pi\Delta f k(T_l + \tau_{pl})\right)$, where T_l is the delay of the l^{th} cluster, while τ_{pl} and β_{pl} are the delay and complex gain of the p^{th} ray within the l^{th} cluster, respectively. By definition in [21], $T_0 = 0$ and $\tau_{0l} = 0$, $l \in \{0, 1, \dots\}$. The arrival time of the clusters and that of the rays within the clusters are modelled as one-dimensional homogeneous Poisson point processes, i.e., $T_l \sim \text{PPP}(\mathbb{R}_+, \Lambda)$ and $\tau_{kl} \sim \text{PPP}(\mathbb{R}_+, \lambda)$ with parameters $\Lambda, \lambda > 0$. The gains β_{kl}, conditioned on T_l and τ_{kl}, are modelled as iid zero-mean complex Gaussian random variables. Their conditional variance is modelled as $\mathbb{E}\left[|\beta_{kl}|^2 | T_l, \tau_{kl}\right] = Q \exp(-T_l/\Gamma) \exp(-\tau_{kl}/\gamma)$, with Q being the average power of the first arriving multipath component, and $\Gamma, \gamma > 0$ being the cluster and ray decay constants, respectively. The expression for the power delay spectrum is given in [13]. To calibrate this model, the parameter vector, $\theta = [Q, \Lambda, \lambda, \Gamma, \gamma]^\top$, should be estimated based on $\mathbf{Y}$.

We generate pseudo-observed log moments, $\mathbf{X}_{\text{true}}$, with $N_{\text{obs}} = 1000$ realizations from the model by setting $\boldsymbol{\theta}$ to a “true" value. We now use the proposed method to calibrate the S-V model using $\mathbf{X}_{\text{true}}$ as observed data $\mathbf{Z}$. The number of samples in $\mathbf{X}$ is set to $N_{\text{sim}} = 100$ as a reasonable compromise considering the trade-off between accuracy and computational cost. We assume uninformative (flat) priors for all the parameters to ensure that their marginal posteriors are unaffected by any prior beliefs. The plots indicating convergence of the algorithm and the marginal posterior distributions for $T = 10$ iterations are shown in Fig. 2. The approximate posterior samples concentrate around the true value for all the parameters. The algorithm ran on a Lenovo ThinkPad with Intel Core i7 processor (24 GB RAM) and had a run-time of around 2 days, which can be improved using parallel computations. The algorithm converges rather quickly and the posteriors taper as the iterations proceed. In principle, the iterations could be stopped after four or five iterations, but we let it run till $T = 10$ for clarity. The algorithm gives a reasonable estimate for the parameters even in the first iteration. See [6] for further experiments on applying the method to another model and testing it on real data.

6 Conclusion

This paper proposed a novel approach for calibrating stochastic radio channel models. The approach is an ABC algorithm based on the MMD with a carefully crafted kernel for the type of data available in this field. The main advantage of the approach is that it is not model specific, and can be used to calibrate a wide range of models with different mathematical structures such as [20, 1] and [24]. This will need to be explored in more detail in future work.

References

[1] R. Adeogun, T. Pedersen, C. Gustafson, and F. Tufvesson. Polarimetric Wireless Indoor Channel Modelling Based on Propagation Graph. *IEEE Trans. on Antennas and Propag.*, 67(10):6585–6595, 2019.

[2] M. A. Beaumont, J.-M. Cornuet, J.-M. Marin, and C. P. Robert. Adaptive approximate bayesian computation. *Biometrika*, 96(4):983–990, Oct 2009.

[3] M. A. Beaumont, W. Zhang, and D. J. Balding. Approximate bayesian computation in population genetics. *Genetics*, 162(4):2025–2035, 2002.

[4] A. Berlinet and C. Thomas-Agnan. *Reproducing Kernel Hilbert Spaces in Probability and Statistics*. Springer Science+Business Media, New York, 2004.

[5] A. Bharti, R. Adeogun, X. Cai, W. Fan, F.-X. Briol, L. Clavier, and T. Pedersen. Joint modeling of received power, mean delay, and delay spread for wideband radio channels. *IEEE Trans. on Antennas and Propag.*, pages 1–1, 2021.

[6] A. Bharti, F.-X. Briol, and T. Pedersen. A general method for calibrating stochastic radio channel models with kernels. *IEEE Transactions on Antennas and Propagation*, pages 1–1, 2021.

[7] A. Bharti, L. Clavier, and T. Pedersen. Joint statistical modeling of received power, mean delay, and delay spread for indoor wideband radio channels. In *14th Eur. Conf. on Antennas and Propag.*, pages 1–5, 2020.

