



Proceeding Paper

Bayesian Surrogate Analysis and Uncertainty Propagation [†]

Sascha Ranftl * and Wolfgang von der Linden

Institute of Theoretical Physics-Computational Physics, Graz University of Technology, Petersgasse 16, 8010 Graz, Austria; vonderlinden@tugraz.at

* Correspondence: ranftl@tugraz.at

† Presented at the 40th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering, online, 4–9 July 2021.

Abstract: The quantification of uncertainties of computer simulations due to input parameter uncertainties is paramount to assess a model's credibility. For computationally expensive simulations, this is often feasible only via surrogate models that are learned from a small set of simulation samples. The surrogate models are commonly chosen and deemed trustworthy based on heuristic measures, and substituted for the simulation in order to approximately propagate the simulation input uncertainties to the simulation output. In the process, the contribution of the uncertainties of the surrogate itself to the simulation output uncertainties is usually neglected. In this work, we specifically address the case of doubtful surrogate trustworthiness, i.e., non-negligible surrogate uncertainties. We find that Bayesian probability theory yields a natural measure of surrogate trustworthiness, and that surrogate uncertainties can easily be included in simulation output uncertainties. For a Gaussian likelihood for the simulation data, with unknown surrogate variance and given a generalized linear surrogate model, the resulting formulas reduce to simple matrix multiplications. The framework contains Polynomial Chaos Expansions as a special case, and is easily extended to Gaussian Process Regression. Additionally, we show a simple way to implicitly include spatio-temporal correlations. Lastly, we demonstrate a numerical example where surrogate uncertainties are in part negligible and in part non-negligible.

Keywords: uncertainty quantification; uncertainty propagation; surrogate models; meta-modeling; Bayesian



Citation: Ranftl, S.; von der Linden, W. Bayesian Surrogate Analysis and Uncertainty Propagation. *Phys. Sci. Forum* **2021**, *3*, 6. <https://doi.org/10.3390/psf2021003006>

Academic Editor: MaxEnt 2021 Scientific Committee

Published: 13 November 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Uncertainty quantification of simulations has gained increasing attention, e.g., in the field of Computational Engineering, in order to address doubtful parameter choices and assess the models' credibility. Surrogate models have become a popular tool to propagate simulation input uncertainties to the simulation output, particularly for modern day applications with high computational cost and many uncertain model parameters. For that, a parametrized surrogate model (synonyms: meta-model, emulator) is learned from a finite set of simulation samples. i.e., the surrogate is a function of the uncertain simulation input parameters that is 'fitted' to the simulation output data. The quality of this fit is then judged by heuristic diagnostics, and the surrogate deemed trustworthy, respectively. A key aspect of this procedure is that the surrogate can be evaluated much faster than the simulation, and still retains a reasonable approximation to the simulation.

The simulation is then substituted with the surrogate model in order to compute the marginal probability density function of the simulation output. The simulation uncertainties are thus inferred from the surrogate model instead of the original simulation model at a significantly reduced computational effort. While this practice allows for obtaining estimates on uncertainties of expensive simulations in the first place, the contribution of the uncertainty of the surrogate itself to the total simulation uncertainty is commonly neglected. In other words, the estimation of the surrogate parameters based on the finite set

of simulation samples entails an additional uncertainty in the sought-for uncertainty of the simulation output. The purpose of this paper is to investigate this surrogate uncertainty as the natural measure for the surrogate's trustworthiness, and how the surrogate uncertainty affects the simulation output uncertainty.

In many cases, the surrogate uncertainty is indeed small if the heuristic diagnostics naively imply so. If the heuristic diagnostics imply that the surrogate is not trustworthy, one may resort to two options: (i) Acquire more simulation data until the surrogate is trustworthy. This is limited by the computational budget and the surrogate's convergence properties. (ii) Shrink the parameter space, e.g., omit a number of uncertain simulation parameters by assuming definite parameter values. However, in some cases, (i) is not feasible and (ii) is not desired. In this contribution, we demonstrate how to include surrogate uncertainties if the user deals with a surrogate model with doubtful trustworthiness.

Popular surrogate models are Polynomial Chaos Expansions [1–3] and Gaussian Process Regression [4,5], the latter of which has had its renaissance recently from within the machine learning community. In this work, we assume a Gaussian likelihood for the simulation data with unknown variance and given a generalized linear surrogate model (i.e., linear in the surrogate parameters) that includes Polynomial Chaos Expansions as a special case and is easily extended to Gaussian Process Regression. Other Bayesian perspectives on Uncertainty Quantification of computer simulations with these popular surrogate models are given in [6–13]. A comprehensive collection of reviews on Uncertainty Quantification, from the point of view of computational engineering and applied mathematics, can be found in [14]. In [7,15], a statistician's perspective is discussed. Here, we will use Bayesian Probability Theory [16].

2. Bayesian Uncertainty Quantification

We start with the general structure of uncertainty propagation problems based on surrogate models in Section 2.1. In Section 2.2, we analyze a generalized linear surrogate model with a Gaussian likelihood for the simulation data with unknown surrogate variance. In Section 2.3, we proceed to use the surrogate model to propagate the input uncertainties to the output, and show how the surrogate's uncertainties too can be included.

2.1. General Structure of the Problem

The goal in this paper is to quantify the uncertainties of the simulation results for the observable $z^{(x)}$ at different measurement points $x = 1, \dots, N_x$ in the simulation domain. e.g., z could be the mechanical stress resulting from a structural analysis with a finite element simulation, where x could denote the location of the measurement probe in or on the analyzed structure. $z^{(x)}$ depends on unknown or uncertain model parameters $\mathbf{a} = \{a_i\}_{i=1}^{N_a}$, which are generally inferred from experimental data \mathbf{d}_{exp} . Based on these data, Bayes' theorem allows for determining the posterior probability density function (pdf) for \mathbf{a} ,

$$p(\mathbf{a} \mid \mathbf{d}_{\text{exp}}, \mathcal{I}), \quad (1)$$

where all background information on the experiment is subsumed in \mathcal{I} . The implications of the background information \mathcal{I} will be discussed later. This pdf will be assumed to be (almost) arbitrary but given in the following considerations. It usually is the result of a statistical data analysis of the foregoing experiment. This experiment could be the measurement of some material property needed for the simulation, e.g., viscosity for a computational fluid dynamics simulation. The uncertainty of the model parameters \mathbf{a} entails an uncertainty in the simulated observable $z^{(x)}$, and the latter is determined by the marginalization rule,

$$p(z^{(x)} \mid \mathbf{d}_{\text{exp}}, \mathcal{I}) = \int p(z^{(x)} \mid \mathbf{a}, \mathbf{d}_{\text{exp}}, \mathcal{I}) p(\mathbf{a} \mid \mathbf{d}_{\text{exp}}, \mathcal{I}) dV_{\mathbf{a}}. \quad (2)$$

In the first pdf, we have struck out d_{exp} because the knowledge of a suffices to perform the simulation to obtain $z^{(x)}$. If a consists of only one or two parameters ($N_a = 1, 2$), then the numerical evaluation of the integral over the model parameters a will typically require a few dozens to hundreds of simulations. The uncertainty propagation is then done and no surrogate is needed. However, this is the trivial case, and usually a will consist of way more parameters. Let us assume that a consists of, e.g., four parameters. That would imply the need for performing simulations at least 10^5 times, which is way too CPU expensive for most real problems. This can be avoided if the simulations are replaced by a surrogate model that approximates the observable z by a suitable parametrized surrogate function $z_{\text{sur}} = g(a|c)$, where c are yet unknown parameters. The simulation may yield the observable $z^{(x)}$ at different sites x in the domain; however, x could also denote the time-instance in non-static problems. Clearly, the parameters of the surrogate model will also depend on those positions. Thus, we actually have

$$z^{(x)} \approx z_{\text{sur}}^{(x)} = g(a|c^{(x)}) . \tag{3}$$

The unknown parameters will be inferred from a suitable training data. To this end, simulations are performed for a finite set of model parameters $A_s = \{a_s^{(i)}\}_{i=1}^{N_s}$ and the corresponding observables $Z_s = \{z_s^{(x),(i)}\}_{i,x=1}^{N_s, N_x}$ are computed and combined in $D_{\text{sim}} = \{A_s, Z_s\}$. The surrogate parameters $c^{(x)}$ are then inferred from D_{sim} , and the surrogate is thus constructed. We now proceed to substitute the simulation for the surrogate, $z^{(x)} \rightarrow z_{\text{sur}}^{(x)}$, in order to solve Equation (2) at a significantly reduced computational cost. This implies that the background information has changed. We will denote this as $\tilde{\mathcal{I}}$, suggesting that we take the observable z entering the integral in Equation (2) from the surrogate model Equation (3) rather than from the expensive simulation. More precisely, instead of Equation (2), we now need to consider

$$p(z^{(x)} | d_{\text{exp}}, D_{\text{sim}}, \tilde{\mathcal{I}}) = \int dV_a p(z^{(x)} | a, D_{\text{sim}}, \cancel{d_{\text{exp}}}, \tilde{\mathcal{I}}) p(a | \cancel{D_{\text{sim}}}, d_{\text{exp}}, \tilde{\mathcal{I}}) . \tag{4}$$

As far as the (second) pdf for the model parameters is concerned, we can omit the information on the training set D_{sim} , as it does not tell us anything about the model parameters. This pdf is actually the same as that in Equation (1), i.e., $p(a | d_{\text{exp}}, \tilde{\mathcal{I}}) = p(a | d_{\text{exp}}, \mathcal{I})$, as it makes no difference for the model parameters how we solve the equations underlying the simulation. In the first pdf, we can omit the information on the experiment d_{exp} , as we only need the simulation data D_{sim} to fix the surrogate model, which in turn defines the observable z . The first pdf can be further specified by the marginalization rule upon introducing the surrogate parameters $C = \{c^{(x)}\}_{x=1}^{N_x}$, where x is an index denoting the measurement points as introduced for $z^{(x)}$,

$$p(z^{(x)} | a, D_{\text{sim}}, \tilde{\mathcal{I}}) = \int dV_C p(z^{(x)} | C, a, \cancel{D_{\text{sim}}}, \tilde{\mathcal{I}}) p(C | \cancel{a}, D_{\text{sim}}, \tilde{\mathcal{I}}) .$$

The first pdf is uniquely fixed by the knowledge of a and C , hence D_{sim} is superfluous. Similarly, in the second pdf, where C is inferred from the training data, additional model parameters without the corresponding observables' values z , are useless. In summary, substituting the latter equation into Equation (4), we have

$$p(z^{(x)} | d_{\text{exp}}, D_{\text{sim}}, \tilde{\mathcal{I}}) = \iint dV_a dV_C p(z^{(x)} | C, a, \tilde{\mathcal{I}}) p(C | D_{\text{sim}}, \tilde{\mathcal{I}}) p(a | d_{\text{exp}}, \mathcal{I}) . \tag{5}$$

The first pdf is rather simple. According to the background information $\tilde{\mathcal{I}}$, we will determine the observable via the surrogate model. Since the necessary parameters $\mathbf{c}^{(x)} \in \mathbf{C}$ and \mathbf{a} are part of the conditional complex, the surrogate model allows only one value

$$z_{\text{sur}}^{(x)} = g(\mathbf{a} \mid \mathbf{c}^{(x)}) .$$

for the observable. This means that $p(\mathbf{z}^{(x)} \mid \mathbf{C}, \mathbf{a}, \tilde{\mathcal{I}})$ is equivalent to the probability density function for $\mathbf{z}^{(x)}$ given $\mathbf{z}^{(x)} = g(\mathbf{a} \mid \mathbf{c}^{(x)})$. Hence, the pdf is a Dirac-delta distribution

$$p(\mathbf{z}^{(x)} \mid \mathbf{C}, \mathbf{a}, \tilde{\mathcal{I}}) = \delta[\mathbf{z}^{(x)} - g(\mathbf{a} \mid \mathbf{c}^{(x)})] . \tag{6}$$

Finally, we have

$$p(\mathbf{z}^{(x)} \mid \mathbf{d}_{\text{exp}}, \mathbf{D}_{\text{sim}}, \tilde{\mathcal{I}}) = \iint dV_{\mathbf{a}} dV_{\mathbf{C}} \delta[\mathbf{z}^{(x)} - g(\mathbf{a} \mid \mathbf{c}^{(x)})] p(\mathbf{C} \mid \mathbf{D}_{\text{sim}}, \tilde{\mathcal{I}}) p(\mathbf{a} \mid \mathbf{d}_{\text{exp}}, \mathcal{I}) . \tag{7}$$

Before we can evaluate this integral, we first need to determine the two terms, which have their own independent significance. The last term is the result of a data analysis of a specific foregoing experiment, and will therefore not be treated here. We will suppress the background information in the following, as ambiguities should no longer occur.

2.2. Bayesian Analysis and Selection of the Surrogate Model

We recall that Equation (7) allows for determining the pdf for the observable based on the pdf for the model parameters, and the pdf for the parameters of the surrogate model $p(\mathbf{C} \mid \mathbf{D}_{\text{sim}})$ that we will determine now. To this end, we have to specify the form of the surrogate model. We use the expansion