[8] F.-X. Briol, A. Barp, A. B. Duncan, and M. Girolami. Statistical inference for generative models with maximum mean discrepancy. *arXiv:1906.05944*, 2019.

[9] B.-E. Chérief-Abdellatif and P. Alquier. Finite sample properties of parametric MMD estimation: robustness to misspecification and dependence. *arXiv:1912.05737*, 2020.

[10] B.-E. Cherief-Abdellatif and P. Alquier. MMD-Bayes: Robust bayesian estimation via maximum mean discrepancy. In C. Zhang, F. Ruiz, T. Bui, A. B. Dieng, and D. Liang, editors, *Proc. of Mach. Learn. Res.*, volume 118, pages 1–21. PMLR, 08 Dec 2020.

[11] G. K. Dziugaite, D. M. Roy, and Z. Ghahramani. Training generative neural networks via maximum mean discrepancy optimization. In *Proc. of 31st Conf. on Uncertain. in Artif. Intell.*, pages 258–267, 2015.

[12] A. Gretton, K. Borgwardt, M. J. Rasch, and B. Scholkopf. A kernel two-sample test. *J. of Mach. Learn. Res.*, 13:723–773, 2012.

[13] M. L. Jakobsen, T. Pedersen, and B. H. Fleury. Analysis of the stochastic channel model by saleh & valenzuela via the theory of point processes. *Int. Zurich Seminar on Commun.*, 2012.

[14] K. Kisamori, M. Kanagawa, and K. Yamazaki. Simulator calibration under covariate shift with kernels. In *Proc.s of the 23rd Int. Conf. on Artif. Intell. and Statistics*, volume 108, pages 1244–1253. PMLR, Aug 2020.

[15] Y. Li, K. Swersky, and R. Zemel. Generative moment matching networks. In *Proc. of the 32nd Int. Conf. on Mach. Learn. - Vol. 37*, ICML'15, page 1718–1727. JMLR.org, 2015.

[16] J. Mitrovic, D. Sejdinovic, and Y. W. Teh. DR-ABC: Approximate Bayesian computation with kernel-based distribution regression. *33rd Int. Conf. on Mach. Learn., ICML*, 3:2209–2218, 2016.

[17] K. Muandet, K. Fukumizu, B. Sriperumbudur, and B. Schölkopf. Kernel mean embedding of distributions: A review and beyond. *Found. and Trends in Mach. Learn.*, 10(1-2):1–141, 2017.

[18] S. Nakagome, K. Fukumizu, and S. Mano. Kernel approximate Bayesian computation in population genetic inferences. *Stat. Appl. in Genet. and Mol. Biol.*, 12(6):667–678, 2013.

[19] M. Park, W. Jitkrittum, and D. Sejdinovic. K2-ABC: approximate Bayesian computation with kernel embeddings. *Proc. of the 19th Int. Conf. on Artif. Intell. and Statistics*, 51:398–407, 2015.

[20] T. Pedersen, G. Steinböck, and B. H. Fleury. Modeling of reverberant radio channels using propagation graphs. *IEEE Trans. Antennas Propag.*, 60(12):5978–5988, Dec 2012.

[21] A. A. M. Saleh and R. Valenzuela. A statistical model for indoor multipath propagation. *IEEE J. Sel. Areas Commun.*, 5(2):128–137, February 1987.

[22] S. Salous, V. D. Esposti, F. Fuschini, R. S. Thomae, R. Mueller, D. Dupleich, K. Haneda, J.-M. M. Garcia-Pardo, J. P. Garcia, D. P. Gaillot, S. Hur, and M. Nekovee. Millimeter-wave propagation: Characterization and modeling toward fifth-generation systems. [wireless corner]. *IEEE Antennas and Propagation Magazine*, 58(6):115–127, dec 2016.

[23] D. J. Sutherland, H.-Y. Tung, H. Strathmann, S. De, A. Ramdas, A. Smola, and A. Gretton. Generative models and model criticism via optimized maximum mean discrepancy. In *Int. Conf. on Learn. Represent.*, 2017.

[24] G. L. Turin, F. D. Clapp, T. L. Johnston, S. B. Fine, and D. Lavry. A statistical model of urban multipath propagation. *IEEE Trans. Veh. Technol.*, 21(1):1–9, Feb 1972.

[25] K. Witrisal, C. Antón-Haro, et al. Whitepaper on new localization methods for 5G wireless systems and the internet-of-things. Technical report, COST CA15104 (IRACON), 2018.