$$z_{\text{sur}} = \sum_{\nu=1}^{N_p} c_{\nu} \Phi_{\nu}(\mathbf{a}) , \tag{8}$$

in terms of basis functions $\Phi_{\nu}(\mathbf{a})$ and expansion coefficients c_{ν} . No further specification is needed at this point. Without loss of generality, we will use multi-variate Legendre polynomials for the numerical examples. This expansion is similar to the frequently used generalized Polynomial Chaos Expansion [1], where the polynomials $\Phi_{\nu}(\mathbf{a})$ are orthogonal with respect to the L^2 inner product with the prior of \mathbf{a} , $p(\mathbf{a})$, as an integration measure. However, here we actually consider a posterior $p(\mathbf{a} \mid \mathbf{d}_{\text{exp}})$ that generally has no standard form, for which no standard orthogonal polynomial chaos basis is known, and for which conditional independence of the model parameters \mathbf{a} does not hold. Polynomial Chaos Expansions have been extended to arbitrary probability measures [17] and dependent parameters [18], but, in the present context; however, these polynomials are not of primary interest and would only complicate the numerical evaluation. Note that the approach presented here does not demand any orthogonality properties for the basis, and thus avoid the practical problems encountered with the construction of such orthogonal bases. Parameters here may have complex dependence structures, and the only requirement for the probability distribution $p(\mathbf{a} \mid \mathbf{d}_{\text{exp}})$ is that the integrals with respect to it exist. As outlined in Section 2.1, N_s simulations are performed for a set of model parameters $\mathbf{A}_s = \{\mathbf{a}_s^{(i)}\}_{i=1}^{N_s}$ and the corresponding observables \mathbf{Z}_s are computed. The theory is so far agnostic to the experimental design of these simulations, and it is therefore not of concern here. Now, we want to determine the pdf for the surrogate parameters \mathbf{C} , which are combined in a matrix with elements $C_{\nu,x}$, where ν enumerates the surrogate basis functions and x enumerates the measurement positions in the domain for which the observables are computed. We abbreviated the simulation data by the quantity $\mathbf{D}_{\text{sim}} = \{\mathbf{A}_s, \mathbf{Z}_s\}$, where the matrix \mathbf{Z}_s has the elements $(\mathbf{Z}_s)_{i,x}$, which represent the observable $z^{(x)}$ at position

x corresponding to the model parameter vector $\mathbf{a}_s^{(i)}$. The sought-for pdf follows from Bayes' theorem

$$p(\mathbf{C} | \mathbf{D}_{\text{sim}}) = p(\mathbf{C} | \mathbf{Z}_s, \mathbf{A}_s) \propto p(\mathbf{Z}_s | \mathbf{C}, \mathbf{A}_s) p(\mathbf{C} | \mathbf{A}_s) \propto p(\mathbf{Z}_s | \mathbf{C}, \mathbf{A}_s).$$

The proportionality constant is not required in the ensuing considerations, and we have assumed an ignorant, uniform prior for the coefficients c , i.e., $p(\mathbf{C} | \mathcal{I}) = \text{const}$. We note that this is also the transformation invariant Riemann prior (see Appendix B). However, any prior that is conjugate to the likelihood will retain analytical tractability. For the likelihood, we need the total misfit, which is given by

$$\begin{aligned} \chi^2 &= \sum_{i=1}^{N_s} \sum_{x=1}^{N_x} \left((\mathbf{Z}_s)_{i,x} - \sum_{v=1}^{N_p} (M_s)_{i,v} (\mathbf{C})_{v,x} \right)^2 = \sum_{i,x} \left(\mathbf{Z}_s - M_s \mathbf{C} \right)_{i,x}^2 \\ &= \text{tr} \left\{ (\mathbf{Z}_s - M_s \mathbf{C})^T (\mathbf{Z}_s - M_s \mathbf{C}) \right\}, \end{aligned} \tag{9}$$

with $(M_s)_{i,v} = \Phi_v(\mathbf{a}_s^{(i)})$ and $N_{sx} = N_s \cdot N_x$. We assume a Gaussian type of likelihood, i.e.,

$$p(\mathbf{Z}_s | \mathbf{C}, \mathbf{A}_s, \Delta) = \frac{\Delta^{-N_{sx}}}{Z'} \exp \left\{ -\frac{\chi^2}{2\Delta^2} \right\}. \tag{10}$$

with normalization Z' . We have mentioned the Δ -dependence of the normalization explicitly, while the rest of of the normalization is irrelevant in the present context. Usually, the misfit entering the likelihood comes from the noise of the data. In the present case, however, there is no noise (merely a tiny numerical error), but the surrogate model is presumably not an exact description of the simulation data and Δ covers the corresponding uncertainty. However, the uncertainty level Δ is not known and has to be marginalized over. Along with the appropriate Jeffreys' prior, $p(\mathbf{C}, \Delta) = p(\mathbf{C}) p(\Delta)$, $p(\Delta) = \frac{1}{\Delta}$, $p(\mathbf{C}) = \text{const}$. (see Appendix B), the integration over Δ yields

$$p(\mathbf{C} | \mathbf{Z}_s, \mathbf{A}_s) = \frac{1}{Z} (\chi^2)^{-\frac{N_{sx}}{2}}. \tag{11}$$

with terms independent of \mathbf{C} subsumed in the normalization Z . For computing the mean, variance, and evidence, we first complete the square in Equation (9) to get a quadratic form in \mathbf{C} , which can then be integrated analytically (see Appendix A). The result is

$$\langle \mathbf{C} \rangle_{\mathbf{a}} = H_s^{-1} M_s^T \mathbf{Z}_s, \quad H_s = M_s^T M_s, \tag{12a}$$

$$\langle \Delta C_{vx} \Delta C_{v'x'} \rangle_{\mathbf{a}} = \frac{\chi_{\min}^2}{(N_s - N_p) N_x - 2} (H_s^{-1})_{v,v'} \delta_{xx'}, \quad \chi_{\min}^2 = \text{tr} \left\{ \mathbf{Z}_s^T (\mathbf{1} - M_s H_s^{-1} M_s^T) \mathbf{Z}_s \right\}. \tag{12b}$$

We argue that the prefactor of H_s^{-1} is the Bayesian estimate for Δ^2 , the variance of the Gaussian in Equation (10). This reasoning is similar to [19]. Note that \mathbf{Z}_s is a matrix of size $N_s \times N_x$, containing the data vectors of length N_s for each measurement site x . As shown in Appendix A, Equation (A6), the evidence for a particular set of surrogate models is computed as

$$p(\{z_{\text{sur}}^{(x)}\}_{x=1}^{N_x} | \mathbf{D}_{\text{sim}}, \tilde{\mathcal{I}}) = Z = \Omega_{N_{px}} |H_s|^{-\frac{1}{2}} (\chi_{\min}^2)^{-\frac{N_{sx} - N_{px}}{2}} \frac{\Gamma(\frac{N_{px}}{2}) \Gamma(\frac{N_{sx} - N_{px}}{2})}{\Gamma(\frac{N_{sx}}{2})}, \tag{13}$$

where $\Omega_{N_{px}}$ is the solid angle in N_{px} dimensions. The evidence is the probability for a surrogate model given the data. Note that this quantity does not depend on $p(\mathbf{a} | \mathbf{d}_{\text{exp}}, \mathcal{I})$. This is reasonable because the analysis of the experimental data should be independent of the analysis of the simulation data. However, $p(\mathbf{a} | \mathbf{d}_{\text{exp}}, \mathcal{I})$ will typically be used for

the experimental design of the simulation data acquisition. By comparing the evidence for different models, the user can choose a particular surrogate model or, if the results do not overwhelmingly suggest one single model, average the results for the surrogate analysis and the following uncertainty propagation over several plausible models. Note that the evidence is the pillar of a Bayesian procedure to select a surrogate model, and is distinct from the procedure of incorporating the trustworthiness or uncertainty of the surrogate in the subsequent uncertainty propagation.

2.3. Bayesian Uncertainty Propagation with Surrogate Models

Now that we have selected the surrogate model and determined the ingredients of Equation (7), we can determine the pdf for the observables in the light of the experimental data and the simulation results of the training set. The form in Equation (5) allows an easy evaluation of the mean value by using Equation (12) (see also Equation (A3)),

$$\begin{aligned} \langle z^{(x)} \rangle &= \iint dV_a dV_C f(\mathbf{a} | \mathbf{c}^{(x)}) p(\mathbf{C} | D_{sim}, \tilde{\mathcal{I}}) p(\mathbf{a} | \mathbf{d}_{exp}, \mathcal{I}) \\ &= \sum_v \int dV_a \Phi_v(\mathbf{a}) \langle C_{vx} \rangle_a p(\mathbf{a} | \mathbf{d}_{exp}, \mathcal{I}) . \end{aligned} \tag{14}$$

Similarly, we obtain (see Equations (12) and (A8))

$$\begin{aligned} \langle z^{(x)} z^{(x')} \rangle &= \iint dV_a dV_C f(\mathbf{a} | \mathbf{c}^{(x)}) f(\mathbf{a} | \mathbf{c}^{(x')}) p(\mathbf{C} | D_{sim}, \tilde{\mathcal{I}}) p(\mathbf{a} | \mathbf{d}_{exp}, \mathcal{I}) \\ &= \sum_{v, v'} \int dV_a \Phi_v(\mathbf{a}) \Phi_{v'}(\mathbf{a}) \langle C_{vx} C_{v'x'} \rangle_a p(\mathbf{a} | \mathbf{d}_{exp}, \mathcal{I}) \\ &= \sum_{v, v'} \int dV_a \Phi_v(\mathbf{a}) \Phi_{v'}(\mathbf{a}) \left(\langle C_{vx} \rangle_a \langle C_{v'x'} \rangle_a + \langle \Delta C_{vx} \Delta C_{v'x'} \rangle_a \right) p(\mathbf{a} | \mathbf{d}_{exp}, \mathcal{I}) . \end{aligned} \tag{15}$$

The covariance then follows from

$$\langle \Delta z^{(x)} \Delta z^{(x')} \rangle = \langle z^{(x)} z^{(x')} \rangle - \langle z^{(x)} \rangle \langle z^{(x')} \rangle . \tag{16}$$

If we neglected the uncertainty of the surrogate, i.e.,

$$\begin{aligned} p(\mathbf{C} | D_{sim}) &= \delta(\mathbf{C} - \hat{\mathbf{C}}) , \\ \hat{\mathbf{C}} &= \langle \mathbf{C} \rangle_a , \end{aligned}$$

then we retain the widely known special case of ‘perfectly trustworthy’ surrogates

$$\langle z^{(x)} z^{(x')} \rangle = \sum_{v, v'} \int dV_a \Phi_v(\mathbf{a}) \Phi_{v'}(\mathbf{a}) \langle C_{vx} \rangle_a \langle C_{v'x'} \rangle_a p(\mathbf{a} | \mathbf{d}_{exp}, \mathcal{I}) .$$

Thus, the first part in the integral of Equation (15) is the uncertainty of the observable due to experimental uncertainties and given the surrogate model, while the second term adds the uncertainty of the surrogate itself. The term $\langle \Delta C_{vx} \Delta C_{v'x'} \rangle_a$ is commonly neglected, but easily computed. This result also suggests a natural measure for the trustworthiness of the surrogate model, which is directly linked to the specific experiment:

$$\frac{\sum_{v, v'} \int dV_a \Phi_v(\mathbf{a}) \Phi_{v'}(\mathbf{a}) \langle \Delta C_{vx} \Delta C_{v'x'} \rangle_a p(\mathbf{a} | \mathbf{d}_{exp}, \mathcal{I})}{\sum_{v, v'} \int dV_a \Phi_v(\mathbf{a}) \Phi_{v'}(\mathbf{a}) \langle C_{vx} \rangle_a \langle C_{v'x'} \rangle_a p(\mathbf{a} | \mathbf{d}_{exp}, \mathcal{I})} < \epsilon . \tag{17}$$

If the surrogate uncertainties are, on average, smaller than the experimental uncertainties by a few orders of magnitude, e.g., $\epsilon = 10^{-3}$, then they may be neglected. However, ϵ is the user’s choice. Note that this result does not spare the user to solve the foregoing surrogate model selection problem by e.g., computing evidence. This work only demonstrates

how surrogate uncertainties can be incorporated and a practical rule when they could be neglected, given that the surrogate model has already been selected before.

3. Numerical Example

Here, we demonstrate an application where surrogate uncertainties were in part negligible and in part non-negligible. We apply our method to a computational fluid dynamics simulation of aortic hemodynamics, i.e., blood flow in an aorta resembled by the simplified geometry of an upside down umbrella stick. The simulation depends on a non-Newtonian viscosity model with four parameters $\mathbf{a} = \{a_1, a_2, a_3, a_4\}$. The model was accompanied by viscosity measurements of human blood samples, thus determining $p(\mathbf{a} \mid \mathbf{d}_{\text{exp}}, \mathcal{I})$. This posterior turned out to have a complex landscape that cannot be reasonably approximated by standard distributions. Particularly, strong correlations and multi-modality were observed, i.e., $p(\mathbf{a} \mid \mathbf{d}_{\text{exp}}, \mathcal{I}) \neq \prod_i p(a_i \mid \mathbf{d}_{\text{exp}}, \mathcal{I})$. This means that vanilla Polynomial Chaos Expansions could not be applied without an undesirable transformation to conditionally independent variables. The posterior is described in detail in [20]. Based on $p(\mathbf{a} \mid \mathbf{d}_{\text{exp}}, \mathcal{I})$, $N_s = 100$ parameter samples \mathbf{a}_s were chosen and the simulation evaluated accordingly. The output, \mathbf{Z}_s , was the absolute values of the wall shear stress that the blood flow exerts on the aortic wall, for $N_x = 10$ measurement probes at different locations, each for $N_t = 101$ time-instances equidistantly spaced over one cardiac cycle (ca. 1 s). Further details on the simulation are not relevant here but are documented [20]. A simulation time on the order of 150 CPU hours per sample suggested to use a surrogate for the inference. For the surrogate's basis functions, $\Phi_v(\mathbf{a})$, we found multi-variate Legendre polynomials of up to order two sufficient. The numerical integrals were computed with Riemannian quadrature and convergence checked with successive grid refinement; however, stochastic integration would work just as well. A sketch example on how to implement this procedure computationally efficient via vectorisation in parameter space can be found at https://github.com/Sranf/BayesianSurrogate_sketch.git (1 January 2021).

In Figure 1, we compare the simulation uncertainty (including surrogate uncertainty) as computed with our Bayesian approach (Equation (15)) to the naive estimate for the simulation uncertainty (without surrogate uncertainty, i.e., neglecting $\langle \Delta C_{vx} \Delta C_{v'x'} \rangle_{\mathbf{a}}$ in Equation (15)). The surrogate uncertainties in the first half (left hand side) are relatively small, comprising only a few percent of the total uncertainty, and could possibly be neglected. In the second half (right-hand side), however, the surrogate uncertainties make up to $\sim 50\%$ of the total uncertainty. This demonstrates that simulation uncertainties inferred via surrogate models can be severely underestimated if the surrogate uncertainties are neglected, and subsequently lead to overconfidence in the simulation model. In practice, one would acquire more data in order to reduce the surrogate uncertainties, e.g., more data at later time-instances in Figure 1 are particularly promising. This was limited here not only by the computational budget, but also the impracticality in that dynamic simulations require the full evaluation of all previous time instances where the surrogate is already reasonably accurate. Thus, the procedure of instead explicitly including the surrogate uncertainties here also has proven to be practical. A similar situation is to be expected for most transient simulations, as uncertainties will usually increase as time progresses.

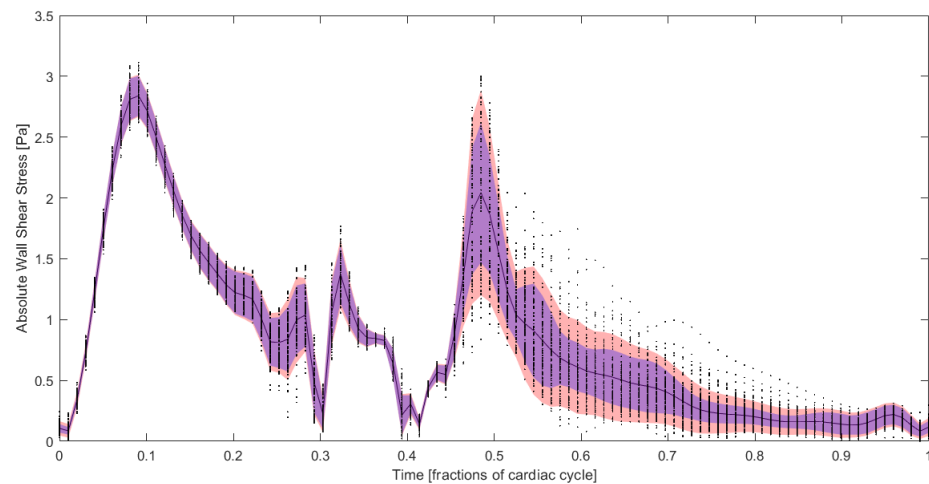


Figure 1. Simulation data (black dots) and simulation uncertainty (1σ) according to our Bayesian approach (red, including surrogate uncertainty) as well as the naive simulation uncertainty (blue, neglecting surrogate uncertainty). The black line is the surrogate mean.

4. Discussion

In this work, we have assumed a Gaussian likelihood for the simulation data, with unknown variance, for a surrogate that is linear in its parameters. Surrogates that are nonlinear in its parameters (e.g., neural networks) may promise higher capacity, however at the expense of losing analytical tractability of the surrogate uncertainty entirely. Other likelihood functions might be useful if further information is available, such as bounds on the observable (Gamma- or Beta-likelihood).

The result is a simple formula to incorporate surrogate uncertainties in the simulation uncertainties. This formula will be particularly useful if ‘convergence’ in the sense of finding the coefficients of e.g., a Polynomial Chaos Expansion is doubtful or not achievable due to a limit to the computational budget. The formula immediately suggests an intrinsic measure for the trustworthiness of the surrogate, distinct from commonly used ad hoc diagnostics. This measure is not to be confused with the evidence and should not be used for model selection because it would not preclude over-fitting, etc. It is merely a measure for the trustworthiness of the already selected surrogate.

Let us now explore the connections of this work to Polynomial Chaos Expansions (PCE) and Gaussian Process Regression (GPR). PCE is a special case of our generalized linear surrogate model, in that the basis functions of the surrogate are chosen such that

$$\int \Phi_v(\mathbf{a})\Phi_{v'}(\mathbf{a})p(\mathbf{a} | \mathbf{d}_{\text{exp}}, \mathcal{I})dV_{\mathbf{a}} := \delta_{v,v'} . \tag{18}$$

The double sum in Equation (15) then contracts to a single sum, and the diagonal of the term for the surrogate uncertainty, $\langle \Delta C_{vx} \Delta C_{v'x'} \rangle_{\mathbf{a}}$, survives. This is expected, in that PCE is defined such that the basis functions are uncorrelated, but still the expansion coefficients must be uncertain to a finite degree, and this must carry over to the simulation uncertainty. A severe limitation of PCE is that it is rather difficult to find basis function sets $\{\Phi_v\}$ that fulfill Equation (18), depending on $p(\mathbf{a} | \mathbf{d}_{\text{exp}}, \mathcal{I})$. For most practical purposes, one demands (i) conditional independence of the simulation parameters, i.e., $p(\mathbf{a} | \mathbf{d}_{\text{exp}}, \mathcal{I}) = \prod_i p(a_i | \mathbf{d}_{\text{exp}}, \mathcal{I})$, as well as (ii) simple standard distributions for $p(a_i | \mathbf{d}_{\text{exp}}, \mathcal{I})$, in order to find a solution (usually a tensor-product) to Equation (18). Known albeit tedious work-arounds are for (i) variable transformations and numerical orthonormalisation [18] and for (ii) PCE-constructions for arbitrary pdfs [17]. Note that also [17] demands (i) conditional independence of the simulation parameters. Note that our approach is not afflicted by above considerations, unless Equation (18) is specifically demanded. In the numerical example above, neither (i) nor (ii) were applicable. Finding a variable transformation

in order to fulfill (i) or numerical construction of orthonormal basis functions can be difficult, and particularly inconvenient if sophisticated priors $p(\mathbf{a} \mid \mathcal{I})$ are being used, e.g., Jeffreys’ generalized prior. An interesting alternative would be to model the input dependencies with vine copulas [21] in order to overcome the limitations of PCE addressed here. Unfortunately, no obvious vine copula was found for the example presented here.

Gaussian Process Regression would correspond to a change in the prior for $z^{(x)}$ in Equation (6) as follows:

$$p(z^{(x)} \mid \mathbf{C}, \mathbf{a}, \theta, \tilde{\mathcal{I}}) = \mathcal{N}\left(g(\mathbf{a} \mid \mathbf{c}^{(x)}) \mid K(\theta)\right). \tag{19}$$

where \mathcal{N} denotes a normal distribution, and K is the prior’s covariance matrix and defined by the parametrized covariance function k , $[K]_{ij} = k(\mathbf{a}^{(i)}, \mathbf{a}^{(j)} \mid \theta)$. This in turn would change $(\mathbf{Z}_s - M_s \mathbf{C})^T (\mathbf{Z}_s - M_s \mathbf{C}) \rightarrow (\mathbf{Z}_s - M_s \mathbf{C})^T K^{-1} (\mathbf{Z}_s - M_s \mathbf{C})$ in Equation (9). By again completing the square and following the same procedure, the corresponding results for mean Equation (12a), variance Equation (12b), and evidence Equation (13) are then retained by a simple substitution of

$$\begin{aligned} H_s &\rightarrow \tilde{H}_s, & \tilde{H}_s &= M_s^T K_s^{-1} M_s, \\ \chi_{min}^2 &\rightarrow \tilde{\chi}_{min}^2, & \tilde{\chi}_{min}^2 &= \text{tr}\left\{\mathbf{Z}_s^T (K_s^{-1} - K_s^{-1} M_s \tilde{H}_s M_s^T K_s^{-1}) \mathbf{Z}_s\right\}, \end{aligned}$$

where $[K_s]_{ij} = k(\mathbf{a}_s^{(i)}, \mathbf{a}_s^{(j)} \mid \theta)$ is the likelihood’s covariance matrix evaluated at A_s for the data set \mathbf{Z}_s at given θ . Equations (14) and (15) would preserve their form with the substitution

$$\begin{aligned} \Phi_V(\mathbf{a}) \langle \mathbf{C} \rangle_{\mathbf{a}} &\rightarrow \Phi_V(\mathbf{a}) \langle \mathbf{C} \rangle_{\mathbf{a}, \theta} + K_*^T K_s^{-1} M_s \langle \mathbf{C} \rangle_{\mathbf{a}, \theta}, \\ \langle \Delta C_{vx} \Delta C_{v'x'} \rangle_{\mathbf{a}} &\rightarrow \langle \Delta C_{vx} \Delta C_{v'x'} \rangle_{\mathbf{a}, \theta} \left(K - K_*^T K_s^{-1} K_* \right), \end{aligned}$$

where the subscript θ acknowledges that the right-hand side now depends on θ , and $[K_*]_{ij} = k(\mathbf{a}^{(i)}, \mathbf{a}_s^{(j)} \mid \theta)$ is the covariance between the training set A_s and the ‘test set’, i.e., the integration variable \mathbf{a} . Note that the additionally introduced hyperparameters θ would require the choice of a prior for θ and marginalization wrt θ in Equation (7), and subsequently also Equations (12)–(15).

We now discuss the implications of $\tilde{\mathcal{I}}$ in contrast to the original background information \mathcal{I} . \mathcal{I} contains, most importantly, that the observable z is uniquely determined by the simulation for a given set of input parameters \mathbf{a} . A prerequisite here was that the simulation is converged. For example, for finite element simulations, this would be a given mesh-converged spatial discretization. The proposition $\tilde{\mathcal{I}}$ additionally assumes Equations (3) and (6), so that it can be used to replace Equation (2) by Equation (4). Formally, this means, to get from Equations (2)–(4), we replace $p(z \mid \mathbf{a}, \mathcal{I}) \rightarrow p(z \mid \mathbf{a}, \mathbf{c}, \tilde{\mathcal{I}})$, where $p(z \mid \mathbf{a}, \mathcal{I}) = \delta(z - z(\mathbf{a}))$, $p(z \mid \mathbf{a}, \mathbf{c}, \tilde{\mathcal{I}}) = \delta(z - z_{sur}(\mathbf{a})) = \delta(z - g(\mathbf{a} \mid \mathbf{c}))$. For example, $\tilde{\mathcal{I}}$ contains in comparison to \mathcal{I} the additional assumption that we can use the value for z as predicted/approximated by the surrogate model. It also means that we introduce additional, artificial, and usually unknown regression parameters \mathbf{c} that need to be marginalized over. The additional uncertainty introduced by this approximation (i.e., the surrogate assumption) is encoded in $p(\mathbf{c} \mid \mathbf{D}_{sim}, \tilde{\mathcal{I}})$, and is correctly incorporated into the simulation observable uncertainties in Equation (15). What is important here is that $p(z \mid \mathbf{d}_{exp}, \mathbf{D}_{sim}, \tilde{\mathcal{I}}) \neq p(z \mid \mathbf{d}_{exp}, \mathbf{D}_{sim}, \mathcal{I})$ in general (the latter is computationally infeasible), but $p(z \mid \mathbf{d}_{exp}, \mathbf{D}_{sim}, \tilde{\mathcal{I}}) \approx p(z \mid \mathbf{d}_{exp}, \mathbf{D}_{sim}, \mathcal{I})$ if Equation (3) holds and $p(\mathbf{C} \mid \mathbf{D}_{sim}) \approx \delta(\mathbf{C} - \hat{\mathbf{C}})$, i.e., if the surrogate is indeed a good approximation and the posterior for the surrogate parameters is sharply peaked at $\hat{\mathbf{C}}$. Very often, this posterior is not sharply peaked; then, we can just gather more data until it is, or, if that is not possible, we

can at least avoid overconfidence induced by neglecting these uncertainties. A numerical example of where this is the case has been demonstrated above.

We have modeled spatial correlations by introducing a location index x , and assumed that the expansion coefficients at different sites, $\mathbf{c}^{(x)}$ and $\mathbf{c}^{(x')}$, are conditionally independent. This assumption is reasonable, in that the expansion coefficients are arbitrary mathematical constructs and no physically motivated model for their correlation is known. The spatial correlation, however, is retained in $z^{(x)}$, as was originally intended. More general models for spatial correlations can easily be implemented by substitution of $\delta_{xx'}$ with a spatial covariance matrix in Equation (12b). Note that this would require an additional marginalization wrt the (typically nonlinear) hyperparameters of the spatial covariance matrix. By introducing a compound index $\tilde{x} = (x, t)$ and substituting $x \rightarrow \tilde{x}$, we find a simple generalization to spatio-temporal correlations. This is equivalent to re-ordering spatial and temporal indices into a single sequence. While this procedure is convenient and requires only minor changes in the numerical implementation, it implicitly assumes conditional independence of spatial and temporal correlations. Analogous to above, general temporal correlations can be modeled by a substitution of $\delta_{\tilde{x}\tilde{x}'}$ in Equation (12b) with a temporal covariance matrix, again requiring an additional marginalization wrt the latter's hyperparameters.

5. Conclusions

We presented a Bayesian analysis of surrogate models and its associated uncertainty propagation problem in the context of uncertainty quantification of computer simulations. The assumptions were a generalized linear surrogate model (linear in its parameters, not the variable) and a Gaussian likelihood with unknown variance. Additionally, spatial and temporal correlations have been discussed. The result suggests a measure of trustworthiness of the surrogate by quantifying the ratio of the surrogate uncertainty to the total uncertainty, in contrast to commonly used heuristic diagnostics. The main result, however, is a rather simple rule to include surrogate uncertainties in the sought-for uncertainties of the simulation output. This is useful particularly for problems where the surrogate's trustworthiness is doubtful and cannot be improved. The connections to Polynomial Chaos Expansions and Gaussian Process Regression have been discussed. A numerical example demonstrated that simulation uncertainties can be significantly underestimated if surrogate uncertainties are neglected.

Funding: This work was funded by Graz University of Technology (TUG) through the LEAD Project "Mechanics, Modeling, and Simulation of Aortic Dissection" (biomechaorta.tugraz.at) and supported by GCCE: Graz Center of Computational Engineering.

Data Availability Statement: All information is contained in the manuscript. Code sketches are available at https://github.com/Sranf/BayesianSurrogate_sketch.git (1 January 2021).

Acknowledgments: The authors are grateful for useful comments from Ali Mohammad-Djafari.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Mathematical Proofs

Here, we want to determine norm, mean, and covariance of the marginalized Gaussian (Student-t distribution) in Equation (11), which is

$$p(\mathbf{C} | \mathbf{Z}_s, \mathbf{A}_s) = \frac{1}{Z} (\chi^2)^{-\frac{N_{sx}}{2}},$$

$$\chi^2 = \text{tr} \left\{ (\mathbf{Z}_s - \mathbf{M}_s \mathbf{C})^T (\mathbf{Z}_s - \mathbf{M}_s \mathbf{C}) \right\}. \quad (\text{A1})$$

In order to perform the integration, we first complete the square to get a quadratic form in C , which can then be integrated analytically, i.e., we bring the misfit χ^2 into a form that elucidates the C -dependence

$$\begin{aligned} \chi^2 &= \chi_{\min}^2 + \text{tr}\left\{ (C - \hat{C})^T H_s (C - \hat{C}) \right\}, & H_s &= M_s^T M_s, \\ \chi_{\min}^2 &= \text{tr}\left\{ Z_s^T (\mathbf{1} - M_s H_s^{-1} M_s^T) Z_s \right\}, & \hat{C} &= H_s^{-1} M_s^T Z_s. \end{aligned}$$

Now, the first moment is easily obtained. Along with the variable transformation under the integral

$$C \rightarrow \hat{C} + X \tag{A2}$$

we obtain

$$\begin{aligned} \langle C \rangle &= \frac{1}{Z} \int dV_C C \left(\text{tr}\left\{ (C - \hat{C})^T H_s (C - \hat{C}) \right\} + \chi_{\min}^2 \right)^{-\frac{N_{sx}}{2}} \\ &= \hat{C} + \underbrace{\frac{1}{Z} \int dV_X X \left(X^T H_s X + \chi_{\min}^2 \right)^{-\frac{N_{sx}}{2}}}_{=0}. \end{aligned} \tag{A3}$$

where we have used the symmetry properties of the likelihood. Next, we transform the expression for normalization based on Equation (A2)

$$Z_{N_{sx}} = \int dV_X \left(\text{tr}\left\{ X^T H_s X \right\} + \chi_{\min}^2 \right)^{-\frac{N_{sx}}{2}}.$$

Now, we combine and reorder the double indices (v, x) into a single index l , which turns the matrix X of dimension $N_p \times N_x$ into a vector x of dimension $N_{px} = N_p \cdot N_x$ and the matrix H of dimension $N_p \times N_p$ into a new block matrix H of dimension $N_{px} \times N_{sx}$ such that

$$[H]_{ll'} = [H_s]_{vv'} \delta_{xx'}. \tag{A4}$$

In this representation, we have

$$Z_{N_{sx}} = \int dV_x \left(x^T H x + \chi_{\min}^2 \right)^{-\frac{N_{sx}}{2}} = |H|^{-\frac{1}{2}} \int dV_y \left(y^T y + \chi_{\min}^2 \right)^{-\frac{N_{sx}}{2}}, \tag{A5}$$

where we substituted $x \rightarrow H^{-\frac{1}{2}} y$. Next, we introduce hyper-spherical coordinates, which leads to

$$Z_{N_{sx}} = \Omega_{N_{px}} |H|^{-\frac{1}{2}} \int_0^\infty \frac{d\rho}{\rho} \rho^{N_{px}} (\rho^2 + \chi_{\min}^2)^{-\frac{N_{sx}}{2}},$$

where $\Omega_{N_{px}}$ is the solid angle in N_{px} dimensions. Finally, based on the substitution $\rho = t \cdot \sqrt{\chi_{\min}^2}$, we recover an identity of the Beta-function, and we obtain

$$Z_{N_{sx}} = \Omega_{N_{px}} |H|^{-\frac{1}{2}} (\chi_{\min}^2)^{-\frac{N_{sx}-N_{px}}{2}} \frac{\Gamma(\frac{N_{px}}{2}) \Gamma(\frac{N_{sx}-N_{px}}{2})}{\Gamma(\frac{N_{sx}}{2})}. \tag{A6}$$

This result is valid only for $N_{sx} > N_{px}$, which is fulfilled in the present application. For future use, we rewrite this as

$$Z_{N_{sx}} = Z_{(N_{sx}-2)} \cdot (\chi_{\min}^2)^{-1} \cdot \frac{N_{sx} - N_{px} - 2}{N_{sx} - 2}. \tag{A7}$$

Finally, we calculate the covariance, based also on the compound index $l = (v, x)$, and by using the variable transformation in Equation (A2):

$$\begin{aligned} \langle \Delta C_l \Delta C_{l'} \rangle &= \frac{1}{Z_{N_{sx}}} \int dV_x x_l x_{l'} \left(\mathbf{x}^T \mathbf{H} \mathbf{x} + \chi_{\min}^2 \right)^{-\frac{N_{sx}}{2}}, \\ &= -\frac{2}{N_{sx} - 2} \cdot \frac{1}{Z_{N_{sx}}} \cdot \frac{\partial}{\partial \mathbf{H}_{ll'}} \int dV_x \left(\mathbf{x}^T \mathbf{H} \mathbf{x} + \chi_{\min}^2 \right)^{-\frac{N_{sx}}{2} + 1}, \\ &= -\frac{2}{N_{sx} - 2} \frac{\chi_{\min}^2 \cdot (N_{sx} - 2)}{Z_{(N_{sx}-2)} \cdot (N_{sx} - N_{px} - 2)} \frac{\partial}{\partial \mathbf{H}_{ll'}} Z_{(N_{sx}-2)}, \\ &= -\frac{2\chi_{\min}^2}{(N_{sx} - N_{px} - 2)} \frac{\partial}{\partial \mathbf{H}_{ll'}} \ln(Z_{(N_{sx}-2)}), \\ &= -\frac{2\chi_{\min}^2}{(N_{sx} - N_{px} - 2)} \underbrace{\frac{\partial}{\partial \mathbf{H}_{ll'}} \ln(|\mathbf{H}|^{-\frac{1}{2}})}_{=-\frac{1}{2}(\mathbf{H}^{-1})_{ll'}}. \end{aligned} \tag{A8}$$

In the last step, we have used that \mathbf{H} is a symmetric matrix. This is a very reasonable result because, if the variance Δ^2 in the Gaussian in Equation (10) would be known, then the covariance is $\Delta^2 \mathbf{H}^{-1}$. Consequently, the prefactor represents the Bayesian estimate for the variance Δ^2 based on the data. Now, we go back to the original meaning of the compound index Equation (A4), i.e., $(\mathbf{H}^{-1})_{ll'} \rightarrow (H_s^{-1})_{vv'} \delta_{xx'}$, and obtain the final result.

Appendix B. The Transformation Invariant Prior for the Surrogate Coefficients

Bayesian probability theory allows for rigorously and consistently incorporating any prior knowledge we have about the experiment before taking a look at the data. This knowledge shall be elicited here. Our inference must not depend on the exact parametrization. e.g., if we re-parametrize the surrogate, re-label, or re-order the surrogate parameters, the surrogate still should describe the same simulation. This is reasonable because the surrogate is a purely mathematical, auxiliary construct. This rescaling-invariance is ensured by Jeffreys' generalized prior and is given by the Riemann metric R (or the determinant of the Fisher information matrix) [16]

$$p(\mathbf{C}) = \frac{1}{Z} |\det(R)|^{1/2} \quad \text{with} \quad R_{ij} = \int p(\mathbf{Z}_s | \mathbf{C}) \frac{\partial^2}{\partial C_i \partial C_j} \ln(p(\mathbf{Z}_s | \mathbf{C})) dV_{Z_s}. \tag{A9}$$

with multi-indices $i, j = (v, x)$. With the likelihood and the generalized surrogate model defined in the manuscript, the result is

$$\begin{aligned} R_{ij} &\propto \sum_{k=1}^{N_s} \frac{\partial g(\mathbf{a}_s^{(k)} | \mathbf{C})}{\partial C_i} \frac{\partial g(\mathbf{a}_s^{(k)} | \mathbf{C})}{\partial C_j} \\ &= \sum_{k=1}^{N_s} \Phi_i(\mathbf{a}_s^{(k)}) \Phi_j(\mathbf{a}_s^{(k)}) \\ &= \text{const.} \end{aligned} \tag{A10}$$

This prior is independent of \mathbf{C} , i.e., a constant.

References

1. Xiu, D.; Karniadakis, G.E. The Wiener-Askey polynomial chaos for stochastic differential equations. *SIAM J. Sci. Comput.* **2005**, *27*, 1118–1139. [[CrossRef](#)]
2. O'Hagan, A. Polynomial Chaos: A Tutorial and Critique from a Statistician's Perspective. 2013. Available online: <http://tonyohagan.co.uk/academic/pdf/Polynomial-chaos.pdf> (accessed on 25 June 2019).
3. Crestaux, T.; Le Maître, O.P.; Martinez, J.-M. Polynomial chaos expansion for sensitivity analysis. *Reliab. Eng. Syst. Saf.* **2009**, *94*, 1161–1172. [[CrossRef](#)]
4. O'Hagan, A. Curve Fitting and Optimal Design for Prediction. *J. R. Stat. Soc. Ser. B* **1978**, *40*, 1–42. [[CrossRef](#)]
5. Rasmussen, C.E.; Williams, C.K. *Gaussian Processes for Machine Learning*; The MIT Press: Cambridge, MA, USA, 2006.
6. Sraj, I.; Le Maître, O.P.; Knio, O.M.; Hoteit, I. Coordinate transformation and Polynomial Chaos for the Bayesian inference of a Gaussian process with parametrized prior covariance function. *Comput. Methods Appl. Mech. Eng.* **2016**, *298*, 205–228. [[CrossRef](#)]
7. O'Hagan, A.; Kennedy, M.C.; Oakley, J.E. Uncertainty analysis and other inference tools for complex computer codes. *Bayesian Stat.* **1999**, *6*, 503–524.
8. Kennedy, M.C.; O'Hagan, A. Predicting the output from a complex computer code when fast approximations are available. *Biometrika* **2000**, *87*, 1–13. [[CrossRef](#)]
9. Arnst, M.; Ghanem, R.G.; Soize, C. Identification of Bayesian posteriors for coefficients of chaos expansions. *J. Comput. Phys.* **2010**, *229*, 3134–3154. [[CrossRef](#)]
10. Madankan, R.; Singla, P.; Singh, T.; Scott, P.D. Polynomial-chaos-based Bayesian approach for state and parameter estimations. *J. Guid. Control Dyn.* **2013**, *36*, 1058–1074. [[CrossRef](#)]
11. Karagiannis, G.; Lin, G. Selection of polynomial chaos bases via Bayesian model uncertainty methods with applications to sparse approximation of PDEs with stochastic inputs. *J. Comput. Phys.* **2014**, *259*, 114–134. [[CrossRef](#)]
12. Lu, F.; Morzfeld, M.; Tu, X.; Chorin, A.J. Limitations of polynomial chaos expansions in the Bayesian solution of inverse problems. *J. Comput. Phys.* **2015**, *282*, 138–147. [[CrossRef](#)]
13. Hwai, M.; Tan, Y. Sequential Bayesian Polynomial Chaos Model Selection for Estimation of Sensitivity Indices. *SIAM/ASA J. Uncertain. Quantif.* **2015**, *3*, 146–168. [[CrossRef](#)]
14. Ghanem, R.G.; Owhadi, H.; Higdon, D. *Handbook of Uncertainty Quantification*; Springer: New York, NY, USA, 2017; [[CrossRef](#)]
15. O'Hagan, A. Bayesian analysis of computer code outputs: A tutorial. *Reliab. Eng. Syst. Saf.* **2006**, *91*, 1290–1300. [[CrossRef](#)]
16. von der Linden, W.; Dose, V.; von Toussaint, U. *Bayesian Probability Theory: Applications in the Physical Sciences*, 1st ed.; Cambridge University Press: Cambridge, UK, 2014; [[CrossRef](#)]
17. Oladyshkin, S.; Nowak, W. Data-driven uncertainty quantification using the arbitrary polynomial chaos expansion. *Reliab. Eng. Syst. Saf.* **2012**, *106*, 179–190. [[CrossRef](#)]
18. Jakeman, J.D.; Franzelin, F.; Narayan, A.; Eldred, M.; Plfüger, D. Polynomial chaos expansions for dependent random variables. *Comput. Methods Appl. Mech. Eng.* **2019**, *351*, 643–666. [[CrossRef](#)]
19. von der Linden, W.; Preuss, R.; Hanke, W. Consistent Application of Maximum Entropy to Quantum-Monte-Carlo Data. *J. Physics: Condens. Matter* **1996**, *8*, 1–13. [[CrossRef](#)]
20. Ranftl, S.; Müller, T.; Windberger, U.; von der Linden, W.; Brenn, G. *Data and Codes for 'A Bayesian Approach to Blood Rheological Uncertainties in Aortic Hemodynamcis'*; Zenodo Digital Repository: Genève, Switzerland, 2021; [[CrossRef](#)]
21. Torre, E.; Marelli, S.; Embrechts, P.; Sudret, B. A general framework for data-driven uncertainty quantification under complex input dependencies using vine copulas. *Probabilistic Eng. Mech.* **2019**, *55*, 1–16. [[CrossRef](#)]